Pacific Journal of Mathematics

Volume 280 No. 1 January 2016

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

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

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

The subscription price for 2016 is US $440/year for the electronic version, and $600/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers 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. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.1

STABLE CAPILLARY HYPERSURFACES IN A WEDGE

JAIGYOUNG CHOE AND MIYUKI KOISO

Let Σ be a compact immersed stable capillary hypersurface in a wedge bounded by two hyperplanes in $\mathbb{R}^{n+1}$. Suppose that Σ meets those two hyperplanes in constant contact angles $\geq \pi/2$ and is disjoint from the edge of the wedge, and suppose that $\partial\Sigma$ consists of two smooth components with one in each hyperplane of the wedge. It is proved that if $\partial\Sigma$ is embedded for $n = 2$, or if each component of $\partial\Sigma$ is convex for $n \geq 3$, then Σ is part of the sphere. The same is true for Σ in the half-space of $\mathbb{R}^{n+1}$ with connected boundary $\partial\Sigma$.

1. Introduction

The isoperimetric inequality says that among all domains of fixed volume in the $(n+1)$-dimensional Euclidean space $\mathbb{R}^{n+1}$ the one with least boundary area is the round ball. What happens if the boundary area is a critical value instead of the minimum? For this question the more general domains enclosed by the immersed hypersurfaces have to be considered, hence one needs to introduce the oriented volume (as defined in (1)). Then the answer to the question is that given a compact immersed hypersurface Σ in $\mathbb{R}^{n+1}$, its area is critical among all variations of Σ preserving the oriented volume enclosed by Σ if and only if Σ has constant mean curvature (CMC).

So, H. Hopf [1989, p. 131] raised the question as to whether there exist closed surfaces with CMC which are not spheres. To this question, W.-Y. Hsiang [1982] obtained a counterexample, a CMC immersion of $\mathbb{S}^3$ in $\mathbb{R}^4$ which is not round, and Wente [1986] constructed a CMC immersion of a torus in $\mathbb{R}^3$.

Is there an extra condition on a CMC surface Σ which guarantees that Σ is a sphere? There are some affirmative results in this regard:

Choe supported in part by NRF, 2011-0030044, SRC-GAIA. Koiso supported in part by Grant-in-Aid for Scientific Research (B) No. 25287012 and Grant-in-Aid for Challenging Exploratory Research No. 26610016 of the Japan Society for the Promotion of Science, and the Kyushu University Interdisciplinary Programs in Education and Projects in Research Development.

MSC2010: primary 49Q10; secondary 53A10.

Keywords: capillary surface, constant mean curvature, stable.

(i) Aleksandrov [1962a; 1962b] showed that every compact *embedded* hypersurface of CMC in $\mathbb{R}^{n+1}$ is a sphere,

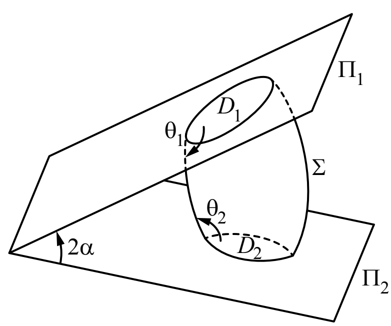
(ii) Hopf himself [1989] proved that an immersed CMC 2-*sphere* is round, and

(iii) Barbosa and do Carmo [1984] showed that the only compact immersed *stable* CMC hypersurface of $\mathbb{R}^{n+1}$ is the sphere.

A CMC hypersurface Σ is said to be stable if the second variation of the n-dimensional area of Σ is nonnegative for all $(n+1)$-dimensional volume-preserving perturbations of Σ.

A CMC surface with nonempty boundary along which it makes a constant contact angle with a prescribed supporting surface is called a capillary surface. It is an equilibrium surface of the sum of the area and the wetting energy on the supporting surface (we call it the total energy of the surface) for volume-preserving variations (see Section 2). Such a surface is said to be stable if the second variation of the total energy is nonnegative for all volume-preserving variations. In this paper, we prove the following uniqueness result (Section 4, Theorem 1) which is a generalization of the theorem by Barbosa and do Carmo [1984] mentioned above:

> *Let Σ be a compact immersed stable capillary hypersurface in a wedge bounded by two hyperplanes in $\mathbb{R}^{n+1}$, $n \geq 2$. Suppose that Σ meets those two hyperplanes in constant contact angles $\geq \pi/2$ and does not hit the edge of the wedge. We also assume that $\partial\Sigma$ consists of two smooth embedded $(n-1)$-dimensional manifolds, one in each hyperplane of the wedge, and that each component of $\partial\Sigma$ is convex when $n \geq 3$* (see figure). *Then Σ is part of the sphere. Also, the same conclusion holds if Σ is in the half-space of $\mathbb{R}^{n+1}$ and $\partial\Sigma$ is connected.*



We emphasize that there is a stable capillary surface between two parallel planes which is not part of the sphere [Vogel 1989]. Our result shows that, if the initial supporting surface is the union of two parallel planes and we consider a stable nonspherical capillary surface, then the configuration changes discontinuously on infinitesimal tilting of one of the planes. Such discontinuity was pointed out already in [Concus et al. 2001] without the stability of the surface.

The idea of our proof is motivated by Wente [1991]. He simplified Barbosa and do Carmo's proof by using the parallel hypersurfaces and the homothetic contraction. We have found that Wente's method carries over nicely to our capillary hypersurfaces in a wedge and in the half-space. On the other hand, the Minkowski inequality for $\partial\Sigma$ is indispensable in our arguments. Wente informed us that recently Marinov [2012] obtained the same result when Σ is in $\mathbb{R}^3$ and $\partial\Sigma$ is in a plane.

Here we mention some additional related results. McCuan [1997] and Park [2005] proved that an embedded annular capillary surface in a wedge in $\mathbb{R}^3$ is necessarily part of the sphere. The question then arises whether one can extend the theorems of Aleksandrov, Hopf, and Barbosa–do Carmo to the case of capillary surfaces in a wedge or in the half-space. That is:

(i) Does there exist no compact embedded capillary surface of genus ≥ 1 in a wedge (or in the half-space) of $\mathbb{R}^3$?

(ii) Is there a compact immersed annular capillary surface of genus 0 (or higher) in a wedge (or in the half-space) which is not part of the sphere?

(iii) Which hypothesis of McCuan's and Park's can be dropped or generalized if the capillary surface is stable?

As mentioned above, in this paper we give an answer to (iii). To question (i), McCuan [1997] gave an affirmative answer with the contact angle condition $\theta_i \leq \pi/2$. In relation to question (ii), Wente [1995] constructed noncompact capillary surfaces bifurcating from the cylinder in a wedge.

Finally, it should be mentioned that the stable capillary surfaces in a ball also have been studied very actively. To begin with, Nitsche [1985] showed that a capillary disk in a ball $\subset \mathbb{R}^3$ is a spherical cap (for a simpler proof, see [Finn and McCuan 2000, Appendix]). Ros and Souam [1997] proved that a stable capillary surface of genus 0 in a ball in $\mathbb{R}^3$ is a spherical cap. They also proved that a stable minimal surface with constant contact angle in a ball $\subset \mathbb{R}^3$ is a flat disk or a surface of genus 1 with at most three boundary components. Moreover, Ros and Vergasta [1995] showed that a stable minimal hypersurface in a ball $B \subset \mathbb{R}^n$ which is orthogonal to ∂B is totally geodesic, and that a stable capillary surface in a ball $\subset \mathbb{R}^3$ and orthogonal to ∂B is a spherical cap or a surface of genus 1 with at most two boundary components.

2. Preliminaries

Let Π_1 and Π_2 be two hyperplanes in $\mathbb{R}^{n+1}$ containing the $(n-1)$-plane $\{x_n = 0, x_{n+1} = 0\}$ and making angles α and $-\alpha$ (with $0 < \alpha < \pi/2$) with the horizontal hyperplane $\{x_{n+1} = 0\}$, respectively. Let $\Omega \subset \{x_n > 0\}$ be the wedge-shaped domain bounded by Π_1 and Π_2. We denote by $\overline{\Omega}$ the closure of Ω. Denote by $X : (\Sigma, \partial\Sigma) \to (\overline{\Omega}, \partial\Omega)$ an immersion of an n-dimensional oriented compact connected C^∞ manifold Σ with nonempty boundary into Ω such that $X(\Sigma^\circ) \subset \Omega$ and $X(\partial\Sigma) \subset \partial\Omega$, where $\Sigma^\circ := \Sigma - \partial\Sigma$. The $(n-1)$-plane

$$\Pi_0 := \Pi_1 \cap \Pi_2 = \{x_n = 0,\, x_{n+1} = 0\}$$

is called the edge of the wedge Ω. In this paper we are concerned only with the immersed surfaces $X(\Sigma)$ which connect Π_1 to Π_2 without intersecting Π_0.

For the immersion $X : (\Sigma, \partial\Sigma) \to (\overline{\Omega}, \partial\Omega)$, the n-dimensional area $\mathcal{H}^n(X)$ is written as

$$\mathcal{H}^n(X) = \int_\Sigma dS,$$

where dS is the volume form of Σ induced by X. The $(n+1)$-dimensional *oriented volume* $V(X)$ enclosed by $X(\Sigma)$ is defined by

$$V(X) = \frac{1}{n+1}\int_\Sigma \langle X, \nu\rangle\, dS, \tag{1}$$

where the Gauss map ν is the unit normal vector field along X with orientation determined as follows. Let $\{e_1, \dots, e_n\}$ be an oriented frame on the tangent space $T_p(\Sigma)$, $p \in \Sigma$. Then $\{dX_p(e_1), \dots, dX_p(e_n), \nu\}$ is a frame of $\mathbb{R}^{n+1}$ with positive orientation.

In this paper $X(\Sigma)$ is immersed while $X(\partial\Sigma)$ is assumed to be embedded. $X(\partial\Sigma)$ influences the area $\mathcal{H}^n(X)$ through the *wetting energy*. Set $C_i = X(\partial\Sigma) \cap \Pi_i$ and let $D_i \subset \Pi_i$ be the domain bounded by C_i. The wetting energy $\mathcal{W}(X)$ of X is defined by

$$\mathcal{W}(X) = \omega_1\mathcal{H}^n(D_1) + \omega_2\mathcal{H}^n(D_2),$$

where ω_i is a constant with $|\omega_i| < 1$ and $\mathcal{H}^n(D_i)$ is the n-dimensional area of D_i. Then we define the *total energy* $E(X)$ of the immersion X by

$$E(X) = \mathcal{H}^n(X) + \mathcal{W}(X).$$

Note that $\Sigma \cup D_1 \cup D_2$ is a piecewise smooth hypersurface without boundary. We can extend $\nu : \Sigma \to S^n$ to the Gauss map $\nu : \Sigma \cup D_1 \cup D_2 \to S^n$. Since the origin of $\mathbb{R}^{n+1}$ is on the edge Π_0 of Ω, $\langle X, \nu\rangle = 0$ on $D_1 \cup D_2$. Hence the oriented volume

$$\widehat{V}(X) = \frac{1}{n+1}\int_{\Sigma\cup D_1\cup D_2} \langle X, \nu\rangle\, dS \tag{2}$$

coincides with $V(X)$.

Let $X_t : (\Sigma, \partial\Sigma) \to (\overline{\Omega}, \partial\Omega)$ be a 1-parameter family of immersions with $X_0 = X$. It is well known [Finn 1986, Chapter 1] that a necessary and sufficient condition for X to be a critical point of the total energy for all variations X_t for which the volume

$\widehat{V}(X_t)$ is constant is that the immersed surface have constant mean curvature H and that the contact angle θ_i of $X(\Sigma)$ with Π_i (measured between $X(\Sigma)$ and D_i) be constant along C_i (see figure on page 2). More precisely,

$$\cos\theta_i = -\omega_i \quad \text{on } C_i.$$

The hypersurface $X(\Sigma)$ of constant mean curvature with constant contact angle along C_i will be called a *capillary* hypersurface. A capillary hypersurface is said to be stable if the second variation of $E(X_t)$ at $t=0$ is nonnegative for all volume-preserving perturbations $X_t : (\Sigma, \partial\Sigma) \to (\overline{\Omega}, \partial\Omega)$ of $X(\Sigma)$.

A capillary hypersurface $X(\Sigma)$ in $\overline{\Omega}$ has a nice property called the *balancing formula* [Choe 2002; Concus et al. 2001; Korevaar et al. 1989]:

Lemma 1. *We have*

$$nH\mathcal{H}^n(D_i) = -(\sin\theta_i)\mathcal{H}^{n-1}(C_i), \quad i = 1, 2. \tag{3}$$

Proof. We first remark the following fact. Let $\hat{\Sigma}$ be an m-dimensional oriented compact connected C^∞ manifold, and $Y : \hat{\Sigma} \to \mathbb{R}^{m+1}$ a continuous map which is a piecewise C^∞ immersion. Also let $\hat{\nu}$ be the Gauss map of Y. Then, by using the divergence theorem, we obtain

$$\int_{\hat{\Sigma}} \hat{\nu}\, dS = 0.$$

Now integrate

$$\Delta_\Sigma X = nH\nu$$

on Σ to get

$$\sum_{i=1}^{2} \int_{C_i} \eta\, ds = nH \int_\Sigma \nu\, d\Sigma,$$

where η is the outward-pointing unit conormal to $\partial\Sigma$ on X. Then, use the above remark to obtain

$$\sum_{i=1}^{2} \int_{C_i} \eta\, ds = -nH \sum_{i=1}^{2} \int_{D_i} \nu\, dS. \tag{4}$$

Denote by N_i the unit normal to Π_i that points outward from Ω. Denote by n_i the inward pointing unit normal to C_i in Π_i. Set

$$\epsilon_i := \begin{cases} 1 & \text{if } \nu = N_i \text{ on } D_i, \\ -1 & \text{if } \nu = -N_i \text{ on } D_i. \end{cases} \tag{5}$$

Then from (4) we obtain

$$\sum_{i=1}^{2}\int_{C_i}\big((\sin\theta_i)\epsilon_i N_i-(\cos\theta_i)n_i\big)\,ds+\sum_{i=1}^{2}nH\mathcal{H}^n(D_i)\epsilon_i N_i=0,$$

that is, for the $(n-1)$-dimensional area $\mathcal{H}^{n-1}(C_i)$,

$$\sum_{i=1}^{2}(\sin\theta_i)\epsilon_i\mathcal{H}^{n-1}(C_i)N_i-\sum_{i=1}^{2}(\cos\theta_i)\int_{C_i}n_i\,ds+\sum_{i=1}^{2}nH\mathcal{H}^n(D_i)\epsilon_i N_i=0.$$

Using the above remark again, we obtain

$$\sum_{i=1}^{2}\big(nH\mathcal{H}^n(D_i)+(\sin\theta_i)\mathcal{H}^{n-1}(C_i)\big)N_i=0.$$

Since N_1 and N_2 are linearly independent, we obtain the formula (3). □

Another tool that will be essential in this paper is the formula for the volume of tubes due to H. Weyl [1939]. Given an immersion X of a compact oriented n-manifold M into $\mathbb{R}^{n+1}$, let $X_t=X+t\nu$ be the one-parameter family of parallel hypersurfaces to X. Thanks to the parallelness of X_t one can easily see that X_t has the same unit normal vector field as X and that the area $\mathcal{H}^n(X_t)$ is a polynomial of degree n in t. Namely, if $k_1,\dots,k_n$ are the principal curvatures of X, then

$$\begin{aligned}\mathcal{H}^n(X_t)&=\int_M\prod_{i=1}^{n}(1-k_i t)\,dS\\&=a_0+a_1t+a_2t^2+\cdots+a_nt^n,\\a_0&=\mathcal{H}^n(X_0),\\a_1&=-\int_M nH\,dS,\\a_2&=\int_M\sum_{i<j}k_ik_j\,dS,\\a_\ell&=(-1)^\ell\int_M\sum_{i_1<\cdots<i_\ell}k_{i_1}k_{i_2}\cdots k_{i_\ell}\,dS.\end{aligned}\tag{6}$$

Moreover, the oriented volume $V(X_t)$ satisfies

$$\frac{d}{dt}V(X_t)=\mathcal{H}^n(X_t).$$

Hence

$$V(X_t)=v_0+v_1t+v_2t^2+\cdots+v_{n+1}t^{n+1},$$
$$v_1=a_0,\quad 2v_2=a_1,\quad\dots.$$

3. Admissible variations

Here we assume that our capillary hypersurface $X : (\Sigma, \partial\Sigma) \to (\bar{\Omega}, \partial\Omega)$ has a nonempty boundary component on each Π_i, $i = 1, 2$. But the case when Σ is in the half-space and $\partial\Sigma$ is connected can be treated similarly.

To check the stability of X one needs to deal with its volume-preserving variations $X_t : (\Sigma, \partial\Sigma) \to (\bar{\Omega}, \partial\Omega)$. The specific variation that we use arises from the parallel hypersurfaces

$$X_t^1 = X + t\nu.$$

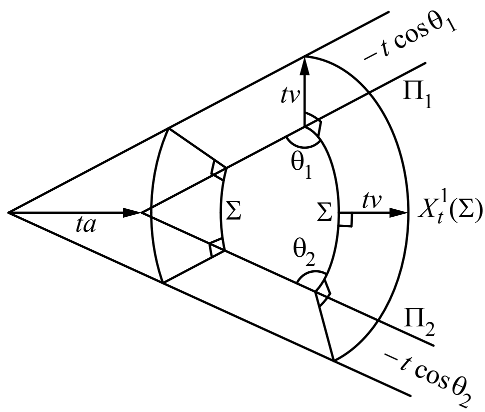
But X_t^1 does not satisfy the boundary condition $X_t^1(\partial\Sigma) \subset \partial\Omega$ unless $\theta_i = \pi/2$. To move the boundary to a desired place in $\partial\Omega$, we apply a translation

$$X_t^2(p) = p + ta$$

for some $a \in \mathbb{R}^{n+1}$. The vector a is determined in such a way that

$$X_t^2 \circ X_t^1(\partial\Sigma) \subset \partial\Omega.$$

Clearly such a vector uniquely exists as can be seen in the figure.



However, $X_t^2 \circ X_t^1$ is not volume-preserving. One way of making it into a volume-preserving variation is to deform it by a homothetic contraction

$$X_t := s(t) X_t^2 \circ X_t^1, \tag{7}$$

where $s(t)$ satisfies

$$\widehat{V}(X_t) = \widehat{V}(X_0) = v_0. \tag{8}$$

In order to compute $\widehat{V}(X_t)$ we first must consider the oriented volume $\widehat{V}(X_t^2 \circ X_t^1)$ enclosed by $X_t^2 \circ X_t^1(\Sigma) \cup D_1^t \cup D_2^t$, where $D_i^t \subset \Pi_i$ is the domain bounded by $\Pi_i \cap X_t^2 \circ X_t^1(\partial\Sigma)$. Note here that since $X_t^2 \circ X_t^1(\Sigma) \cup D_1^t \cup D_2^t$ is closed, the oriented volume $\widehat{V}(X_t^2 \circ X_t^1)$ as computed by (2) is independent of the translation X_t^2. While

t increases by Δt, the oriented volume $\widehat{V}(X_t^2 \circ X_t^1)$ increases by $\mathcal{H}^n(X_t^2 \circ X_t^1)\Delta t$ on $X_t^2 \circ X_t^1(\Sigma)$ and by $-\cos\theta_i \mathcal{H}^n(D_i^t)\Delta t$ on D_i^t. Hence

$$\frac{d}{dt}\widehat{V}(X_t^2 \circ X_t^1) = \mathcal{H}^n(X_t^2 \circ X_t^1) - \sum_i \cos\theta_i \mathcal{H}^n(D_i^t). \tag{9}$$

Calling $-\sum_i \cos\theta_i \mathcal{H}^n(D_i^t)$ the wetting energy $\mathcal{W}(X_t^2 \circ X_t^1)$ of $X_t^2 \circ X_t^1(\Sigma)$, we define the total energy by

$$E(X_t^2 \circ X_t^1) = \mathcal{H}^n(X_t^2 \circ X_t^1) + \mathcal{W}(X_t^2 \circ X_t^1).$$

The tube formula (6) for the capillary hypersurface Σ yields

$$\mathcal{H}^n(X_t^2 \circ X_t^1) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n,$$

$$a_0 = \mathcal{H}^n(\Sigma), \quad a_1 = -nHa_0, \quad a_2 = \int_\Sigma \sum_{i<j} k_i k_j \, dS,$$

$$\frac{d}{dt}\widehat{V}(X_t^2 \circ X_t^1) = E(X_t^2 \circ X_t^1). \tag{10}$$

Recall $C_i = X(\partial\Sigma) \cap \Pi_i$. Since $X_t^2 \circ X_t^1(\Sigma)$ has constant contact angle with $\partial\Omega$ for all t, $X_t^2 \circ X_t^1(C_i)$ are the parallel hypersurfaces of $p_{\Pi_i}(X_t^2(C_i))$, where p_{Π_i} denotes the projection of $\mathbb{R}^{n+1}$ onto Π_i. Also recall $\partial D_i = C_i$, $D_i = D_i^0$. The distance between $X_t^2 \circ X_t^1(C_i)$ and $p_{\Pi_i}(X_t^2(C_i))$ is $t\sin\theta_i$. Hence again by the tube formula for $\mathcal{H}^{n-1}(X_t^2 \circ X_t^1(C_i))$, we obtain

$$\begin{aligned}\mathcal{H}^n(D_i^t) = \mathcal{H}^n(D_i) + \mathcal{H}^{n-1}(C_i)\, t\sin\theta_i - \frac{1}{2}\left(\int_{C_i}(n-1)\overline{H}\, d\overline{S}\right)t^2\sin^2\theta_i \\ + \cdots + (-1)^{n-1}\frac{1}{n}\left(\int_{C_i}\bar{k}_1\bar{k}_2\cdots\bar{k}_{n-1}\, d\overline{S}\right)t^n\sin^n\theta_i,\end{aligned}$$

where $\overline{H}$ and $\bar{k}_i$ are, respectively, the mean curvature and the principal curvature of C_i in Π_i with respect to the outward unit normal, and $d\overline{S}$ is the $(n-1)$-dimensional volume form of C_i.

Then (9) gives

$$\begin{aligned}\frac{d}{dt}\widehat{V}(X_t^2 \circ X_t^1) = a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i) - \left(nHa_0 + \sum_i \cos\theta_i \sin\theta_i\, \mathcal{H}^{n-1}(C_i)\right)t \\ + \left(\int_\Sigma \sum_{i<j} k_i k_j \, dS + \frac{1}{2}\sum_i \cos\theta_i \sin^2\theta_i \int_{C_i}(n-1)\overline{H}\, d\overline{S}\right)t^2 + \cdots.\end{aligned}$$

Hence if we write

$$E(X_t^2 \circ X_t^1) = e_0 + e_1 t + \cdots + e_n t^n,$$

then (10) yields

$$
\begin{aligned}
e_0 &= a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i), \\
e_1 &= -nHa_0 - \sum_i \cos\theta_i \sin\theta_i \mathcal{H}^{n-1}(C_i), \qquad (11) \\
e_2 &= \int_\Sigma \sum_{i<j} k_i k_j \, dS + \frac{1}{2} \sum_i \cos\theta_i \sin^2\theta_i \int_{C_i} (n-1)\overline{H}\, d\overline{S}.
\end{aligned}
$$

On the other hand, if we let

$$\widehat{V}(X_t^2 \circ X_t^1) = v_0 + v_1 t + v_2 t^2 + \cdots + v_{n+1} t^{n+1},$$

then it follows from (7), (8), and the binomial series that

$$
\begin{aligned}
s(t)^n &= v_0^{n/(n+1)} (v_0 + v_1 t + v_2 t^2 + \cdots + v_{n+1} t^{n+1})^{-n/(n+1)} \\
&= 1 - \frac{n}{n+1}\Big(\frac{v_1}{v_0}\Big) t + \bigg(\frac{n(2n+1)}{2(n+1)^2}\Big(\frac{v_1}{v_0}\Big)^2 - \frac{n}{n+1}\Big(\frac{v_2}{v_0}\Big)\bigg) t^2 + \cdots .
\end{aligned}
$$

Thus

$$
\begin{aligned}
E(X_t) &= s(t)^n E(X_t^2 \circ X_t^1(\Sigma)) \qquad (12) \\
&= e_0 + \bigg(e_1 - \frac{n}{n+1}\Big(\frac{v_1}{v_0}\Big) e_0\bigg) t \\
&\quad + \bigg(e_2 - \frac{n}{n+1}\Big(\frac{v_1}{v_0}\Big) e_1 + \frac{n(2n+1)}{2(n+1)^2}\Big(\frac{v_1}{v_0}\Big)^2 e_0 - \frac{n}{n+1}\Big(\frac{v_2}{v_0}\Big) e_0\bigg) t^2 \\
&\quad + \cdots .
\end{aligned}
$$

From (10) we have

$$v_1 = e_0, \quad 2v_2 = e_1, \qquad (13)$$

and the fact that $E'(0) = 0$ in (12) implies

$$v_0 = \frac{n}{n+1} \frac{e_0^2}{e_1}. \qquad (14)$$

Substituting the identities of (13) and (14) into the coefficient of t^2 in (12) yields

$$E''(0)/2 = \frac{1}{2ne_0}\big(2ne_0 e_2 - (n-1)e_1^2\big).$$

Hence from (11) we get

$$
\begin{aligned}
ne_0 E''(0) = 2n\Big(a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big) \\
\times \Big(\int_\Sigma \sum_{i<j} k_i k_j \, dS + \frac{1}{2}\sum_i \cos\theta_i \sin^2\theta_i \int_{C_i} (n-1)\overline{H}\, d\overline{S}\Big) \\
-(n-1)\Big(nHa_0 + \sum_i \cos\theta_i \sin\theta_i \, \mathcal{H}^{n-1}(C_i)\Big)^2.
\end{aligned}
$$

Then the balancing formula (3) yields

$$
\Big(nHa_0 + \sum_i \cos\theta_i \sin\theta_i \, \mathcal{H}^{n-1}(C_i)\Big)^2 = n^2H^2\Big(a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big)^2.
$$

Therefore,

$$
\begin{aligned}
ne_0 E''(0) &= \Big(a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big) \\
&\quad \times \Big(2n\int_\Sigma \sum_{i<j} k_i k_j \, dS + n\sum_i \cos\theta_i \sin^2\theta_i \int_{C_i}(n-1)\overline{H}\, d\overline{S} \\
&\quad - \int_\Sigma n^2(n-1)H^2\, dS + n^2(n-1)H^2 \sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big) \\
&= \Big(a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big) \\
&\quad \times \Big(-\int_\Sigma \sum_{i<j}(k_i-k_j)^2\, dS + n\sum_i \cos\theta_i \sin^2\theta_i \int_{C_i}(n-1)\overline{H}\, d\overline{S} \\
&\quad + n^2(n-1)H^2\sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big).
\end{aligned}
$$

Applying the balancing formula (3) again, this gives

$$
\begin{aligned}
ne_0 E''(0) = \Big(a_0 - \sum_i \cos\theta_i \mathcal{H}^n(D_i)\Big)\Big(-\int_\Sigma \sum_{i<j}(k_i-k_j)^2\, dS \\
+ (n-1)\sum_i \cos\theta_i \sin^2\theta_i \Big(n\int_{C_i}\overline{H}\, d\overline{S} + \frac{\mathcal{H}^{n-1}(C_i)^2}{\mathcal{H}^n(D_i)}\Big)\Big).
\end{aligned}
\tag{15}
$$

We shall see in the next section that

$$
n\int_{\partial D_i}\overline{H}\, d\overline{S} + \frac{\mathcal{H}^{n-1}(\partial D_i)^2}{\mathcal{H}^n(D_i)} \geq 0.
$$

4. Theorem

We are now ready to state the theorem of this paper.

Theorem 1. *Let W be a wedge in $\mathbb{R}^{n+1}$ bounded by two hyperplanes Π_1 and Π_2. Let $\Sigma \subset W$ be a compact oriented immersed hypersurface that is disjoint from the edge $\Pi_1 \cap \Pi_2$ of W, having smooth embedded boundary $\partial\Sigma \subset \Pi_1 \cup \Pi_2$, and satisfying $\partial\Sigma \cap \Pi_i = \partial D_i$ for a nonempty bounded connected domain D_i in Π_i. Suppose that Σ is a stable capillary hypersurface in W. In other words, Σ is an immersed constant mean curvature hypersurface making a constant contact angle $\theta_i \geq \pi/2$ with D_i such that for all volume-preserving perturbations (for the oriented volume enclosed by $\Sigma \cup D_1 \cup D_2$), the second variation of the total energy*

$$E(\Sigma) = \mathcal{H}^n(\Sigma) - \cos\theta_1 \mathcal{H}^n(D_1) - \cos\theta_2 \mathcal{H}^n(D_2)$$

is nonnegative.

(i) *If $n = 2$, then Σ is part of the 2-sphere.*

(ii) *If $n \geq 3$ and D_1 and D_2 are convex, then Σ is part of the n-sphere.*

Conversely, if Σ is part of the n-sphere, then it is stable.

Moreover, the same conclusion holds when Σ is in the half-space of $\mathbb{R}^{n+1}$ and $\partial\Sigma$ is connected.

Proof. We prove the theorem for Σ in a wedge, and the proof for Σ in the half-space is similar.

When $n = 2$, (15) becomes

$$2e_0 E''(0) = \Big(a_0 - \sum_i \cos\theta_i \mathcal{H}^2(D_i)\Big)\Big(-\int_\Sigma (k_1 - k_2)^2\, dS + \sum_i \cos\theta_i \sin^2\theta_i \Big(2\int_{\partial D_i} k\, ds + \frac{\mathcal{H}^1(\partial D_i)^2}{\mathcal{H}^2(D_i)}\Big)\Big),$$

where k is the geodesic curvature of ∂D_i with respect to the outward unit normal along ∂D_i. Note that on the smooth Jordan curve ∂D_i, $\int_{\partial D_i} k\, ds = -2\pi$. Hence the isoperimetric inequality of D_i and the angle condition $\cos\theta_i \leq 0$ yield

$$E''(0) \leq 0.$$

Therefore Σ needs to be umbilic everywhere if it is stable.

When $n \geq 3$, Minkowski showed that for a convex domain $D \subset \mathbb{R}^n$ with mean curvature H on ∂D,

$$n\int_{\partial D} |H|\, dS \leq \frac{\mathcal{H}^{n-1}(\partial D)^2}{\mathcal{H}^n(D)}$$

[Osserman 1978, p. 1191]. Hence it follows from (15) that the stable Σ is all umbilic.

If Σ is part of the n-sphere, then Σ is the minimizer of the energy E among all embedded hypersurfaces in Ω enclosing the same volume [Zia et al. 1988]. The proof is similar to that of Theorem 4.1 in [Koiso and Palmer 2007]; the method is essentially the same as in [Winterbottom 1967]. Hence Σ is stable for all $n \geq 2$. □

Remark 1. Our contact angle condition $\theta_i \geq \pi/2$ is quite natural because McCuan [1997] proved the nonexistence of embedded capillary surfaces with $\theta_i \leq \pi/2$ in a wedge of $\mathbb{R}^3$. Also it had been experimentally observed that a wedge forces the liquid drops (bridges) with $\theta_i \leq \pi/2$ to move toward its edge.

5. Minkowski's inequality

The Minkowski inequality is not well known among geometers and its proof is not easily available in the literature. So in this section we sketch a proof of it. First we need to introduce the mixed volume [Schneider 1993].

The *Minkowski sum* of two sets A and B in $\mathbb{R}^n$ is the set

$$A + B = \{a + b \in \mathbb{R}^n : a \in A, b \in B\}.$$

Given convex bodies $K_1, \dots, K_r$ in $\mathbb{R}^n$, the volume of the Minkowski sum $\lambda_1 K_1 + \cdots + \lambda_r K_r$ (for $\lambda_i \geq 0$) of the scaled convex bodies $\lambda_i K_i$ of K_i is a homogeneous polynomial of degree n given by

$$\mathcal{H}^n(\lambda_1 K_1 + \cdots + \lambda_r K_r) = \sum_{j_1,\dots,j_n=1}^{r} V(K_{j_1}, \dots, K_{j_n}) \lambda_{j_1} \cdots \lambda_{j_n}.$$

$V(K_{j_1}, \dots, K_{j_n})$ is called the *mixed volume* of $K_{j_1}, \dots, K_{j_n}$. The mixed volume is uniquely determined by the following three properties:

(i) $V(K, \dots, K) = \mathcal{H}^n(K)$, (ii) V is symmetric, (iii) V is multilinear.

A remarkable property of the mixed volume is the *Aleksandrov–Fenchel inequality*:

$$V(K_1, K_2, K_3, \dots, K_n)^2 \geq V(K_1, K_1, K_3, \dots, K_n) \cdot V(K_2, K_2, K_3, \dots, K_n).$$

For a convex body $K \subset \mathbb{R}^n$ and a unit ball $B \subset \mathbb{R}^n$, the mixed volume

$$W_j(K) := V(\overbrace{K, K, \dots, K}^{n-j \text{ times}}, \overbrace{B, B, \dots, B}^{j \text{ times}})$$

is called the j-th *quermassintegral* of K. The Steiner formula says that the quermassintegrals of K determine the volume of the parallel bodies of K:

$$\mathcal{H}^n(K + tB) = \sum_{j=0}^{n} \binom{n}{j} W_j(K) t^j.$$

Comparing the Steiner formula for a convex domain $D \subset \mathbb{R}^n$ with its tube formula, one can obtain

$$\begin{aligned} W_0(D) &= \mathcal{H}^n(D), \\ nW_1(D) &= \mathcal{H}^{n-1}(\partial D), \\ nW_2(D) &= \int_{\partial D} |H|\, dS, \\ n(n-1)(n-2)W_3(D) &= 2\int_{\partial D} \sum_{i<j} k_i k_j\, dS. \end{aligned}$$

The Aleksandrov–Fenchel inequality for the quermassintegrals yields

$$\begin{aligned} W_1(D)^2 &\geq W_0(D)W_2(D), \\ W_2(D)^2 &\geq W_1(D)W_3(D). \end{aligned}$$

Consequently,

$$n\int_{\partial D} |H|\, dS \leq \frac{\mathcal{H}^{n-1}(\partial D)^2}{\mathcal{H}^n(D)}, \tag{16}$$

$$\begin{aligned} \int_{\partial D} \sum_{i<j} k_i k_j\, dS &\leq \frac{(n-1)(n-2)}{2} \frac{\left(\int_{\partial D} |H|\, dS\right)^2}{\mathcal{H}^{n-1}(\partial D)} \\ &\leq \frac{(n-1)(n-2)}{2n^2} \frac{\mathcal{H}^{n-1}(\partial D)^3}{\mathcal{H}^n(D)^2}, \end{aligned} \tag{17}$$

where (16) is the desired Minkowski inequality.

Remark 2. We note that (16) is the isoperimetric inequality when D is a domain in $\mathbb{R}^2$, and so is (17) when $D \subset \mathbb{R}^3$, because

$$\int_{\partial D \subset \mathbb{R}^2} |k|\, ds = 2\pi \quad \text{and} \quad \int_{\partial D \subset \mathbb{R}^3} k_1 k_2\, dS = 4\pi.$$

Remark 3. Let $D_t \subset \mathbb{R}^n$ be the parallel domain with distance t to D. Then (16) is equivalent to

$$n\, \frac{\mathcal{H}^{n-1}(\partial D_t)'}{\mathcal{H}^{n-1}(\partial D_t)} \leq \frac{(n-1)\mathcal{H}^n(D_t)'}{\mathcal{H}^n(D_t)},$$

or equivalently,

$$\left(\frac{\mathcal{H}^{n-1}(\partial D_t)^n}{\mathcal{H}^n(D_t)^{n-1}}\right)' \leq 0.$$

Hence the isoperimetric quotient $\mathcal{H}^{n-1}(\partial D_t)^n / \mathcal{H}^n(D_t)^{n-1}$ decreases as t increases. Indeed, the parallel domain D_t becomes rounder and rounder as t increases.

Acknowledgement

We thank Professor Monika Ludwig for referring us to the Aleksandrov–Fenchel inequality.

References

[Aleksandrov 1962a] A. D. Aleksandrov, "Uniqueness theorems for surfaces in the large, I", *Amer. Math. Soc. Transl.* (2) **21** (1962), 341–354. MR 27 #698a Zbl 0122.39601

[Aleksandrov 1962b] A. D. Aleksandrov, "Uniqueness theorems for surfaces in the large, II", *Amer. Math. Soc. Transl.* (2) **21** (1962), 354–388. MR 27 #698b Zbl 0122.39601

[Barbosa and do Carmo 1984] J. L. Barbosa and M. do Carmo, "Stability of hypersurfaces with constant mean curvature", *Math. Z.* **185**:3 (1984), 339–353. MR 85k:58021c Zbl 0513.53002

[Choe 2002] J. Choe, "Sufficient conditions for constant mean curvature surfaces to be round", *Math. Ann.* **323**:1 (2002), 143–156. MR 2003f:53008 Zbl 1016.53007

[Concus et al. 2001] P. Concus, R. Finn, and J. McCuan, "Liquid bridges, edge blobs, and Scherk-type capillary surfaces", *Indiana Univ. Math. J.* **50**:1 (2001), 411–441. MR 2002g:76023 Zbl 0996.76014

[Finn 1986] R. Finn, *Equilibrium capillary surfaces*, Grundlehren der Mathematischen Wissenschaften **284**, Springer, New York, 1986. MR 88f:49001 Zbl 0583.35002

[Finn and McCuan 2000] R. Finn and J. McCuan, "Vertex theorems for capillary drops on support planes", *Math. Nachr.* **209** (2000), 115–135. MR 2000k:53058 Zbl 0962.76014

[Hopf 1989] H. Hopf, *Differential geometry in the large*, 2nd ed., Lecture Notes in Mathematics **1000**, Springer, Berlin, 1989. MR 90f:53001 Zbl 0669.53001

[Hsiang 1982] W.-y. Hsiang, "Generalized rotational hypersurfaces of constant mean curvature in the Euclidean spaces, I", *J. Differential Geom.* **17**:2 (1982), 337–356. MR 84h:53009 Zbl 0493.53043

[Koiso and Palmer 2007] M. Koiso and B. Palmer, "Anisotropic capillary surfaces with wetting energy", *Calc. Var. Partial Differential Equations* **29**:3 (2007), 295–345. MR 2008d:53007 Zbl 1136.76011

[Korevaar et al. 1989] N. J. Korevaar, R. Kusner, and B. Solomon, "The structure of complete embedded surfaces with constant mean curvature", *J. Differential Geom.* **30**:2 (1989), 465–503. MR 90g:53011 Zbl 0726.53007

[Marinov 2012] P. I. Marinov, "Stability of capillary surfaces with planar boundary in the absence of gravity", *Pacific J. Math.* **255**:1 (2012), 177–190. MR 2923699 Zbl 1242.49090

[McCuan 1997] J. McCuan, "Symmetry via spherical reflection and spanning drops in a wedge", *Pacific J. Math.* **180**:2 (1997), 291–323. MR 98m:53013 Zbl 0885.53009

[Nitsche 1985] J. C. C. Nitsche, "Stationary partitioning of convex bodies", *Arch. Rational Mech. Anal.* **89**:1 (1985), 1–19. MR 86j:53013 Zbl 0572.52005

[Osserman 1978] R. Osserman, "The isoperimetric inequality", *Bull. Amer. Math. Soc.* **84**:6 (1978), 1182–1238. MR 58 #18161 Zbl 0411.52006

[Park 2005] S.-h. Park, "Every ring type spanner in a wedge is spherical", *Math. Ann.* **332**:3 (2005), 475–482. MR 2006h:53008 Zbl 1102.53007

[Ros and Souam 1997] A. Ros and R. Souam, "On stability of capillary surfaces in a ball", *Pacific J. Math.* **178**:2 (1997), 345–361. MR 98c:58029 Zbl 0930.53007

[Ros and Vergasta 1995] A. Ros and E. Vergasta, "Stability for hypersurfaces of constant mean curvature with free boundary", *Geom. Dedicata* **56**:1 (1995), 19–33. MR 96h:53013 Zbl 0912.53009

[Schneider 1993] R. Schneider, *Convex bodies: The Brunn–Minkowski theory*, Encyclopedia of Mathematics and its Applications **44**, Cambridge Univ. Press, 1993. MR 94d:52007 Zbl 0798.52001

[Vogel 1989] T. I. Vogel, "Stability of a liquid drop trapped between two parallel planes, II: General contact angles", *SIAM J. Appl. Math.* **49**:4 (1989), 1009–1028. MR 90k:53013 Zbl 0691.53007

[Wente 1986] H. C. Wente, "Counterexample to a conjecture of H. Hopf", *Pacific J. Math.* **121**:1 (1986), 193–243. MR 87d:53013 Zbl 0586.53003

[Wente 1991] H. C. Wente, "A note on the stability theorem of J. L. Barbosa and M. Do Carmo for closed surfaces of constant mean curvature", *Pacific J. Math.* **147**:2 (1991), 375–379. MR 92g:53010 Zbl 0715.53041

[Wente 1995] H. C. Wente, "The capillary problem for an infinite trough", *Calc. Var. Partial Differential Equations* **3**:2 (1995), 155–192. MR 97f:53006 Zbl 0960.53011

[Weyl 1939] H. Weyl, "On the volume of tubes", *Amer. J. Math.* **61**:2 (1939), 461–472. MR 1507388 Zbl 0021.35503

[Winterbottom 1967] W. L. Winterbottom, "Equilibrium shape of a small particle in contact with a foreign substrate", *Acta Metal.* **15**:2 (1967), 303–310.

[Zia et al. 1988] R. K. P. Zia, J. E. Avron, and J. E. Taylor, "The summertop construction: Crystals in a corner", *J. Statist. Phys.* **50**:3-4 (1988), 727–736. MR 89g:82058 Zbl 1084.82582

Received June 4, 2014. Revised May 18, 2015.

JAIGYOUNG CHOE
SCHOOL OF MATHEMATICS
KOREA INSTITUTE FOR ADVANCED STUDY
207-43 CHEONGNYANGNI 2-DONG
DONGDAEMUN-GU
SEOUL 130-722
SOUTH KOREA
choe@kias.re.kr

MIYUKI KOISO
INSTITUTE OF MATHEMATICS FOR INDUSTRY
KYUSHU UNIVERSITY
744, MOTOOKA, NISHI-KU
FUKUOKA 819-0395
JAPAN
koiso@math.kyushu-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.17

THE CHERN–SIMONS INVARIANTS FOR THE DOUBLE OF A COMPRESSION BODY

DAVID L. DUNCAN

Given a 3-manifold that can be written as the double of a compression body, we compute the Chern–Simons critical values for arbitrary compact connected structure groups. We also show that the moduli space of flat connections is connected when there are no reducibles.

1. Introduction

Let G be a Lie group with Lie algebra $\mathfrak{g}$. Given a principal G-bundle $P \to Y$ over a closed, oriented 3-manifold Y, one can define the Chern–Simons function

$$\mathcal{CS} : \mathcal{A}(P) \to \mathbb{R}/\mathbb{Z},$$

where $\mathcal{A}(P)$ is the space of connections on P. The set of critical points of $\mathcal{CS}$ is the space of flat connections $\mathcal{A}_{\text{flat}}(P) \subset \mathcal{A}(P)$, and the critical values are topological invariants of Y. In general, computing the critical values of $\mathcal{CS}$ is fairly difficult. Nevertheless, various techniques have been developed to handle certain classes of 3-manifolds; for example, see [Kirk and Klassen 1993; Auckly 1994; Reznikov 1996; Nishi 1998; Neumann and Yang 1995; Dostoglou and Salamon 1994; Wehrheim 2006]. Most of these techniques are specific to the choice of Lie group G, common examples being $\mathrm{SU}(2)$, $\mathrm{Sp}(1)$ and $\mathrm{SL}_{\mathbb{C}}(2)$.

In the present paper we compute the Chern–Simons critical values for any 3-manifold Y that can be written as a *double*

$$Y = \overline{H} \cup_{\partial H} H,$$

MSC2010: primary 53C07; secondary 53C05.
Keywords: Chern–Simons invariants.

where H is a compression body, $\overline{H}$ is a copy of H with the opposite orientation, and the identity map on ∂H is used to glue $\overline{H}$ and H; see Figure 1. For us, the term *compression body* means that

- H is a compact, connected, oriented cobordism between surfaces Σ_-, Σ_+,
- H admits a Morse function $f\colon H \to [-1, 1]$ with critical points of index 0 or 1,
- all critical values of f are in the interior of $(-1, 1)$, and
- $f^{-1}(\pm 1) = \Sigma_\pm$.

It follows that, up to homotopy, H can be obtained from Σ_+ by attaching 2-handles. These topological assumptions imply that Σ_+ is connected; there is no bound on the number of components of Σ_-. (Note that not every 3-manifold can be realized as the double of a compression body; the Poincaré homology sphere is a simple counterexample.)

Throughout this paper we work with an arbitrary compact, connected Lie group G, and we assume the bundle P is obtained by doubling a bundle over H in the obvious way.

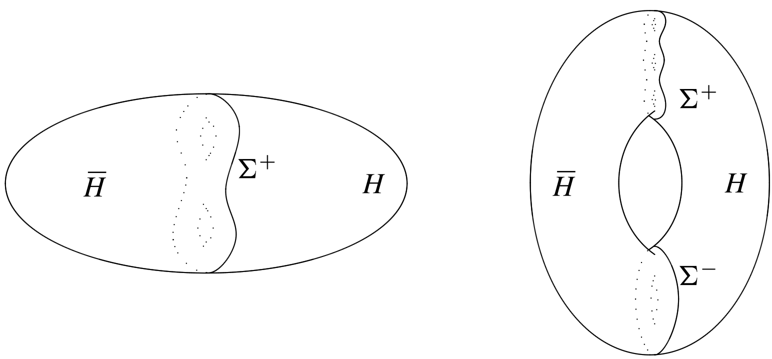


Figure 1. Pictured above are two possibilities for Y. The first has $\Sigma_- = \varnothing$, making H a handlebody. In the second figure, Σ_- is connected.

Before stating the main result, we mention that the definition of $\mathcal{CS}$ requires a choice of normalization. When G is simple this choice can be made in an essentially unique way. However, for arbitrary compact G the situation is not as simple. It turns out that, in general, this normalization can be fixed by choosing a faithful unitary representation $\rho_0 : G \to \mathrm{U}(W)$, where W is a finite-dimensional Hermitian vector space. One upshot of this approach is that certain computations reduce to the case where G is a classical group; see Remark 2.2. It is convenient to phrase the main result in terms of a lift $\mathcal{CS}_a : \mathcal{A}(P) \to \mathbb{R}$ of the Chern–Simons function $\mathcal{CS}$; this lift can be defined by fixing a flat reference connection $a \in \mathcal{A}_{\mathrm{flat}}(P)$. See Section 2B for more details.

Theorem 1.1. *Let G be a compact, connected Lie group. There is a positive integer N_G such that if H, P, Y, ρ_0 are as above, then all critical values of $\mathcal{CS}_a : \mathcal{A}(P) \to \mathbb{R}$ are integer multiples of $1/N_G$.*

The dependence of these critical values on the choice of ρ_0 is only up to an overall integer multiple. In particular, if the representation ρ_0 has image in $\mathrm{SU}(W) \subset \mathrm{U}(W)$, then all critical values are multiples of $2/N_G$. If ρ_0 is the complexification of a faithful orthogonal representation of G (see Remark 2.2), then all critical values are multiples of $4/N_G$.

Following Wehrheim [2006], the integer N_G appearing in Theorem 1.1 can be defined explicitly as follows. Consider the integer

$$n_G := \sup_{G' \leq G} \{|\pi_0(C(G'))|\},$$

where the supremum is over all subgroups of G, and $C(G')$ denotes the centralizer in G. Then n_G is finite since G is compact. We define N_G to be the least common multiple of $\{1, 2, \ldots, n_G\}$. Thus $N_G \geq 1$ is an integer depending only on G.

The definition of N_G can often be refined if one has certain knowledge about G or P. In particular, the proof will show that we can take $N_G = 1$ provided the following hypothesis holds.

Hypothesis 1. *For each connected component $S \subset \Sigma_-$, the identity component of the gauge group acts trivially on $\mathcal{A}_{\mathrm{flat}}(P|_S)$.*

For example, Hypothesis 1 holds trivially when Σ_- is empty. When Σ_- is nonempty, the hypothesis holds when $G = \mathrm{SO}(3)$ and the restriction of P to each component of Σ_- is nontrivial. More generally, this hypothesis is satisfied if $G = \mathrm{U}(r)$ or $\mathrm{PU}(r)$ and the integer $c_1(P)[S]$ is coprime to r for all connected components $S \subset \Sigma_-$; see [Wehrheim and Woodward 2009]. On the other hand, Hypothesis 1 is never satisfied if the bundle P is trivial, due to the trivial connection. That being said, it is perhaps worth mentioning that there are other hypotheses that allow one to replace N_G by 1. For example, an argument by Wehrheim [2006] can be used in our proof below to show that when $G = \mathrm{SU}(2)$, one can always replace $N_{\mathrm{SU}(2)}$ by 1 in the statement of Theorem 1.1. We also point out that Hypothesis 1 is not assumed in Theorem 1.1; our primary motivation for introducing this hypothesis is to simplify the discussion at various times.

Motivated by the techniques of [Dostoglou and Salamon 1994, page 633] and [Wehrheim 2006], our strategy for proving Theorem 1.1 is to show that all flat connections are gauge equivalent to a connection in a certain canonical form. As a consequence, Theorem 1.1 can be viewed as a statement about the connected components of $\mathcal{A}_{\mathrm{flat}}(P)$. For example, we arrive at the following corollary; see Remark 3.5.

Corollary 1.2. *Let $P \to Y$ be as in Theorem 1.1. Assume Hypothesis 1 is satisfied and either*

- $G = \mathrm{U}(r)$ *or* $\mathrm{SU}(r)$ *and* ρ_0 *is the standard representation, or*
- $G = \mathrm{PU}(r)$ *and* ρ_0 *is the adjoint representation.*

If $a, a' \in \mathcal{A}_{\mathrm{flat}}(P)$, then there is a gauge transformation u such that u^*a and a' lie in the same component of $\mathcal{A}_{\mathrm{flat}}(P)$. Moreover, two flat connections a, a' lie in the same component of $\mathcal{A}_{\mathrm{flat}}(P)$ if and only if $\mathcal{CS}(a) = \mathcal{CS}(a')$.

Our proof also identifies precisely when flat connections on P exist. To state this, consider the commutator subgroup $[G, G] \subseteq G$. Then the quotient $P/[G, G]$ is a torus bundle over Y. For example, if G is semisimple then $P/[G, G] = Y$, and if $G = \mathrm{U}(r)$ then this quotient is the determinant $U(1)$-bundle. The next result follows from the proof of Proposition 3.3 below.

Corollary 1.3. *Let $P \to Y$ be as in Theorem 1.1. The space $\mathcal{A}_{\mathrm{flat}}(P)$ of flat connections is nonempty if and only if (i) the restriction $P/[G, G]\,|_{\partial H}$ is the trivial bundle, and (ii) for any spherical component $S^2 \subseteq \partial H$, the restriction $P|_{S^2}$ is the trivial bundle.*

The author's primary interest in Theorem 1.1 is due to its implications for the instanton energy values on certain noncompact 4-manifolds; see [Duncan 2013b]. These 4-manifolds are those of the form $\mathbb{R} \times H^\infty$, where

$$H^\infty := H \cup_{\partial H} ([0, \infty) \times \partial H) \tag{1}$$

is obtained from a Riemannian 3-manifold H by attaching a cylindrical end on its boundary. Given a principal G-bundle $P \to H$, define $P^\infty \to H^\infty$ similarly. Then the "manifold at infinity" of $\mathbb{R} \times H^\infty$ is the double of H (see Section 3C).

Corollary 1.4. *Suppose G is a compact, connected Lie group and H is a compact, oriented 3-manifold with boundary. Let A be any finite-energy instanton on $\mathbb{R} \times P^\infty \to \mathbb{R} \times H^\infty$, with the instanton equation defined using the product metric. Then there is a flat connection $a_\flat$ on $\overline{H} \cup_{\partial H} H$ such that the energy of A is $\mathcal{CS}_a(a_\flat)$.*

Note that the assumptions on G and H are very general. Corollary 1.4 is proved in Section 3C using an extension of a standard argument; see [Taubes 1982; Dostoglou and Salamon 1994; Salamon 1995; Wehrheim 2006; 2005; Nishinou 2010]. See also [Yeung 1991; Etesi 2013] for similar results on instanton energies and characteristic numbers for noncompact manifolds.

2. Background

Given a vector bundle $E \to X$, we will write $\Omega^\bullet(X, E) := \oplus_k \Omega^k(X, E)$ for the space of differential forms on X with values in E. We use the wedge product given by $\mu \wedge \nu = \mu \otimes \nu - \nu \otimes \mu$ for real-valued 1-forms μ, ν.

Let G be a compact Lie group, and $\rho_0 : G \to \mathrm{U}(W)$ the faithful unitary representation from the introduction. Then define a bilinear form $\langle \cdot, \cdot \rangle$ on the Lie algebra $\mathfrak{g}$ by setting

$$\langle \mu, \nu \rangle := -\frac{1}{2\pi^2} \operatorname{Tr}((\rho_0)_* \mu \cdot (\rho_0)_* \nu), \quad \forall \mu, \nu \in \mathfrak{g}, \tag{2}$$

where the trace is the one on $\mathfrak{u}(W)$. (The normalizing factor $1/2\pi^2$ is chosen so that the quantities (4) and (6) below are integers. If ρ_0 has image in $\mathrm{SU}(W)$ then the more familiar $1/4\pi^2$ can be used.) Since we have assumed ρ_0 is faithful, it follows that $\langle \cdot, \cdot \rangle$ is nondegenerate, and so this defines an Ad-invariant inner product on $\mathfrak{g}$.

Suppose $\pi : P \to X$ is a principal G-bundle over a smooth n-manifold X; we assume G acts on P on the right. Given a right action $\rho : G \to \mathrm{Diff}(F)$ of G on a manifold F we will denote the associated bundle by $P \times_G F := (P \times F)/G$. If $F = V$ is a vector space and $G \to \mathrm{Diff}(V)$ has image in $\mathrm{GL}(V) \subset \mathrm{Diff}(V)$, then $P \times_G V$ is a vector bundle and we will write $P(V) := P \times_G V$. Pullback by π induces an injection

$$\pi^* : \Omega^\bullet(X, P(V)) \hookrightarrow \Omega^\bullet(P, P \times V)$$

with image the space of forms that are equivariant and horizontal.

We will write $P(\mathfrak{g})$ for the *adjoint bundle* associated to the adjoint representation $G \to \mathrm{GL}(\mathfrak{g})$. The Lie bracket $[\cdot, \cdot]$ on $\mathfrak{g}$ is Ad-invariant, and so this combines with the wedge to define a bilinear map $\mu \otimes \nu \mapsto [\mu \wedge \nu]$ on $\Omega^\bullet(X, P(\mathfrak{g}))$, endowing $\Omega^\bullet(X, P(\mathfrak{g}))$ with the structure of a graded algebra. Similarly, the Ad-invariance of the inner product $\langle \cdot, \cdot \rangle$ implies that it induces a fiberwise inner product on the vector bundle $P(\mathfrak{g})$. This combines with the wedge to give a graded bilinear map

$$\Omega^k(X, P(\mathfrak{g})) \otimes \Omega^l(X, P(\mathfrak{g})) \to \Omega^{k+l}(X), \quad \mu \otimes \nu \mapsto \langle \mu \wedge \nu \rangle.$$

2A. ***Gauge theory.*** We denote by $\mathcal{A}(P)$ the set of all connections on P. By definition, $\mathcal{A}(P)$ consists of the elements of $\Omega^1(P, P \times \mathfrak{g})$ that are both G-equivariant and vertical. It follows that $\mathcal{A}(P)$ is an affine space modeled on $\pi^*\Omega^1(X, P(\mathfrak{g})) \cong \Omega^1(X, P(\mathfrak{g}))$. We will write $\mathcal{A}^1(P)$ for the completion of $\mathcal{A}(P)$ with respect to the H^1-Sobolev norm; we will always assume $\mathcal{A}^1(P)$ is equipped with the H^1-topology. The space $\mathcal{A}^1(P)$ is well-defined when X is compact; when X is noncompact the H^1-norm depends on the choice of a smooth reference connection at infinity.

Given any representation $\rho : G \to \mathrm{GL}(V)$, each connection $A \in \mathcal{A}(P)$ determines a covariant derivative

$$d_{A,\rho} : \Omega^\bullet(X, P(V)) \to \Omega^{\bullet+1}(X, P(V)), \quad \mu \mapsto (\pi^*)^{-1}\big(d(\pi^*\mu) + \rho_*(A) \wedge \pi^*\mu\big),$$

where d is the trivial connection on $P \times V$. When considering the adjoint representation, we will write $d_A := d_{A,\mathrm{Ad}}$. The *curvature endomorphism* $\mathrm{curv}(d_{A,\rho}) \in$

$\Omega^2(X, \mathrm{End}(P(V)))$ is defined by the relation

$$d_{A,\rho} \circ d_{A,\rho}\mu = \mathrm{curv}(d_{A,\rho}) \wedge \mu$$

for all $\mu \in \Omega^\bullet(X, P(V))$. We define the *curvature (2-form)* of A by

$$F_A = (\pi^*)^{-1}\big(dA + \tfrac{1}{2}[A \wedge A]\big) \in \Omega^2(X, P(\mathfrak{g})).$$

The curvature 2-form F_A recovers the curvature endomorphism $\mathrm{curv}(d_{A,\rho})$ in any representation ρ in the sense that

$$\rho_* F_A = \mathrm{curv}(d_{A,\rho}). \tag{3}$$

Taking $\rho = \mathrm{Ad}$, we therefore have $\mathrm{curv}(d_A) \wedge \mu = [F_A \wedge \mu]$ for all $\mu \in \Omega^\bullet(X, P(\mathfrak{g}))$. Given any $A \in \mathcal{A}(P)$, the covariant derivative and curvature satisfy

$$d_{A+\mu} = d_A + [\mu \wedge \cdot\,], \qquad F_{A+\mu} = F_A + d_A\mu + \tfrac{1}{2}[\mu \wedge \mu],$$

for all $\mu \in \Omega^1(X, P(\mathfrak{g}))$. We also have the Bianchi identity $d_A F_A = 0$. A connection A is *flat* if $F_A = 0$, and we denote the set of all smooth (resp. H^1) flat connections on P by $\mathcal{A}_{\mathrm{flat}}(P)$ (resp. $\mathcal{A}^1_{\mathrm{flat}}(P)$).

Suppose X is a closed, oriented 4-manifold. Then associated to the fixed representation $\rho_0 : G \to \mathrm{U}(W)$ from the introduction, we obtain a complex vector bundle $P(W)$ equipped with a Hermitian inner product. In particular, this has well-defined Chern classes $c_i := c_i(P(W)) \in H^{2i}(X, \mathbb{Z})$. The usual Chern–Weil formula says

$$\kappa(P) = \kappa(P; \rho_0) := (c_1^2 - 2c_2)[X] = -\frac{1}{4\pi^2}\int_X \mathrm{Tr}(\mathrm{curv}(d_{A,\rho_0}) \wedge \mathrm{curv}(d_{A,\rho_0})) \in \mathbb{Z},$$

for any connection $A \in \mathcal{A}(P)$; the Bianchi identity shows this is independent of the choice of A. Here $\mathrm{Tr}(\mu \wedge \nu)$ is obtained by combining the wedge with the trace on $\mathfrak{u}(W)$. Then equations (2) and (3) show

$$\kappa(P) = \frac{1}{2}\int_X \langle F_A \wedge F_A \rangle. \tag{4}$$

Remark 2.1. This characteristic number can be equivalently defined as follows. Let $B\mathrm{U}(W)$ be the classifying space for the unitary group, and let $\kappa \in H^4(B\mathrm{U}(W), \mathbb{Z})$ be given by the square of the first Chern class minus two times the second Chern class. Then $\kappa(P) \in H^4(X, \mathbb{Z}) \cong \mathbb{Z}$ is obtained by pulling back κ under the map $X \to BG \to B\mathrm{U}(W)$; here the first arrow is the classifying map for P, and the second is induced by the representation $\rho_0 : G \to \mathrm{U}(W)$.

It follows immediately from the definition that $\kappa(P)$ is even if the mod-2 reduction of c_1 vanishes. Now suppose ρ_0 is obtained by complexifying a (real) orthogonal representation $G \to \mathrm{O}(V)$. Then $P(W) = P(V)_{\mathbb{C}}$ is the complexification of the real vector bundle $P(V)$ and so $c_1 = 0$ vanishes. If, in addition, $X = S^1 \times Y$ is a product, then a characteristic class argument shows that c_2 is even (e.g., see [Duncan 2013a, Section 4.3]), and so $\kappa(P)$ is a multiple of 4.

For example, consider the case where $G = \mathrm{SO}(r)$ with $r \geq 2$, and $\rho_0 = \mathrm{Ad}_{\mathbb{C}}$ is the complexified adjoint representation. Then $\kappa(P) = 2(r-2)p_1(P(\mathbb{R}^r))[X]$, where $p_1(P(\mathbb{R}^4))$ is the Pontryagin class of the vector bundle associated to the standard representation of $\mathrm{SO}(r)$.

As a second example, consider $G = \mathrm{SU}(r)$. Then the integers κ coming from the complexified adjoint and standard representations are related by

$$\kappa(P; \mathrm{Ad}_{\mathbb{C}}) = 2r\,\kappa(P; \text{standard}).$$

A *gauge transformation* on P is a G-equivariant bundle map $P \to P$ covering the identity. The set $\mathcal{G}(P)$ of gauge transformations on P forms a group, called the *gauge group*. One may equivalently view the gauge group as the set of G-equivariant maps $P \to G$. Here G acts on itself by conjugation of the inverse, making it a *right* action. A third equivalent way to view $\mathcal{G}(P)$ is as the space of sections of the bundle $P \times_G G \to X$, where $P \times_G G$ is formed using the same action of G on itself.

Denote by $\mathcal{G}_0 = \mathcal{G}_0(P)$ the connected component of the identity in $\mathcal{G}(P)$. We need to specify a topology on $\mathcal{G}(P)$ for the term "connected component" to be meaningful, and we do this by viewing $\mathcal{G}(P)$ as a subspace of the space of functions $P \to G$, equipped with the H^2-topology (however, any other Hölder or Sobolev topology would determine the same connected components). We denote by $\mathcal{G}^2(P)$ the completion of $\mathcal{G}(P)$ in the H^2-topology. Note that this depends on a choice of faithful representation of G (see [Wehrheim 2004, Appendix B]), and we take ρ_0 for this choice.

The gauge group acts on $\Omega^\bullet(P, P \times \mathfrak{g})$ and $\mathcal{A}(P) \subset \Omega^\bullet(P, P \times \mathfrak{g})$ by pullback. When the dimension of X is three or less, this action is smooth with the specified topologies [Wehrheim 2004, Appendix A]. We note that the action of a gauge transformation u on a connection A can be expressed as

$$u^*A = u^{-1}Au + u^{-1}du, \tag{5}$$

where the concatenation on the right is matrix multiplication and du is the linearization of $u : P \to G$. In dimensions three or less, Equation (5) combines with the Sobolev multiplication theorem to show that if u, A and u^*A are all of Sobolev class H^1, then u is actually of Sobolev class H^2.

The group $\mathcal{G}(P)$ also acts on $\Omega^\bullet(X, P(\mathfrak{g}))$ by the pointwise adjoint action $(\xi, u) \mapsto \mathrm{Ad}(u^{-1})\xi$. In particular, the curvature of $A \in \mathcal{A}(P)$ transforms under $u \in \mathcal{G}(P)$ by

$$F_{u^*A} = \mathrm{Ad}(u^{-1})F_A.$$

We introduce a notation convention that is convenient when the dimension of the underlying space X is relevant. If $\dim X = 4$, then we use A, U for connections and gauge transformations; if $\dim X = 3$, then we use a, u for connections and gauge transformations; if $\dim X = 2$, then we use α, μ for connections and gauge transformations. For example, this provides an effective way to distinguish between a path of gauge transformations $\mu : I \to \mathcal{G}(P)$ on a surface X, and its associated gauge transformation $u \in \mathcal{G}(I \times P)$ on the 3-manifold $I \times X$ defined by $u|_{\{t\}\times P} = \mu(t)$.

2B. *The Chern–Simons functional.* Fix a closed, connected, oriented 3-manifold Y, as well as a principal G-bundle $P \to Y$. The space of connections admits a natural 1-form $\lambda \in \Omega^1(\mathcal{A}(P), \mathbb{R})$ defined at $a \in \mathcal{A}(P)$ by

$$\lambda_a : T_a\mathcal{A}(P) \to \mathbb{R}, \qquad v \mapsto \int_Y \langle v \wedge F_a \rangle.$$

The Bianchi identity shows that this is a closed 1-form. Since $\mathcal{A}(P)$ is contractible it follows that λ is exact. Fixing a reference connection a_0, this exact 1-form can therefore be integrated along paths from a_0 to obtain a real-valued function $\mathcal{CS}_{a_0} : \mathcal{A}(P) \to \mathbb{R}$. One can compute that $\mathcal{CS}_{a_0}$ is given by the formula

$$\mathcal{CS}_{a_0}(a) := \int_Y \langle F_{a_0} \wedge v \rangle + \tfrac{1}{2}\langle d_{a_0} v \wedge v \rangle + \tfrac{1}{6}\langle [v \wedge v] \wedge v \rangle,$$

where we have set $v := a - a_0 \in \Omega^1(Y, P(\mathfrak{g}))$. We will typically choose a_0 to be flat, but this is not always convenient. In general, however, changing a_0 changes $\mathcal{CS}_{a_0}$ by a constant. Projecting $\mathcal{CS}_{a_0}$ to the circle $\mathbb{R}/\mathbb{Z}$, one obtains the Chern–Simons function $\mathcal{CS} : \mathcal{A}(P) \to \mathbb{R}/\mathbb{Z}$ from the introduction; we will refer to the lift $\mathcal{CS}_{a_0}$ as the *Chern–Simons functional*. Moreover, $\mathcal{CS}_{a_0}$ has a smooth extension from the smooth connections $\mathcal{A}(P)$ to the H^1-completion $\mathcal{A}^1(P)$.

Suppose $a, a' \in \mathcal{A}(P)$. Any path $a(\cdot) : [0, 1] \to \mathcal{A}(P)$ from a to a' can be interpreted as a connection A on $[0, 1] \times P \to [0, 1] \times Y$ by requiring that it restricts to $a(t)$ on $\{t\} \times Y$. It follows from the definitions that

$$\mathcal{CS}_{a_0}(a') - \mathcal{CS}_{a_0}(a) = \frac{1}{2}\int_{I\times Y} \langle F_A \wedge F_A \rangle.$$

In the special case where $a' = u^*a$, with $u \in \mathcal{G}(P)$, the connection A descends to a connection on the mapping torus

$$P_u := I \times P/(0, u(q)) \sim (1, q),$$

which is a bundle over $S^1 \times Y$. Then the above gives

$$\mathcal{CS}_{a_0}(u^*a) - \mathcal{CS}_{a_0}(a) = \frac{1}{2}\int_{S^1\times Y} \langle F_A \wedge F_A \rangle = \kappa(P_u) \in \mathbb{Z}, \tag{6}$$

where we used (4) in the second equality. It follows that the value of this depends only on the path component of u in $\mathcal{G}(P)$. Equation (6) also shows that $\mathcal{CS}_{a_0}$ is invariant under the subgroup of gauge transformations u with $\kappa(P_u)=0$ (the 'degree zero' gauge transformations), and that the circle-valued function $\mathcal{CS}:\mathcal{A}(P)\to\mathbb{R}/\mathbb{Z}$ is invariant under the full gauge group $\mathcal{G}(P)$.

Remark 2.2. The discussion following Equation (4) shows that if the mod-2 reduction of $c_1(P_u(W))$ vanishes, then (6) is even. Similarly, if the fixed representation ρ_0 is the complexification of a real representation, then (6) is a multiple of 4.

For completeness we show that the space of flat connections on P is locally path-connected. This implies, for example, that the Chern–Simons critical values are always isolated since the moduli space $\mathcal{A}_{\text{flat}}(P)/\mathcal{G}(P)$ is compact and $\mathcal{CS}_{a_0}$ is constant on the path components of $\mathcal{A}_{\text{flat}}(P)$.

Proposition 2.3. *The space $\mathcal{A}^1_{\text{flat}}(P)$ of flat connections is locally path-connected. In particular, the path components are the connected components.*

Proof. Råde [1992] used the heat flow associated to the Yang–Mills equations to show that there is some $\epsilon_P>0$ such that if $a\in\mathcal{A}^1(P)$ is a connection with $\|F_a\|_{L^2}\le\epsilon_P$, then there is a nearby flat connection

$$\mathrm{Heat}(a)\in\mathcal{A}^1_{\text{flat}}(P).$$

Råde shows that the map $a\mapsto\mathrm{Heat}(a)$ is continuous, gauge equivariant and restricts to the identity on $\mathcal{A}^1_{\text{flat}}(P)$.

Let $a_0,a_1\in\mathcal{A}^1_{\text{flat}}(P)$. We want to show that if a_0 and a_1 are close enough (in H^1), then they are connected by a path in $\mathcal{A}^1_{\text{flat}}(P)$. Consider the straight-line path $a(t)=a_0+t(a_1-a_0)$. Then

$$F_{a(t)}=t\,d_{a_0}(a_1-a_0)+\frac{t^2}{2}[a_1-a_0\wedge a_1-a_0],$$

and so

$$\|F_{a(t)}\|_{L^2}\le\|d_{a_0}(a_1-a_0)\|_{L^2}+\|a_1-a_0\|^2_{L^4}\le C\big(\|a_1-a_0\|_{H^1}+\|a_1-a_0\|^2_{H^1}\big),$$

where we have used the Sobolev embedding $H^1\hookrightarrow L^4$. Then $a(t)$ is in the realm of Råde's heat flow map for all $t\in[0,1]$, provided $\|a_1-a_0\|_{H^1}<\min\{1,\epsilon_P/2C\}$. When this is the case, $t\mapsto\mathrm{Heat}(a(t))\in\mathcal{A}^1_{\text{flat}}(P)$ is a path from a_0 to a_1, as desired. □

3. Chern–Simons values and instantons

We prove Theorem 1.1 and Corollary 1.4 in Sections 3B and 3C, respectively. We take a TQFT approach to the proof of Theorem 1.1 in the sense that we treat each connection on $Y=\overline{H}\cup_\partial H$ as a pair of connections on H that agree on the boundary.

This reduces the problem to a study of the flat connections on H and ∂H, which is the content of Section 3A.

3A. ***The components of the gauge group and the space of flat connections.*** In this section we fix a principal G-bundle $P \to X$, where X is a manifold with (possibly empty) boundary. Unfortunately, the action of the gauge group is rarely free. To account for this, it is convenient to consider the *based gauge group* $\mathcal{G}_p = \mathcal{G}_p(P)$ defined as the kernel of the map $\mathcal{G}(P) \to G$ given by evaluating $u : P \to G$ at some fixed point $p \in P$. If X is connected, then $\mathcal{G}_p$ acts freely on $\mathcal{A}(P)$ (in general, the stabilizer in $\mathcal{G}(P)$ of a connection $A \in \mathcal{A}(P)$ can be identified with the image in G of the evaluation map $u \to u(p)$).

Let $\widetilde{G} \to G$ be the universal cover. We will be interested in the subgroup $\mathcal{H} = \mathcal{H}(P)$ of gauge transformations $u : P \to G$ that lift to G-equivariant maps $\tilde{u} : P \to \widetilde{G}$, where the (right) action of G on $\widetilde{G}$ is induced by the conjugation action of $\widetilde{G}$ on itself.

Lemma 3.1. *The subgroup $\mathcal{H}$ is a union of connected components of $\mathcal{G}(P)$. In particular, $\mathcal{H}$ contains the identity component $\mathcal{G}_0$ of $\mathcal{G}(P)$.*

Proof. Consider the aforementioned right action of G on $\widetilde{G}$. Use this action to define a bundle $P \times_G \widetilde{G} \to X$, and consider the natural projection $P \times_G \widetilde{G} \to P \times_G G$. Viewing a gauge transformation u as a section of $P \times_G G \to X$, the defining condition of $\mathcal{H}$ is equivalent to the existence of a section $\tilde{u} : X \to P \times_G \widetilde{G}$ lifting u. It follows from the homotopy lifting property for the covering space $P \times_G \widetilde{G} \to P \times_G G$ that u is an element of $\mathcal{H}$ if and only if u can be connected by a path to an element of $\mathcal{H}$. □

Lemma 3.2. *Suppose G is compact and connected, and that X has the homotopy type of a connected 2-dimensional CW complex. Then $\mathcal{H} \cap \mathcal{G}_p$ is connected, and the inclusion $\mathcal{G}_p \subseteq \mathcal{G}(P)$ induces a bijection $\pi_0(\mathcal{G}_p) \cong \pi_0(\mathcal{G}(P))$. Consequently, $\mathcal{H}$ is the identity component $\mathcal{G}_0$ of $\mathcal{G}(P)$.*

Proof. First we show that $\mathcal{H} \cap \mathcal{G}_p$ is connected. For $u \in \mathcal{H}$, let $\tilde{u}$ be a lift as above. Note that if $u \in \mathcal{G}_p$, then $\tilde{u}(p) \in Z(\widetilde{G})$ is in the center and so $\tilde{u}(p)^{-1}\tilde{u}$ is another equivariant lift of u. In particular, by replacing $\tilde{u}$ with $\tilde{u}(p)^{-1}\tilde{u}$, we may assume $\tilde{u}$ has been chosen so that $\tilde{u}(p) = e \in \widetilde{G}$. Moreover, by homotoping u we may assume that u (hence $\tilde{u}$) restricts to the identity on $\pi^{-1}(B)$, where $B \subset X$ is some open coordinate ball around $x = \pi(p)$. The topological assumptions imply that B can be chosen so the complement $X - B$ deformation retracts to its 1-skeleton. Since G is connected, the restriction $P|_{X-B} \to X - B$ is trivializable. By equivariance, we may therefore view $\tilde{u}$ simply as a map

$$\tilde{u} : (X - B, \partial B) \to (\widetilde{G}, e).$$

Now we show $\pi_0(\mathcal{G}_p) \cong \pi_0(\mathcal{G}(P))$. We may homotope any gauge transformation $u : P \to G$ so that it is constant on $\pi^{-1}(B) \subset P$, with B as above. Just as above $P|_{\bar{B}} \to \bar{B}$ is the trivial bundle, so gauge transformations on $P|_{\bar{B}}$ are exactly maps $\bar{B} \to G$. Since G is connected, we can obviously find a homotopy rel ∂B of $u : (\bar{B}, \partial B) \to (G, u(p))$ to a map that sends $x \in B$ to the identity. This shows that u can be homotoped to an element of $\mathcal{G}_p$.

Finally, by Lemma 3.1 we have $\mathcal{G}_0 \subseteq \mathcal{H}$, while the reverse inclusion follows from the conclusions of the previous two paragraphs. □

Fix $x \in X$ as well as a point $p \in P$ over x. It is well-known that the holonomy provides a map $\mathrm{hol} : \mathcal{A}_{\mathrm{flat}}(P) \to \hom(\pi_1(X, x), G)$. This intertwines the action of $\mathcal{G}(P)$ on $\mathcal{A}_{\mathrm{flat}}(P)$ with the conjugation action of G on itself in the sense that if $\gamma : (S^1, 1) \to (X, x)$ is a smooth loop, then

$$\mathrm{hol}_{u^*A}(\gamma) = u(p)^{-1}\,\mathrm{hol}_A(\gamma)u(p)$$

for all gauge transformations $u \in \mathcal{G}(P)$ and flat connections A; see [Kobayashi and Nomizu 1963, Proposition 4.1] and [Atiyah and Bott 1983]. Moreover, the holonomy descends to a topological embedding

$$\mathcal{A}_{\mathrm{flat}}(P)/\mathcal{G}_p \hookrightarrow \hom(\pi_1(X, x), G)$$

with image a union of connected components that are determined by the topological type of the bundle P. To determine this set of image components for a given bundle P, it is useful to consider the following variation dating back to Atiyah and Bott [1983]. Let $j : G \to P$ denote the embedding $g \mapsto p \cdot g^{-1}$ (recall G acts on P on the right), and let j_* denote the induced map on π_1. Consider the universal cover $\widetilde{G} \to G$ and denote by $\iota : \pi_1(G) \hookrightarrow Z(\widetilde{G})$ the natural inclusion into the center of $\widetilde{G}$. Then there is a homeomorphism

$$\mathcal{A}_{\mathrm{flat}}(P)/(\mathcal{H} \cap \mathcal{G}_p) \cong \left\{\rho \in \hom(\pi_1(P, p), \widetilde{G}) \,\middle|\, \rho \circ j_* = \iota\right\}. \tag{7}$$

We defer a proof of (7) until the end of this section.

Proposition 3.3. *Assume G is compact and connected. Suppose X is either a closed, connected, oriented surface, or $X = H$ is a compression body. Then the space of flat connections $\mathcal{A}_{\mathrm{flat}}(P)$ is connected when it is nonempty.*

Proof. By Lemma 3.2, the group $\mathcal{H} \cap \mathcal{G}_p = \mathcal{G}_0 \cap \mathcal{G}_p$ is connected. Moreover, it acts freely on $\mathcal{A}_{\mathrm{flat}}(P)$ since this is the case with $\mathcal{G}_p$. We will show the space on the right-hand side of (7) is connected. The proposition follows immediately by the homotopy exact sequence for the bundle $\mathcal{A}_{\mathrm{flat}}(P) \to \mathcal{A}_{\mathrm{flat}}(P)/(\mathcal{H} \cap \mathcal{G}_p)$.

First assume X is a surface of genus $g \geq 0$. For $g = 0$, the space $\mathcal{A}_{\mathrm{flat}}(P)/(\mathcal{G}_0 \cap \mathcal{G}_p)$ is either a single point or empty, depending on whether P is trivial or not. We may therefore assume $g \geq 1$. The bundle $P \to X$ is determined up to bundle isomorphism by some $\delta \in \pi_1(G) \subset Z(\widetilde{G})$. Since $\widetilde{G}$ is simply-connected, it follows that $\widetilde{G} = G_1 \times \ldots \times G_k \times \mathbb{R}^l$ for some simple, connected, simply-connected Lie groups $G_1, \ldots, G_k$. Write $\delta = (\delta_1, \ldots, \delta_k, r)$ according to this decomposition.

Now we compute $\pi_1(P, p)$. Let U be the complement in X of a point y, and let V be a small disk around y. Applying the Seifert–van Kampen theorem to the sets $P|_U, P|_V \subset P$, one finds a presentation for $\pi_1(P, p)$ that consists of generators and relations coming from $\pi_1(G)$, as well as additional generators $\alpha_1, \beta_1, \ldots, \alpha_g, \beta_g$ subject to the relation

$$\prod_{j=1}^{g} [\alpha_j, \beta_j] = \delta, \tag{8}$$

as well as further relations asserting that each element of $\{\alpha_i, \beta_i\}_i$ commutes with each generator coming from $\pi_1(G)$. Alternatively, the relation (8) can be viewed as arising when one compares trivializations of $P|_U$ and $P|_V$ on the overlap $U \cap V$. It follows that $\mathcal{A}_{\text{flat}}(P)/\mathcal{G}_0 \cap \mathcal{G}_p$ can be identified with the set of tuples $(A_{ij}, B_{ij})_{i,j}$, for $1 \le i \le k+1$ and $1 \le j \le g$, where

(i) $A_{ij}, B_{ij} \in G_i$, and $\prod_{j=1}^{g} [A_{ij}, B_{ij}] = \delta_i$ for $1 \le i \le k$;

(ii) $A_{kj}, B_{kj} \in \mathbb{R}^l$, and $\prod_{j=1}^{g} [A_{kj}, B_{kj}] = r$.

Since $\mathbb{R}^l$ is abelian, the tuples $(A_{kj}, B_{kj})_j$ appearing in (ii) can only exist if $r = 0$. This shows that $\mathcal{A}_{\text{flat}}(P)$ is empty if $r \neq 0$, so we may assume $r = 0$. (Note that $r = 0$ if and only if the torus bundle $P/[G, G]$, from the introduction, is the trivial bundle.)

For $1 \le i \le k$, given any $\delta_i \in \widetilde{G}$ it can be shown that (a) there always exist tuples $(A_{ij}, B_{ij})_j \subset G_i^{2g}$ satisfying $\prod_{j=1}^{g} [A_{ij}, B_{ij}] = \delta_i$, and (b) the set of such $(A_{ij}, B_{ij})_j$ is always connected; see [Alekseev et al. 1998], [Ramadas et al. 1989, Section 2.1] or [Ho and Liu 2003, Fact 3]. It follows that $\mathcal{A}_{\text{flat}}(P)/\mathcal{G}_0 \cap \mathcal{G}_p$ is a product of connected spaces and is therefore connected. This finishes the proof in the case where X is a surface.

Now suppose $X = H$ is a compression body. Then there is a homotopy equivalence $H \simeq (\bigvee_{i=1}^{s} \Sigma_i) \vee (\bigvee_{i=1}^{t} S^1)$ onto a wedge sum of closed, connected, oriented surfaces Σ_i and circles; note that the surfaces can be identified with the components of the incoming end $\Sigma_- \subset \partial H$. It follows from (7) that $\mathcal{A}_{\text{flat}}(P)/\mathcal{G}_0 \cap \mathcal{G}_p$ is homeomorphic to

$$\left\{\rho \in \hom(\pi_1(P_1), \widetilde{G}) \,\middle|\, \rho \circ j_* = \iota\right\} \times \ldots \times \left\{\rho \in \hom(\pi_1(P_s), \widetilde{G}) \,\middle|\, \rho \circ j_* = \iota\right\} \times (\widetilde{G})^t,$$

where $P_i \to \Sigma_i$ is the restriction of P to the surface $\Sigma_i \subset H$. By the previous paragraph this is a product of connected spaces, and so is itself connected. □

Remark 3.4. The above proof shows that, when H is a compression body, restricting to the incoming end $\Sigma_- \subset \partial H$ yields a surjective map

$$\frac{\mathcal{A}_{\text{flat}}(P)}{\mathcal{G}_0 \cap \mathcal{G}_p} \to \frac{\mathcal{A}_{\text{flat}}(P_1)}{\mathcal{G}_0(P_1) \cap \mathcal{G}_{p_1}(P_1)} \times \ldots \times \frac{\mathcal{A}_{\text{flat}}(P_s)}{\mathcal{G}_0(P_s) \cap \mathcal{G}_{p_s}(P_s)}$$

that is a (trivial) principal $\widetilde{G}^t$-bundle. Similarly, restricting to the outgoing end $\Sigma_+ \subset \partial H$ yields an injection

$$\frac{\mathcal{A}_{\text{flat}}(P)}{\mathcal{G}_0 \cap \mathcal{G}_p} \hookrightarrow \frac{\mathcal{A}_{\text{flat}}(P_+)}{\mathcal{G}_0(P_+) \cap \mathcal{G}_{p_+}(P_+)},$$

where $P_+ \to \Sigma_+$ is the restriction of P. In particular, a flat connection on $P \to H$ is determined uniquely, up to $\mathcal{G}_0(P) \cap \mathcal{G}_p(P)$, by its value on the boundary component Σ_+, and hence by its value on ∂H.

Now we verify (7). This can be viewed as arising from the $\widetilde{G}$*-valued holonomy*, which we now describe. Let $A \in \mathcal{A}(P)$ be a connection. Given a smooth loop $\gamma : S^1 = \mathbb{R}/\mathbb{Z} \to P$, consider the induced loop in the base $\pi \circ \gamma : S^1 \to X$. Use this to pull P back to a bundle over the circle $(\pi \circ \gamma)^* P \to S^1$. The standard ($G$-valued) holonomy determines a lift $\text{hol}_A(\pi \circ \gamma)$ of the quotient map $[0, 1] \hookrightarrow S^1 = \mathbb{R}/\mathbb{Z}$:

(9)
$$\begin{tikzcd} & (\pi \circ \gamma)^* P \arrow[d] \\ {[0,1]} \arrow[ur, dashed, "\text{hol}_A(\pi\circ\gamma)"] \arrow[r, hook] & S^1 \end{tikzcd}$$

and this lift is unique if we require that it sends 0 to $\gamma(0) \in (\pi \circ \gamma)^* P$. On the other hand, γ determines a trivialization of this pullback bundle

$$(\pi \circ \gamma)^* P \cong S^1 \times G, \quad \gamma(t) \mapsto (t, e).$$

Compose the lift in (9) with this isomorphism and then with the projection to the G-factor in $S^1 \times G$ to get a map

(10)
$$\text{hol}_A(\pi \circ \gamma) : [0, 1] \to G,$$

which we denote by the same symbol we used for the standard holonomy. Then the map in (10) sends 0 to the identity $e \in G$ and the value at 1 recovers the standard holonomy for A around γ. Viewing $\widetilde{G} \to G$ as a covering space, $\text{hol}_A(\pi \circ \gamma)$ lifts to a unique map $\widetilde{\text{hol}}_A(\pi \circ \gamma) : [0, 1] \to \widetilde{G}$ that sends 0 to e. Then we declare the $\widetilde{G}$*-valued holonomy* of A around γ to be the value at 1:

$$\text{hol}_A^{\widetilde{G}}(\gamma) := \widetilde{\text{hol}}_A(\pi \circ \gamma)(1) \in \widetilde{G}.$$

As with the standard holonomy, one can check that this is multiplicative under concatenation of paths γ. Similarly, this is equivariant in the following sense.

Suppose $u \in \mathcal{H}$ and so u lifts to a G-equivariant map $\tilde{u} : P \to \widetilde{G}$. Setting $g := \tilde{u}(p)$, we have

$$\mathrm{hol}^{\widetilde{G}}_{u^*A}(\gamma) = g^{-1}\,\mathrm{hol}^{\widetilde{G}}_{A}(\gamma) g.$$

Next, suppose A is a flat connection. Then $\mathrm{hol}^{\widetilde{G}}_{A}(\gamma)$ depends only on the homotopy class of γ. It follows from the above observations that the $\widetilde{G}$-valued holonomy defines a map $\mathcal{A}_{\mathrm{flat}}(P) \to \hom(\pi_1(P, p), \widetilde{G})$, and this intertwines the actions of $\mathcal{H}$ and $\widetilde{G}$. Moreover, from the definition of $\mathcal{G}_p$ we have that the $\widetilde{G}$-valued holonomy is invariant under the action of $\mathcal{H} \cap \mathcal{G}_p$. We therefore have a well-defined map $\mathcal{A}_{\mathrm{flat}}(P)/\mathcal{H} \cap \mathcal{G}_p \to \hom(\pi_1(P), \widetilde{G})$. It follows from the definitions above that the image lies in the right-hand side of (7). That this map is a homeomorphism follows from the analogous argument for the standard holonomy, together with the commutativity of the following diagram.

$$\begin{tikzcd}
\mathcal{A}_{\mathrm{flat}}(P)/(\mathcal{H} \cap \mathcal{G}_p) \arrow[r] \arrow[d] & \{\rho \in \hom(\pi_1(P), \widetilde{G}) \mid \rho \circ j_* = \iota\} \arrow[d] \\
\mathcal{A}_{\mathrm{flat}}(P)/\mathcal{G}(P) \arrow[r, hook] & \hom(\pi_1(X), G)/G
\end{tikzcd}$$

3B. ***Proof of Theorem 1.1.*** Write $Y = \overline{H} \cup_{\partial H} H$, where H is a compression body. Fix a collar neighborhood $[0, \epsilon) \times \partial H \hookrightarrow H$ for ∂H, and use this to define the smooth structure on Y. This smooth structure is independent, up to diffeomorphism, of the choice of collar neighborhood, see [Milnor 1965, Theorem 1.4]. The product structure of this collar neighborhood can be used to define a vector field ν on Y that is normal to ∂H and that does not vanish at ∂H. Moreover, we assume ν has support near ∂H, and so ν lifts to an equivariant vector field on P that we denote by the same symbol.

Restriction to each of the H factors in $Y = \overline{H} \cup_{\partial H} H$ determines an embedding

$$\mathcal{A}^1_{\mathrm{flat}}(P) \hookrightarrow \left\{ (b, c) \in \mathcal{A}^1_{\mathrm{flat}}(P|_H) \times \mathcal{A}^1_{\mathrm{flat}}(P|_H) \;\middle|\; \begin{array}{rcl} b|_{\partial H} &=& c|_{\partial H}, \\ -\iota_\nu b|_{\partial H} &=& \iota_\nu c|_{\partial H} \end{array} \right\}$$

given by

$$a \mapsto (a|_{\overline{H}},\ a|_H). \tag{11}$$

A few comments about the defining conditions in the codomain are in order: (i) we are treating $\nu = \nu|_H$ as a vector field on H, viewed as the second factor in $\overline{H} \cup_{\partial H} H$; (ii) the negative sign is due to the reversed orientation of the first factor; and (iii) restriction to the hypersurface $\partial H \subset Y$ extends to a bounded linear map $H^1(Y) \to L^2(\partial H)$ (see, e.g., [Adams 1975, Theorem 6.3]), and so these equalities should be treated as equalities in the L^2 sense.

Suppose (b, c) is in the codomain of (11). These define a connection a on Y by setting $a|_{\overline{H}} = b$ and $a|_H = b$. It is straightforward to check that if b and c are both smooth, then a is continuous and of Sobolev class H^1 on Y. Since the smooth connections are dense in $\mathcal{A}^1$, it follows that (11) is surjective, and so we may treat (11) as an identification.

The bijection (11) singles out a preferred subspace that we call the *diagonal*

$$\big\{(b, b) \in \mathcal{A}^1_{\text{flat}}(P|_H) \times \mathcal{A}^1_{\text{flat}}(P|_H) \,\big|\, \iota_\nu b|_{\partial H} = 0\big\} \subset \mathcal{A}^1_{\text{flat}}(P). \tag{12}$$

It is convenient to consider a slightly larger space $\mathcal{C} \subset \mathcal{A}^1_{\text{flat}}(P)$ defined to be the set of flat connections that can be connected by a path to an element of the diagonal (12).

Claim. *The diagonal (12) is path-connected. In particular, $\mathcal{C}$ is also path-connected.*

To see this, consider diagonal elements (b_0, b_0), (b_1, b_1). It suffices to prove the claim under the assumption that b_0, b_1 are both smooth and satisfy

$$\iota_\nu b_0|_U = \iota_\nu b_1|_U = 0 \tag{13}$$

on some neighborhood U of ∂H (this is because the H^1-completion of the space of these connections recovers (12) and the path-components are stable under completion). By Proposition 3.3 there is a path of flat connections $t \mapsto b_t \in \mathcal{A}_{\text{flat}}(P|_H)$ connecting b_0 and b_1. We will be done if we can ensure that $\iota_\nu b_t|_{\partial H} = 0$ for all $t \in [0, 1]$. We will accomplish this by putting b_t in a suitable 'ν-temporal gauge', as follows. Restrict attention to the bicollar neighborhood $(-\epsilon, \epsilon) \times \partial H \subset Y$ obtained by doubling the collar neighborhood from the beginning of this section. Let s denote the variable in the $(-\epsilon, \epsilon)$-direction and fix a bump function β for U that is equal to 1 on ∂H. For each $t \in [0, 1]$, define a gauge transformation u_t at $(s, h) \in (-\epsilon, \epsilon) \times \partial H$ by the formula

$$u_t(s, h) := \exp\Big(-\int_0^s \iota_{\beta\nu(\sigma,h)} b_t(\sigma, h)\, d\sigma\Big).$$

Then u_t depends smoothly on all variables, and a computation shows

$$\iota_{\beta\nu}(u_t^* b_t) = 0.$$

Moreover, it follows from (13) that u_t is the identity gauge transformation when $t = 0, 1$. The claim follows by extending u_t to all of Y using a bump function.

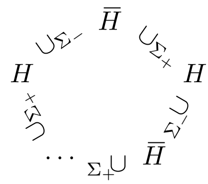
It follows from the claim that the Chern–Simons functional is constant on $\mathcal{C}$, since $\mathcal{CS}_{a_0}$ is *locally* constant on its critical set $\mathcal{A}^1_{\text{flat}}(P)$. Suppose Hypothesis 1 holds. We will show that every flat connection in $\mathcal{A}^1_{\text{flat}}(P)$ is gauge equivalent to one in $\mathcal{C}$; Theorem 1.1 will then follow immediately from Remark 2.2. In fact, by another density argument, it suffices to show that every *smooth* flat connection is gauge equivalent to one in $\mathcal{C}$. So we fix $a \in \mathcal{A}_{\mathrm{flat}}(P)$. As in the proof of the claim, by applying a suitable gauge transformation, we may assume that $\iota_\nu a = 0$. Use (11) to identify a with a pair $(b, c) \in \mathcal{A}_{\mathrm{flat}}(P|_H) \times \mathcal{A}_{\mathrm{flat}}(P|_H)$. Then b, c agree on the boundary, so by Remark 3.4, there is some gauge transformation $u \in \mathcal{G}_0(P|_H) \cap \mathcal{G}_p(P|_H)$ for which $u^*c = b$. Here we have chosen $p \in H$ to lie in $\Sigma_+ \subset \partial H$, and we are thinking of the H that appears here as the second factor in $Y = \overline{H} \cup_{\partial H} H$. Our immediate goal is to show that u restricts to the identity gauge transformation on the boundary $\partial H = \Sigma_+ \cup \Sigma_-$. Since $p \in \Sigma_+$, it follows that the restriction $u|_{\Sigma_+}$ lies in $\mathcal{G}_p(P|_{\Sigma_+})$, which acts freely. Since b and c agree on Σ_+, it must be the case that $u|_{\Sigma_+}$ is the identity. Turning attention to Σ_-, for each component $\Sigma' \subset \Sigma_-$, the restriction $u|_{\Sigma'}$ lies in the identity component of the gauge group. In particular, by Hypothesis 1 we have that $u|_{\partial H} = e$ is the identity. At this point we have that u is a gauge transformation on $H \subset Y$ that is the identity on all of ∂H. Then u extends over $\overline{H} \subset Y$ by the identity to define a continuous gauge transformation $u^{(1)} = (e, u)$ on P. This is of Sobolev class H^1. We also have $(u^{(1)})^*a \in \mathcal{C}$, since under (11) the connection $(u^{(1)})^*a$ corresponds to the pair $(b, b) = (b, u^*c)$ and we have assumed $\iota_\nu a = 0$. Finally, since $u^{(1)}$, a and $(u^{(1)})^*a$ are all H^1, it follows from (5) that $u^{(1)}$ is H^2. This finishes the proof of Theorem 1.1 under Hypothesis 1.

Remark 3.5. Continue to assume Hypothesis 1, and suppose a, a' are flat connections. Then the construction of the previous paragraph shows that there is a gauge transformation $w \in \mathcal{H}(P)$ such that w^*a and a' lie in the same path component. If we further assume that $\mathcal{CS}_{a_0}(a) = \mathcal{CS}_{a_0}(a')$, then it follows that $\kappa(P_w) = 0$. In many cases, if $w \in \mathcal{H}$ and $\kappa(P_w) = 0$, then w necessarily lies in the identity component. For example, this is well-known when $G = \mathrm{U}(r)$ or $\mathrm{SU}(r)$ and $\rho_0 : G \to \mathrm{U}(\mathbb{C}^r)$ is the standard representation [Freed and Uhlenbeck 1991, page 79], or if $G = \mathrm{PU}(r)$ and ρ_0 is the adjoint representation [Duncan 2013a]. In such cases, it follows that a and a' lie in the same component of $\mathcal{A}_{\mathrm{flat}}(P)$.

To prove the theorem without Hypothesis 1, we follow a strategy of Wehrheim [2006]. Let n_G be as in the definition of N_G. Without Hypothesis 1 it may not be the case that $u \in \mathcal{G}(P|_H)$ restricts to the identity on Σ_-. Write $\Sigma_- = \Sigma_1 \cup \cdots \cup \Sigma_s$ in terms of its connected components and write P_i for the restriction of P to $\Sigma_i \subset \partial H$. Since G is compact, the stabilizer subgroup in $\mathcal{G}(P_i)$ of each restriction $b|_{\Sigma_i}$ has only finitely many components, and so there is some integer $n \leq n_G$ for which $u^n|_{\Sigma_i}$ lies in the identity component of the stabilizer group for $b|_{\Sigma_i}$. For simplicity we assume $u^n|_{\Sigma_i} = e$ is the identity for each i; one can check that the following argument can be easily reduced to this case.

View H as a cobordism from Σ_- to Σ_+ (we may assume Σ_- is not empty, otherwise Hypothesis 1 is satisfied), and define a manifold $Y^{(n)}$ by gluing H to itself $2n$ times:

(14)



Define a bundle $P^{(n)} \to Y^{(n)}$ similarly. Then $a = (b, c)$ determines a continuous flat connection on $P^{(n)}$ by the formula

$$a^{(n)} := (b, c, b, c, \dots, b, c).$$

The notation means that the k-th component lies in the k-th copy of H in (14). Similarly, the reference connection a_0 defines a reference connection $a_0^{(n)}$ on $P^{(n)}$, and the gauge transformation u determines a continuous gauge transformation on $P^{(n)}$ by

$$u^{(n)} := (e, u, u, u^2, u^2, \dots, u^{n-1}, u^{n-1}, u^n).$$

Let $\mathcal{CS}^{(n)}$ denote the Chern–Simons functional for $P^{(n)}$ defined using $a_0^{(n)}$. Then (6) and the additivity of the integral over its domain give

$$\mathcal{CS}^{(n)}\big((u^{(n)})^* a^{(n)}\big) = \mathcal{CS}^{(n)}\big(a^{(n)}\big) + \kappa(P_{u^{(n)}}) = n\mathcal{CS}_{a_0}(a) + \kappa(P_{u^{(n)}}).$$

In addition, the pullback of $a^{(n)}$ by $u^{(n)}$ is $(b, b, u^*b, u^*b, \dots, (u^{n-1})^*b, (u^{n-1})^*b)$, and so

$$\mathcal{CS}^{(n)}\big((u^{(n)})^* a^{(n)}\big) = n\mathcal{CS}_{a_0}(a') + k_n, \quad k_n := \tfrac{1}{2}n(n-1)\kappa(P_u) \in \mathbb{Z},$$

where $a' \in \mathcal{C}$ is the connection corresponding to (b, b) under (11). Combining these gives $\mathcal{CS}_{a_0}(a) - \mathcal{CS}_{a_0}(a') \in \frac{1}{n}\mathbb{Z} \subseteq \frac{1}{N_G}\mathbb{Z}$.

3C. ***The energies of instantons.*** Let $P^\infty \to H^\infty$ be as in Corollary 1.4, and let g be the cylindrical end metric on H^∞. Equip the 4-manifold $\mathbb{R} \times H^\infty$ with the product metric, and denote by $Q \to \mathbb{R} \times H^\infty$ the pullback of P^∞ under the projection $\mathbb{R} \times H^\infty \to H^\infty$. The *energy* of a connection $A \in \mathcal{A}(Q)$ is defined to be

$$\frac{1}{2}\|F_A\|_{L^2(\mathbb{R}\times H^\infty)} = \frac{1}{2}\int_{\mathbb{R}\times H^\infty} \langle F_A \wedge *F_A\rangle,$$

where $*$ is the Hodge star coming from the metric. We will always assume the energy of A is finite. We say that A is an *instanton* if $*F_A = \pm F_A$. It follows that the energy of any instanton is given, up to a sign, by

$$\frac{1}{2}\int_{\mathbb{R}\times H^\infty} \langle F_A \wedge F_A\rangle. \tag{15}$$

In this section we will prove Corollary 1.4 by showing that (15) is equal to $\mathcal{CS}_{a_0}(a_\flat)$ for some flat connections $a_\flat, a_0$ on $Y := \overline{H} \cup_{\partial H} H$. First we introduce some notation.

Recalling the decomposition (1), there is a projection

$$\mathbb{R} \times H^{\infty} \to \mathbb{H} \tag{16}$$

to the upper half-plane, sending $\{s\} \times H$ to $(s, 0) \in \mathbb{H}$, and sending each element of $\{(s, t)\} \times \partial H$ to (s, t). (This projection is continuous, but *not* differentiable.) Note that for each $\tau \in (0, \infty)$, the inverse image under (16) of the semicircle

$$\{(\tau \cos(\theta), \tau \sin(\theta)) \mid \theta \in [0, \pi]\} \subset \mathbb{H}$$

is the closed 3-manifold

$$Y_\tau := \overline{H} \cup_{\{0\} \times \partial H} ([0, \tau\pi] \times \partial H) \cup_{\{\tau\pi\} \times \partial H} H.$$

In the degenerate case $\tau = 0$, we declare Y_0 to be the inverse image under (16) of the origin; so $Y_0 = \{0\} \times H$. Then we have

$$\mathbb{R} \times H^{\infty} = \bigcup_{\tau \geq 0} Y_\tau.$$

Moreover, for each $\tau > 0$, there is an identification $Y_\tau \cong Y_1$ induced from the obvious linear map $[0, \tau\pi] \cong [0, \pi]$. This identification is continuous, but when $\tau \neq 1$ this identification is not smooth due to the directions transverse to $\{0, \tau\pi\} \times \partial H$ in Y_τ. We note also that we can identify Y_1 with the double Y; however we find it convenient to work with Y_1 rather than Y at this stage. In summary, we have defined a continuous embedding

$$\Pi : (0, \infty) \times Y_1 \to \mathbb{R} \times H^{\infty}$$

with image the complement of Y_0; this map is not smooth. We think of Π as providing certain "polar coordinates" on $\mathbb{R} \times H^{\infty}$.

Fix a connection A. Then we can write the pullback under Π as

$$\Pi^* A = a(\tau) + p(\tau)\, d\tau,$$

where τ is the coordinate on $(0, \infty)$, $a(\cdot)$ is a path of connections on Y_1, and $p(\cdot)$ is a path of 0-forms on Y_1. Fixing τ, the failure of Π to be smooth implies that the connection $a(\tau)$ will not be continuous on Y_1, unless

$$\iota_\nu A = 0. \tag{17}$$

Here, ν is the normal vector to the hypersurface $\mathbb{R} \times \partial H \subset \mathbb{R} \times H^{\infty}$. However, by performing a suitable gauge transformation to A, we can always achieve (17). (See the previous section for a similar construction; also note that the action of the gauge group on A does not change the value of (15).) When (17) holds it follows that the connection $a(\tau)$

- is continuous everywhere on Y_1,
- is smooth away from the hypersurface $\{0, \pi\} \times \partial H \subset Y_1$, and
- has bounded derivative near this hypersurface.

In particular, $a(\tau)$ is of Sobolev class H^1 on Y_1.

Now we introduce a convenient reference connection a_0 on Y_1 with which we will define $\mathcal{CS}_{a_0}$. This reference connection will depend on the given connection A; we continue to assume that (17) holds. Define a_0 on the first copy of H in Y_1 by declaring it to equal $A|_{Y_0}$, where we are identifying Y_0 with H in the obvious way. Define a_0 on the second copy of H to also equal $A|_{Y_0}$. It remains to define a_0 on the cylinder $[0, \pi] \times \partial H$, and there is a unique way to do this if we require that a_0 is (i) continuous and (ii) constant in the $[0, \pi]$-direction. It follows from (17) that a_0 is of Sobolev class H^1. Moreover,

$$\lim_{\tau \to 0^+} a(\tau) = a_0,$$

where this limit is in the H^1-topology on Y_1 (this is basically just the statement that A is continuous at $Y_0 \subset \mathbb{R} \times H^\infty$). Note that this choice of a_0 may not be flat. However, it turns out that $\mathcal{CS}_{a_0} = \mathcal{CS}_{a_1}$ for some flat connection a_1 (in fact, any flat connection in the diagonal (12) will do); see Remark 3.6.

Now we prove Corollary 1.4. At this stage the argument follows essentially as in [Wehrheim 2006, Theorem 1.1]; we recall the details for convenience. Let A be any finite energy connection on $\mathbb{R} \times H^\infty$, and assume it has been put in a gauge so that (17) holds. Use the identity $F_{\Pi^* A} = F_a + d\tau \wedge (\partial_\tau a - d_a p)$ to get

$$\tfrac{1}{2}\Pi^* \langle F_A \wedge F_A \rangle = d\tau \wedge \langle F_a \wedge (\partial_\tau a - d_a p)\rangle.$$

Integrate both sides and use the fact that the image of Π has full measure in $\mathbb{R} \times H^\infty$ to get

$$\begin{aligned}\frac{1}{2}\int_{\mathbb{R}\times H^\infty} \langle F_A \wedge F_A \rangle &= \int_0^\infty \int_{Y_1} d\tau \wedge \langle F_a \wedge \partial_\tau a\rangle \\ &= \int_0^\infty \frac{d}{d\tau}\mathcal{CS}_{a_0}(a(\tau))\, d\tau \\ &= \lim_{\tau\to\infty} \mathcal{CS}_{a_0}(a(\tau)) - \lim_{\tau \to 0^+} \mathcal{CS}_{a_0}(a(\tau)),\end{aligned} \tag{18}$$

where we used the Bianchi identity to kill off the $d_a p$-term, and then used the definition of $\mathcal{CS}_{a_0}$. From the definition of a_0, we have

$$\lim_{\tau \to 0^+} \mathcal{CS}_{a_0}(a(\tau)) = \mathcal{CS}_{a_0}(a_0) = 0,$$

so it suffices to consider the limit at ∞.

Notice that (18) shows that $\lim_{\tau\to\infty} \mathcal{CS}_{a_0}(a(\tau))$ exists. The goal now is to show that this limit equals $\mathcal{CS}_{a_0}(a_\flat)$ for some flat connection $a_\flat$. Endow Y_1 with the metric induced from $ds^2 + g$ via the inclusion $Y_1 \subset \mathbb{R} \times H^\infty$. Then it follows from the definitions that

$$\int_1^\infty \|F_{a(\tau)}\|^2_{L^2(Y_1)} \leq \|F_A\|^2_{L^2(\mathbb{R}\times H^\infty)}.$$

Since the energy of A is finite, the integral over $[1,\infty)$ on the left converges and so there is a sequence $\tau_i \in \mathbb{R}$ with

$$\|F_{a(\tau_i)}\|^2_{L^2(Y_1)} \xrightarrow{i} 0 \quad \text{and} \quad \tau_i \xrightarrow{i} \infty.$$

By Uhlenbeck's weak compactness theorem [1982], we can find

- a subsequence of the $\{a(\tau_i)\}$, denoted by $\{a_i\}$,
- a sequence of gauge transformations $\{u_i\}$, and
- a flat connection a_∞,

for which $\{u_i^* a_i\}$ converges to a_∞ weakly in H^1 and hence strongly in L^4. This convergence is enough to put each $u_i^* a_i$ in Coulomb gauge with respect to a_∞ [Wehrheim 2004, Theorem 8.1], so by redefining each u_i we may assume this is the case. Then $u_i^* a_i$ converges to a_∞ strongly in H^1. Since $\mathcal{CS}_{a_0}$ is continuous in the H^1-topology, we have

$$\lim_{i\to\infty} \mathcal{CS}_{a_0}(u_i^* a_i) = \mathcal{CS}_{a_0}(a_\infty).$$

On the other hand,

$$\mathcal{CS}_{a_0}(u_i^* a_i) - \mathcal{CS}_{a_0}(a_i) = \kappa(P_{u_i}) \in \mathbb{Z}$$

for all i. Since $\mathcal{CS}_{a_0}(u_i^* a_i)$ and $\mathcal{CS}_{a_0}(a_i)$ both converge, it follows that $\kappa(P_{u_i})$ is constant for all but finitely many i. By passing to yet another subsequence, we may assume that $\kappa(P_{u_i})$ is constant for all i. Then there is some gauge transformation u such that $\kappa(P_u) = \kappa(P_{u_i})$ for all i (just take u to be one of the u_i). This gives

$$\begin{aligned}\frac{1}{2}\int_{\mathbb{R}\times H^\infty} \langle F_A \wedge F_A\rangle = \lim_{i\to\infty} \mathcal{CS}_{a_0}(a_i) &= \lim_{i\to\infty} \mathcal{CS}_{a_0}(u_i^* a_i) - \kappa(P_{u_i}) \\ &= \mathcal{CS}_{a_0}(a_\infty) - \kappa(P_u) = \mathcal{CS}_{a_0}((u^{-1})^* a_\infty).\end{aligned}$$

So taking $a_\flat := (u^{-1})^* a_\infty$ finishes the proof.

Remark 3.6. Here we address the fact that the reference connection a_0, constructed in the proof above, may not be a flat connection. We address this from two different angles. First of all, the quantity (15) is independent of the choice of connection A, provided that one restricts to connections with the same asymptotic behavior at infinity. In particular, one can always modify the connection A so that its restriction to Y_0 is flat. This forces a_0 to be flat.

Secondly, the argument of the previous paragraph suggests that the value $\mathcal{CS}_{a_0}(a)$ is somehow independent of a_0. It is interesting to see this explicitly without modifying the original connection A. There is an obvious $\mathbb{Z}_2$ action on $Y = \overline{H} \cup_{\partial H} H$ given by interchanging the two H-factors. Call a form or connection on Y *symmetric* if it is fixed by this action. For example, all elements of the diagonal (12) are symmetric. The key observation here is that a_0 is symmetric. Then we claim that function $\mathcal{CS}_{a_0}$ is independent of the choice of a_0 from the class of symmetric connections. Indeed, suppose a_1 is a second connection that is symmetric. We want to show that $\mathcal{CS}_{a_0}(a) = \mathcal{CS}_{a_1}(a)$ for all connections a. From the definition of the Chern–Simons functional we have

$$\mathcal{CS}_{a_0}(a) - \mathcal{CS}_{a_1}(a) = -\mathcal{CS}_a(a_0) + \mathcal{CS}_a(a_1).$$

Note that the right-hand side is actually independent of a, since changing the connection a changes $\mathcal{CS}_a$ by a constant. We can therefore replace a with a_0 on the right-hand side to get

$$\mathcal{CS}_{a_0}(a) - \mathcal{CS}_{a_1}(a) = \mathcal{CS}_{a_0}(a_1) = \int_Y \langle F_{a_0} \wedge v\rangle + \tfrac{1}{2}\langle d_{a_0} v \wedge v\rangle + \tfrac{1}{6}\langle [v \wedge v] \wedge v\rangle,$$

where $v := a_1 - a_0$. Let $cs_{a_0}(v)$ denote the integrand on the right. Now use the following facts: (i) Y decomposes into two copies of H, (ii) the two copies of H have opposite orientations, and (iii) $cs_{a_0}(v)$ is symmetric (it is made up of the symmetric a_0, a_1). These allow us to compute

$$\mathcal{CS}_{a_0}(a) - \mathcal{CS}_{a_1}(a) = \int_{\overline{H}} cs_{a_0}(v) + \int_H cs_{a_0}(v) = -\int_H cs_{a_0}(v) + \int_H cs_{a_0}(v) = 0.$$

Acknowledgments

Duncan would like to thank Chris Woodward, Sai Kee Yeung, Sushmita Venugopalan, Dan Freed, and Tom Parker for valuable discussions, comments and suggestions. This work was partially supported by NSF Grant DMS 0739208.

References

[Adams 1975] R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics, Academic Press, New York, 1975. MR 56 #9247 Zbl 0314.46030

[Alekseev et al. 1998] A. Alekseev, A. Malkin, and E. Meinrenken, "Lie group valued moment maps", *J. Differential Geom.* **48**:3 (1998), 445–495. MR 99k:58062 Zbl 0948.53045

[Atiyah and Bott 1983] M. F. Atiyah and R. Bott, "The Yang–Mills equations over Riemann surfaces", *Philos. Trans. Roy. Soc. London Ser. A* **308**:1505 (1983), 523–615. MR 85k:14006 Zbl 0509.14014

[Auckly 1994] D. R. Auckly, "Chern–Simons invariants of 3-manifolds which fiber over S^1", *Internat. J. Math.* **5**:2 (1994), 179–188. MR 95b:57030 Zbl 0854.57012

[Dostoglou and Salamon 1994] S. Dostoglou and D. A. Salamon, "Self-dual instantons and holomorphic curves", *Ann. of Math.* (2) **139**:3 (1994), 581–640. MR 95g:58050 Zbl 0812.58031

[Duncan 2013a] D. Duncan, "On the components of the gauge group for PU(r)-bundles", preprint, 2013. arXiv 1311.5611

[Duncan 2013b] D. L. Duncan, *Compactness results for the quilted Atiyah–Floer conjecture*, Ph.D. thesis, Rutgers The State University of New Jersey - New Brunswick, 2013, Available at http://search.proquest.com/docview/1441954671.

[Etesi 2013] G. Etesi, "On the energy spectrum of Yang–Mills instantons over asymptotically locally flat spaces", *Comm. Math. Phys.* **322**:1 (2013), 1–17. MR 3073155 Zbl 1270.81136

[Freed and Uhlenbeck 1991] D. S. Freed and K. K. Uhlenbeck, *Instantons and four-manifolds*, 2nd ed., Mathematical Sciences Research Institute Publications **1**, Springer, New York, 1991. MR 91i:57019 Zbl 0968.57502

[Ho and Liu 2003] N.-K. Ho and C.-C. M. Liu, "Connected components of the space of surface group representations", *Int. Math. Res. Not.* **2003**:44 (2003), 2359–2372. MR 2004h:53116 Zbl 1043.53064

[Kirk and Klassen 1993] P. Kirk and E. Klassen, "Chern–Simons invariants of 3-manifolds decomposed along tori and the circle bundle over the representation space of T^2", *Comm. Math. Phys.* **153**:3 (1993), 521–557. MR 94d:57042 Zbl 0789.57011

[Kobayashi and Nomizu 1963] S. Kobayashi and K. Nomizu, *Foundations of differential geometry, I*, Wiley, New York, 1963. MR 97c:53001a

[Milnor 1965] J. Milnor, *Lectures on the h-cobordism theorem*, Notes by L. Siebenmann and J. Sondow, Princeton University Press, Princeton, 1965. MR 32 #8352

[Neumann and Yang 1995] W. D. Neumann and J. Yang, "Invariants from triangulations of hyperbolic 3-manifolds", *Electron. Res. Announc. Amer. Math. Soc.* **1**:2 (1995), 72–79. MR 97d:57017 Zbl 0851.57013

[Nishi 1998] H. Nishi, "SU(n)-Chern–Simons invariants of Seifert fibered 3-manifolds", *Internat. J. Math.* **9**:3 (1998), 295–330. MR 99i:57053 Zbl 0911.57010

[Nishinou 2010] T. Nishinou, "Convergence of adiabatic family of anti-self-dual connections on products of Riemann surfaces", *J. Math. Phys.* **51**:2 (2010), 022306, 10. MR 2011b:53048

[Råde 1992] J. Råde, "On the Yang–Mills heat equation in two and three dimensions", *J. Reine Angew. Math.* **431** (1992), 123–163. MR 94a:58041 Zbl 0760.58041

[Ramadas et al. 1989] T. R. Ramadas, I. M. Singer, and J. Weitsman, "Some comments on Chern–Simons gauge theory", *Comm. Math. Phys.* **126**:2 (1989), 409–420. MR 90m:58218 Zbl 0686.53066

[Reznikov 1996] A. Reznikov, "Rationality of secondary classes", *J. Differential Geom.* **43**:3 (1996), 674–692. MR 98b:14006 Zbl 0874.32009

[Salamon 1995] D. Salamon, "Lagrangian intersections, 3-manifolds with boundary, and the Atiyah–Floer conjecture", pp. 526–536 in *Proceedings of the International Congress of Mathematicians* (Zürich, 1994), vol. 1, edited by S. D. Chatterji, Birkhäuser, 1995. MR 97e:57040 Zbl 0854.58021

[Taubes 1982] C. H. Taubes, "Self-dual Yang–Mills connections on non-self-dual 4-manifolds", *J. Differential Geom.* **17**:1 (1982), 139–170. MR 83i:53055 Zbl 0484.53026

[Uhlenbeck 1982] K. K. Uhlenbeck, "Connections with L^p bounds on curvature", *Comm. Math. Phys.* **83**:1 (1982), 31–42. MR 83e:53035 Zbl 0499.58019

[Wehrheim 2004] K. Wehrheim, *Uhlenbeck compactness*, EMS Series of Lectures in Mathematics **1**, European Mathematical Society, Zürich, 2004. MR 2004m:53045 Zbl 1055.53027

[Wehrheim 2005] K. Wehrheim, "Lagrangian boundary conditions for anti-self-dual instantons and the Atiyah–Floer conjecture", *J. Symplectic Geom.* **3**:4 (2005), 703–747. MR 2007j:53114 Zbl 0559.57001

[Wehrheim 2006] K. Wehrheim, "Energy identity for anti-self-dual instantons on $\mathbb{C} \times \Sigma$", *Math. Res. Lett.* **13**:1 (2006), 161–166. MR 2006h:53019 Zbl 1111.53023

[Wehrheim and Woodward 2009] K. Wehrheim and C. Woodward, "Floer field theory", Preprint, 2009, Available at https://math.berkeley.edu/~katrin/papers/fielda.pdf.

[Yeung 1991] S.-K. Yeung, "Integrality of characteristic numbers on complete Kähler manifolds", *Math. Ann.* **289**:3 (1991), 491–516. MR 92b:32015

Received October 7, 2014. Revised March 25, 2015.

DAVID L. DUNCAN
DEPARTMENT OF MATHEMATICS
MICHIGAN STATE UNIVERSITY
619 RED CEDAR ROAD
EAST LANSING, MI 48824
UNITED STATES

duncan42@math.msu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.41

COMPACTNESS AND THE PALAIS–SMALE PROPERTY FOR CRITICAL KIRCHHOFF EQUATIONS IN CLOSED MANIFOLDS

EMMANUEL HEBEY

We prove the Palais–Smale property and the compactness of solutions for critical Kirchhoff equations using solely energy arguments in the situation where no sign assumption is made on the solutions. We then prove the existence of a mountain-pass solution to the equation, discuss its ground-states structure, and, in extreme cases, prove uniqueness of this solution.

The Kirchhoff equation [1883] was proposed as an extension of the classical wave equation of D'Alembert for the vibration of elastic strings. The model takes into account the small vertical vibrations of a stretched elastic string when the tension is variable but the ends of the string are fixed. The equation in [loc. cit.] was written as

$$\rho \frac{\partial^2 u}{\partial t^2} - \left(\frac{P_0}{h} + \frac{E}{2L}\int_0^L \left|\frac{\partial u}{\partial x}\right|^2 dx\right)\frac{\partial^2 u}{\partial x^2} = 0,$$

where L is the length of the string, h is the area of the cross-section, E is the young modulus of the material (also referred to as the elastic modulus — it measures the string's resistance to being deformed elastically), ρ is the mass density, and P_0 is the initial tension. Almost one century later, Jacques-Louis Lions [1978] returned to the equation and proposed a general Kirchhoff equation in arbitrary dimension with external force term which was written as

$$\frac{\partial^2 u}{\partial t^2} + \left(a + b\int_\Omega |\nabla u|^2\, dx\right)\Delta u = f(x, u),$$

where

$$\Delta = -\sum \frac{\partial^2}{\partial x_i^2}$$

is the Laplace–Beltrami Euclidean Laplacian. We investigate in this paper the stationary version of this equation, in the case of closed manifolds, and when f is the critical pure power nonlinearity. We prove the surprising result that the equation satisfies the Palais–Smale property when a and b are large (in a sense to be made

MSC2010: 58J05.
Keywords: compactness, ground-states, Kirchhoff equation, mountain-pass solution, Palais–Smale property.

precised in Theorem 1 below). As usual, solutions of the stationary equation (with the square of the phase added as a potential) correspond to standing wave solutions of the evolution equation.

In what follows, we let (M, g) be a closed n-dimensional Riemannian manifold of dimension $n \geq 4$, $a, b > 0$ be positive real numbers, and $h \in C^1(M, \mathbb{R})$. The Kirchhoff equation we investigate is written as

$$\left(a + b\int_M |\nabla u|^2 \, dv_g\right)\Delta_g u + hu = |u|^{2^\star - 2}u, \tag{1}$$

where $2^\star = 2n/(n-2)$ is the critical Sobolev exponent. It is an appealing mathematical model because of its nonlocal nature and its integrodifferential structure. It has been paid much attention over the past years. Among other possible references (the following list is far from being exhaustive), we mention Figueiredo [2013], Figueiredo, Ikoma, and Santos [Figueiredo et al. 2014], Figueiredo and Santos [2012], He and Zou [2012], and the references in these papers. The case of positive solutions in the curved setting of closed manifolds has been investigated in Hebey and Thizy [2015a; 2015b]. We treat here the case where absolutely no sign assumption is made on the solutions. As a remark, the equation always has a pair of constant solutions if $h > 0$ is constant.

In what follows, we let H^1 be the Sobolev space of functions in L^2 with one derivative in L^2. We let also $I : H^1 \to \mathbb{R}$ be the functional

$$I(u) = \frac{a}{2}\int_M |\nabla u|^2 \, dv_g + \frac{b}{4}\left(\int_M |\nabla u|^2 \, dv_g\right)^2 + \frac{1}{2}\int_M hu^2 \, dv_g - \frac{1}{2^\star}\int_M |u|^{2^\star} dv_g. \tag{2}$$

As is easily checked, critical points of I are solutions of (1). In particular, (1) has a variational structure. A sequence $(u_\alpha)_\alpha$ in H^1 is said to be a *Palais–Smale sequence* for I if the sequence $(I(u_\alpha))_\alpha$ is bounded with respect to α, and $I'(u_\alpha) \to 0$ in $(H^1)'$ as $\alpha \to +\infty$. Following standard terminology, we say that I satisfies the *Palais–Smale property* if Palais–Smale sequences for I converge, up to a subsequence, in H^1. Let S_n be the sharp Euclidean Sobolev constant given by $S_n = \frac{1}{4}n(n-2)\omega_n^{2/n}$, where ω_n is the volume of the unity n-sphere. We define the dimensional constant $C(n)$ by

$$C(n) = \frac{2(n-4)^{(n-4)/2}}{(n-2)^{(n-2)/2} S_n^{n/2}}. \tag{3}$$

The main result of this paper provides very simple criteria on a and b for the equation to be compact and I to satisfy the Palais–Smale property. Our main result is stated as follows.

Theorem 1. *Suppose that (M, g) is a closed n-dimensional Riemannian manifold of dimension $n \geq 4$, that $a, b > 0$ are positive real numbers, and that $h \in C^1(M, \mathbb{R})$ makes $\Delta_g + h/a$ positive. Assume that $b \gg 1$ when $n = 4$, and that $a^{(n-4)/2} b > C(n)$*

when $n \geq 5$, where $C(n)$ is as in (3). Then, I satisfies the Palais–Smale property and the set of solutions of (1) is compact in the C^2-topology.

It is very surprising that such a compactness result, in strong topologies, for an equation with critical nonlinearity, can be obtained without the whole machinery of strong pointwise estimates (see Hebey [2014] for a reference in book form on this machinery). Moreover, no assumption of positiveness is made on the solutions in Theorem 1.

Proof of Theorem 1. (i) We prove that Palais–Smale sequences for I are bounded in H^1, assuming that $b \gg 1$ when $n = 4$. Let $(u_\alpha)_\alpha$ be a Palais–Smale sequence for I. Then, we get that $I(u_\alpha) = O(1)$ and $I'(u_\alpha) . (u_\alpha) = o(\|u_\alpha\|_{H^1})$, where $\|\cdot\|_{H^1}$ is the H^1-norm given for $u \in H^1$ by

$$\|u\|_{H^1}^2 = \|\nabla u\|_{L^2}^2 + \|u\|_{L^2}^2.$$

In particular,

$$a\int_M\left(|\nabla u_\alpha|^2 + \frac{h}{a}u_\alpha^2\right)dv_g + b\left(\int_M |\nabla u_\alpha|^2\,dv_g\right)^2 = \int_M |u_\alpha|^{2^\star}dv_g + o(\|u_\alpha\|_{H^1}) \tag{4}$$

and that

$$\frac{a}{2}\int_M\left(|\nabla u_\alpha|^2 + \frac{h}{a}u_\alpha^2\right)dv_g + \frac{b}{4}\left(\int_M |\nabla u_\alpha|^2 dv_g\right)^2 = \frac{1}{2^\star}\int_M |u_\alpha|^{2^\star}dv_g + O(1). \tag{5}$$

By the Sobolev–Poincaré inequality, there exist $C_1, C_2 > 0$ such that

$$\|u_\alpha\|_{L^{2^\star}}^{2^\star} \leq C_1\|\nabla u_\alpha\|_{L^2}^{2^\star} + C_2|\bar{u}_\alpha|^{2^\star} \tag{6}$$

for all α, where

$$\bar{u}_\alpha = \frac{1}{V_g}\int_M u_\alpha\,dv_g$$

is the average of u_α, and by the Poincaré inequality,

$$\|u_\alpha - \bar{u}_\alpha\|_{L^2}^2 \leq \frac{1}{\lambda_1}\|\nabla u_\alpha\|_{L^2}^2 \tag{7}$$

for all α, where $\lambda_1 = \lambda_1(M, g) > 0$ is the first nonzero eigenvalue of Δ_g. It clearly follows from the positivity of $\Delta_g + h/a$, (5), and (6) that if either $n = 4$ and $b > C_1$ or $n \geq 5$ and if $\bar{u}_\alpha = O(1)$, then $\|u_\alpha\|_{H^1} = O(1)$. We may therefore assume that $\bar{u}_\alpha \to +\infty$ as $\alpha \to +\infty$. Then, still by the positivity of $\Delta_g + h/a$, (5), and (6),

$$\int_M |\nabla u_\alpha|^2\,dv_g = \begin{cases} \frac{1}{b}O(\bar{u}_\alpha^2) & \text{if } n = 4, \\ o(\bar{u}_\alpha^2) & \text{if } n \geq 5, \end{cases} \tag{8}$$

where we assume that $b > C_1$ when $n = 4$. Now, we write that

$$u_\alpha = \bar{u}_\alpha(1 + \varphi_\alpha). \tag{9}$$

Then, $\int_M \varphi_\alpha \, dv_g = 0$ and

$$\bar{u}_\alpha^2 \int_M |\nabla \varphi_\alpha|^2 \, dv_g = \int_M |\nabla u_\alpha|^2 \, dv_g. \tag{10}$$

It follows from (8), (10), the Poincaré inequality, (7), and (10) that

$$\|\varphi_\alpha\|_{H^1}^2 = \begin{cases} O\big(\frac{1}{b}\big) & \text{if } n = 4, \\ o(1) & \text{if } n \geq 5. \end{cases} \tag{11}$$

In particular, by (9) and (11),

$$\int_M \Big(|\nabla u_\alpha|^2 + \frac{h}{a} u_\alpha^2\Big) \, dv_g = \bar{u}_\alpha^2 (1 + A_\alpha) \quad \text{and} \quad \int_M |u_\alpha|^{2^\star} \, dv_g = \bar{u}_\alpha^{2^\star} (1 + B_\alpha), \tag{12}$$

where $A_\alpha = O\big(\frac{1}{b}\big)$ and $B_\alpha = O\big(\frac{1}{b}\big)$ if $n = 4$, and $A_\alpha = o(1)$ and $B_\alpha = o(1)$ if $n \geq 5$. Subtracting $\frac{1}{4}$ of (4) from (5) yields

$$\frac{a}{4} \int_M \Big(|\nabla u_\alpha|^2 + \frac{h}{a} u_\alpha^2\Big) \, dv_g = \Big(\frac{1}{2^\star} - \frac{1}{4}\Big) \int_M |u_\alpha|^{2^\star} dv_g + O(1) + O(\|u_\alpha\|_{H^1}). \tag{13}$$

Picking $b \gg 1$ when $n = 4$, the contradiction follows by combining (12) and (13). This proves that $(u_\alpha)_\alpha$ is bounded in H^1.

(ii) We prove that I satisfies the Palais–Smale property assuming that $b \gg 1$ when $n = 4$, and that $a^{(n-4)/2} b > C(n)$ when $n \geq 5$. We let

$$K_\alpha = a + b \int_M |\nabla u_\alpha|^2 \, dv_g, \tag{14}$$

$h_\alpha = K_\alpha^{-1} h$, and

$$v_\alpha = \Big(\frac{1}{K_\alpha}\Big)^{\frac{1}{2^\star - 2}} u_\alpha. \tag{15}$$

We define $I_\alpha : H^1 \to \mathbb{R}$ by

$$I_\alpha(u) = \frac{1}{2} \int_M \big(|\nabla u|^2 + h_\alpha u^2\big) \, dv_g - \frac{1}{2^\star} \int_M |u|^{2^\star} dv_g. \tag{16}$$

According to (i), and up to passing to a subsequence, $K_\alpha \to K_\infty$ as $\alpha \to +\infty$ for some $K_\infty \geq a$. In particular, $(h_\alpha)_\alpha$ converges in C^k for all k, and $(v_\alpha)_\alpha$ is bounded in H^1. This implies that $I_\alpha(v_\alpha) = O(1)$, and, as one can check,

$$I'_\alpha(v_\alpha) \,.\, (\varphi) = \Big(\frac{1}{K_\alpha}\Big)^{\frac{2^\star - 1}{2^\star - 2}} I'(u_\alpha) \,.\, (\varphi)$$

for all $\varphi \in H^1$. Then $(v_\alpha)_\alpha$ is a Palais–Smale sequence for the family $(I_\alpha)_\alpha$ (in the sense of Hebey [2014]). In particular the H^1-decomposition as in Struwe [1984] applies (see Druet, Hebey, and Robert [Druet et al. 2004], Hebey [2014], and Vétois

[2007] for the closed setting with varying potentials), and we get that there exists $v_\infty \in H^1$, $k \in \mathbb{N}$, and $k+1$ sequences $(B_{1,\alpha})_\alpha, \dots, (B_{k,\alpha})_\alpha$, $(R_\alpha)_\alpha$ in H^1 such that

$$v_\alpha = v_\infty + \sum_{i=1}^{k} B_{i,\alpha} + R_\alpha \text{in } M \tag{17}$$

and

$$\int_M |\nabla v_\alpha|^2 \, dv_g = \int_M |\nabla v_\infty|^2 \, dv_g + \sum_{i=1}^{k} \int_M |\nabla B_{i,\alpha}|^2 \, dv_g + o(1) \tag{18}$$

for all α, $R_\alpha \to 0$ in H^1 as $\alpha \to +\infty$ and the "bubbles" $(B_{i,\alpha})_\alpha$ satisfy the following properties for any $i = 1, \dots, k$:

(a) $B_{i,\alpha} \to 0$ in L^2 as $\alpha \to +\infty$,

(b) $\|B_{i,\alpha}\| = O(1)$, and

(c) $\int_M |\nabla B_{i,\alpha}|^2 \, dv_g \geq S_n^{n/2} + o(1)$ for all α,

where S_n is the sharp Euclidean constant as in (3). In (c), there is equality if each $B_{i,\alpha}$ is positive. Then, since $(u_\alpha)_\alpha$ is bounded in H^1, and by (17)–(18), we get that, up to passing to a subsequence,

$$\begin{aligned} K_\alpha &= a + b \int_M |\nabla u_\alpha|^2 \, dv_g \\ &= a + b K_\alpha^{2/(2^\star - 2)} \int_M |\nabla v_\alpha|^2 \, dv_g \\ &= a + b K_\alpha^{2/(2^\star - 2)} \int_M |\nabla v_\infty|^2 \, dv_g + b K_\alpha^{2/(2^\star - 2)} \sum_{i=1}^{k} \int_M |\nabla B_{i,\alpha}|^2 \, dv_g + o(1) \\ &= a + b K_\infty^{2/(2^\star - 2)} \int_M |\nabla v_\infty|^2 \, dv_g + b C K_\infty^{2/(2^\star - 2)} + o(1), \end{aligned} \tag{19}$$

where $C \geq k S_n^{n/2}$. In particular, by (19),

$$K_\infty = a + b K_\infty^{2/(2^\star - 2)} \int_M |\nabla v_\infty|^2 \, dv_g + b C K_\infty^{2/(2^\star - 2)}. \tag{20}$$

When $n = 4$, we have $2/(2^\star - 2) = 1$, and (20) implies that $k = 0$ in (17) as soon as $b \gg 1$. In particular, the sequence $(u_\alpha)_\alpha$ converges strongly in H^1, and I satisfies the Palais–Smale property. When $n \geq 5$, we define

$$f(x) = bk S_n^{n/2} x^{(n-2)/2} - x + a.$$

By (20), and since $C \geq kS_n^{n/2}$, we have that $f(K_\infty) \leq 0$. Assuming that $k \geq 1$, noting that f is minimum at x_0, where

$$x_0 = \left(\frac{2}{(n-2)bkS_n^{n/2}}\right)^{2/(n-4)},$$

we compute that

$$f(x_0) = -\frac{n-4}{n-2}\left(bkS_n^{n/2}\right)^{-2/(n-4)}\left(\frac{2}{n-2}\right)^{2/(n-4)} + a. \tag{21}$$

If $f(K_\infty) \leq 0$, then $f(x_0) \leq 0$, and by (21), $bka^{(n-4)/2} \leq C(n)$. Since by assumption $a^{(n-4)/2}b > C(n)$, it must be the case that $k = 0$ in (17). In particular, the sequence $(u_\alpha)_\alpha$ converges strongly in H^1, and I satisfies the Palais–Smale property also when $n \geq 5$.

(iii) We prove the compactness of (1) assuming that $b \gg 1$ when $n = 4$ and that $a^{(n-4)/2}b > C(n)$ when $n \geq 5$. Noting that a bounded sequence in H^1 of solutions of (1) is a Palais–Smale sequence for I, according to what we proved above, it suffices to prove that if $(u_\alpha)_\alpha$ is a sequence of solutions of (1), then $(u_\alpha)_\alpha$ is bounded in H^1 when $n \geq 5$ and when $n = 4$ and $b \gg 1$. By the Palais–Smale property we would indeed get that, up to passing to a subsequence, $(u_\alpha)_\alpha$ converges in H^1, and by standard elliptic theory, this actually implies that the sequence converges in C^2. Now, we multiply the equation by u_α and integrate over M, yielding

$$a\int_M \left(|\nabla u_\alpha|^2 + \frac{h}{a}u_\alpha^2\right) dv_g + b\left(\int_M |\nabla u_\alpha|^2\, dv_g\right)^2 = \int_M |u_\alpha|^{2^\star} dv_g \tag{22}$$

for all α. We clearly get from (6) and (22) that $(u_\alpha)_\alpha$ is bounded in H^1 if the sequence $(\bar{u}_\alpha)_\alpha$ is bounded (and $b \gg 1$ when $n = 4$). We may thus assume that $\bar{u}_\alpha \to +\infty$ as $\alpha \to +\infty$. By (6) and (22), we get that (8) holds. Writing (9), we then get that (12) holds and also that

$$\int_M |u_\alpha|\, dv_g = |\bar{u}_\alpha|(1 + C_\alpha), \tag{23}$$

where $C_\alpha = O\left(\frac{1}{b}\right)$ if $n = 4$, and $C_\alpha = o(1)$ if $n \geq 5$. Integrating the equation,

$$\int_M hu_\alpha\, dv_g = \int_M |u_\alpha|^{2^\star - 2} u_\alpha\, dv_g. \tag{24}$$

The contradiction follows from (12), (23), and (24). This proves the above claim that $(u_\alpha)_\alpha$ is bounded in H^1. This also proves that the set of solutions of (1) is compact in the C^2-topology. □

At this point we define a mountain-pass solution of (1) as a solution which we obtain from I by the use of the mountain-pass lemma. We easily get from Theorem 1 that the following existence result holds true.

Proposition 2. *Suppose that* (M, g) *is a closed Riemannian manifold of dimension* $n \geq 4$, *that* a *and* b *are positive real numbers, and that* $h \in C^1(M, \mathbb{R})$ *is such that* $\Delta_g + h/a$ *is positive. Assume that* $b \gg 1$ *when* $n = 4$, *and that* $a^{(n-4)/2}b > C(n)$ *when* $n \geq 5$, *where* $C(n)$ *is as in* (3). *Then,* (1) *possesses a nontrivial mountain-pass solution.*

Proof of Proposition 2. Let $u_0 \equiv 1$. Then, I is C^1, $I(0) = 0$, $I(Tu_0) < 0$ for $T \gg 1$, and by the coercivity of $\Delta_g + h/a$, there exist $C_1, C_2 > 0$ such that $I(u) \geq C_1\|u\|_{H^1}^2 - C_2\|u\|_{H^1}^{2^\star}$ for all u. Then, we can apply the mountain-pass lemma of Ambrosetti and Rabinowitz [1973] and we get that there exists a sequence $(u_\alpha)_\alpha$ in H^1 such that $I(u_\alpha) = c + o(1)$ and $I'(u_\alpha).(\psi) = o(\|\psi\|_{H^1})$ for all $\psi \in H^1$, where

$$c = \inf_{\gamma \in \Gamma} \sup_{u \in \gamma} I(u),$$

and Γ is the set of continuous paths from 0 to Tu_0. Obviously, $c > 0$. By Theorem 1, up to passing to a subsequence, $(u_\alpha)_\alpha$ converges in H^1. Let u_∞ be the limit in H^1 of the sequence u_α. Then $I(u_\infty) = c$, $u_\infty \not\equiv 0$, and by passing to the limit in the equation $I'(u_\alpha).(\varphi) = o(1)$ for all $\varphi \in H^1$, we get that u_∞ solves (1). $\square$

It is easily seen that the mountain-pass solution u_∞ obtained in Proposition 2 has a nice ground-state structure when $n = 4$. We define the Nehari manifold $\mathcal{N}$ attached to I by

$$\mathcal{N} = \{u \in H^1 \setminus \{0\} \mid I'(u).(u) = 0\}. \tag{25}$$

The following 4-dimensional ground-state characterization of the solution obtained in Proposition 2 holds true.

Proposition 3. *Suppose that* (M, g) *is a closed* 4*-dimensional Riemannian manifold, that* a *and* b *are positive real numbers, and that* $h \in C^1(M, \mathbb{R})$ *is such that* $\Delta_g + h/a$ *is positive. Assume that* $b \gg 1$. *Then, the mountain-pass solution* u_∞ *obtained in Proposition 2 has a ground-state structure given by*

$$I(u_\infty) = \inf_{u \in \mathcal{N}} I(u), \tag{26}$$

where $\mathcal{N}$ *is the Nehari manifold attached to* I *given by* (25).

Proof of Proposition 3. We obviously have that $u_\infty \in \mathcal{N}$, and thus there holds that $I(u_\infty) \geq \inf_{u \in \mathcal{N}} I(u)$. Given $\tilde{u} \in H^1 \setminus \{0\}$, we define the mountain-pass energy level $c_{\tilde{u}}$ by

$$c_{\tilde{u}} = \inf_{\gamma \in \Gamma_{\tilde{u}}} \sup_{u \in \gamma} I(u),$$

where $\Gamma_{\tilde{u}}$ is the set of continuous paths from 0 to $\tilde{u}$. Let $u_0 \equiv 1$ be as in the proof of Proposition 2. Let $T_0 \gg 1$ be such that $I(T_0u_0) < 0$. By construction (see the proof of Proposition 2), it holds that $I(u_\infty) = c_{T_0 u_0}$. Let $u \in \mathcal{N}$. Then $I(u) = I(|u|)$,

$$a\int_M \left(|\nabla u|^2 + \frac{h}{a}u^2\right) dv_g + b\left(\int_M |\nabla u|^2\, dv_g\right)^2 = \int_M u^4\, dv_g$$

and for $t \geq 0$,

$$\begin{aligned}(27)\quad I(t|u|) &= \frac{at^2}{2}\int_M \left(|\nabla u|^2 + \frac{h}{a}u^2\right) dv_g + \frac{bt^4}{4}\left(\int_M |\nabla u|^2\, dv_g\right)^2 - \frac{t^4}{4}\int_M u^4\, dv_g \\ &= \frac{at^2(2-t^2)}{4}\int_M \left(|\nabla u|^2 + \frac{h}{a}u^2\right) dv_g.\end{aligned}$$

In particular, $I(T_1|u|) < 0$ for $T_1 > \sqrt{2}$. Let $u_1 = |u|$ and $T_1 \gg 1$. It is easily checked (since u_0 is constant) that

$$I(tT_1u_1 + (1-t)T_0u_0) \leq t^2 I(T_1u_1) - \frac{(1-t)^2 T_0^2 u_0^2 V_g}{4} < 0$$

for all $0 \leq t \leq 1$, where V_g is the volume of (M, g). In particular, $c_{T_0u_0} = c_{T_1u_1}$ since T_0u_0 and T_1u_1 can be connected by a continuous path along which I is everywhere negative. So,

$$(28)\qquad c_{T_0u_0} \leq \sup_{0\leq t\leq T_1} I(tu_1).$$

By (27) we see that $t \to I(tu_1)$ is maximal at $t = 1$, and thus $c_{T_0u_0} \leq I(u)$ by (28). This proves that $I(u_\infty) \leq I(u)$ for all $u \in \mathcal{N}$, and thus that (26) holds. □

Balancing Proposition 2 we prove that the following uniqueness result, in the sense of Brézis and Li [2006], holds.

Proposition 4. *Suppose that (M, g) is a closed Riemannian manifold of dimension $n \geq 4$ and that h is a positive constant. Let $\varepsilon_0 > 0$ arbitrary. For $a, b \gg 1$ when $n = 4$, and $a \gg 1$, $b \geq \varepsilon_0$ when $n \geq 5$, the sole nontrivial pair of solutions of (1) is the pair $(-u, u)$ of constant solutions, where $u = h^{(n-2)/4}$.*

Proof of Proposition 4. Let $\varepsilon_0 > 0$ be given arbitrarily small. We prove the result by contradiction. We assume that there exist sequences $(a_\alpha)_\alpha$, $(b_\alpha)_\alpha$ of positive real numbers, and a sequence $(u_\alpha)_\alpha$ of nonconstant solutions of

$$(29)\qquad \left(a_\alpha + b_\alpha\int_M |\nabla u_\alpha|^2\, dv_g\right)\Delta_g u_\alpha + hu_\alpha = |u_\alpha|^{2^\star-2}u_\alpha$$

for all α such that $a_\alpha \to +\infty$ and $b_\alpha \to +\infty$ as $\alpha \to +\infty$ when $n = 4$, and such that $a_\alpha \to +\infty$ as $\alpha \to +\infty$ and $b_\alpha \geq \varepsilon_0$ for all α when $n \geq 5$. As in the proof of Theorem 1, this implies that $\|u_\alpha\|_{H^1} = O(1)$. Suppose that K_α is as in (14), $h_\alpha = K_\alpha^{-1}h$, and v_α is as in (15). Then,

$$\Delta_g v_\alpha + h_\alpha v_\alpha = |v_\alpha|^{2^\star-2} v_\alpha, \tag{30}$$

and $K_\alpha \to +\infty$ since $a_\alpha \to +\infty$ as $\alpha \to +\infty$. Then, by elliptic regularity, $v_\alpha \to 0$ in C^0. Multiplying (30) by $v_\alpha - \bar{v}_\alpha$, and integrating over M,

$$\begin{aligned}\lambda_1 \int_M (v_\alpha - \bar{v}_\alpha)^2 \, dv_g &\le \int_M (v_\alpha - \bar{v}_\alpha)\big(|v_\alpha|^{2^\star-2} v_\alpha - |\bar{v}_\alpha|^{2^\star-2}\bar{v}_\alpha\big)\, dv_g \\ &\le C \|v_\alpha\|_{L^\infty}^{2^\star-2} \int_M (v_\alpha - \bar{v}_\alpha)^2 \, dv_g\end{aligned} \tag{31}$$

for all α, where $C > 0$ is independent of α, and $\lambda_1 > 0$ is the first nontrivial eigenvalue of Δ_g. Since $v_\alpha \to 0$ in C^0, (31) implies that $v_\alpha = \bar{v}_\alpha$, and we get a contradiction. □

Acknowledgements

The author would like to thank Pierre-Damien Thizy for helpful and valuable comments on the manuscript, and for pointing out the very nice Proposition 4.

References

[Ambrosetti and Rabinowitz 1973] A. Ambrosetti and P. H. Rabinowitz, "Dual variational methods in critical point theory and applications", *J. Functional Analysis* **14** (1973), 349–381. MR 51 #6412 Zbl 0273.49063

[Brezis and Li 2006] H. Brezis and Y. Li, "Some nonlinear elliptic equations have only constant solutions", *J. Partial Differential Equations* **19**:3 (2006), 208–217. MR 2007f:35076 Zbl 1174.35368

[Druet et al. 2004] O. Druet, E. Hebey, and F. Robert, *Blow-up theory for elliptic PDEs in Riemannian geometry*, Mathematical Notes **45**, Princeton University Press, 2004. MR 2005g:53058 Zbl 1059.58017

[Figueiredo 2013] G. M. Figueiredo, "Existence of a positive solution for a Kirchhoff problem type with critical growth via truncation argument", *J. Math. Anal. Appl.* **401**:2 (2013), 706–713. MR 3018020 Zbl 1307.35110

[Figueiredo and Santos 2012] G. M. Figueiredo and J. R. Santos, Jr., "Multiplicity of solutions for a Kirchhoff equation with subcritical or critical growth", *Differential Integral Equations* **25**:9-10 (2012), 853–868. MR 2985683 Zbl 1274.35087

[Figueiredo et al. 2014] G. M. Figueiredo, N. Ikoma, and J. R. Santos, Jr., "Existence and concentration result for the Kirchhoff type equations with general nonlinearities", *Arch. Ration. Mech. Anal.* **213**:3 (2014), 931–979. MR 3218834 Zbl 1302.35356

[He and Zou 2012] X. He and W. Zou, "Existence and concentration behavior of positive solutions for a Kirchhoff equation in $\mathbb{R}^3$", *J. Differential Equations* **252**:2 (2012), 1813–1834. MR 2853562 Zbl 1235.35093

[Hebey 2014] E. Hebey, *Compactness and stability for nonlinear elliptic equations*, European Mathematical Society, Zürich, 2014. MR 3235821 Zbl 1305.58001

[Hebey and Thizy 2015a] E. Hebey and P. D. Thizy, "Stationary Kirchhoff systems in closed 3-dimensional manifolds", *Calc. Var. Partial Differential Equations* (online publication April 2015).

[Hebey and Thizy 2015b] E. Hebey and P. D. Thizy, "Stationary Kirchhoff systems in closed high dimensional manifolds", *Commun. Contemp. Math.* (online publication April 2015).

[Kirchhoff 1883] G. Kirchhoff, *Vorlesungen über Mechanik*, 3rd ed., Vorlesungen über Mathematische Physik **1**, Teubner, Leipzig, 1883. JFM 09.0597.01

[Lions 1978] J.-L. Lions, "On some questions in boundary value problems of mathematical physics", pp. 284–346 in *Contemporary developments in continuum mechanics and partial differential equations* (Rio de Janeiro, 1977), edited by G. M. de la Penha and L. A. J. Medeiros, North-Holland Math. Stud. **30**, North-Holland, Amsterdam, 1978. MR 82b:35020 Zbl 0404.35002

[Struwe 1984] M. Struwe, "A global compactness result for elliptic boundary value problems involving limiting nonlinearities", *Math. Z.* **187**:4 (1984), 511–517. MR 86k:35046 Zbl 0535.35025

[Vétois 2007] J. Vétois, "Multiple solutions for nonlinear elliptic equations on compact Riemannian manifolds", *Internat. J. Math.* **18**:9 (2007), 1071–1111. MR 2009f:53051 Zbl 1148.35026

Received April 17, 2015. Revised May 3, 2015.

EMMANUEL HEBEY
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ DE CERGY-PONTOISE
2 AVENUE ADOLPHE CHAUVIN
95302 CERGY-PONTOISE
FRANCE

Emmanuel.Hebey@math.u-cergy.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.51

ON THE EQUIVALENCE OF THE DEFINITIONS OF VOLUME OF REPRESENTATIONS

SUNGWOON KIM

Let G be a rank-1 simple Lie group and let M be a connected, orientable, aspherical, tame manifold. Assume that each end of M has amenable fundamental group. There are several definitions of volume of representations of $\pi_1(M)$ into G. We give a new definition of volume of representations and furthermore show that all definitions so far are equivalent.

1. Introduction

Let G be a semisimple Lie group and let $\mathcal{X}$ be the associated symmetric space of dimension n. Let M be a connected, orientable, aspherical, tame manifold of the same dimension as $\mathcal{X}$. First assume that M is compact. To each representation $\rho : \pi_1(M) \to G$, one can associate a volume of ρ in the following way. First, associate a flat bundle E_ρ over M with fiber $\mathcal{X}$ to ρ. Since $\mathcal{X}$ is contractible, there always exists a section $s : M \to E_\rho$. Let $\omega_{\mathcal{X}}$ be the Riemannian volume form on $\mathcal{X}$. One may think of $\omega_{\mathcal{X}}$ as a closed differential form on E_ρ by spreading $\omega_{\mathcal{X}}$ over the fibers of E_ρ. Then the volume of ρ is defined by

$$\mathrm{Vol}(\rho) = \int_M s^*\omega_{\mathcal{X}}.$$

Since any two sections are homotopic to each other, the volume $\mathrm{Vol}(\rho)$ does not depend on the choice of section.

The volume of representations has been used to characterize discrete faithful representations. Let Γ be a uniform lattice in G. Then the volume of representations satisfies a Milnor–Wood type inequality. More precisely, for any representation $\rho : \Gamma \to G$, we have

$$|\mathrm{Vol}(\rho)| \leq \mathrm{Vol}(\Gamma \backslash \mathcal{X}). \tag{1}$$

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2012R1A1A2040663).
MSC2010: 53C24, 53C35.
Keywords: Lie groups, representation, volume, symmetric spaces, disjoint cone.

Furthermore, equality holds in (1) if and only if ρ is discrete and faithful. This is the so-called *volume rigidity theorem*. Goldman [1982] proved the volume rigidity theorem in the higher rank case and Besson, Courtois and Gallot [Besson et al. 2007] proved the theorem in the rank-1 case.

Now assume that M is noncompact. Then the definition of volume of representations as above is not valid anymore since some problems of integrability arise. So far, three definitions of volume of representations have been given under some conditions on M. Let us first fix the following notation throughout the paper.

Setup. Let M be a noncompact, connected, orientable, aspherical, tame manifold. Denote by $\bar{M}$ the compact manifold with boundary whose interior is homeomorphic to M. Assume that each connected component of $\partial\bar{M}$ has amenable fundamental group. Let G be a rank-1 semisimple Lie group with trivial center and no compact factors. Let $\mathcal{X}$ be the associated symmetric space of dimension n. Assume that M has the same dimension as $\mathcal{X}$.

First of all, Dunfield [1999] introduced the notion of pseudodeveloping map to define the volume of representations of a nonuniform lattice Γ in $\mathrm{SO}(3, 1)$. It was successful in making an invariant associated with a representation $\rho:\Gamma\to \mathrm{SO}(3, 1)$ but he did not prove that the volume of representations does not depend on the chosen pseudodeveloping map. After that, Francaviglia [2004] proved the well-definedness of the volume of representations. Then Francaviglia and Klaff [2006] extended the definition of volume of representations and the volume rigidity theorem to general nonuniform hyperbolic lattices. We call the definition of volume of representations via pseudodeveloping map D1. For more detail about D1, see [Francaviglia and Klaff 2006] or Section 4.

The second definition D2 of volume of representations was given by Bucher, Burger and Iozzi [Bucher et al. 2013] and generalizes the one introduced in [Burger et al. 2010] for noncompact surfaces. They used the theory of bounded cohomology to make an invariant associated with a representation. Given a representation $\rho:\pi_1(M)\to G$, one cannot get any information from the pullback map in degree n in continuous cohomology, $\rho^*_c : H^n_c(G,\mathbb{R})\to H^n(\pi_1(M),\mathbb{R})$, since $H^n(\pi_1(M),\mathbb{R})\cong H^n(M,\mathbb{R})$ is trivial. However, the situation is different in continuous bounded cohomology. Not only may the pullback map $\rho^*_b : H^n_{c,b}(G,\mathbb{R})\to H^n_b(\pi_1(M),\mathbb{R})$ be nontrivial but it also encodes subtle algebraic and topological properties of a representation such as injectivity and discreteness. Bucher, Burger and Iozzi gave a proof of the volume rigidity theorem for representations of hyperbolic lattices from the point of view of bounded cohomology. We refer the reader to [Bucher et al. 2013] or Section 2 for further discussion about D2.

Recently, S. Kim and I. Kim [2014] gave a new definition, called D3, of volume of representations in the case that M is a complete Riemannian manifold with finite Lipschitz simplicial volume. See [Kim and Kim 2014] or Section 5 for the exact definition of D3. In D3, it is not necessary that each connected component of $\partial \bar{M}$ has amenable fundamental group, while the amenable condition on $\partial \bar{M}$ is necessary in D2. They only use the bounded cohomology and ℓ^1-homology of M. It is quite useful to define the volume of representations in the case that the amenable condition on $\partial \bar{M}$ does not hold. They give a proof of the volume rigidity theorem for representations of lattices in an arbitrary semisimple Lie group in their setting.

In this note, we will give another definition of volume of representations, called D4. In D4, ρ-equivariant maps are involved as in D1 and the bounded cohomology of M is involved as in D2 and D3. In fact, D4 seems to be a kind of definition connecting the other definitions D1, D2 and D3. Eventually we show that all the definitions are equivalent.

Theorem 1.1. *Let G be a rank-1 simple Lie group with trivial center and no compact factors. Let M be a noncompact, connected, orientable, aspherical, tame manifold. Suppose that each end of M has amenable fundamental group. Then all definitions D1, D2 and D3 of volume of representations of $\pi_1(M)$ into G are equivalent. Furthermore, if M admits a complete Riemannian metric with finite Lipschitz simplicial volume, all definitions D1, D2, D3 and D4 are equivalent.*

The paper is organized as follows. For our proof, we recall the definitions of volume of representations in the order D2, D4, D1, D3. In Section 2, we first recall definition D2. In Section 3, we give definition D4 and then prove that D2 and D4 are equivalent. In Section 4, after recalling definition D1, we show the equivalence of D1 and D4. Finally in Section 5, we complete the proof of Theorem 1.1 by proving that D3 and D4 are equivalent.

2. Bounded cohomology and definition D2

We choose the appropriate complexes for the continuous cohomology and continuous bounded cohomology of G for our purpose. Consider the complex $C_c^*(\mathcal{X}, \mathbb{R})_{\text{alt}}$ with the homogeneous coboundary operator, where

$$C_c^k(\mathcal{X}, \mathbb{R})_{\text{alt}} = \{f : \mathcal{X}^{k+1} \to \mathbb{R} \mid f \text{ is continuous and alternating}\}.$$

The action of G on $C_c^k(\mathcal{X}, \mathbb{R})_{\text{alt}}$ is given by

$$g \cdot f(x_0, \dots, x_k) = f(g^{-1}x_0, \dots, g^{-1}x_k).$$

Then the continuous cohomology $H_c^*(G, \mathbb{R})$ can be isomorphically computed by the cohomology of the G-invariant complex $C_c^*(\mathcal{X}, \mathbb{R})_{\text{alt}}^G$ (see [Guichardet 1980, Chapitre III]). According to the Van Est isomorphism [Borel and Wallach 2000, Proposition IX.5.5], the continuous cohomology $H_c^*(G, \mathbb{R})$ is isomorphic to the set of G-invariant differential forms on $\mathcal{X}$. Hence, in degree n, $H_c^n(G,\mathbb{R})$ is generated by the Riemannian volume form $\omega_{\mathcal{X}}$ on $\mathcal{X}$.

Let $C_{c,b}^k(\mathcal{X},\mathbb{R})_{\mathrm{alt}}$ be a subcomplex of continuous, alternating, bounded real-valued functions on $\mathcal{X}^{k+1}$. The continuous bounded cohomology $H_{c,b}^*(G,\mathbb{R})$ is obtained by the cohomology of the G-invariant complex $C_{c,b}^*(\mathcal{X},\mathbb{R})_{\mathrm{alt}}^G$ (see [Monod 2001, Corollary 7.4.10]). The inclusion of complexes $C_{c,b}^*(\mathcal{X},\mathbb{R})_{\mathrm{alt}}^G \subset C_c^*(\mathcal{X},\mathbb{R})_{\mathrm{alt}}^G$ induces a comparison map $H_{c,b}^*(G,\mathbb{R}) \to H_c^*(G,\mathbb{R})$.

Let Y be a countable CW-complex. Denote by $C_b^k(Y,\mathbb{R})$ the complex of bounded real-valued k-cochains on Y. For a subspace $B \subset Y$, let $C_b^k(Y,B,\mathbb{R})$ be the subcomplex of those bounded k-cochains on Y that vanish on simplices with image contained in B. The complexes $C_b^*(Y,\mathbb{R})$ and $C_b^*(Y,B,\mathbb{R})$ define the bounded cohomologies $H_b^*(Y,\mathbb{R})$ and $H_b^*(Y,B,\mathbb{R})$ respectively. For our convenience, we give another complex which computes the bounded cohomology $H_b^*(Y,\mathbb{R})$ of Y. Let $C_b^k(\widetilde{Y},\mathbb{R})_{\mathrm{alt}}$ denote the complex of bounded, alternating real-valued Borel functions on $(\widetilde{Y})^{k+1}$. The $\pi_1(Y)$-action on $C_b^*(\widetilde{Y},\mathbb{R})_{\mathrm{alt}}$ is defined as the G-action on $C_c^*(\mathcal{X},\mathbb{R})$. Ivanov [1985] proved that the $\pi_1(Y)$-invariant complex $C_b^*(\widetilde{Y},\mathbb{R})_{\mathrm{alt}}^{\pi_1(Y)}$ defines the bounded cohomology of Y.

Bucher, Burger and Iozzi [Bucher et al. 2013] used bounded cohomology to define the volume of representations. Let $\bar{M}$ be a connected, orientable, compact manifold with boundary. Suppose that each component of $\partial\bar{M}$ has amenable fundamental group. In that case, it is proved in [Bucher et al. 2012; Kim and Kuessner 2015] that the natural inclusion $i:(\bar{M},\varnothing) \to (\bar{M},\partial\bar{M})$ induces an isometric isomorphism in bounded cohomology,

$$i_b^*: H_b^*(\bar{M},\partial\bar{M},\mathbb{R}) \to H_b^*(\bar{M},\mathbb{R}),$$

in degrees $* \geq 2$. Noting the remarkable result of Gromov [1982, Section 3.1] that the natural map $H_b^n(\pi_1(\bar{M}),\mathbb{R}) \to H_b^n(\bar{M},\mathbb{R})$ is an isometric isomorphism in bounded cohomology, for a given representation $\rho:\pi_1(M) \to G$ we have a map

$$\rho_b^*: H_{c,b}^n(G,\mathbb{R}) \to H_b^n(\pi_1(\bar{M}),\mathbb{R}) \cong H_b^n(\bar{M},\mathbb{R}) \cong H_b^n(\bar{M},\partial\bar{M},\mathbb{R}).$$

The G-invariant Riemannian volume form $\omega_{\mathcal{X}}$ on $\mathcal{X}$ gives rise to a continuous bounded cocycle $\Theta:\mathcal{X}^{n+1} \to \mathbb{R}$ defined by

$$\Theta(x_0,\ldots,x_n) = \int_{[x_0,\ldots,x_n]} \omega_{\mathcal{X}},$$

where $[x_0,\ldots,x_n]$ is the geodesic simplex with ordered vertices $x_0,\ldots,x_n$ in $\mathcal{X}$. The boundedness of Θ is due to the fact that the volume of geodesic simplices in $\mathcal{X}$ is uniformly bounded from above [Inoue and Yano 1982]. Hence the cocycle Θ induces a continuous cohomology class $[\Theta]_c \in H_c^n(G,\mathbb{R})$ and, moreover, a continuous bounded cohomology class $[\Theta]_{c,b} \in H_{c,b}^n(G,\mathbb{R})$. The image of $((i_b^*)^{-1} \circ \rho_b^*)[\Theta]_{c,b}$

via the comparison map $c : H^n_b(\bar{M}, \partial\bar{M}, \mathbb{R}) \to H^n(\bar{M}, \partial\bar{M}, \mathbb{R})$ is an ordinary relative cohomology class. Its evaluation on the relative fundamental class $[\bar{M}, \partial\bar{M}]$ gives an invariant associated with ρ.

Definition 2.1 (D2). For a representation $\rho : \pi_1(M) \to G$, define the invariant

$$\mathrm{Vol}_2(\rho) = \langle (c \circ (i^*_b)^{-1} \circ \rho^*_b)[\Theta]_{c,b}, [\bar{M}, \partial\bar{M}] \rangle.$$

In definition D2, a specific continuous bounded volume class $[\Theta]_{c,b}$ in $H^n_{c,b}(G, \mathbb{R})$ is involved. The question is naturally raised as to whether, if another continuous bounded volume class is used in D2 instead of $[\Theta]_{c,b}$, the value of the volume of representations changes or not. One could expect that definition D2 does not depend on the choice of continuous bounded volume class but it does not seem easy to get an answer directly. It turns out that D2 is independent of the choice of continuous bounded volume class. For a proof, see Section 5.

Proposition 2.2. *Definition D2 does not depend on the choice of continuous bounded volume class. That is, for any two continuous bounded volume classes* $\omega_b, \omega'_b \in H^n_{c,b}(G, \mathbb{R})$,

$$\langle (c \circ (i^*_b)^{-1} \circ \rho^*_b)(\omega_b), [\bar{M}, \partial\bar{M}] \rangle = \langle (c \circ (i^*_b)^{-1} \circ \rho^*_b)(\omega'_b), [\bar{M}, \partial\bar{M}] \rangle.$$

Bucher, Burger and Iozzi proved the volume rigidity theorem for hyperbolic lattices as follows.

Theorem 2.3 [Bucher et al. 2013]. *Let $n \geq 3$. Let $i : \Gamma \hookrightarrow \mathrm{Isom}^+(\mathbb{H}^n)$ be a lattice embedding and let $\rho : \Gamma \to \mathrm{Isom}^+(\mathbb{H}^n)$ be any representation. Then*

$$|\mathrm{Vol}_2(\rho)| \leq |\mathrm{Vol}_2(i)| = \mathrm{Vol}(\Gamma\backslash\mathbb{H}^n),$$

with equality if and only if ρ is conjugated to i by an isometry.

3. New definition D4

In this section we give a new definition of volume of representations. It will turn out that the new definition is useful in proving that all the definitions of volume of representations are equivalent.

End compactification. Let $\widehat{M}$ be the end compactification of M obtained by adding one point for each end of M. Let $\widetilde{M}$ denote the universal cover of M. Define $\widehat{\widetilde{M}}$ to be the space obtained by adding to $\widetilde{M}$ one point for each lift of each end of M. The points added to M are called *ideal points* of M and the points added to $\widetilde{M}$ are called *ideal points* of $\widetilde{M}$. Denote by $\partial\widehat{M}$ the set of ideal points of M and by $\partial\widehat{\widetilde{M}}$ the set of ideal points of $\widetilde{M}$. Let $p : \widetilde{M} \to M$ be the universal covering map. The covering map $p : \widetilde{M} \to M$ extends to a map $\hat{p} : \widehat{\widetilde{M}} \to \widehat{M}$ and, moreover, the action of $\pi_1(M)$

on $\widetilde{M}$ by covering transformations induces an action on $\widehat{\widetilde{M}}$. The action on $\widehat{\widetilde{M}}$ is not free because each point of $\partial\widehat{\widetilde{M}}$ is stabilized by some peripheral subgroup of $\pi_1(M)$.

Note that $\widehat{M}$ can be obtained by collapsing each connected component of $\partial\bar{M}$ to a point. Similarly, $\widehat{\widetilde{M}}$ can be obtained by collapsing each connected component of $\bar{p}^{-1}(\partial\bar{M})$ to a point where $\bar{p} : \widetilde{\bar{M}} \to \bar{M}$ is the universal covering map. We denote the collapsing map by $\pi : \widetilde{\bar{M}} \to \widehat{\widetilde{M}}$.

One advantage of $\widehat{M}$ is the existence of a fundamental class in singular homology. While the top dimensional singular homology of M vanishes, the top dimensional singular homology of $\widehat{M}$ with coefficients in $\mathbb{Z}$ is isomorphic to $\mathbb{Z}$. Moreover, it can be easily seen that $H_*(\widehat{M}, \mathbb{R})$ is isomorphic to $H_*(\bar{M}, \partial\bar{M}, \mathbb{R})$ in degree $* \geq 2$. Hence the fundamental class of $\widehat{M}$ is well-defined and we denote it by $[\widehat{M}]$.

The cohomology groups. Let Y be a topological space and suppose that a group L acts continuously on Y. Then the cohomology group $H^*(Y; L, \mathbb{R})$ associated with Y and L is defined in the following way. Our main reference for this cohomology is [DuPre 1968].

For $k > 0$, define

$$F^k_{\mathrm{alt}}(Y, \mathbb{R}) = \{f : Y^{k+1} \to \mathbb{R} \mid f \text{ is alternating}\}.$$

Let $F^k_{\mathrm{alt}}(Y, \mathbb{R})^L$ denote the subspace of L-invariant functions, where the action of L on $F^k_{\mathrm{alt}}(Y, \mathbb{R})$ is given by

$$(g \cdot f)(y_0, \dots, y_k) = f(g^{-1}y_0, \dots, g^{-1}y_k),$$

for $f \in F^k_{\mathrm{alt}}(Y)$, $g \in L$. Define a coboundary operator $\delta_k : F^k_{\mathrm{alt}}(Y, \mathbb{R}) \to F^{k+1}_{\mathrm{alt}}(Y, \mathbb{R})$ by the usual

$$(\delta_k f)(y_0, \dots, y_{k+1}) = \sum_{i=0}^{k+1} (-1)^i f(y_0, \dots, \hat{y}_i, \dots, y_{k+1}).$$

The coboundary operator restricts to the complex $F^*_{\mathrm{alt}}(Y, \mathbb{R})^L$. The cohomology $H^*(Y; L, \mathbb{R})$ is defined as the cohomology of this complex. Define $F^*_{\mathrm{alt},b}(Y, \mathbb{R})$ as the subspace of $F^*_{\mathrm{alt}}(Y, \mathbb{R})$ consisting of bounded alternating functions. Clearly the coboundary operator restricts to the complex $F^*_{\mathrm{alt},b}(Y, \mathbb{R})^L$ and so it defines a cohomology, denoted by $H^*_b(Y; L, \mathbb{R})$. In particular, for a manifold M, the cohomology $H^*(\widetilde{M}; \pi_1(M), \mathbb{R})$ is actually isomorphic to the group cohomology $H^*(\pi_1(M), \mathbb{R})$, and $H^*_b(\widetilde{M}; \pi_1(M), \mathbb{R})$ is isomorphic to the bounded cohomology $H^*_b(\pi_1(M), \mathbb{R})$.

Remark 3.1. Let L and L' be groups acting continuously on topological spaces Y and Y', respectively. Given a homomorphism $\rho : L \to L'$, any ρ-equivariant continuous map $P : Y \to Y'$ defines a chain map,

$$P^* : F^*_{\mathrm{alt}}(Y', \mathbb{R})^{L'} \to F^*_{\mathrm{alt}}(Y, \mathbb{R})^L.$$

Thus it gives a morphism in cohomology. Let $Q : Y \to Y'$ be another ρ-equivariant map. For each $k > 0$, one may define

$$H_k(y_0, \ldots, y_k) = \sum_{i=0}^{i} (-1)^k (P(y_0), \ldots, P(y_i), Q(y_i), \ldots, Q(y_k)).$$

Then by a straightforward computation,

$$(\partial_{k+1} H_k + H_{k-1}\partial_k)(y_0, \ldots, y_k) = (P(y_0), \ldots, P(y_k)) - (Q(y_0), \ldots, Q(y_k)).$$

It follows from the above identity that, for any cocycle $f \in F^k_{\mathrm{alt}}(Y', \mathbb{R})^{L'}$,

$$(P^* f - Q^* f)(y_0, \ldots, y_k) = \delta_k(f \circ H_{k-1})(y_0, \ldots, y_k).$$

From this usual process in cohomology theory, one could expect that P and Q induce the same morphism in cohomology. However, since $f \circ H_{k-1}$ may not be alternating, P and Q may not induce the same morphism in cohomology.

Since $\Theta : \mathcal{X}^{n+1} \to \mathbb{R}$ is a G-invariant continuous bounded alternating cocycle, it yields a bounded cohomology class $[\Theta]_b \in H^n_b(\mathcal{X}; G, \mathbb{R})$. Let $\bar{\mathcal{X}}$ be the compactification of $\mathcal{X}$ obtained by adding the ideal boundary $\partial\mathcal{X}$. Extending the G-action on $\mathcal{X}$ to $\bar{\mathcal{X}}$, we can define a cohomology $H^*(\bar{\mathcal{X}}; G, \mathbb{R})$ and bounded cohomology $H^*_b(\bar{\mathcal{X}}; G, \mathbb{R})$. In the rank-1 case, since the geodesic simplex is well-defined for any $(n+1)$-tuple of points of $\bar{X}$, the cocycle Θ can be extended to a G-invariant alternating bounded cocycle $\bar{\Theta} : \bar{X}^{n+1} \to \mathbb{R}$. Hence $\bar{\Theta}$ determines a cohomology class $[\bar{\Theta}] \in H^n(\bar{\mathcal{X}}; G, \mathbb{R})$ and $[\bar{\Theta}]_b \in H^n_b(\bar{\mathcal{X}}; G, \mathbb{R})$.

Let $\widehat{D} : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$ be a ρ-equivariant continuous map whose restriction to $\widetilde{M}$ is a ρ-equivariant continuous map from $\widetilde{M}$ to $\mathcal{X}$. We will consider only such kinds of equivariant maps throughout the paper. Denote by $D : \widetilde{M} \to \mathcal{X}$ the restriction of $\widehat{D}$ to $\widetilde{M}$. Then $\widehat{D}$ induces a homomorphism in cohomology,

$$\widehat{D}^* : H^n(\bar{\mathcal{X}}; G, \mathbb{R}) \to H^n(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R}).$$

Note that the action of $\pi_1(M)$ on $\widehat{\widetilde{M}}$ is not free and hence $H^*(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R})$ may not be isomorphic to $H^*(\widehat{M}, \mathbb{R})$. Let $H^*_{\mathrm{simp}}(\widehat{M}, \mathbb{R})$ denote the simplicial cohomology induced from a simplicial structure on $\widehat{M}$. Then there is a natural restriction map $H^*(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R}) \to H^*_{\mathrm{simp}}(\widehat{M}, \mathbb{R}) \cong H^*(\widehat{M}, \mathbb{R})$. Thus we regard the cohomology class $\widehat{D}^*[\bar{\Theta}]$ as a cohomology class of $H^n(\widehat{M}, \mathbb{R})$. Let $[\widehat{M}]$ be the fundamental cycle in $H_n(\widehat{M}, \mathbb{R}) \cong \mathbb{R}$.

Definition 3.2 (D4). Let $D : \widetilde{M} \to \mathcal{X}$ be a ρ-equivariant continuous map which is extended to a ρ-equivariant map $\widehat{D} : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$. Then we define the invariant

$$\mathrm{Vol}_4(\rho, D) = \langle \widehat{D}^*[\bar{\Theta}], [\widehat{M}] \rangle.$$

As observed before, $\widehat{D}^*[\bar{\Theta}]$ may depend on the choice of ρ-equivariant map. However, it turns out that the value $\mathrm{Vol}_4(\rho, D)$ is independent of the choice of ρ-equivariant continuous map as follows.

Proposition 3.3. *Let $\rho : \pi_1(M) \to G$ be a representation. Then*

$$\mathrm{Vol}_2(\rho) = \mathrm{Vol}_4(\rho, D).$$

Proof. Since the continuous bounded cohomology $H^*_{c,b}(G, \mathbb{R})$ can be computed isomorphically from the complex $C^*_{c,b}(\mathcal{X}, \mathbb{R})_{\mathrm{alt}}$, there is the natural inclusion $C^*_{c,b}(\mathcal{X}, \mathbb{R})_{\mathrm{alt}} \subset F^*_{\mathrm{alt},b}(\mathcal{X}, \mathbb{R})$. Denote the homomorphism in cohomology induced from the inclusion by $i_G : H^k_{c,b}(G, \mathbb{R}) \to H^k_b(\mathcal{X}; G, \mathbb{R})$. Clearly, $i_G([\Theta]_{c,b}) = [\Theta]_b$.

The bounded cohomology $H^*_b(\pi_1(M), \mathbb{R})$ is obtained by the cohomology of the complex $C^*_b(\widetilde{M}, \mathbb{R})^{\pi_1(M)}_{\mathrm{alt}}$. Since $C^*_b(\widetilde{M}, \mathbb{R})_{\mathrm{alt}} = F^*_{\mathrm{alt},b}(\widetilde{M}, \mathbb{R})$, the induced map $i_M : H^k_b(\pi_1(M), \mathbb{R}) \to H^k_b(\widetilde{M}; \pi_1(M), \mathbb{R})$ is the identity map. Let $\widehat{D} : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$ be a ρ-equivariant map which maps $\widetilde{M}$ to $\mathcal{X}$. Then consider the following commutative diagram, where $\pi : \widetilde{\widetilde{M}} \to \widehat{\widetilde{M}}$ is the collapsing map:

$$\begin{tikzcd}
H^n(\bar{\mathcal{X}}; G, \mathbb{R}) \arrow[r, "\widehat{D}^*"] & H^n(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R}) \arrow[dr, "\pi^*"] & \\
H^n_b(\bar{\mathcal{X}}; G, \mathbb{R}) \arrow[u, "\bar{c}"] \arrow[r, "\widehat{D}^*_b"] \arrow[d, "\mathrm{res}_{\mathcal{X}}"] & H^n_b(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R}) \arrow[u, "\hat{c}"] \arrow[d, "\mathrm{res}_M"] \arrow[dr, "\pi^*_b"] & H^n(\bar{M}, \partial \bar{M}, \mathbb{R}) \\
H^n_b(\mathcal{X}; G, \mathbb{R}) \arrow[r, "D^*_b"] & H^n_b(\widetilde{M}; \pi_1(M), \mathbb{R}) & H^n_b(\bar{M}, \partial \bar{M}, \mathbb{R}) \arrow[l, "i^*_b"'] \arrow[u, "c"] \\
H^n_{c,b}(G, \mathbb{R}) \arrow[u, "i_G"] \arrow[r, "\rho^*_b"] & H^n_b(\pi_1(M), \mathbb{R}) \arrow[u, "i_M"] &
\end{tikzcd}$$

Note that the map ρ^*_b in the bottom of the diagram is actually induced from the restriction map $D : \widetilde{M} \to \mathcal{X}$. However, it does not depend on the choice of equivariant map but only on the homomorphism ρ. In other words, any continuous equivariant map from $\widetilde{M}$ to $\mathcal{X}$ gives rise to the same map, $\rho^*_b : H^*_{c,b}(G, \mathbb{R}) \to H^*_b(\pi_1(M), \mathbb{R})$. For this reason, we denote it by ρ^*_b instead of $D^*_{c,b}$.

Note that π induces a map $\pi^* : F^*_{\mathrm{alt}}(\widehat{\widetilde{M}}, \mathbb{R}) \to F^*_{\mathrm{alt}}(\widetilde{\widetilde{M}}, \mathbb{R})$. It follows from the alternating property that the image of π^* is contained in $C^*(\bar{M}, \partial \bar{M}, \mathbb{R})$. Hence the map $\pi^* : H^n(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R}) \to H^n(\bar{M}, \partial \bar{M}, \mathbb{R})$ makes sense. One can understand $\pi_b^* : H_b^n(\widehat{\widetilde{M}}; \pi_1(M), \mathbb{R}) \to H_b^n(\bar{M}, \partial \bar{M}, \mathbb{R})$ in a similar way.

Noting that $\bar{c}([\bar{\Theta}]_b) = [\bar{\Theta}]$ and $\mathrm{res}_{\mathcal{X}}([\bar{\Theta}]_b) = [\Theta]_b$, it follows from the above commutative diagram that

$$\begin{aligned}((i_b^*)^{-1} \circ i_M \circ \rho_b^*)[\Theta]_{c,b} &= ((i_b^*)^{-1} \circ D_b^* \circ i_G)[\Theta]_{c,b} = ((i_b^*)^{-1} \circ D_b^* \circ \mathrm{res}_{\mathcal{X}})[\bar{\Theta}]_b \\ &= ((i_b^*)^{-1} \circ \mathrm{res}_M \circ \widehat{D}_b^*)[\bar{\Theta}]_b = (\pi_b^* \circ \widehat{D}_b^*)[\bar{\Theta}]_b.\end{aligned}$$

Hence

$$\begin{aligned}\mathrm{Vol}_2(\rho) &= \langle (c \circ (i_b^*)^{-1} \circ i_M \circ \rho_b^*)[\Theta]_{c,b}, [\bar{M}, \partial \bar{M}] \rangle \\ &= \langle (c \circ \pi_b^* \circ \widehat{D}_b^*)[\bar{\Theta}]_b, [\bar{M}, \partial \bar{M}] \rangle = \langle (\pi^* \circ \widehat{D}^* \circ \bar{c})[\bar{\Theta}]_b, [\bar{M}, \partial \bar{M}] \rangle \\ &= \langle (\pi^* \circ \widehat{D}^*)[\bar{\Theta}], [\bar{M}, \partial \bar{M}] \rangle = \langle \widehat{D}^*[\bar{\Theta}], \pi_*[\bar{M}, \partial \bar{M}] \rangle \\ &= \langle \widehat{D}^*[\bar{\Theta}], [\widehat{M}] \rangle = \mathrm{Vol}_4(\rho, D).\end{aligned}$$ □

Proposition 3.3 implies that the value $\mathrm{Vol}_4(\rho, D)$ does not depend on the choice of continuous equivariant map. Hence from now on we will use the notation $\mathrm{Vol}_4(\rho) := \mathrm{Vol}(\rho, D)$. Furthermore, Proposition 3.3 allows us to interpret the invariant $\mathrm{Vol}_2(\rho)$ in terms of a pseudodeveloping map via $\mathrm{Vol}_4(\rho)$ in the next section. Note that a pseudodeveloping map for ρ is a specific kind of ρ-equivariant continuous map $\widehat{\widetilde{M}} \to \bar{\mathcal{X}}$.

4. Pseudodeveloping map and definition D1

Dunfield [1999] introduced the notion of pseudodeveloping map in order to define the volume of representations $\rho : \pi_1(M) \to \mathrm{SO}(3, 1)$ for a noncompact complete hyperbolic 3-manifold M of finite volume. We start by recalling the definition of pseudodeveloping map.

Definition 4.1 (cone map). Let $\mathcal{A}$ be a set, let $t_0 \in \mathbb{R}$, and let $\mathrm{cone}(\mathcal{A})$ be the cone obtained from $\mathcal{A} \times [t_0, \infty]$ by collapsing $\mathcal{A} \times \{\infty\}$ to a point, called ∞. A map $\widehat{D} : \mathrm{cone}(\mathcal{A}) \to \bar{\mathcal{X}}$ is a *cone map* if $\widehat{D}(\mathrm{cone}(\mathcal{A})) \cap \partial \mathcal{X} = \{\widehat{D}(\infty)\}$ and for all $a \in \mathcal{A}$ the map $\widehat{D}|_{a \times [t_0, \infty]}$ is either the constant to $\widehat{D}(\infty)$ or the geodesic ray from $\widehat{D}(a, t_0)$ to $\widehat{D}(\infty)$, parametrized in such a way that the parameter $(t - t_0)$, $t \in [t_0, \infty]$, is the arc length.

For each ideal point v of M, fix a product structure $T_v \times [0, \infty)$ on the end relative to v. The fixed product structure induces a cone structure on a neighborhood of v in $\widehat{M}$, which is obtained from $T_v \times [0, \infty]$ by collapsing $T_v \times \{\infty\}$ to a point v. We lift such structures to the universal cover. Let $\tilde{v}$ be an ideal point of $\widetilde{M}$ that projects to the ideal point v. Denote by $E_{\tilde{v}}$ the cone at $\tilde{v}$ that is homeomorphic to $P_{\tilde{v}} \times [0, \infty]$, where $P_{\tilde{v}}$ covers T_v and $P_{\tilde{v}} \times \{\infty\}$ is collapsed to $\tilde{v}$.

Definition 4.2 (pseudodeveloping map). Let $\rho : \pi_1(M) \to G$ be a representation. A *pseudodeveloping map* for ρ is a piecewise-smooth ρ-equivariant map $D : \widetilde{M} \to \mathcal{X}$. Moreover, D is required to extend to a continuous map $\widehat{D} : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$ with the property that there exists a $t \in \mathbb{R}^+$ such that, for each end $E_{\tilde{v}} = P_{\tilde{v}} \times [0, \infty]$ of $\widehat{\widetilde{M}}$, the restriction of $\widehat{D}$ to $P_{\tilde{v}} \times [t, \infty]$ is a cone map.

Definition 4.3. A *triangulation* of $\widehat{M}$ is an identification of $\widehat{M}$ with a complex obtained by gluing together with simplicial attaching maps. It is not required for the complex to be simplicial, but it is required that open simplices embed.

Note that a triangulation of $\widehat{M}$ always exists and it lifts uniquely to a triangulation of $\widehat{\widetilde{M}}$. Given a triangulation of $\widehat{M}$, one can define the straightening of pseudodeveloping maps as follows.

Definition 4.4 (straightening map). Let $\widehat{M}$ be triangulated. Let $\rho : \pi_1(M) \to G$ be a representation and $D : \widetilde{M} \to \mathcal{X}$ a pseudodeveloping map for ρ. A straightening of D is a continuous piecewise-smooth ρ-equivariant map $\mathrm{Str}(D) : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$ such that

- for each simplex σ of the triangulation, $\mathrm{Str}(D)$ maps $\widetilde{\sigma}$ to $\mathrm{Str}(D \circ \widetilde{\sigma})$,
- for each end $E_{\tilde{v}} = P_{\tilde{v}} \times [0, \infty]$, there exists a $t \in \mathbb{R}$ such that $\mathrm{Str}(D)$ restricted to $P_{\tilde{v}} \times [t, \infty]$ is a cone map,

where $\widetilde{\sigma}$ is a lift of σ to $\widehat{\widetilde{M}}$ and $\mathrm{Str}(D \circ \widetilde{\sigma})$ is the geodesic straightening of the map $D \circ \widetilde{\sigma} : \Delta^n \to \bar{\mathcal{X}}$.

Note that any straightening of a pseudodeveloping map is also a pseudodeveloping map.

Lemma 4.5. *Let $\widehat{M}$ be triangulated. Let $\rho : \pi_1(M) \to G$ be a representation and $D : \widetilde{M} \to \mathcal{X}$ a pseudodeveloping map for ρ. Then a straightening* $\mathrm{Str}(D)$ *of D exists and, furthermore,* $\mathrm{Str}(D) : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$ *is always equivariantly homotopic to $\widehat{D}$ via a homotopy that fixes the vertices of the triangulation.*

Proof. First, set $\mathrm{Str}(D)(V) = f(V)$ for every vertex V of the triangulation. Then extend $\mathrm{Str}(D)$ to a map which is piecewise-straight with respect to the triangulation. This is always possible because $\mathcal{X}$ is contractible. Note that $\widehat{D}$ and $\mathrm{Str}(D)$ agree on the ideal vertices of $\widehat{\widetilde{M}}$ and are equivariantly homotopic via the straight-line homotopy between them. Hence it can be easily seen that the extension is a straightening of D. □

For any pseudodeveloping map $D : \widetilde{M} \to \mathcal{X}$, for ρ,

$$\int_M D^* \omega_{\mathcal{X}}$$

is always finite. This can be seen as follows. We stick to the notation used in Definition 4.2. We may assume that the restriction of $\widehat{D}$ to each $E_{\tilde{v}} = P_{\tilde{v}} \times [0, \infty]$

is a cone map. Choose a fundamental domain F_0 of T_v in $P_{\tilde{v}}$. Then there exists a $t \in \mathbb{R}^+$ such that

$$\left| \int_{T_v \times [t,\infty)} D^*\omega_{\mathcal{X}} \right| = \mathrm{Vol}_n\big(\mathrm{cone}(D(F_0 \times \{t\}))\big) \leq \frac{1}{n-1} \mathrm{Vol}_{n-1}(D(F_0 \times \{t\})),$$

where Vol_{n-1} denotes the $(n-1)$-dimensional volume. The last inequality holds for any Hadamard manifold with sectional curvature at most -1. See [Gromov 1982, Section 1.2]. Hence the integral of $D^*\omega_{\mathcal{X}}$ over M is finite.

Definition 4.6 (D1). Let $D : \widetilde{M} \to \mathcal{X}$ be a pseudodeveloping map for a representation $\rho : \pi_1(M) \to G$. Define the invariant

$$\mathrm{Vol}_1(\rho, D) = \int_M D^*\omega_{\mathcal{X}}.$$

In the case that $G = \mathrm{SO}(n, 1)$, Francaviglia [2004] showed that definition D1 does not depend on the choice of pseudodeveloping map. We give a self-contained proof for this in the rank-1 case.

Proposition 4.7. *Let $\rho : \pi_1(M) \to G$ be a representation. Then, for any pseudodeveloping map $D : \widetilde{M} \to \mathcal{X}$,*

$$\mathrm{Vol}_1(\rho, D) = \mathrm{Vol}_4(\rho).$$

Thus, $\mathrm{Vol}_1(\rho, D)$ does not depend on the choice of pseudodeveloping map.

Proof. Let $\mathcal{T}$ be a triangulation of $\widehat{M}$ with simplices $\sigma_1, \dots, \sigma_N$. Then the triangulation gives rise to a fundamental cycle $\sum_{i=1}^N \sigma_i$ of $\widehat{M}$. Let $\mathrm{Str}(D)$ be a straightening of D with respect to the triangulation $\mathcal{T}$. Since $\mathrm{Str}(D)$ is a ρ-equivariant continuous map, we have

$$\begin{aligned}\mathrm{Vol}_4(\rho) := \mathrm{Vol}_4(\rho, D) = \langle \mathrm{Str}(D)^*[\bar{\Theta}], [\widehat{M}]\rangle = \langle \bar{\Theta}, \sum_{i=1}^N \mathrm{Str}(\widehat{D}(\sigma_i))\rangle \\ = \sum_{i=1}^N \int_{\mathrm{Str}(\widehat{D}(\sigma_i))} \omega_{\mathcal{X}} = \int_M \mathrm{Str}(D)^*\omega_{\mathcal{X}}.\end{aligned}$$

Since both $\mathrm{Str}(D)$ and $\widehat{D}$ are pseudodeveloping maps for ρ that agree on the ideal points of $\widehat{\widetilde{M}}$, it can be proved, using the same arguments as the proof of [Dunfield 1999, Lemma 2.5.1], that

$$\int_M \mathrm{Str}(D)^*\omega_{\mathcal{X}} = \int_M D^*\omega_{\mathcal{X}} = \mathrm{Vol}_1(\rho, D). \qquad \square$$

Remark 4.8. While D1 is defined with only a pseudodeveloping map, definition D4 is defined with any equivariant map. This is one advantage of definition D4. By Proposition 4.7, the notation $\mathrm{Vol}_1(\rho) := \mathrm{Vol}_1(\rho, D)$ makes sense.

5. Lipschitz simplicial volume and definition D3

In this section, M is assumed to be a Riemannian manifold with finite Lipschitz simplicial volume. Gromov [1982, Section 4.4] introduced the Lipschitz simplicial volume of Riemannian manifolds. One can define the Lipschitz constant for each singular simplex in M by giving the Euclidean metrics on the standard simplices. Then the Lipschitz constant of a locally finite chain c of M is defined as the supremum of the Lipschitz constants of all singular simplices occurring in c. The Lipschitz simplicial volume of M is defined by the infimum of the ℓ^1-norms of all locally finite fundamental cycles with finite Lipschitz constant. Let $[M]^{\ell^1}_{\mathrm{Lip}}$ be the set of all locally finite fundamental cycles of M with finite ℓ^1-seminorm and finite Lipschitz constant. If $[M]^{\ell^1}_{\mathrm{Lip}} = \varnothing$, the Lipschitz simplicial volume of M is infinite.

In the case that $[M]^{\ell^1}_{\mathrm{Lip}} \neq \varnothing$, we gave a new definition of volume of representations in [Kim and Kim 2014] as follows. A representation $\rho : \pi_1(M) \to G$ induces a canonical pullback map $\rho_b^* : H^*_{c,b}(G, \mathbb{R}) \to H^*_b(\pi_1(M), \mathbb{R}) \cong H^*_b(M, \mathbb{R})$ in continuous bounded cohomology. Hence, for any continuous bounded volume class $\omega_b \in H^n_{c,b}(G, \mathbb{R})$, we obtain a bounded cohomology class $\rho_b^*(\omega_b) \in H^n_b(M, \mathbb{R})$. Then, the bounded cohomology class $\rho_b^*(\omega_b)$ can be evaluated on ℓ^1-homology classes in $H^{\ell^1}_n(M, \mathbb{R})$ by the Kronecker products,

$$\langle \cdot , \cdot \rangle : H^*_b(M, \mathbb{R}) \otimes H^{\ell^1}_*(M, \mathbb{R}) \to \mathbb{R}.$$

For more details about this, see [Kim and Kim 2014].

Definition 5.1 (D3). We define the invariant

$$\mathrm{Vol}_3(\rho) = \inf \langle \rho_b^*(\omega_b), \alpha \rangle,$$

where the infimum is taken over all $\alpha \in [M]^{\ell^1}_{\mathrm{Lip}}$ and all $\omega_b \in H^n_{c,b}(G, \mathbb{R})$ with $c(\omega_b) = \omega_{\mathcal{X}}$.

One advantage of D3 is that the isomorphism $H^n_b(\bar{M}, \partial \bar{M}, \mathbb{R}) \to H^n_b(\bar{M}, \mathbb{R})$ is not needed. When M admits the isomorphism above, we will verify that definition D3 is eventually equivalent to the other definitions of volume of representations.

Lemma 5.2. *Suppose that M is a noncompact, connected, orientable, aspherical, tame Riemannian manifold with finite Lipschitz simplicial volume and that each end of M has amenable fundamental group. Then, for any $\alpha \in [M]^{\ell^1}_{\mathrm{Lip}}$ and any continuous bounded volume class ω_b,*

$$\langle \rho_b^*(\omega_b), \alpha \rangle = \langle (c \circ (i_b^*)^{-1} \circ \rho_b^*)(\omega_b), [\bar{M}, \partial \bar{M}] \rangle.$$

Proof. When M is a 2-dimensional manifold, the proof is given in [Kim and Kim 2014]. Actually the proof in the general case is the same. We sketch the proof here for the reader's convenience. Let K be a compact core of M. Note that K is a compact submanifold with boundary that is a deformation retract of M. Consider the following commutative diagram, where every map is the map induced from the canonical inclusion:

$$\begin{tikzcd} C_b^*(M,\mathbb{R}) & C_b^*(\bar{M},\mathbb{R}) \arrow[l, "j_b^*"'] & C_b^*(\bar{M},\partial\bar{M},\mathbb{R}) \arrow[l, "i_b^*"'] \\ & C_b^*(\bar{M},\bar{M}-K,\mathbb{R}) \arrow[u, "l_b^*"] \arrow[ur, "q_b^*"'] & \end{tikzcd}$$

Every map in the diagram induces an isomorphism in bounded cohomology in $* \geq 2$. Thus, there exists a cocycle $z_b \in C_b^n(\bar{M}, \bar{M}-K, \mathbb{R})$ such that $l_b^*([z_b]) = \rho_b^*(\omega_b)$.

Let $c = \sum_{i=1}^{\infty} a_i \sigma_i$ be a locally finite fundamental ℓ^1-cycle with finite Lipschitz constant representing $\alpha \in [M]_{\mathrm{Lip}}^{\ell^1}$. Then we have

$$\langle \rho_b^*(\omega_b), \alpha\rangle = \langle l_b^*([z_b]), \alpha\rangle = \langle z_b, c\rangle.$$

Since z_b vanishes on simplices with image contained in $\bar{M}-K$, we have the equality $\langle z_b, c\rangle = \langle z_b, c|_K\rangle$, where $c|_K = \sum_{\mathrm{im}\,\sigma_i \cap K \neq \varnothing} a_i\sigma_i$. It is a standard fact that the sum $c|_K$ represents the relative fundamental class $[\bar{M}, \bar{M}-K]$ in $H_n(\bar{M}, \bar{M}-K, \mathbb{R})$ (see [Löh 2007, Theorem 5.3]). On the other hand, we have

$$\begin{aligned} \langle (c \circ (i_b^*)^{-1} \circ \rho_b^*)(\omega_b), [\bar{M}, \partial\bar{M}]\rangle &= \langle (c \circ q_b^*)([z_b]), [\bar{M}, \partial\bar{M}])\rangle \\ &= \langle [z_b], q_*[\bar{M}, \partial\bar{M}]\rangle \\ &= \langle [z_b], [\bar{M}, \bar{M}-K]\rangle = \langle z_b, c|_K\rangle. \qquad \square \end{aligned}$$

By Lemma 5.2 we can reformulate definition D3 as

$$\mathrm{Vol}_3(\rho) = \inf_{\omega_b} \langle (c \circ (i_b^*)^{-1} \circ \rho_b^*)(\omega_b), [\bar{M}, \partial\bar{M}]\rangle,$$

where the infimum is taken over all continuous bounded volume classes. Noting that $[\Theta]_{c,b} \in H_{c,b}^n(G,\mathbb{R})$ is a continuous bounded volume class, it is clear that

$$\mathrm{Vol}_3(\rho) \leq \mathrm{Vol}_2(\rho).$$

It is conjecturally true that the comparison map $H_{c,b}^n(G,\mathbb{R}) \to H_c^n(G,\mathbb{R})$ is an isomorphism for any connected semisimple Lie group G with finite center. Hence, conjecturally, $\mathrm{Vol}_2(\rho) = \mathrm{Vol}_3(\rho)$. In spite of the absence of a proof of the conjecture, we will give a proof for $\mathrm{Vol}_2(\rho) = \mathrm{Vol}_3(\rho)$ by using definition D4.

Lemma 5.3. *Let $\omega_b \in H_{c,b}^n(G,\mathbb{R})$ be a continuous bounded volume class, and let $f_b : \mathcal{X}^{n+1} \to \mathbb{R}$ be a continuous bounded alternating G-invariant cocycle representing ω_b. Then f_b is extended to a bounded alternating G-invariant cocycle $\bar{f}_b : \bar{\mathcal{X}}^{n+1} \to \mathbb{R}$. Furthermore, $\bar{f}_b$ is uniformly continuous on $\mathcal{X}^n \times \{\xi\}$ for any $\xi \in \partial\mathcal{X}$.*

Proof. For any $(\bar{x}_0, \dots, \bar{x}_n) \in \bar{\mathcal{X}}^{n+1}$, define

$$\bar{f}_b(\bar{x}_0, \dots, \bar{x}_n) = \lim_{t\to\infty} f_b(c_0(t), \dots, c_n(t)),$$

where each $c_i(t)$ is a geodesic ray toward $\bar{x}_i$. Here, for $x \in \mathcal{X}$, we say that $c : [0, \infty) \to \mathcal{X}$ is a geodesic ray toward x if there exists a $t \in [0, \infty)$ such that the restriction map $c|_{[0,t]}$ of c to $[0, t]$ is a geodesic with $c(t) = x$ and $c|_{[t,\infty)}$ is constant to x. Then it is clear that $\bar{f}_b(x_0, \dots, x_n) = f_b(x_0, \dots, x_n)$ for $(x_0, \dots, x_n) \in \mathcal{X}^{n+1}$. To see the well-definedness of $\bar{f}_b$, we need to show that, for other geodesic rays $c'_i(t)$ toward $\bar{x}_i$,

$$\lim_{t\to\infty} f_b(c_0(t), \dots, c_n(t)) = \lim_{t\to\infty} f_b(c'_0(t), \dots, c'_n(t)). \tag{2}$$

Note that the limit always exists because f_b is bounded. In the rank-1 case, the distance between two geodesic rays with the same endpoint decays exponentially to 0 as they go to the endpoint. Moreover, since f_b is G-invariant and G transitively acts on $\mathcal{X}$, we have that f_b is uniformly continuous on $\mathcal{X}^{n+1}$. Thus, for any $\epsilon > 0$, there exists some number $T > 0$ such that

$$|f_b(c_0(t), \dots, c_n(t)) - f_b(c'_0(t), \dots, c'_n(t))| < \epsilon$$

for all $t > T$. This implies (2) and hence $\bar{f}_b$ is well-defined.

The alternating property of $\bar{f}_b$ actually comes from f_b. Due to the alternating property of f_b, we have

$$\begin{aligned}
\bar{f}_b(\bar{x}_0, \dots, \bar{x}_i, \dots, \bar{x}_j, \dots, \bar{x}_n) &= \lim_{t\to\infty} f_b(c_0(t), \dots, c_i(t), \dots, c_j(t), \dots, c_n(t)) \\
&= \lim_{t\to\infty} -f_b(c_0(t), \dots, c_j(t), \dots, c_i(t), \dots, c_n(t)) \\
&= -\bar{f}_b(\bar{x}_0, \dots, \bar{x}_j, \dots, \bar{x}_i, \dots, \bar{x}_n).
\end{aligned}$$

Therefore, we conclude that $\bar{f}_b$ is alternating. The boundedness and G-invariance of $\bar{f}_b$ immediately follows from the boundedness and G-invariance of f_b. Furthermore, it is easy to check that $\bar{f}_b$ is a cocycle by a direct computation.

Now it remains to prove that $\bar{f}_b$ is uniformly continuous on $\mathcal{X}^n \times \{\xi\}$. It is obvious that $\bar{f}_b$ is continuous on $\mathcal{X}^n \times \{\xi\}$. Noting that the parabolic subgroup of G stabilizing ξ acts on $\mathcal{X}$ transitively, it can be easily seen that $\bar{f}_b$ is uniformly continuous on $\mathcal{X}^n \times \{\xi\}$. □

The existence of $\bar{f}_b$ allows us to reformulate Vol_3 in terms of Vol_4. Following the proof of Proposition 3.3, we get

$$\langle (c \circ (i_b^*)^{-1} \circ \rho_b^*)(\omega_b), [\bar{M}, \partial \bar{M}] \rangle = \langle \widehat{D}^*[\bar{f}_b], [\widehat{M}] \rangle. \tag{3}$$

The last term $\langle \widehat{D}^*[\bar{f}_b], [\widehat{M}] \rangle$ above is computed by $\langle \widehat{D}^*\bar{f}_b, \hat{c} \rangle$ for any equivariant map $\widehat{D}$ and fundamental cycle $\hat{c}$ of $\widehat{M}$. By choosing the proper equivariant map and fundamental cycle, we will show that $\langle \widehat{D}^*[\bar{f}_b], [\widehat{M}]\rangle$ does not depend on the choice of continuous bounded volume class.

Proposition 5.4. *Let ω_b and ω_b' be continuous bounded volume classes, and let $\bar{f}_b$ and $\bar{f}_b'$ be the bounded alternating cocycles in $F^n_{\mathrm{alt}}(\bar{X}; G, \mathbb{R})$ associated with ω_b and ω_b' respectively, as in Lemma 5.3. Then*

$$\langle \widehat{D}^*[\bar{f}_b], [\widehat{M}]\rangle = \langle \widehat{D}^*[\bar{f}_b'], [\widehat{M}]\rangle.$$

Proof. It suffices to prove that, for some ρ-equivariant map $\widehat{D} : \widehat{\widetilde{M}} \to \bar{\mathcal{X}}$ and fundamental cycle $\hat{c}$ of $\widehat{M}$,

$$\langle \widehat{D}^*\bar{f}_b, \hat{c}\rangle = \langle \widehat{D}^*\bar{f}_b', \hat{c}\rangle.$$

To show this, we will prove that, for some sequence $(\hat{c}_k)_{k\in\mathbb{N}}$ of fundamental cycles of $\widehat{M}$,

$$\lim_{k\to\infty} (\langle \widehat{D}^*\bar{f}_b, \hat{c}_k\rangle - \langle \widehat{D}^*\bar{f}_b', \hat{c}_k\rangle) = 0.$$

Let $v_1, \dots, v_s$ be the ideal points of M. As in Section 4, fix a product structure $T_{v_i} \times [0, \infty]$ on the end relative to v_i for each $i = 1, \dots, s$ and then lift such structures to the universal cover. We stick to the notation used in Section 4. Set

$$M_k = M - \bigcup_{i=1}^{s} T_{v_i} \times (k, \infty].$$

Then $(M_k)_{k\in\mathbb{N}}$ is an exhausting sequence of compact cores of M. The boundary ∂M_k of M_k consists of $\bigcup_{i=1}^s T_{v_i} \times \{k\}$. Let $\mathcal{T}_0$ be a triangulation of M_0. Then we extend it to a triangulation on $\widehat{M}$ as follows. First note that $\mathcal{T}_0$ induces a triangulation on each T_{v_i}. Let τ be an $(n-1)$-simplex of the induced triangulation on T_{v_i} for some $i \in \{1, \dots, s\}$. Then we attach $\pi(\tau \times [0, \infty])$ to $T_{v_i} \times \{0\}$ along $\tau \times \{0\}$, where $\pi : \bar{M} \to \widehat{M}$ is the collapsing map. Since π is an embedding on $\tau \times [0, \infty)$ and π maps $\tau \times \{\infty\}$ to the ideal point v_i, it can be easily seen that $\mathrm{cone}(\tau) := \pi(\tau \times [0, \infty])$ is an n-simplex. Hence we can obtain a triangulation of $\widehat{M}$ by attaching each $\mathrm{cone}(\tau)$ to ∂M_0, which is denoted by $\widehat{\mathcal{T}_0}$.

Next, we extend $\mathcal{T}_0$ to a triangulation of M_k. In fact, M_k is decomposed as

$$M_k = M_0 \cup \bigcup_{i=1}^{s} T_{v_i} \times [0, k].$$

Hence we can attach each $\tau \times [0, k]$ to M_0 along $\tau \times \{0\}$ and then triangulate $\tau \times [0, k]$ by using the prism operator [Hatcher 2002, Chapter 2.1]. Via this process, we obtain a triangulation of M_k, denoted by $\mathcal{T}_k$. Note that $\mathcal{T}_0$ and $\mathcal{T}_k$ induce the same triangulation on each T_{v_i}. In addition, one can obtain a triangulation $\widehat{\mathcal{T}_k}$ of $\widehat{M}$ from $\mathcal{T}_k$ similarly to how $\widehat{\mathcal{T}_0}$ is obtained from $\mathcal{T}_0$ above.

Let c_k be the relative fundamental class of $(M_k, \partial M_k)$ induced from $\mathcal{T}_k$. Then it can be seen that

$$\hat{c}_k = c_k + (-1)^{n+1}\,\mathrm{cone}(\partial c_k)$$

is the fundamental cycle of $\widehat{M}$ induced from $\widehat{\mathcal{T}_k}$. Any simplex occurring in c_k is contained in M_k. Now we choose a pseudodeveloping map $\widehat{D} : \widetilde{\widehat{M}} \to \bar{X}$. Let $\tilde{v}_i$ be a lift of v_i to $\widetilde{\widehat{M}}$. Let $P_{\tilde{v}_i} \times [0, \infty]$ be the cone structure of a neighborhood of $\tilde{v}_i$, where $P_{\tilde{v}_i}$ covers T_{v_i} and $P_{\tilde{v}_i} \times \{\infty\}$ is just the ideal point $\tilde{v}_i$. We may assume that $\widehat{D}$ is a cone map on each $P_{\tilde{v}_i} \times [0, \infty]$. Let $\tilde{c}_k$ be a lift of c_k to a cochain in $\widetilde{M}$ and let $\widetilde{\partial c_k}$ be a lift of ∂c_k. Let $\tau \times \{0\}$ be an $(n-1)$-simplex in $T_{v_i} \times \{0\}$ occurring in ∂c_0 and let $\tilde{\tau}$ be a lift of τ to $P_{\tilde{v}_i}$. Then $\tilde{\tau} \times \{k\}$ is a lift of $\tau \times \{k\} \in \partial c_k$. Since $\widehat{D}$ is a cone map on $P_{\tilde{v}_i} \times [0, \infty]$, we have that $D(\tilde{\tau} \times [0, \infty])$ is the geodesic cone over $\tilde{\tau} \times \{0\}$ with top point $\tilde{v}_i$ in $\bar{\mathcal{X}}$. Hence the diameter of $D(\tilde{\tau} \times \{k\})$ decays exponentially to 0 as $k \to \infty$ for each τ.

By a direct computation, we have

$$\begin{aligned}\langle \widehat{D}^*\bar{f}_b - \widehat{D}^*\bar{f}'_b, \hat{c}_k\rangle &= \langle \widehat{D}^*\bar{f}_b - \widehat{D}^*\bar{f}'_b, \tilde{c}_k\rangle + (-1)^{n+1}\langle \widehat{D}^*\bar{f}_b - \widehat{D}^*\bar{f}'_b, \mathrm{cone}(\widetilde{\partial c_k})\rangle \\ &= \langle \bar{f}_b - \bar{f}'_b, \widehat{D}_*(\tilde{c}_k)\rangle + (-1)^{n+1}\langle \bar{f}_b - \bar{f}'_b, \widehat{D}_*(\mathrm{cone}(\widetilde{\partial c_k}))\rangle \\ &= \langle f_b - f'_b, D_*(\tilde{c}_k)\rangle + (-1)^{n+1}\langle \bar{f}_b - \bar{f}'_b, \widehat{D}_*(\mathrm{cone}(\widetilde{\partial c_k}))\rangle.\end{aligned}$$

The last equality comes from the fact that $\widehat{D}_*(\tilde{c}_k)$ is a singular chain in $\mathcal{X}$. Since f_b and f'_b are continuous bounded alternating cocycles representing the continuous volume class $\omega_{\mathcal{X}} \in H^n_c(G, \mathbb{R})$, there is a continuous alternating G-invariant function $\beta : \mathcal{X}^n \to \mathbb{R}$ such that $f_b - f'_b = \delta\beta$. Hence

$$\langle f_b - f'_b, D_*(\tilde{c}_k)\rangle = \langle \delta\beta, D_*(\tilde{c}_k)\rangle = \langle \beta, \partial D_*(\tilde{c}_k)\rangle = \langle \beta, D_*(\widetilde{\partial c_k})\rangle.$$

As observed before, since the diameter of all simplices occurring in $D_*(\widetilde{\partial c_k})$ decays to 0 as $k \to \infty$ and, moreover, β is uniformly continuous on $\mathcal{X}$, we have

$$\lim_{k\to\infty} \langle \beta, D_*(\widetilde{\partial c_k})\rangle = 0.$$

Note that $D(\mathrm{cone}(\tilde{\tau} \times \{k\}))$ is the geodesic cone over $D(\tilde{\tau} \times \{k\})$ with top point $\tilde{v}_i$. By Lemma 5.3, both $\bar{f}_b$ and $\bar{f}'_b$ are uniformly continuous on $\mathcal{X}^n \times \{\tilde{v}_i\}$. Since the diameter of $D(\tilde{\tau} \times \{k\})$ decays to 0 as $k \to \infty$,

$$\lim_{k\to\infty} \langle \bar{f}_b, D(\mathrm{cone}(\tilde{\tau} \times \{k\}))\rangle = \lim_{k\to\infty} \langle \bar{f}'_b, D(\mathrm{cone}(\tilde{\tau} \times \{k\}))\rangle = 0.$$

Applying this to each τ, we can conclude that

$$\lim_{k\to\infty} \langle \bar{f}_b, D_*(\mathrm{cone}(\widetilde{\partial c_k}))\rangle = \lim_{k\to\infty} \langle \bar{f}'_b, D_*(\mathrm{cone}(\widetilde{\partial c_k}))\rangle = 0.$$

In the end, it follows that

$$\lim_{k\to\infty} \langle \widehat{D}^*\bar{f}_b - \widehat{D}^*\bar{f}'_b, \hat{c}_k \rangle = 0.$$

As we mentioned, the value on the left-hand side does not depend on $\hat{c}_k$. Thus we can conclude that $\langle \widehat{D}^*\bar{f}_b - \widehat{D}^*\bar{f}'_b, \hat{c}_k \rangle = 0$. This implies that $\langle \widehat{D}^*\bar{f}_b, \hat{c} \rangle = \langle \widehat{D}^*\bar{f}'_b, \hat{c} \rangle$ for any fundamental cycle $\hat{c}$ of $\widehat{M}$, which completes the proof. □

Combining Proposition 5.4 with (3), Proposition 2.2 immediately follows.

Proposition 5.5. *The definitions of D3 and D4 are equivalent.*

Proof. By Lemma 5.2 and Proposition 3.3, we have

$$\begin{aligned}\mathrm{Vol}_3(\rho) &= \inf\{\langle \rho_b^*(\omega_b), \alpha \rangle \mid c(\omega_b) = \omega_{\mathcal{X}} \text{ and } \alpha \in [M]^{\ell^1}_{\mathrm{Lip}}\} \\ &= \inf\{\langle (c \circ (i_b^*)^{-1} \circ \rho_b^*)(\omega_b), [\bar{M}, \partial \bar{M}] \rangle \mid c(\omega_b) = \omega_{\mathcal{X}}\} \\ &= \inf\{\langle \widehat{D}^*[\bar{f}_b], [\widehat{M}] \rangle \mid c(\omega_b) = \omega_{\mathcal{X}}\} = \langle \widehat{D}^*[\bar{\Theta}], [\widehat{M}] \rangle = \mathrm{Vol}_4(\rho). \quad \square\end{aligned}$$

References

[Besson et al. 2007] G. Besson, G. Courtois, and S. Gallot, "Inégalités de Milnor–Wood géométriques", *Comment. Math. Helv.* **82**:4 (2007), 753–803. MR 2009e:53055 Zbl 1143.53040

[Borel and Wallach 2000] A. Borel and N. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, 2nd ed., Mathematical Surveys and Monographs **67**, American Mathematical Society, Providence, RI, 2000. MR 2000j:22015 Zbl 0980.22015

[Bucher et al. 2012] M. Bucher, M. Burger, R. Frigerio, A. Iozzi, C. Pagliantini, and M. B. Pozzetti, "Isometric properties of relative bounded cohomology", preprint, 2012. arXiv 1205.1022

[Bucher et al. 2013] M. Bucher, M. Burger, and A. Iozzi, "A dual interpretation of the Gromov–Thurston proof of Mostow rigidity and volume rigidity for representations of hyperbolic lattices", pp. 47–76 in *Trends in harmonic analysis* (Rome, 2011), edited by M. A. Picardello, Springer INdAM Ser. **3**, Springer, Milan, 2013. MR 3026348 Zbl 1268.53056

[Burger et al. 2010] M. Burger, A. Iozzi, and A. Wienhard, "Surface group representations with maximal Toledo invariant", *Ann. of Math.* (2) **172**:1 (2010), 517–566. MR 2012j:22014 Zbl 1208.32014

[Dunfield 1999] N. M. Dunfield, "Cyclic surgery, degrees of maps of character curves, and volume rigidity for hyperbolic manifolds", *Invent. Math.* **136**:3 (1999), 623–657. MR 2000d:57022 Zbl 0928.57012

[DuPre 1968] A. M. DuPre, III, "Real Borel cohomology of locally compact groups", *Trans. Amer. Math. Soc.* **134** (1968), 239–260. MR 38 #2240 Zbl 0201.55603

[Francaviglia 2004] S. Francaviglia, "Hyperbolic volume of representations of fundamental groups of cusped 3-manifolds", *Int. Math. Res. Not.* **2004**:9 (2004), 425–459. MR 2004m:57032 Zbl 1088.57015

[Francaviglia and Klaff 2006] S. Francaviglia and B. Klaff, "Maximal volume representations are Fuchsian", *Geom. Dedicata* **117** (2006), 111–124. MR 2007d:51019 Zbl 1096.51004

[Goldman 1982] W. M. Goldman, "Characteristic classes and representations of discrete subgroups of Lie groups", *Bull. Amer. Math. Soc.* (*N.S.*) **6**:1 (1982), 91–94. MR 83b:22012 Zbl 0493.57011

[Gromov 1982] M. Gromov, "Volume and bounded cohomology", *Inst. Hautes Études Sci. Publ. Math.* **56** (1982), 5–99. MR 84h:53053 Zbl 0516.53046

[Guichardet 1980] A. Guichardet, *Cohomologie des groupes topologiques et des algèbres de Lie*, Textes Mathématiques **2**, CEDIC, Paris, 1980. MR 83f:22004 Zbl 0464.22001

[Hatcher 2002] A. Hatcher, *Algebraic topology*, Cambridge University Press, 2002. MR 2002k:55001 Zbl 1044.55001

[Inoue and Yano 1982] H. Inoue and K. Yano, "The Gromov invariant of negatively curved manifolds", *Topology* **21**:1 (1982), 83–89. MR 82k:53091 Zbl 0469.53038

[Ivanov 1985] N. V. Ivanov, "Основания теории ограниченных когомологий", *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* (*LOMI*) **143** (1985), 69–109. Translated as "Foundations of the theory of bounded cohomology" in *J. Soviet Math.* **37**:3 (1987), 1090–1115. MR 87b:53070 Zbl 0573.55007

[Kim and Kim 2014] S. Kim and I. Kim, "Volume invariant and maximal representations of discrete subgroups of Lie groups", *Math. Z.* **276**:3-4 (2014), 1189–1213. MR 3175177 Zbl 1292.22003

[Kim and Kuessner 2015] S. Kim and T. Kuessner, "Simplicial volume of compact manifolds with amenable boundary", *J. Topol. Anal.* **7**:1 (2015), 23–46. MR 3284388 Zbl 1310.53040

[Löh 2007] C. Löh, *ℓ^1-homology and simplicial volume*, thesis, Westfälische Wilhelms-Universität, Münster, 2007, available at http://nbn-resolving.de/urn:nbn:de:hbz:6-37549578216. Zbl 1152.57304

[Monod 2001] N. Monod, *Continuous bounded cohomology of locally compact groups*, Lecture Notes in Mathematics **1758**, Springer, Berlin, 2001. MR 2002h:46121 Zbl 0967.22006

Received January 6, 2015.

SUNGWOON KIM
DEPARTMENT OF MATHEMATICS
JEJU NATIONAL UNIVERSITY
102 JEJUDAEHAK-RO
JEJU 63243
REPUBLIC OF KOREA
sungwoon@jejunu.ac.kr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.69

STRONGLY POSITIVE REPRESENTATIONS OF EVEN GSPIN GROUPS

YEANSU KIM

We obtain a classification of strongly positive representations of split even general spin groups over a p-adic field F using the Jacquet module method (Tadić's structure formula). Furthermore, we study discrete series representations of split even general spin groups over F.

1. Introduction

The classifications of (strongly positive) discrete series representations of metaplectic groups, classical groups, and odd GSpin groups over a p-adic field have been studied by several authors [Kim 2015b; Matić 2011; Mœglin 2002; Mœglin and Tadić 2002; Zelevinsky 1980]. The main purpose of this paper is to obtain a classification of strongly positive representations of split even GSpin groups over a nonarchimedean local field F of characteristic different from two, assuming the uniqueness of the nonnegative rank one reducibility point (see Remark 1.2 for more details about this assumption). Our results generalize Matić's algebraic approach [Matić 2011] to the case of even GSpin groups. Our results for even GSpin groups can be applied to even special orthogonal groups to classify strongly positive representations of SO_{2n}. In addition, the results are parallel to those for odd GSpin groups [Kim 2015b]. However, parts of their proofs are quite different because of differences in the group structures. For example, there are two nonconjugate standard parabolic subgroups whose Levi subgroups are of the form $\mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k}$ in the even case; therefore, we classify $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ instead of $D(\rho; \sigma_{\mathrm{cusp}})$ in Section 4B, where c is an outer automorphism on the Dynkin diagram of even GSpin groups that permutes the last two simple roots.

To explain our results more precisely, let $\boldsymbol{G}_n := \mathbf{GSpin}_{2n}$ denote the split even general spin group of semisimple rank n over F, and $\mathbf{GL}_m$ the general linear group of semisimple rank m. Let G_n and GL_m denote the groups of F-points of $\boldsymbol{G}_n$ and $\mathbf{GL}_m$, respectively. Let R and R_{GL} denote the Grothendieck groups of the category of all admissible representations of finite length of even GSpin and GL groups. Note that R contains two inequivalent representations of the group

MSC2010: primary 20C11; secondary 11F70.
Keywords: strongly positive representations, discrete series representations, Jacquet module.

$\mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k}$ (e.g., $\rho_1 \otimes \cdots \otimes \rho_k \otimes (1 \otimes e)$ and $\rho_1 \otimes \cdots \otimes \rho_k \otimes (1 \otimes c)$, following the notation in Section 2).

Let SP denote the set of all strongly positive representations in R, and let LJ denote the set of pairs (Jord, σ'), where

$$\mathrm{Jord} = \bigcup_{i=1}^{k} \bigcup_{j=1}^{k_i} \{(\rho_i, b_j^{(i)})\}$$

and σ' is an irreducible supercuspidal representation in R such that

- $\{\rho_1, \rho_2, \ldots, \rho_k\} \subset R_{\mathrm{GL}}$ is a (possibly empty) set of mutually nonisomorphic, irreducible and essentially self-dual supercuspidal unitary representations such that $\nu^{a_{\rho_i}} \rho_i \rtimes \sigma'$ is reducible for $a_{\rho_i} > 0$ (this defines a_{ρ_i} due to the uniqueness of the reducibility point; see Remark 1.2 for more details),
- $k_i = \lceil a_{\rho_i} \rceil$, and
- for each $i = 1, 2, \ldots, k$, the sequence $b_1^{(i)}, b_2^{(i)}, \ldots, b_{k_i}^{(i)}$ of real numbers is such that $-1 < b_1^{(i)} < b_2^{(i)} < \cdots < b_{k_i}^{(i)}$ and $a_{\rho_i} - b_j^{(i)} \in \mathbb{Z}$ for each $j = 1, \ldots, k_i$.

Remark 1.1. In the set LJ, the condition of "being essentially self-dual" on the representation ρ_i for each $i = 1, \ldots, k$ is due to certain Weyl group actions on the induced representation $\nu^a \rho_i \rtimes \sigma'$ (see Corollary 4.6). In the case of even special orthogonal groups, after a minor change to the set LJ, we can construct an SO version of LJ that corresponds to the set of strongly positive representations of even SO. One minor change would be the condition of "being self-dual" on the corresponding representation ρ_i in the case of even special orthogonal groups.

Remark 1.2 [Silberger 1980]. Let ρ and σ denote irreducible unitary supercuspidal representations of GL_n and G_n, respectively. In this paper, we assume that there exists a unique nonnegative real number a such that $\nu^a \rho \rtimes \sigma$ reduces. This number a is called the nonnegative rank one reducibility point determined by ρ and σ.

We construct a bijective mapping as follows (see Theorem 4.16).

Theorem A. *There exists a bijective mapping* Φ *between* SP *and* LJ. *More precisely, consider* $\sigma \in$ SP *to be the unique irreducible subrepresentation of the induced representation of the form*

$$\left(\prod_{i=1}^{k} \prod_{j=1}^{k_i} \delta\big([\nu^{a_{\rho_i} - k_i + j} \rho_i, \nu^{b_j^{(i)}} \rho_i]\big) \right) \rtimes \sigma'.$$

Then, we may define $\Phi(\sigma)$ *as*

$$\left(\bigcup_{i=1}^{k} \bigcup_{j=1}^{k_i} \{(\rho_i, b_j^{(i)})\}, \sigma' \right) \in \mathrm{LJ}.$$

To construct Φ, we first classify the special case $D(\rho; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}}))$ (see Section 4B), which is the set of strongly positive representations whose supercuspidal supports are the representations σ_{cusp}, $c(\sigma_{\text{cusp}})$ of G_n and twists of the representation ρ of GL. More precisely, in Theorem 4.9 we construct a bijective mapping between $D(\rho; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}}))$ and the set of induced representations of the form

$$\delta\big([\nu^{a_1}\rho, \nu^{b_1}\rho]\big) \times \delta\big([\nu^{a_2}\rho, \nu^{b_2}\rho]\big) \times \cdots \times \delta\big([\nu^{a_k}\rho, \nu^{b_k}\rho]\big) \rtimes c'(\sigma_{\text{cusp}}),$$

where $a_i = a - k + i$, $b_1 < \cdots < b_k$, $k \leq \lceil a \rceil$ and $c' \in C := \{e, c\}$. Here, a is the reducibility point determined by ρ and σ_{cusp}, i.e., $\nu^s\rho \rtimes \sigma_{\text{cusp}}$ is reducible if and only if $|s| = a$. We then generalize this result to the set of strongly positive representations. The classification for even GSpin groups needs more work than those for odd GSpin groups since we consider two different representations σ_{cusp}, $c(\sigma_{\text{cusp}})$ in $D(\rho; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}}))$. (In the odd case, we only need to consider $D(\rho; \sigma_{\text{cusp}})$.)

As an application, our main results give rise to a proof of the equality of L-functions through the local Langlands correspondence in the case of GSpin groups [Kim 2015a]. Furthermore, the equality of L-functions also has an application in the proof of the generic Arthur packet conjecture in our case. Briefly, the generic Arthur packet conjecture states that if the L-packet attached to the Arthur parameter has a generic member, then it is tempered [Shahidi 2011]. This conjecture can be considered a local version of the Generalized Ramanujan Conjecture.

The second purpose of the paper is to explicitly construct Tadić's structure formula for even GSpin groups. Tadić's structure formula studies the Jacquet modules of parabolically induced representations, and it is one of the main tools in the proof of Theorem A. We apply and adapt the ideas from papers of Ban [1999a] and Jantzen [2006] (Tadić's structure formula for even special orthogonal groups) to our case.

The paper is organized as follows. In Section 2, we recall the standard notation and preliminaries. In Section 3, we construct Tadić's structure formula for G_n (Theorem 3.4), which gives the explicit structure of the Jacquet modules of the parabolically induced representations of G_n. In Section 4, we construct the classification of strongly positive representations for G_n (Theorem A). In Section 5, we describe embeddings of the general discrete series representations using Casselman's square integrability criterion [Kim 2009] (Theorem 5.1). This embedding of discrete series representations, together with Theorem A, has an interesting application in the proof of the equality of L-functions through the local Langlands correspondence for GSpin groups [Kim 2015a].

2. Notation and preliminaries

Let F be a nonarchimedean local field of characteristic different from two. For a connected reductive group $\boldsymbol{G}$ defined over F, we let G be the group of F-points of

$\boldsymbol{G}$. Let $\boldsymbol{G}_n$ be a split even general spin group of semisimple rank n defined over F, as defined by Asgari [2002]. A split even GSpin group $\boldsymbol{G}_n := \mathbf{GSpin}_{2n}$ is a split reductive linear algebraic group of type D_n whose derived group is a double covering of a split special orthogonal group and whose connected component of the Langlands dual group is ${}^L G^0 = \mathrm{GSO}_{2n}(\mathbb{C})$. Therefore, the based root datum of G_n is the dual based root datum of GSO_{2n}.

The following proposition describes the structure of GSpin groups as studied by Asgari [2002].

Proposition 2.1. *The root datum* (X^*, R^*, X_*, R_*) *of* G_n *can be described as*

$$X^* = \mathbb{Z}e_0 \oplus \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_n,$$
$$X_* = \mathbb{Z}e_0^* \oplus \mathbb{Z}e_1^* \oplus \cdots \oplus \mathbb{Z}e_n^*$$

(*there is a standard* $\mathbb{Z}$*-pairing* $\langle\ ,\ \rangle$ *on* $X^* \times X_*$), *with* R^* *and* R_* *generated by*

$$\Delta^* = \{\alpha_1 = e_1 - e_2, \alpha_2 = e_2 - e_3, \ldots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = e_{n-1} + e_n\},$$
$$\Delta_* = \{\alpha_1^\vee = e_1^* - e_2^*, \alpha_2^\vee = e_2^* - e_3^*, \ldots, \alpha_{n-1}^\vee = e_{n-1}^* - e_n^*, \alpha_n^\vee = e_{n-1}^* + e_n^* - e_0^*\}.$$

Let $\boldsymbol{s} = (n_1, n_2, \ldots, n_k; n')$ be an ordered partition of n. Let $P_{\boldsymbol{s}} = M_{\boldsymbol{s}} N_{\boldsymbol{s}}$ denote the standard parabolic subgroup of G_n that corresponds to the partition $\boldsymbol{s}$. The Levi factor $M_{\boldsymbol{s}}$ is isomorphic to $\mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k} \times G_{n'}$ [Asgari 2002, Theorem 2.7]. When $n' = 0$ and $n_k > 0$, we have two nonconjugate standard parabolic subgroups whose Levi subgroups are of the form $\mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k}$. In this case, the corresponding set of simple roots contains exactly one of α_{n-1}, α_n. The corresponding Levi factor is denoted by $M_{(n_1,\ldots,n_k;0)} = \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k}$, if the corresponding set of simple roots contains α_{n-1}; or by $c(M_{(n_1,\ldots,n_k;0)})$ otherwise. Here, we let c be the outer automorphism on the Dynkin diagram of G_n that permutes α_{n-1} and α_n and fixes other simple roots.

For representations $\rho_1, \ldots, \rho_k$ of $\mathrm{GL}_{n_1}, \ldots, \mathrm{GL}_{n_k}$, we let $\rho_1 \otimes \cdots \otimes \rho_k \otimes (1 \otimes e)$ denote a representation of $M_{(n_1,\ldots,n_k;0)}$, and $\rho_1 \otimes \cdots \otimes \rho_k \otimes (1 \otimes c))$ a representation of $c(M_{(n_1,\ldots,n_k;0)})$. Let ν be a character of GL_n defined by $|\det|_F$. We denote the induced representation $\mathrm{Ind}_P^{G_n}(\rho_1 \otimes \cdots \otimes \rho_k \otimes \sigma)$ by

$$\rho_1 \times \cdots \times \rho_k \rtimes \sigma$$

where each ρ_i is a representation of some GL_{n_i}, and σ is a representation of G_n. We also write $r_{\boldsymbol{s}}(\sigma)$ for the normalized Jacquet module of the representation σ with respect to $P_{\boldsymbol{s}}$. In other words, $r_{\boldsymbol{s}}$ is a functor from admissible representations of G_n to admissible representations of $M_{\boldsymbol{s}}$. In particular, for a subquotient σ of $\rho \rtimes \sigma_{\mathrm{cusp}}$, where ρ is a representation of GL_k and σ_{cusp} is a supercuspidal representation of G_n, we define $r_{\mathrm{GL}}(\sigma) = r_{(k;n)}(\sigma)$.

In the case of GL, for $P' = M'N'$ the standard parabolic subgroup of GL_n with $M' \cong \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k}$, we denote the induced representation $\mathrm{Ind}_{P'}^{\mathrm{GL}_n}(\rho_1 \otimes \cdots \otimes \rho_k)$ by

$$\rho_1 \times \cdots \times \rho_k,$$

with each ρ_i a representation of some GL_{n_i}.

We also follow the notation introduced in [Bernstein and Zelevinsky 1977]. Let ρ be an irreducible unitary supercuspidal representation of some GL_p. We define the segment $\Delta := [\nu^a\rho, \nu^{a+k}\rho] = \{\nu^a\rho, \nu^{a+1}\rho, \ldots, \nu^{a+k}\rho\}$, where $a \in \mathbb{R}$ and $k \in \mathbb{Z}_{\geq 0}$. If $a > 0$, we call the segment Δ strongly positive. We note that the induced representation $\nu^{a+k}\rho \times \nu^{a+k-1}\rho \times \cdots \times \nu^a\rho$ has a unique irreducible subrepresentation, which we denote by $\delta(\Delta)$. Then $\delta(\Delta)$ is an essentially square-integrable representation attached to Δ (Section 3.1 of [Zelevinsky 1980]).

Let us briefly review the Langlands classification for general linear groups. For every irreducible essentially square-integrable representation δ of some GL_n, there exists a unique $e(\delta) \in \mathbb{R}$ such that the representation $\nu^{-e(\delta)}\delta$ is unitarizable. When $\delta = \delta(\Delta)$, we simply denote $e(\delta(\Delta))$ by $e(\Delta)$. Suppose $\delta_1, \delta_2, \ldots, \delta_k$ are irreducible essentially square-integrable representations of $\mathrm{GL}_{n_1}, \mathrm{GL}_{n_2}, \ldots, \mathrm{GL}_{n_k}$ with $e(\delta_1) \leq e(\delta_2) \leq \cdots \leq e(\delta_k)$. Then, the induced representation $\delta_1 \times \delta_2 \times \cdots \times \delta_k$ has a unique irreducible subrepresentation, which we denote by $L(\delta_1, \delta_2, \ldots, \delta_k)$. This irreducible subrepresentation is called the Langlands subrepresentation, and it appears with the multiplicity one in $\delta_1 \times \delta_2 \times \cdots \times \delta_k$. Every irreducible representation ρ of GL_n is isomorphic to some $L(\delta_1, \delta_2, \ldots, \delta_k)$. Given ρ, the representations $\delta_1, \delta_2, \ldots, \delta_k$ are unique up to a permutation. If $i_1, i_2, \ldots, i_k$ is a permutation of $1, 2, \ldots, k$ such that the representations $\delta_{i_1} \times \cdots \times \delta_{i_k}$ and $\delta_1 \times \cdots \times \delta_k$ are isomorphic, we also write $L(\delta_{i_1}, \delta_{i_2}, \ldots, \delta_{i_k})$ for $L(\delta_1, \delta_2, \ldots, \delta_k)$.

The Grothendieck group of the category of all finite length admissible representations of G_n (resp. GL_n), a free abelian group over the set of all irreducible representations of G_n (resp. GL_n), is denoted by $R(n)$ (resp. $R_{\mathrm{GL}}(n)$). We set

$$R = \bigoplus_{n \geq 0} R(n),$$

$$R_{\mathrm{GL}} = \bigoplus_{n \geq 0} R_{\mathrm{GL}}(n).$$

The strongly positive representations of G_n are defined as follows.

Definition 2.2 (strongly positive). An irreducible representation σ of G_n is called strongly positive if for each representation $\nu^{s_1}\rho_1 \times \nu^{s_2}\rho_2 \times \cdots \times \nu^{s_k}\rho_k \rtimes \sigma_{\mathrm{cusp}}$, where each ρ_i ($i = 1, 2, \ldots, k$) is an irreducible supercuspidal unitary representation of some GL_{n_i}, $\sigma_{\mathrm{cusp}} \in R$ is an irreducible supercuspidal representation of $G_{n'}$ and

$s_i \in \mathbb{R}$ $(i = 1, 2, \ldots, k)$ is such that

$$\sigma \hookrightarrow \nu^{s_1}\rho_1 \times \nu^{s_2}\rho_2 \times \cdots \times \nu^{s_k}\rho_k \rtimes \sigma_{\text{cusp}},$$

we have $s_i > 0$ for each i.

Finally, the next proposition recalls the properties of discrete series representations studied in [Asgari 2002].

Proposition 2.3. *Let M be $\mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \cdots \times \mathrm{GL}_{n_k} \times G_{n'} \subset G_n$. Let ρ_i be a supercuspidal representation of GL_{n_i} and σ a supercuspidal representation of $G_{n'}$. Write $\rho_i = \nu^{e(\rho_i)}\rho_i^u$, where $e(\rho_i) \in \mathbb{R}$ and ρ_i^u is a unitary supercuspidal representation for $i = 1, \ldots, k$. If $\rho_1 \times \cdots \times \rho_k \rtimes \sigma$ has a discrete series subrepresentation, then $\rho_i^u \cong \tilde{\rho}_i^u \otimes (\omega_\sigma \circ \det)$ for $i = 1, \ldots, k$.*

3. Tadić's structure formula for even GSpin groups

Tadić's structure formulae for Sp_{2n}, SO_{2n+1}, SO_{2n} and GSpin_{2n+1} in [Tadić 1995; Jantzen 2006; Kim 2015b] are based on the geometric lemma (2.11 in [Bernstein and Zelevinsky 1977] or Section 6 in [Casselman 1995]). Briefly, the geometric lemma explicitly calculates the Jacquet modules of induced representations ($r_{G,N} \circ i_{G,N}$ in [Bernstein and Zelevinsky 1977]) and it depends on the double coset representative Weyl group elements ($W^{M,N}$ in [Bernstein and Zelevinsky 1977]) and their actions on the simple roots and representations of Levi subgroups. In this section, we explicitly construct the structure of Jacquet modules of parabolically induced representations of G_n (Tadić's structure formula for G_n, Theorem 3.4) using the geometric lemma. We will adapt and follow the results in [Ban 1999a; Jantzen 2006], i.e., the case of SO_{2n}. Ban characterizes the double coset representative Weyl group elements ($[W_{\Omega_{i_1}} \setminus W / W_{\Omega_{i_2}}]$ in [Ban 1999a, Section 5]) and its action on the simple roots in the cases of D_n-type groups, which include SO_{2n} and G_n. Jantzen [2006] constructs Tadić's structure formula for SO_{2n} using Ban's results. Therefore, once we calculate the Weyl group action on the representations in our case, we are ready to adapt Jantzen's calculation [2006] to construct Tadić's structure formula for G_n.

Let $(p, \epsilon) \in S_n \rtimes \{\pm 1\}^n$ be an element in the Weyl group W_{G_n} with $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{\pm 1\}^n$ such that $\prod_{i=1}^n \epsilon_i = 1$. We can identify $(p, \epsilon) \in W_{G_n}$ with $p \cdot \epsilon \in W_{\mathrm{SO}_{2n}}$ where the action by conjugation of p and $\epsilon \in W_{\mathrm{SO}_{2n}}$ on the standard maximal torus in SO_{2n} can be defined by

$$p \cdot \mathrm{diag}(x_1, \ldots, x_n, x_n^{-1}, \ldots, x_1^{-1}) = \mathrm{diag}(x_{p^{-1}(1)}, \ldots, x_{p^{-1}(n)}, x_{p^{-1}(n)}^{-1}, \ldots, x_{p^{-1}(1)}^{-1})$$
$$\epsilon \cdot \mathrm{diag}(x_1, \ldots, x_n, x_n^{-1}, \ldots, x_1^{-1}) = \mathrm{diag}(x_1^{\epsilon_1}, \ldots, x_n^{\epsilon_n}, x_n^{-\epsilon_n}, \ldots, x_1^{-\epsilon_1}).$$

We can get the action of those on the roots (see also [Hundley and Sayag 2012]).

Lemma 3.1. *Let $e_0, e_1, \dots, e_n$ and $e'_0, e'_1, \dots, e'_n$ be the standard bases of the character lattice and the cocharacter lattice of G_n as in Proposition 2.1. Let $(p, \epsilon) \in S_n \rtimes \{\pm 1\}^n$ be as above.*

Then

$$(p,\epsilon)\cdot e_i = \begin{cases} e_{p(i)} & \text{for } i > 0,\ \epsilon_i = 1, \\ -e_{p(i)} & \text{for } i > 0,\ \epsilon_i = -1, \\ e_0 + \sum_{\epsilon_i=-1} e_{p(i)} & \text{for } i = 0, \end{cases}$$

$$(p,\epsilon)\cdot e'_i = \begin{cases} e'_{p(i)} & \text{for } i > 0,\ \epsilon_i = 1, \\ e'_0 - e'_{p(i)} & \text{for } i > 0,\ \epsilon_i = -1, \\ e'_0 & \text{for } i = 0. \end{cases}$$

Proof. The calculations of $(p, \epsilon)\cdot e_i$ for $i > 0$ are done in [Ban 1999a, Section 5]. We can also calculate $(p, \epsilon)\cdot e'_i$ directly from the matrix calculation since $e'_0, \dots, e'_n$ comprise the character lattice of GSO. For $(p, \epsilon)\cdot e_0$, we need to use the duality of e_i and e'_i. □

Let $\Delta^* = \{\alpha_1 = e_1 - e_2, \alpha_2 = e_2 - e_3, \dots, \alpha_{n-1} = e_{n-1} - e_n, \alpha_n = e_{n-1} + e_n\}$ be a simple root for G_n as explained in Section 2. From Lemma 3.1, we can calculate the action of (p, ϵ) on the set of simple roots in R^* (see also [Ban 1999a]).

Corollary 3.2. *With notation as in Lemma 3.1,*

$$(p,\epsilon)\cdot \alpha_i = \begin{cases} \epsilon_i e_{p(i)} - \epsilon_{i+1} e_{p(i+1)} & \text{for } 0 \le i \le n-1, \\ \epsilon_{n-1} e_{p(n-1)} + \epsilon_n e_{p(n)} & \text{for } i = n. \end{cases}$$

Now we are ready to construct Tadić's structure formula for G_n. Let ρ_i be an irreducible smooth representation of GL_{n_i} for $i = 1, 2, 3, 4$. Let σ be an irreducible smooth representation in R and let ω_σ be the central character of σ. For any element $c_1 \in C = \{e, c\}$, we define $\tilde{\rtimes}$ as follows:

$$(\rho_1 \otimes \rho_2 \otimes \rho_3 \otimes c_1)\tilde{\rtimes}(\rho_4 \otimes \sigma) = (\tilde{\rho}_1 \otimes (\omega_\sigma \circ \det)) \times \rho_2 \times \rho_4 \otimes \rho_3 \rtimes c_1(\sigma). \tag{3-1}$$

One extends $\tilde{\rtimes}$ to a $\mathbb{Z}$-bilinear mapping

$$\tilde{\rtimes} : (R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \otimes \mathbb{Z}[C]) \times (R_{\mathrm{GL}} \otimes R) \to R_{\mathrm{GL}} \otimes R.$$

We denote by $\boldsymbol{m}$ the linear extension to $R_{\mathrm{GL}} \otimes R_{\mathrm{GL}}$ of parabolic induction from a maximal parabolic subgroup.

Let

$$\Omega_k = \begin{cases} \Delta & \text{if } k = 0, \\ \Delta \setminus \{\alpha_k\} & \text{if } k \le n-2, \\ \Delta \setminus \{\alpha_{n-1}, \alpha_n\} & \text{if } k = n-1, \\ \Delta \setminus \{\alpha_n\} & \text{if } k = n. \end{cases}$$

and $\bar{\Omega}_n = \Delta \setminus \{\alpha_{n-1}\} = c(\Omega_n)$. We define $\mu^*(\sigma)$ as follows. For $0 \le k \le n$, write

$$r_{M_{\Omega_k},G_n}(\sigma) = \sum_{i \in I_k} \rho_{i,k} \otimes \sigma_{i,k},$$
$$r_{M_{\bar{\Omega}_n},G_n}(\sigma) = \sum_{j \in J} \rho_j \otimes (1 \otimes c),$$

the normalized Jacquet modules of σ with respect to the standard maximal parabolic subgroups $P_{\Omega_k} = M_{\Omega_k} N_{\Omega_k}$ and $P_{\bar{\Omega}_n} = M_{\bar{\Omega}_n} N_{\bar{\Omega}_n}$, respectively. For such σ, we can define $\mu^*(\sigma) \in R_{\mathrm{GL}} \otimes R$ as

$$\mu^*(\sigma) = \sum_{k=0}^{n} \sum_{i \in I_k} \rho_{i,k} \otimes \sigma_{i,k} + \sum_{j \in J} \rho_j \otimes (1 \otimes c).$$

Using Jacquet modules with respect to the maximal parabolic subgroups of GL_n, we can also define

$$\boldsymbol{m}^*(\rho) = \sum_{k=0}^{n} s.s.(r_k(\rho)) \in R_{\mathrm{GL}} \otimes R_{\mathrm{GL}}$$

for an irreducible representation ρ of GL_n, and then extend $\boldsymbol{m}^*$ linearly to the whole of R_{GL}. Here, $r_k(\rho)$ denotes the Jacquet module of the representation ρ with respect to the parabolic subgroup whose Levi subgroup is $\mathrm{GL}_k \times \mathrm{GL}_{n-k}$, and $s.s.$ denotes so-called semisimplification, i.e., a canonical map from objects of the category of all smooth finite length representations of GL into the Grothendieck group of this category. We define $\boldsymbol{s} : R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \to R_{\mathrm{GL}} \otimes R_{\mathrm{GL}}$ by $\boldsymbol{s}(x \otimes y) = y \otimes x$. Let $\mathfrak{M}^*_C : R_{\mathrm{GL}} \to R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \otimes \mathbb{Z}[C]$ be the map $(1 \otimes \boldsymbol{m}^*_C) \circ \boldsymbol{s} \circ \boldsymbol{m}^*$, where

$$1 \otimes \boldsymbol{m}^*_C : R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \to R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \otimes R_{\mathrm{GL}} \otimes \mathbb{Z}[C]$$
$$\rho_1 \otimes \rho_2 \mapsto \rho_1 \otimes \boldsymbol{m}^*(\rho_2) \otimes c^{n_1},$$

for representations ρ_1 of GL_{n_1} and ρ_2 of GL_{n_2}.

Remark 3.3. In our case of even GSpin groups, we have c^{n_1} when we calculate μ^*, while we don't have such action in the case of odd GSpin groups. This is due to the differences of the Weyl group actions on the simple roots. In our case, the corresponding Weyl group element acts on e_n by $(-1)^{n_1}$. Therefore, if n_1 is odd, we need to permute $\alpha_{n-1} = e_{n-1} - e_n$ and $\alpha_n = e_{n-1} + e_n$. In other words, we have the outer automorphism c on the representation σ.

The following theorem is called Tadić's structure formula for even GSpin groups.

Theorem 3.4. *For $\rho \in R_{\mathrm{GL}}(i)$ and $\sigma \in R(n-i)$, we have*

$$\mu^*(\rho \rtimes \sigma) = \mathfrak{M}^*_C(\rho) \mathbin{\widetilde{\rtimes}} \mu^*(\sigma).$$

Proof. We sketch the proof and explain how we can adapt the approach in [Jantzen 2006] to our case. Let us explicitly calculate the left hand side of the equation in the theorem. Since GSpin_{2n} is also of D_n-type, we can apply the double coset representative Weyl group elements, which are studied in [Ban 1999a], to our case. Therefore, we have $\mu^*(\rho \rtimes \sigma)$ as in [Jantzen 2006, pp. 811–812]. Now, we calculate the action of the double coset representative Weyl group elements (q_n in [Jantzen 2006]) on the representations. The actions of q_n produce contragredient of $\tau_s^{(3)}(d)$ with $(\omega_\theta \circ \det)$, i.e., $(\omega_\theta \circ \det)\tilde{\tau}_s^{(3)}(d)$ (see p. 812 for the notation of $\tau_s^{(3)}(d)$ and θ). Accordingly, we need to define $\tilde{\rtimes}$ as (3-1). This forces $\mu^*(\rho \rtimes \sigma)$ to be equal to $\mathfrak{M}_C^*(\rho)\tilde{\rtimes}\mu^*(\sigma)$ after changing index several times as in the proof of [Jantzen 2006, Chapter 4]. □

4. Classification of strongly positive representations for even GSpin groups

We give the classification of strongly positive representations of even GSpin groups in this section. We apply and adapt some proofs from Matić's results [2011] for metaplectic groups to our case. Therefore, we mostly focus on the following cases which are quite different from [Matić 2011]. For example, when the reducibility point is $\frac{1}{2}$, we follow the idea of [Kim 2015b] instead of [Matić 2011]. We also emphasize the difference between the even case and the odd case, and omit the proof if it is similar to the case of either metaplectic groups or odd GSpin groups. For example, in the even case, we need to add $c(\sigma_{\mathrm{cusp}})$ when we classify the special case $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ (in the odd case, we classify $D(\rho; \sigma_{\mathrm{cusp}})$).

4A. *Several lemmas.* We recall several lemmas which are essential in this section. Let us first recall the half integer conjecture in the case of GSpin groups. Let σ be an irreducible supercuspidal representation of $G_{n'}$ and let ρ be an irreducible supercuspidal unitary representation of GL_k. The following is a recent result of Mœglin which is called the half integer conjecture for GSpin groups [Mœglin 2014, Theorem 3.1.1]:

Lemma 4.1 [Mœglin 2014]. *Let $a \in \mathbb{R}$ be a nonnegative real number such that $\nu^a\rho \rtimes \sigma$ reduces. Then, $a \in \frac{1}{2}\mathbb{Z}$.*

Remark 4.2. In [Kim 2015b], we classified the strongly positive representations of odd GSpin groups assuming the half integer conjecture. Due to Mœglin's results (Lemma 4.1), we can completely remove the assumption in the odd case.

Remark 4.3. When we further assume that σ is a generic representation, Shahidi [1990] proved that $a \in \{0, \frac{1}{2}, 1\}$.

Let us calculate $\mu^*\big(\big(\prod_{j=1}^n \delta\big([\nu^{a_j}\rho_j, \nu^{b_j}\rho_j]\big)\big) \rtimes \sigma\big)$, where each ρ_j is an irreducible supercuspidal representation of GL_k, the real numbers a_j and b_j are such that $b_j - a_j \in \mathbb{Z}_{\geq 0}$ for each $j = 1, \ldots, n$ and σ is an irreducible supercuspidal representation of finite length of $G_{n'}$.

Example 4.4. Let us first calculate the case when $n = 1$. Since σ is supercuspidal, $\mu^*(\sigma) = 1 \otimes \sigma$. Following the definition of $\mathfrak{M}^*_C$, we have

$$\mathfrak{M}^*_C\big(\delta\big([\nu^a\rho_1, \nu^b\rho_1]\big)\big)$$
$$= \sum_{i=a-1}^{b}\sum_{j=i}^{b} \delta\big([\nu^a\rho_1, \nu^i\rho_1]\big) \otimes \delta\big([\nu^{j+1}\rho_1, \nu^b\rho_1]\big) \otimes \delta\big([\nu^{i+1}\rho_1, \nu^j\rho_1]\big) \otimes c^{k(i-a+1)}.$$

Therefore,

$$\mu^*\big(\delta\big([\nu^a\rho_1, \nu^b\rho_1]\big) \rtimes \sigma\big) = \sum_{i=a-1}^{b}\sum_{j=i}^{b} \delta\big([\nu^{-i}\tilde{\rho}_1 \otimes (\omega_\sigma \circ \det), \nu^{-a}\tilde{\rho}_1 \otimes (\omega_\sigma \circ \det)]\big)$$
$$\times \delta\big([\nu^{j+1}\rho_1, \nu^b\rho_1]\big) \otimes \delta\big([\nu^{i+1}\rho_1, \nu^j\rho_1]\big) \rtimes c^{k(i-a+1)}(\sigma).$$

We omit $\delta\big([\nu^x\rho_1, \nu^y\rho_1]\big)$ if $x > y$. Then, to calculate

$$\mu^*\Bigg(\bigg(\prod_{j=1}^{n} \delta\big([\nu^{a_j}\rho_j, \nu^{b_j}\rho_j]\big)\bigg) \rtimes \sigma\Bigg),$$

we use (1.3) of [Tadić 1998]:

$$m^*\bigg(\prod_{j=1}^{n} \delta\big([\nu^{a_j}\rho_j, \nu^{b_j}\rho_j]\big)\bigg) = \prod_{j=1}^{n}\bigg(\sum_{i_j=a_j-1}^{b_j} \delta\big([\nu^{i_j+1}\rho_j, \nu^{b_j}\rho_j]\big) \otimes \delta\big([\nu^{a_j}\rho_j, \nu^{i_j}\rho_j]\big)\bigg).$$

The Weyl group elements are essential objects when we define the intertwining operators between two induced representations [Shahidi 2010, Chapter 4]. We recall the action of the Weyl group elements on the induced representations. Let M_θ be a Levi subgroup isomorphic to $\mathrm{GL}_k \times G_{n-k}$ for $\theta = \Delta \setminus \alpha_k$. There is a unique Weyl group element w_0 such that $w_0(\alpha_k) < 0$ and $w_0(\theta) \subset \Delta$.

Lemma 4.5. *Let ρ and σ be irreducible supercuspidal representations of GL_k and G_{n-k}, respectively. Then*

$$(\rho \otimes \sigma)^{w_0} = (\tilde{\rho} \otimes (\omega_\sigma \circ \det)) \otimes c^k(\sigma),$$

where ω_σ is the central character of σ.

Proof. Since $w_0(\alpha_k) < 0$ and $w_0(\theta) \subset \Delta$, we can explicitly calculate its action on the simple roots. Let us identify w_0 as $(p, \epsilon) \in S_n \rtimes \{\pm 1\}^n$ with $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{\pm 1\}^n$ such that $\prod_{i=1}^n \epsilon_i = 1$. Then

$$p(i)=\begin{cases}k+1-i & \text{for } 1\le i\le k,\\ i & \text{for } k+1\le i\le n,\end{cases}\quad \text{and}\quad \epsilon_i=\begin{cases}-1 & \text{for } 1\le i\le k,\\ 1 & \text{for } k+1\le i\le n,\\ (-1)^k & \text{for } i=n,\end{cases}$$

for $k\neq n-1, n$, and

$$p(i)=n+1-i\quad \text{and}\quad \epsilon_i=\begin{cases}(-1)^{k-1} & \text{for } i=1,\\ -1 & \text{for } 2\le i\le n-1,\\ (-1)^{k+n-1} & \text{for } i=n,\end{cases}$$

for $k=n-1, n$. Using this identification, the lemma follows from Lemma 3.1. □

Corollary 4.6. *Let ρ and σ be as in Lemma 4.5. Then $\rho\rtimes\sigma$ and $(\tilde\rho\otimes(\omega_\sigma\circ\det))\rtimes c^k(\sigma)$ are associate. Therefore, Lemma* 5.4 (iii) *of* [Bernstein et al. 1986] *implies that the sets of irreducible composition factors of $\rho\rtimes\sigma$ and $(\tilde\rho\otimes(\omega_\sigma\circ\det))\rtimes c^k(\sigma)$ are the same. Furthermore, if we assume that $\rho\rtimes\sigma$ is irreducible, then $\rho\rtimes\sigma\cong(\tilde\rho\otimes(\omega_\sigma\circ\det))\rtimes c^k(\sigma)$.*

Now we show that strongly positive representations can be embedded into parabolically induced representations of special type. More precisely, we consider parabolically induced representations of the form

$$\delta(\Delta_1)\times\delta(\Delta_2)\times\cdots\times\delta(\Delta_k)\rtimes c'(\sigma_{\text{cusp}}), \tag{4-1}$$

where $\Delta_1, \Delta_2, \dots, \Delta_k$ is a sequence of strongly positive segments satisfying $0<e(\Delta_1)\le e(\Delta_2)\le\cdots\le e(\Delta_k)$ (we allow $k=0$ here), σ_{cusp} is an irreducible supercuspidal representation of $G_{n'}$ and $c'\in C$. Note that the idea of such embeddings of representations was initiated in [Muić 2006] and further refined in [Hanzer and Muić 2008].

Lemma 4.7. *Let $\Delta_1, \dots, \Delta_k$ and σ_{cusp} be as above. Then the induced representation $\delta(\Delta_1)\times\delta(\Delta_2)\times\cdots\times\delta(\Delta_k)\rtimes c'(\sigma_{\text{cusp}})$ has a unique irreducible subrepresentation, which we denote by $\delta(\Delta_1, \dots, \Delta_k; c'(\sigma_{\text{cusp}}))$.*

Proof. We briefly explain the main ideas of the proof and how we adapt the proof from [Matić 2011] to the case of even GSpin groups. The case $k=0$ is clear. Let $j_1<j_2<\cdots<j_s$ be the positive integers such that

$$e(\Delta_1)=\cdots=e(\Delta_{j_1})<e(\Delta_{j_1+1})=\cdots=e(\Delta_{j_2})<\cdots<e(\Delta_{j_s+1})=\cdots=e(\Delta_k).$$

Then Theorem 3.4 implies that an irreducible representation

$$\delta(\Delta_1)\times\cdots\times\delta(\Delta_{j_1})\otimes\delta(\Delta_{j_1+1})\times\cdots\times\delta(\Delta_{j_2})\otimes\cdots\otimes c'(\sigma_{\text{cusp}})$$

appears with multiplicity one in the Jacquet module of $\delta(\Delta_1)\times\delta(\Delta_2)\times\cdots\times\delta(\Delta_k)\rtimes c'(\sigma_{\text{cusp}})$ with respect to the appropriate parabolic subgroup. Therefore, since this irreducible representation is contained in the Jacquet module of any subrepresentation of $\delta(\Delta_1) \times \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \rtimes c'(\sigma_{\mathrm{cusp}})$ with respect to the appropriate parabolic subgroup, it follows that the induced representation $\delta(\Delta_1) \times \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \rtimes c'(\sigma_{\mathrm{cusp}})$ has a unique irreducible subrepresentation. □

Now, as in the odd case [Kim 2015b], we can show that strongly positive representations can be embedded into induced representations of the form (4-1).

Lemma 4.8. *Let $\sigma \in R$ denote a strongly positive representation. Then σ can be embedded into certain induced representations of the form (4-1).*

4B. ***Classification of strongly positive representations: the $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ case.*** Let ρ be an essentially self-dual irreducible supercuspidal representation of GL_{n_ρ} and σ_{cusp} an irreducible supercuspidal representation of $G_{n'}$. Also, let $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ be the set of strongly positive representations whose supercuspidal supports are the representations σ_{cusp}, $c(\sigma_{\mathrm{cusp}})$ and twists of the representation ρ. We assume that there exists a unique nonnegative real number a such that $\nu^a\rho \rtimes \sigma_{\mathrm{cusp}}$ reduces [Silberger 1980]. The half integer conjecture for GSpin groups (Lemma 4.1) implies that $a \in \frac{1}{2}\mathbb{Z}$.

In this section, we construct the classification of strongly positive representations in $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$. Lemma 4.8 implies that there exists a mapping from the set of strongly positive representations of G_n into the set of induced representations of the form (4-1). We first refine the image of this map when we restrict the map to $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$.

Theorem 4.9. *Suppose that σ is an irreducible strongly positive representation in $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$, taken as the unique irreducible subrepresentation of an induced representation of the form (4-1). Write $\Delta_i = [\nu^{a_i}\rho, \nu^{b_i}\rho]$ for $i = 1, \ldots, k$. Then $a_i = a - k + i$ for each i, $b_1 < \cdots < b_k$ and $k \le \lceil a \rceil$.*

Proof. We consider the case when $a = \frac{1}{2}$. We first show, by induction on k, that $a_i = a$ for each $i = 1, \ldots, k$ and that $b_1 \le \cdots \le b_k$ when $a = \frac{1}{2}$.

For the case $k = 1$, note that if $a_i \ne a$, then

$$\nu^{a_i}\rho \rtimes \sigma_{\mathrm{cusp}} \cong (\nu^{-a_i}\tilde{\rho} \otimes (\omega_{\sigma_{\mathrm{cusp}}} \circ \det)) \rtimes c^{n_\rho}(\sigma_{\mathrm{cusp}}) \cong \nu^{-a_i}\rho \rtimes \sigma_{\mathrm{cusp}}.$$

Therefore, we have the embedding

$$\sigma \hookrightarrow \nu^{b_1}\rho \times \cdots \times \nu^{a_1+1}\rho \times \nu^{-a_1}\rho \rtimes \sigma_{\mathrm{cusp}}$$

which contradicts the strong positivity of σ.

Now suppose the theorem holds for all $m \in \mathbb{Z}$ such that $0 \le m < k$. We prove it for k. As in the case when $k = 1$, we show $a_k = a$. We know that σ embeds in $\delta(\Delta_1) \rtimes \delta(\Delta_2, \ldots, \Delta_k; \sigma_{\mathrm{cusp}})$, since σ is the unique irreducible subrepresentation of $\delta(\Delta_1) \times \cdots \delta(\Delta_k) \rtimes \sigma_{\mathrm{cusp}}$. This implies that each $a_i = a$ for $2 \le i \le k$ and $b_2 \le \cdots \le b_k$. It remains to show that $a_1 = \frac{1}{2}$ and $b_1 \le b_2$.

Suppose that $a_1 \notin \frac{1}{2}+\mathbb{Z}$. Then $\nu^{a_1}\rho \times \delta(\Delta_i)$ is irreducible for $i \geq 2$. Therefore, we have the embedding

$$\begin{aligned}\sigma \hookrightarrow\ & \delta([\nu^{a_1+1}\rho, \nu^{b_1}\rho]) \times \nu^{a_1}\rho \times \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \rtimes \sigma_{\text{cusp}} \\ &\cong \delta([\nu^{a_1+1}\rho, \nu^{b_1}\rho]) \times \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \times \nu^{a_1}\rho \rtimes \sigma_{\text{cusp}} \\ &\cong \delta([\nu^{a_1+1}\rho, \nu^{b_1}\rho]) \times \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \times \nu^{-a_1}\rho \rtimes \sigma_{\text{cusp}}\end{aligned}$$

which contradicts the strong positivity of σ. Therefore, $a_1 \in \frac{1}{2}+\mathbb{Z}$. If $a_1 \neq \frac{1}{2}$, then $a_1 \geq \frac{3}{2}$. This implies that $\delta(\Delta_1) \times \delta(\Delta_i)$ is irreducible for $i \geq 2$ since $b_1 \leq b_i$. Therefore, we have the embedding

$$\begin{aligned}\sigma \hookrightarrow\ & \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \times \delta(\Delta_1) \rtimes \sigma_{\text{cusp}} \\ &\cong \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \times \delta([\nu^{a_1+1}\rho, \nu^{b_1}\rho]) \times \nu^{a_1}\rho \rtimes \sigma_{\text{cusp}} \\ &\cong \delta(\Delta_2) \times \cdots \times \delta(\Delta_k) \times \delta([\nu^{a_1+1}\rho, \nu^{b_1}\rho]) \times \nu^{-a_1}\rho \rtimes \sigma_{\text{cusp}}\end{aligned}$$

which again contradicts the strong positivity of σ. Thus, $a_1 = \frac{1}{2}$. Also, $b_1 \leq b_2$ follows from $e(\Delta_1) \leq e(\Delta_2)$. The following lemma finishes the proof of the theorem in the case when $a = \frac{1}{2}$.

Lemma 4.10. *The unique irreducible subrepresentation of*

$$\delta([\nu^{1/2}\rho, \nu^{b'_1}\rho]) \times \delta([\nu^{1/2}\rho, \nu^{b'_2}\rho]) \times \cdots \times \delta([\nu^{1/2}\rho, \nu^{b'_k}\rho]) \rtimes \sigma_{\text{cusp}},$$

denoted $\sigma^*_{(b'_1,\dots,b'_k;1/2)}$, *is not strongly positive when* $k \geq 2$.

Proof. We first show the case when $k=2$. The embedding of $\nu^{1/2}\rho \rtimes \delta(\nu^{1/2}\rho, \sigma_{\text{cusp}})$ into $\nu^{1/2}\rho \times \nu^{1/2}\rho \rtimes \sigma_{\text{cusp}}$ implies the embedding

$$\sigma^*_{(1/2,1/2;1/2)} \hookrightarrow \nu^{1/2}\rho \rtimes \delta(\nu^{1/2}\rho, \sigma_{\text{cusp}}).$$

Using Lemma 3.8 (b) as well as Remark 3.2 of [Tadić 1998], we can show that $\delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \rtimes \sigma_{\text{cusp}}$ is a direct sum of two irreducible nonisomorphic representations, say τ_1 and τ_2, in the same way as in the proof of Sublemma 5.8 of [Kim 2015b]. From Frobenius reciprocity and

$$\begin{aligned}& r_{\text{GL}}\big(\delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \rtimes \sigma_{\text{cusp}}\big) = \\ &\qquad 2\delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \otimes \sigma_{\text{cusp}} + \nu^{1/2}\rho \times \nu^{1/2}\rho \otimes c^{n_\rho}(\sigma_{\text{cusp}}),\end{aligned}$$

it follows that $r_{\text{GL}}(\tau_1) = \delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \otimes \sigma_{\text{cusp}} + \nu^{1/2}\rho \times \nu^{1/2}\rho \otimes c^{n_\rho}(\sigma_{\text{cusp}})$ and $r_{\text{GL}}(\tau_2) = \delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \otimes \sigma_{\text{cusp}}$. We also obtain that $\sigma^*_{(1/2,1/2;1/2)} \cong \tau_1$, which is a subrepresentation of $\delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \rtimes \sigma_{\text{cusp}}$, in the same way as in the proof of Sublemma 5.9 of [Kim 2015b]. Therefore, we have the embedding

$$\sigma^*_{(1/2,1/2;1/2)} \hookrightarrow \delta([\nu^{-1/2}\rho, \nu^{1/2}\rho]) \rtimes \sigma_{\text{cusp}} \hookrightarrow \nu^{1/2}\rho \times \nu^{-1/2}\rho \rtimes \sigma_{\text{cusp}}.$$

We conclude that $\sigma^*_{(1/2,1/2;1/2)}$ is not strongly positive.

Next we consider the case $k \geq 2$. Suppose that $\sigma^*_{(b'_1,\dots,b'_k;1/2)}$ is strongly positive. Since each representation $\nu^{1/2}\rho \times \delta([\nu^{1/2}\rho, \nu^{b'_i}\rho])$ is irreducible for all $i = 1, \dots, k$, we have an embedding of $\sigma^*_{(b'_1,\dots,b'_k;1/2)}$ into

$$\delta([\nu^{3/2}\rho, \nu^{b'_1}\rho]) \times \delta([\nu^{3/2}\rho, \nu^{b'_2}\rho]) \times \\ \cdots \times \delta([\nu^{3/2}\rho, \nu^{b'_k}\rho]) \times \nu^{1/2}\rho \times \cdots \times \nu^{1/2}\rho \rtimes \sigma_{\mathrm{cusp}}.$$

Furthermore, since we know $\sigma^*_{(1/2,1/2;1/2)}$ is the unique irreducible subrepresentation of $\nu^{1/2}\rho \times \nu^{1/2}\rho \rtimes \sigma_{\mathrm{cusp}}$, we have the embedding

$$\sigma^*_{(b'_1,\dots,b'_k;1/2)} \hookrightarrow \delta([\nu^{3/2}\rho, \nu^{b'_1}\rho]) \times \cdots \times \nu^{1/2}\rho \rtimes \sigma^*_{(1/2,1/2;1/2)}.$$

This contradicts the strong positivity of $\sigma^*_{(1/2,1/2;1/2)}$. □

Returning to the proof of Theorem 4.9, it remains to prove the case when $a \neq \frac{1}{2}$. This case is similar to the proof in [Matić 2011] and we skip the proof here. □

We also show that the map from $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ to the set of induced representations of the form (4-1) is well defined in the following theorem.

Theorem 4.11. *Suppose that σ is an irreducible strongly positive representation in $D(\rho; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$. Then there exist a unique set of strongly positive segments $\Delta_1, \Delta_2, \dots, \Delta_k$ with $0 < e(\Delta_1) \leq e(\Delta_2) \leq \cdots \leq e(\Delta_k)$, and a unique irreducible supercuspidal representation $\sigma' \in R$ such that $\sigma \simeq \delta(\Delta_1, \Delta_2, \dots, \Delta_k; \sigma')$.*

Proof. We first show the uniqueness of the partial supercuspidal support σ'. Suppose that there are two sets of strongly positive segments and representations in R, $\{\Delta_1, \Delta_2, \dots, \Delta_k, \sigma_{\mathrm{cusp}}\}$ and $\{\Delta'_1, \Delta'_2, \dots, \Delta'_l, c(\sigma_{\mathrm{cusp}})\}$, which satisfy the conditions in Theorem 4.9. Then we have the two embeddings

$$\sigma \hookrightarrow \left(\prod_{i=1}^{k} \delta(\Delta_i)\right) \rtimes \sigma_{\mathrm{cusp}}, \tag{4-2}$$

$$\sigma \hookrightarrow \left(\prod_{j=1}^{l} \delta(\Delta'_j)\right) \rtimes c(\sigma_{\mathrm{cusp}}). \tag{4-3}$$

These embeddings imply that

$$\left(\prod_{j=1}^{l} \delta(\Delta'_j)\right) \otimes c(\sigma_{\mathrm{cusp}}) \leq r_{\mathrm{GL}}(\sigma) \leq r_{\mathrm{GL}}\left(\left(\prod_{i=1}^{k} \delta(\Delta_i)\right) \rtimes \sigma_{\mathrm{cusp}}\right).$$

However, Theorem 3.4 implies that $r_{\mathrm{GL}}\big(\big(\prod_{i=1}^{k} \delta(\Delta_i)\big) \rtimes \sigma_{\mathrm{cusp}}\big)$ can contain the support $c(\sigma_{\mathrm{cusp}})$ only if the corresponding GL part, $\prod_{j=1}^{l} \delta(\Delta'_j)$, has negative exponent. This is a contradiction since each Δ'_i is strongly positive for all i.

It remains to show the uniqueness of strongly positive segments. The situation becomes similar to the odd case [Kim 2015b] since we show the uniqueness of σ'. Therefore, we can apply the idea of [Matić 2011] (for $a \neq \frac{1}{2}$) and [Kim 2015b] (for $a = \frac{1}{2}$) to complete the proof, which we omit here. □

In Theorem 4.9 and Theorem 4.11, we construct an injective mapping from $D(\rho; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}}))$ into the set of induced representations of the form

$$\delta\big([\nu^{a-k+1}\rho, \nu^{b'_1}\rho]\big) \times \delta\big([\nu^{a-k+2}\rho, \nu^{b'_2}\rho]\big) \times \cdots \times \delta\big([\nu^{a}\rho, \nu^{b'_k}\rho]\big) \rtimes c'(\sigma_{\text{cusp}})$$

((4-1) with refinement on the unitary exponents as in Theorem 4.9). In other words, if we let $\text{Jord}^{c'}_{(\rho,a)}$ be the set of $(c'; b_1, b_2, \ldots, b_{k_\rho})$, where $c' \in C$ and $b_i \in \mathbb{R}$ are such that $b_i - a + k_\rho - i \in \mathbb{Z}_{\geq 0}$ for $i = 1, \ldots, k_\rho$ and $-1 < b_1 < b_2 < \cdots < b_{k_\rho}$, we construct the injective mapping

$$D(\rho; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}})) \hookrightarrow \text{Jord}^{e}_{(\rho,a)} \cup \text{Jord}^{c}_{(\rho,a)}\,.$$

It remains to show that this injective mapping is also surjective. For an element $(c'; b_1, b_2, \ldots, b_{k_\rho}) \in \text{Jord}^{e}_{(\rho,a)} \cup \text{Jord}^{c}_{(\rho,a)}$, let $\sigma_{(c'; b_1, \ldots, b_{k_\rho}; a)}$ be a unique irreducible subrepresentation of

$$\delta\big([\nu^{a-k_\rho+1}\rho, \nu^{b_1}\rho]\big) \times \delta\big([\nu^{a-k_\rho+2}\rho, \nu^{b_2}\rho]\big) \times \cdots \times \delta\big([\nu^{a}\rho, \nu^{b_{k_\rho}}\rho]\big) \rtimes c'(\sigma_{\text{cusp}}).$$

To show the surjectivity, we apply the induction argument in [Matić 2011] to show that the above subrepresentation $\sigma_{(c'; b_1, \ldots, b_{k_\rho}; a)}$ is strongly positive. We don't repeat the argument here.

Theorem 4.12. *The representation $\sigma_{(c'; b_1, \ldots, b_{k_\rho}; a)}$ is strongly positive.*

Remark 4.13. In the case of odd GSpin groups, we classify the special case $D(\rho; \sigma_{\text{cusp}})$ in [Kim 2015b]. In the even case, we need to consider $c(\sigma_{\text{cusp}})$ as a support in $D(\rho; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}}))$ as well, since the action of certain Weyl group elements makes σ into $c(\sigma)$ (e.g., Lemma 4.5).

4C. ***Classification of strongly positive representations.*** Let ρ_i be an essentially self-dual irreducible supercuspidal representation of $\text{GL}_{n_{\rho_i}}$ for $i = 1, \ldots, k$, and σ_{cusp} an irreducible supercuspidal representation of $G_{n'}$. We consider the set $D(\rho_1, \rho_2, \ldots, \rho_k; \sigma_{\text{cusp}}, c(\sigma_{\text{cusp}}))$ of strongly positive representations whose supercuspidal supports are the representations σ_{cusp}, $c(\sigma_{\text{cusp}})$ and twists of the representations $\rho_1, \ldots, \rho_k$. We assume that there exists a unique nonnegative real number a_{ρ_i} such that $\nu^{a_{\rho_i}}\rho_i \rtimes \sigma_{\text{cusp}}$ reduces for each i [Silberger 1980]. The half integer conjecture for GSpin groups (i.e., Lemma 4.1) implies that each $a_{\rho_i} \in \frac{1}{2}\mathbb{Z}$.

Theorem 4.14. *Let σ be a strongly positive representation in $D(\rho_1, \rho_2, \dots, \rho_k; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$. Then σ can be considered to be the unique irreducible subrepresentation of the induced representation*

$$\left(\prod_{i=1}^{k}\prod_{j=1}^{k_i}\delta\big([\nu^{a_{\rho_i}-k_i+j}\rho_i, \nu^{b_j^{(i)}}\rho_i]\big)\right)\rtimes c'(\sigma_{\mathrm{cusp}}), \tag{4-4}$$

where $c' \in C = \{e, c\}$ and for $i = 1, \dots, k$ and $j = 1, \dots, k_i$, each $k_i \in \mathbb{Z}_{\geq 0}$ satisfies $k_i \leq \lceil a_{\rho_i}\rceil$ and each $b_j^{(i)} > 0$ is such that $b_j^{(i)} - a_{\rho_i} \in \mathbb{Z}_{\geq 0}$. Also, $b_j^{(i)} < b_{j+1}^{(i)}$ for $1 \leq j \leq k_i - 1$.

Proof. The proof is exactly the same as the odd case, and so is omitted. □

We also show that the map from $D(\rho_1, \rho_2, \dots, \rho_k; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ to the set of induced representations of the form (4-4) is well defined in the following theorem.

Theorem 4.15. *Suppose that the representation σ is the unique irreducible subrepresentations of both representations*

$$\left(\prod_{i=1}^{k}\prod_{j=1}^{k_i}\delta\big([\nu^{a_{\rho_i}-k_i+j}\rho_i, \nu^{b_j^{(i)}}\rho_i]\big)\right)\rtimes \sigma_{\mathrm{cusp}},$$

$$\left(\prod_{i=1}^{k'}\prod_{j=1}^{k'_i}\delta\big([\nu^{a_{\rho'_i}-k'_i+j}\rho'_i, \nu^{c_j^{(i)}}\rho'_i]\big)\right)\rtimes c'(\sigma_{\mathrm{cusp}})$$

as in Theorem 4.14. Then $k = k'$, $\sigma_{\mathrm{cusp}} \cong c'(\sigma_{\mathrm{cusp}})$ and

$$\left\{\prod_{j=1}^{k_i}\delta\big([\nu^{a_{\rho_i}-k_i+j}\rho_i, \nu^{b_j^{(i)}}\rho_i]\big)\ \Big|\ i = 1, \dots, k\right\}$$

is a permutation of

$$\left\{\prod_{j=1}^{k'_i}\delta\big([\nu^{a_{\rho'_i}-k'_i+j}\rho'_i, \nu^{c_j^{(i)}}\rho'_i]\big)\ \Big|\ i = 1, \dots, k\right\}.$$

Proof. The arguments of the proof follow the same lines as those in the proof of Theorem 4.11. We, therefore, omit the proof here. □

Theorem 4.14 and Theorem 4.15 imply that there exists an injective mapping from $D(\rho_1, \rho_2, \dots, \rho_k; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ into the set of induced representations of the form (4-4). Finally, it remains to show that this mapping is surjective.

Theorem 4.16. *The map described above gives a bijective correspondence between the set $D(\rho_1, \rho_2, \dots, \rho_k; \sigma_{\mathrm{cusp}}, c(\sigma_{\mathrm{cusp}}))$ and the set of induced representations of the form (4-4).*

Proof. Let σ be the unique irreducible subrepresentation of the form (4-4), i.e.,

$$\left(\prod_{i=1}^{k}\prod_{j=1}^{k_i}\delta\big([\nu^{a_{\rho_i}-k_i+j}\rho_i,\nu^{b_j^{(i)}}\rho_i]\big)\right)\rtimes c'(\sigma_{\text{cusp}}).$$

Since $\rho_p \ncong \rho_q$ for $p \neq q$, we have for any $l = 1,\dots,k$ the embedding

$$\sigma \hookrightarrow \left(\prod_{i\neq l}\prod_{j=1}^{k_i}\delta\big([\nu^{a_{\rho_i}-k_i+j}\rho_i,\nu^{b_j^{(i)}}\rho_i]\big)\right)$$
$$\rtimes\delta\big([\nu^{a_{\rho_l}-k_l+l}\rho_l,\nu^{b_l^{(l)}}\rho_l],\dots,[\nu^{a_{\rho_l}}\rho_l,\nu^{b_{k_l}^{(l)}}\rho_l];c'(\sigma_{\text{cusp}})\big).$$

Theorem 4.12 implies that $\delta\big([\nu^{a_{\rho_l}-k_l+l}\rho_l,\nu^{b_l^{(l)}}\rho_l],\dots,[\nu^{a_{\rho_l}}\rho_l,\nu^{b_{k_l}^{(l)}}\rho_l];c'(\sigma_{\text{cusp}})\big)$ is strongly positive. Since l can be arbitrary, σ always has positive unitary exponents in the Jacquet module with respect to the Levi subgroup $\mathrm{GL}_{n_{\rho_1}}\times\mathrm{GL}_{n_{\rho_1}}\times\cdots\times\mathrm{GL}_{n_{\rho_k}}\times G_{n'}$. Therefore, σ is strongly positive. $\square$

Since any strongly positive representation in R can be considered an element of $D(\rho_1',\rho_2',\dots,\rho_k';\sigma'_{\text{cusp}},c'(\sigma'_{\text{cusp}}))$ for some ρ_i' and σ'_{cusp}, we can extend the bijective mapping constructed in Theorem 4.16 to any strongly positive representation in R. In sum, let SP and LJ be as defined in Section 1. Then we have a bijective correspondence between SP and LJ.

Remark 4.17. The ideas used in this section for the results of GSpin groups can be applied to even special orthogonal groups. Let us also remark that it is easier to work with even special orthogonal groups than even GSpin groups due to the results of Ban [1999a; 1999b] and Jantzen [2006]. For example, in the case of even special orthogonal groups, the Weyl group actions on the simple roots and induced representations are studied in [Ban 1999a; 1999b] and Tadić's structure formula is constructed in [Jantzen 2006, Theorem 3.4] (see also [Jantzen and Liu 2014, Theorem 3.1]). Let us remark that the classification of discrete series representations of SO_{2n} is first proved by C. Jantzen [2011] using the results for O_{2n} in [Mœglin 2002; Mœglin and Tadić 2002]. Our approach is different from [Jantzen 2011] and we generalize Matić's idea to the case of even special orthogonal groups.

5. Applications

It is easy to see that the strongly positive representations are special kinds of discrete series due to Casselman's square integrability criterion in [Kim 2009]. Furthermore, the strongly positive representations can be considered basic building blocks for discrete series representations (Theorem 5.1). The proof of the following embedding theorem is exactly the same as the case of odd GSpin groups, since the main idea of the proof depends on a slight variation of Casselman's square integrability criterion for even GSpin groups (Proposition 3.2 in [Kim 2009]). Hence the proof is omitted. Let us remark that this idea originally comes from the results for metaplectic groups by Matić and we adapt some proofs from [Matić 2012, Chapter 3] to our situation.

Theorem 5.1. *Let σ denote a discrete series representation of G_n. Then there exists an embedding of the form*

$$\sigma \hookrightarrow \delta([\nu^{a_1}\rho_1, \nu^{b_1}\rho_1]) \times \delta([\nu^{a_2}\rho_2, \nu^{b_2}\rho_2]) \times \cdots \times \delta([\nu^{a_r}\rho_r, \nu^{b_r}\rho_r]) \rtimes \sigma_{\mathrm{sp}}$$

where $a_i \le 0$, $a_i + b_i > 0$ and $\rho_i \in R_{\mathrm{GL}}$ is an irreducible unitary supercuspidal representation for $i = 1, \dots, r$ (we allow $r = 0$); and $\sigma_{\mathrm{sp}} \in R$ is a strongly positive representation.

Theorem 5.1, together with our main result Theorem 4.16 giving the classification of strongly positive representations, imply the embedding

$$\text{(5-1)}\quad \sigma \hookrightarrow \delta([\nu^{a_1}\rho_1, \nu^{b_1}\rho_1]) \times \cdots \times \delta([\nu^{a_r}\rho_r, \nu^{b_r}\rho_r]) \times \left(\prod_{i=1}^{k}\prod_{j=1}^{k_i} \delta([\nu^{a_{\rho_i'}-k_i+j}\rho_i', \nu^{b_j^{(i)}}\rho_i'])\right) \rtimes \sigma_{\mathrm{cusp}}.$$

where a_i, b_i and ρ_i are as in Theorem 5.1; $a_{\rho_i'}$, $b_j^{(i)}$, k_i and ρ_i' are as in Theorem 4.14; and σ_{cusp} is an irreducible supercuspidal representation of G_n.

This embedding has an interesting application in the proof of the equality of L-functions from the Langlands–Shahidi method and Artin L-functions through local Langlands correspondence (see [Shahidi 2010] for the Langlands–Shahidi method). More precisely, in [Kim 2015a] we used the following filtration of admissible representations to prove the equality of L-functions in the case of GSpin groups:

$$\text{supercuspidal} \subseteq \text{discrete series} \subseteq \text{tempered} \subseteq \text{admissible}.$$

We first showed the equality of L-functions in the supercuspidal case. Then, the above embedding (5-1) was used to generalize that result to the case of discrete series representations. Finally, that result was generalized via Langlands classification and properties of tempered representations to the cases of tempered representations and admissible representations. Furthermore, the equality of L-functions also has an interesting application in the proof of the generic Arthur packet conjecture in the case of GSpin groups [Shahidi 2011]. This conjecture can be considered a local version of the generalized Ramanujan conjecture.

Acknowledgements

This paper is based on part of my thesis. I would like to express my deepest gratitude to my advisor, Professor Shahidi for his constant encouragement and help. I would also like to thank Professor Tadić and Professor Matić for answering many questions and for their suggestions. This paper could not be done without their help. I also thank Professor Mœglin for answering many questions including the question about the half integer conjecture (Lemma 4.1). Thanks are also due to Professors Goldberg, Asgari and Jantzen for very helpful comments. I would also like to thank the referee, who read the paper very carefully, corrected a number of errors and gave very helpful comments.

References

[Asgari 2002] M. Asgari, "Local L-functions for split spinor groups", *Canad. J. Math.* **54**:4 (2002), 673–693. MR 2003i:11062 Zbl 1011.11034

[Ban 1999a] D. Ban, "Parabolic induction and Jacquet modules of representations of $\mathrm{O}(2n, F)$", *Glas. Mat. Ser. III* **34(54)**:2 (1999), 147–185. MR 2001m:22033 Zbl 0954.22013

[Ban 1999b] D. Ban, "Self-duality in the case of $\mathrm{SO}(2n, F)$", *Glas. Mat. Ser. III* **34(54)**:2 (1999), 187–196. MR 2002a:22022 Zbl 0954.22012

[Bernstein and Zelevinsky 1977] J. 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 58 #28310 Zbl 0992.22015

[Bernstein et al. 1986] J. Bernstein, P. Deligne, and D. Kazhdan, "Trace Paley–Wiener theorem for reductive p-adic groups", *J. Analyse Math.* **47** (1986), 180–192. MR 88g:22016 Zbl 0634.22011

[Casselman 1995] W. Casselman, "Introduction to the theory of admissible representations of p-adic reductive groups", Lecture notes, 1995, Available at http://www.math.ubc.ca/∼cass/research/pdf/p-adic-book.pdf.

[Hanzer and Muić 2008] M. Hanzer and G. Muić, "On an algebraic approach to the Zelevinsky classification for classical p-adic groups", *J. Algebra* **320**:8 (2008), 3206–3231. MR 2009k:20107 Zbl 1166.22011

[Hundley and Sayag 2012] J. Hundley and E. Sayag, "Descent construction for GSpin groups", 2012. To appear in *Mem. Amer. Math. Soc.* arXiv 1110.6788

[Jantzen 2006] C. Jantzen, "Jacquet modules of induced representations for p-adic special orthogonal groups", *J. Algebra* **305**:2 (2006), 802–819. MR 2007m:22014 Zbl 1104.22019

[Jantzen 2011] C. Jantzen, "Discrete series for p-adic $\mathrm{SO}(2n)$ and restrictions of representations of $\mathrm{O}(2n)$", *Canad. J. Math.* **63**:2 (2011), 327–380. MR 2012f:22030 Zbl 1219.22016

[Jantzen and Liu 2014] C. Jantzen and B. Liu, "The generic dual of p-adic split SO_{2n} and local Langlands parameters", *Israel J. Math.* **204**:1 (2014), 199–260. MR 3273456 Zbl 1303.22008

[Kim 2009] W. Kim, "Square integrable representations and the standard module conjecture for general spin groups", *Canad. J. Math.* **61**:3 (2009), 617–640. MR 2010e:11048 Zbl 1258.11062

[Kim 2015a] Y. Kim, "Langlands–Shahidi L-functions for GSpin groups and the generic Arthur packet conjecture", preprint, 2015. arXiv 1507.07156

[Kim 2015b] Y. Kim, "Strongly positive representations of GSpin_{2n+1} and the Jacquet module method", *Math. Z.* **279**:1-2 (2015), 271–296. MR 3299853 Zbl 06393879

[Matić 2011] I. Matić, "Strongly positive representations of metaplectic groups", *J. Algebra* **334** (2011), 255–274. MR 2012d:20011 Zbl 1254.22010

[Matić 2012] I. Matić, "Theta lifts of strongly positive discrete series: the case of $(\widetilde{\mathrm{Sp}}(n), O(V))$", *Pacific J. Math.* **259**:2 (2012), 445–471. MR 2988500 Zbl 1277.22017

[Mœglin 2002] C. Mœglin, “Sur la classification des séries discrètes des groupes classiques p-adiques: paramètres de Langlands et exhaustivité”, *J. Eur. Math. Soc.* **4**:2 (2002), 143–200. MR 2003g:22021 Zbl 1002.22009

[Mœglin 2014] C. Mœglin, “Paquets stables des séries discrètes accessibles par endoscopie tordue; leur paramètre de Langlands”, pp. 295–336 in *Automorphic forms and related geometry: assessing the legacy of I. I. Piatetski-Shapiro*, edited by J. W. Cogdell et al., Contemp. Math. **614**, American Mathematical Society, Providence, RI, 2014. MR 3220932 Zbl 1298.22019

[Mœglin and Tadić 2002] C. Mœglin and M. Tadić, “Construction of discrete series for classical p-adic groups”, *J. Amer. Math. Soc.* **15**:3 (2002), 715–786. MR 2003g:22020 Zbl 0992.22015

[Muić 2006] G. Muić, “On the non-unitary unramified dual for classical p-adic groups”, *Trans. Amer. Math. Soc.* **358**:10 (2006), 4653–4687. MR 2007j:22029 Zbl 1102.22014

[Shahidi 1990] F. Shahidi, “A proof of Langlands’ conjecture on Plancherel measures; complementary series for p-adic groups”, *Ann. of Math.* (2) **132**:2 (1990), 273–330. MR 91m:11095 Zbl 0780.22005

[Shahidi 2010] F. Shahidi, *Eisenstein series and automorphic L-functions*, American Mathematical Society Colloquium Publications **58**, American Mathematical Society, Providence, RI, 2010. MR 2012d:11119 Zbl 1215.11054

[Shahidi 2011] F. Shahidi, “Arthur packets and the Ramanujan conjecture”, *Kyoto J. Math.* **51**:1 (2011), 1–23. MR 2784745 Zbl 1238.11060

[Silberger 1980] A. J. Silberger, “Special representations of reductive p-adic groups are not integrable”, *Ann. of Math.* (2) **111**:3 (1980), 571–587. MR 82k:22015 Zbl 0437.22015

[Tadić 1995] M. Tadić, “Structure arising from induction and Jacquet modules of representations of classical p-adic groups”, *J. Algebra* **177**:1 (1995), 1–33. MR 97b:22023 Zbl 0874.22014

[Tadić 1998] M. Tadić, “On reducibility of parabolic induction”, *Israel J. Math.* **107** (1998), 29–91. MR 2001d:22012 Zbl 1166.22011

[Zelevinsky 1980] A. V. Zelevinsky, “Induced representations of reductive p-adic groups. II: On irreducible representations of $\mathrm{GL}(n)$”, *Ann. Sci. École Norm. Sup.* (4) **13**:2 (1980), 165–210. MR 83g:22012 Zbl 0441.22014

Received March 10, 2015. Revised April 30, 2015.

YEANSU KIM
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF IOWA
14 MACLEAN HALL
IOWA CITY, IA 52242
UNITED STATES

yeansu-kim@uiowa.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.89

AN ORLIK–RAYMOND TYPE CLASSIFICATION OF SIMPLY CONNECTED 6-DIMENSIONAL TORUS MANIFOLDS WITH VANISHING ODD-DEGREE COHOMOLOGY

SHINTARÔ KUROKI

Dedicated to Professor Mikiya Masuda on his 60th birthday.

The aim of this paper is to classify simply connected 6-dimensional torus manifolds with vanishing odd-degree cohomology. It is shown that there is a one-to-one correspondence between equivariant diffeomorphism types of these manifolds and 3-valent labelled graphs, called torus graphs, introduced by Maeda, Masuda and Panov. Using this correspondence and combinatorial arguments, we prove that a simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M) = 0$ is equivariantly diffeomorphic to the 6-dimensional sphere S^6 or an equivariant connected sum of copies of 6-dimensional quasitoric manifolds or S^4-bundles over S^2.

1. Introduction

Let M be a $2n$-dimensional closed, connected, oriented manifold with an effective n-dimensional (i.e., half-dimensional) torus T^n-action. We call M, or (M, T), a *torus manifold* if $M^T \neq \varnothing$ (see [Hattori and Masuda 2003]), where M^T is the set of fixed points. A toric manifold (i.e., a nonsingular, complete toric variety viewed as a complex analytic space) with restricted T^n-action is a typical example of a torus manifold. Recall that a toric manifold is a complex $(\mathbb{C}^*)^n$-manifold with a dense orbit (see [Oda 1988; Fulton 1993]), and T^n is the maximal compact subgroup of $(\mathbb{C}^*)^n$. A fundamental result of toric geometry tells us that there is a one-to-one correspondence between toric manifolds and combinatorial objects called fans. Thus, topological (more precisely, geometric) invariants of toric manifolds can be described in terms of combinatorial invariants of fans, such as equivariant cohomology rings, equivariant characteristic classes and other topological invariants.

The author was partially supported by JSPS KAKENHI grant number 15K17531, 24224002 and the JSPS Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation "Deepening and Evolution of Mathematics and Physics, Building of International Network Hub based on OCAMI".

MSC2010: primary 57S25; secondary 94C15.

Keywords: torus manifold, torus graph, GKM graph, equivariant connected sum.

Hattori and Masuda [2003] introduced a torus manifold as the topological generalization of a toric manifold. They also introduced the combinatorial objects called multifans (see [Masuda 1999; Hattori and Masuda 2003]), and computed topological invariants (such as equivariant characteristic classes or Todd genus for unitary torus manifolds) in terms of multifans. However, unlike the case for toric geometry, a multifan does not contain enough information to determine some topological invariants of torus manifolds (e.g., equivariant cohomology). So, in 2007, Maeda, Masuda and Panov introduced combinatorial objects called torus graphs, which were motivated by the GKM graphs introduced by Guillemin and Zara [2001]. The combinatorial information of torus graphs can completely determine the equivariant cohomology rings of torus manifolds with vanishing odd-degree cohomology, i.e., $H^{\mathrm{odd}}(M;\mathbb{Z})=0$ (in this paper, we only consider integer coefficients); see [Masuda and Panov 2006; Maeda et al. 2007], and see also Section 3 in this paper about torus graphs. However, in general, there is no one-to-one correspondence between torus manifolds with $H^{\mathrm{odd}}(M)=0$ and torus graphs.

So, we are naturally led to ask the following two questions: (1) Which subclasses of torus manifolds are completely determined by combinatorial objects (like multifans or torus graphs)? (2) If we find such a subclass of torus manifolds, how can we classify such torus manifolds? Several mathematicians have answered the first question: for example, Davis and Januszkiewicz [1991] for the subclass called *quasitoric manifolds* (see [Buchstaber and Panov 2002] or Section 4C in this paper), Ishida, Fukukawa and Masuda [2013] for the subclass called *topological toric manifolds*, and Wiemeler [2013] for the class of simply connected 6-dimensional torus manifolds with $H^{\mathrm{odd}}(M)=0$ (see Theorem 2.7). The aim of this paper is to answer the second question for the class of simply connected 6-dimensional torus manifolds with $H^{\mathrm{odd}}(M)=0$ using torus graphs.

Let us briefly recall the classification results for torus manifolds with lower dimensions. If T^1 acts on a compact 2-dimensional manifold M, then M is the 2-dimensional sphere S^2, the 2-dimensional real projective space $\mathbb{R}P^2$, the 2-dimensional torus T^2 or the Klein bottle. Because $M^T\neq\varnothing$ and M is oriented, M must be equivariantly diffeomorphic to S^2 with T^1-action (see [Kawakubo 1991]). When $\dim M=4$, by Orlik and Raymond's theorem [1970], we have the following:

Theorem 1.1 (Orlik–Raymond). *Let M be a 4-dimensional simply connected torus manifold. Then, M is equivariantly diffeomorphic to the 4-sphere S^4 or an equivariant connected sum of copies of complex projective spaces $\mathbb{C}P^2$, $\overline{\mathbb{C}P^2}$ (reversed orientation) or a Hirzebruch surface H_k.*

Here a Hirzebruch surface H_k ($k\in\mathbb{Z}$) is a manifold which is defined by the projectivization of the complex 2-dimensional vector bundle $\gamma^{\otimes k}\oplus\epsilon$ over $\mathbb{C}P^1$, where γ and ϵ are the tautological and the trivial complex line bundles over $\mathbb{C}P^1$.

In this paper, we prove an Orlik–Raymond type theorem similar to Theorem 1.1 for simply connected 6-dimensional torus manifolds with $H^{\mathrm{odd}}(M)=0$. Before we state our main results, we introduce the result for torus manifolds that are not simply connected. One of the consequences of Masuda and Panov's theorem (see Theorem 2.2 in Section 2B) is the following proposition (see also [Wiemeler 2013]).

Proposition 1.2. *Let W be a 6-dimensional torus manifold with $H^{\mathrm{odd}}(W)=0$ (it might not be simply connected). Then, there are a simply connected 6-dimensional torus manifold M with $H^{\mathrm{odd}}(M)=0$ and a homology 3-sphere hS^3 such that*

$$W \cong M \#_T (hS^3 \times T^3)$$

up to equivariant diffeomorphism.

Here the product manifold $hS^3 \times T^3$ is the product of hS^3 and the 3-dimensional torus T^3 with the free T^3-action on the second factor, and the symbol $\#_T$ represents the equivariant gluing along two free orbits of M and $hS^3 \times T^3$.

In Proposition 1.2, because the fundamental groups $\pi_1(W)$ and $\pi_1(hS^3)$ are isomorphic, W is simply connected if and only if hS^3 is simply connected, i.e., the standard sphere. Our main theorem is a classification of the simply connected torus manifolds that appear in Proposition 1.2.

Theorem 1.3. *Let M be a simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M)=0$. Then, M is equivariantly diffeomorphic to the 6-sphere S^6 or obtained by an equivariant connected sum of copies of 6-dimensional quasitoric manifolds or S^4-bundles over S^2 equipped with the structure of a torus manifold.*

This type of classification, i.e., classification by equivariant connected sum, may be regarded as the 6-dimensional analogue of Orlik and Raymond's classification in Theorem 1.1. So, in this paper, we call this theorem an *Orlik–Raymond type classification* (see [McGavran 1976; Kuroki 2008]).

Remark 1.4. Izmestiev [2001] proved an Orlik–Raymond type classification for a class of 3-dimensional small covers (i.e., the real analogue of quasitoric manifolds; see Section 4B), called a linear model (see also [Lü and Yu 2011; Nishimura 2012]).

The organization of this paper is as follows. In Section 2, we recall the basic facts about torus manifolds. In Section 3, we do the same for torus graphs. In particular, Corollary 3.5 is the key fact used to prove Theorem 1.3. In Section 4, we introduce the torus graphs of S^6, quasitoric manifolds and S^4-bundles over S^2. These torus graphs will be the basic graphs used to classify simply connected 6-dimensional torus manifolds with $H^{\mathrm{odd}}(M)=0$. In Section 5, we introduce the "oriented" torus graphs and translate the equivariant connected sum around fixed points of torus manifolds to the connected sum around vertices of oriented torus graphs. In Sections 6 and 7, we prove Theorem 1.3. A brief outline of the proof is as follows. By

Corollary 3.5, there is a one-to-one correspondence between 6-dimensional simply connected torus manifolds with $H^{\mathrm{odd}}(M) = 0$ and 3-valent torus graphs. Therefore, to prove Theorem 1.3, it is enough to prove that an (oriented) torus graph can be decomposed into basic torus graphs in Section 4 by the connected sum. We prove this using combinatorial arguments.

2. Orbit spaces of torus manifolds

In this section, we recall some basic facts about torus manifolds (see [Masuda 1999] or [Hattori and Masuda 2003] for details).

2A. ***Torus manifolds.*** A $2n$-dimensional torus manifold M is said to be *locally standard* if every point in M has a T-invariant open neighborhood U which is weakly equivariantly homeomorphic to an open subset $\Omega_U \subset \mathbb{C}^n$ invariant under the standard T^n-action on $\mathbb{C}^n$, where two group actions (U, T) and (Ω_U, T) are said to be *weakly equivariantly homeomorphic* if there is an equivariant homeomorphism from U to Ω_U up to an automorphism on T^n (see, e.g., [Kuroki 2011, Section 2.1] for details).

Let M_i, $i = 1, \ldots, m$, be a codimension-2 torus submanifold in a $2n$-dimensional torus manifold M which is fixed by some circle subgroup T_i in T. Such an M_i is a $(2n-2)$-dimensional torus manifold with T/T_i-action, called a *characteristic submanifold*. Because a torus manifold M is compact, the cardinality of all characteristic submanifolds in M is finite. If M is locally standard, each characteristic submanifold is also locally standard.

An *omniorientation* $\mathbb{O}$ of M is a choice of orientation for the torus manifold M as well as for each characteristic submanifold. If there are just m characteristic submanifolds in M, there are exactly 2^{m+1} omniorientations (see [Buchstaber and Panov 2002; Hattori and Masuda 2003]). If M has a T-invariant almost complex structure J (in this case, M is automatically locally standard), then there exists the canonical omniorientation $\mathbb{O}_J$ determined by J. We call the torus manifold M with a fixed omniorientation $\mathbb{O}$ an *omnioriented torus manifold* and denote it by $(M, \mathbb{O})$.

2B. ***Orbit spaces of locally standard torus manifolds.*** The orbit space M/T of a locally standard torus manifold M naturally admits the structure of a “topological” manifold with corners. We next recall the basic facts about a topological manifold with corners (cf. the definition of a smooth manifold with corners in [Lee 2013]) and introduce the structure on M/T.

We will use the notation

$$[n] = \{0, 1, \ldots, n\}$$

and

$$\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, \ldots, n\}.$$

Let Q^n be an n-dimensional topological manifold with boundary. A *chart with corners* for Q^n is a pair (V, ψ_V), where V is an open subset of Q^n and

$$\psi_V : V \to \mathbb{R}^n_+$$

is homeomorphic from V to a (relatively) open subset $\Omega_V \subset \mathbb{R}^n_+$. Two charts with corners (V, ψ_V) and (W, ψ_W) are said to be (*topologically*) *compatible* if the composition of functions $\psi_V \circ \psi_W^{-1} : \psi_W(V \cap W) \to \psi_V(V \cap W)$ is a strata-preserving homeomorphism. This implies that if $\psi_W(p) \in \mathbb{R}^n_+$ contains exactly k zero-coordinates then $\psi_V(p) \in \mathbb{R}^n_+$ also contains exactly k zero-coordinates for $0 \le k \le n$. We call the collection of compatible charts with corners $\{(V, \psi_V)\}$ whose domains cover Q^n an *atlas*. Then, its maximal atlas is called a *structure with corners* of Q^n. A topological manifold with boundary together with a structure with corners is called a (*topological*) *manifold with corners*. Let $p \in Q^n$ be a point of an n-dimensional manifold with corners Q^n. For a chart (V, ψ_V) with corners such that $p \in V$, we define $d(p) \in [n]$ to be the number of zero-coordinates of $\psi_V(p) \in \mathbb{R}^n_+$. By the compatibility of charts, this number is independent of the choice of a chart with corners which contains p. Therefore, the map $d : Q^n \to [n]$ is well defined. The number $d(p)$ is called the *depth* of p. We call the closure of a connected component of $d^{-1}(k)$, $0 \le k \le n$, a *codimension-k face*. In particular, the codimension-0 face is Q^n itself. Moreover, codimension-1, codimension-$(n-1)$ and codimension-n faces are called *facets*, *edges* and *vertices*, respectively. The set of all edges and vertices is called a *one-skeleton of* Q^n (or a *graph of* Q^n). By restricting the structure with corners on Q^n to faces, we may regard each codimension-k face as an $(n-k)$-dimensional (sub)manifold with corners.

Definition 2.1 (manifold with faces). An n-dimensional manifold with corners Q is said to be a *manifold with faces* (or a *nice manifold with corners*) if Q satisfies the following conditions:

(1) For every $k \in [n]$, there exists a codimension-k face.

(2) For each codimension-k face H, there are exactly k facets $F_1, \dots, F_k$ such that H is a connected component of $\bigcap_{i=1}^k F_i$; moreover, $H \cap F \neq H$ for any facet $F \neq F_i$ $(i = 1, \dots, k)$.

Let (M, T) be a torus manifold. When (M, T) is locally standard, by the differentiable slice theorem, the orbit space M/T has the structure of an n-dimensional manifold with faces. On the other hand, when M satisfies $H^{\mathrm{odd}}(M) = 0$, its orbit space M/T satisfies a stronger condition by the following theorem (see [Masuda and Panov 2006, Lemma 2.1 and Theorem 2]).

Theorem 2.2 (Masuda–Panov). *Let M be a $2n$-dimensional torus manifold. Then, the following conditions are equivalent*:

(1) $H^{\mathrm{odd}}(M) = 0$.

(2) *The T-action on M is locally standard and its orbit space M/T has the structure of an n-dimensional face acyclic manifold with corners.*

Here, an n-dimensional *face acyclic* manifold with corners Q is an n-dimensional manifold with faces such that all faces F of Q (including Q) are acyclic, i.e., $H_*(F) \simeq H_0(F) \simeq \mathbb{Z}$. For example, if Q is a simply connected 3-dimensional face acyclic manifold with corners, then it is easy to check that the boundary of Q is homeomorphic to the 2-sphere S^2. Moreover, in this case, we can also check that Q itself is homeomorphic to the 3-dimensional disk D^3. Therefore, as one of the consequences of Theorem 2.2, we have the following corollary.

Corollary 2.3. *Let M be a simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M) = 0$. Then, its orbit space M/T is homeomorphic to the 3-dimensional disk.*

By the definition of a manifold with faces Q, we can define a simplicial poset (partially ordered set) $\mathcal{P}(Q)$, called a *face poset* of Q (see [Masuda 2005]), to be the set of faces in Q with the empty set $\varnothing$ ordered by inclusion, where $\varnothing$ is the smallest element under this ordering, say $\preceq$. We often denote the face poset structure of Q by $(\mathcal{P}(Q), \preceq)$. Let Q_1 and Q_2 be n-dimensional manifolds with faces. We say Q_1 and Q_2 are *combinatorially equivalent* if their face posets $(\mathcal{P}(Q_1), \preceq_1)$ and $(\mathcal{P}(Q_2), \preceq_2)$ are isomorphic as posets (i.e., there is an order-preserving bijection between them). We denote the equivalence by $Q_1 \approx_c Q_2$. By the definition of weakly equivariant homeomorphism, if two locally standard torus manifolds M_1 and M_2 are weakly equivariantly homeomorphic then $M_1/T \approx_c M_2/T$.

2C. ***Characteristic functions.*** Let M be a $2n$-dimensional locally standard torus manifold. By the argument demonstrated in Section 2B, the orbit map $\pi : M \to M/T = Q$ may be regarded as the projection onto some manifold with faces Q. Let $\mathcal{F}(Q) = \{F_1, \dots, F_m\} \subset \mathcal{P}(Q)$ be the set of all facets in Q. By the definition of facet $F_i \in \mathcal{F}(Q)$, its preimage $\pi^{-1}(F_i)$ is a characteristic submanifold M_i. Then, there exists the circle subgroup T_i $(\subset T)$ fixing $M_i = \pi^{-1}(F_i)$ (recall that $\dim M_i = 2n-2$). Recall that T_i is determined by a primitive element in $\mathfrak{t}_{\mathbb{Z}} \simeq \mathbb{Z}^n$ (the lattice of the Lie algebra of T). Therefore, using this primitive element (up to sign) in $\mathfrak{t}_{\mathbb{Z}}$, we can define the map

$$\lambda : \mathcal{F}(Q) \to \mathfrak{t}_{\mathbb{Z}}/\{\pm 1\},$$

where $\mathfrak{t}_{\mathbb{Z}}/\{\pm 1\}$ represents the quotient of $\mathfrak{t}_{\mathbb{Z}}$ by signs. We call λ a *characteristic function.*

Now the choice of omniorientation $\mathbb{O}$ of M determines the sign of λ as follows. Fix an omniorientation $\mathbb{O}$ of M. Namely, we fix the orientation of the tangent bundle of M (resp. M_i), say τ (resp. τ_i). Restricting τ to the submanifold M_i, say $\tau|_{M_i}$, we obtain the T^n-equivariant decomposition $\tau|_{M_i} \simeq \tau_i \oplus \nu_i$, where ν_i is the T_i-equivariant normal bundle of M_i. Therefore, because we fix the orientation of $\tau|_{M_i}$ (induced from the orientation of τ) and that of τ_i, we may choose an orientation of ν_i such that the orientation of $\tau|_{M_i}$ coincides with that of $\tau_i \oplus \nu_i$ (thus, we may regard ν_i as the complex line bundle over M_i). Because T_i acts on ν_i, we may choose an orientation of T_i such that the T_i-action preserves the orientation of ν_i. This orientation of T_i determines the sign of $\lambda(F_i)$ for $i = 1, \ldots, m$. In this way, we have the function

$$\lambda_{\mathbb{O}} : \mathcal{F}(Q) \to \mathfrak{t}_{\mathbb{Z}}.$$

In this paper, this is called an *omnioriented characteristic function* (of $(M, \mathbb{O})$).

Remark 2.4. The characteristic function defined in [Wiemeler 2013] may be regarded as the characteristic function λ above. On the other hand, the characteristic function defined in [Davis and Januszkiewicz 1991] may be regarded as the characteristic function $\lambda_{\mathbb{O}}$ above by taking an appropriate omniorientation (see also [Buchstaber and Panov 2002, Section 5.2]).

Let $p \in M^T$. We define the subset $I_p \subset [m]$ by

$$I_p = \{i \in [m] \mid p \in M_i\}.$$

By the differentiable slice theorem around $p \in M^T$, we have that its cardinality $|I_p|$ equals n for every $p \in M^T$. Put $I_p = \{i_1, \ldots, i_n\}$. Because the T-action on M is effective, $\{\lambda(F_{i_1}), \ldots, \lambda(F_{i_n})\}$ spans $\mathfrak{t}^*_{\mathbb{Z}}/\{\pm 1\}$, i.e., the determinant of the induced $(n \times n)$-matrix

$$(\lambda(F_{i_1}) \cdots \lambda(F_{i_n}))$$

satisfies

$$\det(\lambda(F_{i_1}) \cdots \lambda(F_{i_n})) = \pm 1. \tag{2-1}$$

Similarly, we have

$$\det(\lambda_{\mathbb{O}}(F_{i_1}) \cdots \lambda_{\mathbb{O}}(F_{i_n})) = \pm 1 \tag{2-2}$$

for each set of n facets such that $\bigcap_{j=1}^n F_{i_j} = \{p\}$ for some vertex $p \in Q$ (called the *facets around a vertex*).

Motivated by the above observations, we may abstractly define the characteristic function on a manifold with faces as follows (see [Buchstaber and Panov 2002; Davis and Januszkiewicz 1991] for simple polytopes and [Masuda and Panov 2006; Wiemeler 2013] for manifolds with faces).

Definition 2.5. Let Q be an n-dimensional manifold with faces and $\mathcal{F}(Q)$ be the set of its facets. Let $\mathfrak{t}_{\mathbb{Z}}$ be the lattice of the Lie algebra of T^n and $\mathfrak{t}_{\mathbb{Z}}/\{\pm 1\}$ be its quotient by $\{\pm 1\}$. A function $\lambda : \mathcal{F}(Q) \to \mathfrak{t}_{\mathbb{Z}}/\{\pm 1\}$ is said to be a *characteristic function* if λ satisfies (2-1) for the facets around every vertex, and a function $\lambda_{\mathbb{O}} : \mathcal{F}(Q) \to \mathfrak{t}_{\mathbb{Z}}$ is said to be an *omnioriented characteristic function* if $\lambda_{\mathbb{O}}$ satisfies (2-2) for the facets around every vertex.

We denote an n-dimensional manifold with faces Q with its characteristic function λ (resp. omnioriented characteristic function $\lambda_{\mathbb{O}}$) by (Q, λ) (resp. $(Q, \lambda_{\mathbb{O}})$).

Let Q_1 and Q_2 be manifolds with faces, and let λ_1 and λ_2 be their characteristic functions and $\lambda_{\mathbb{O}_1}$ and $\lambda_{\mathbb{O}_2}$ be their omnioriented characteristic functions, respectively. Assume that $Q_1 \approx_c Q_2$, induced by the bijective map $\tilde{f} : \mathcal{P}(Q_1) \to \mathcal{P}(Q_2)$. Denote its restriction onto the set of facets by

$$f = \tilde{f}|_{\mathcal{F}(Q_1)} : \mathcal{F}(Q_1) \to \mathcal{F}(Q_2).$$

We say that (Q_1, λ_1) and (Q_2, λ_2) are *combinatorially equivalent* if the following diagram commutes:

$$\begin{tikzcd}
\mathcal{F}(Q_1) \arrow[r, "\lambda_1"] \arrow[d, "f"] & \mathfrak{t}_{\mathbb{Z}}/\{\pm 1\} \arrow[d, "\mathrm{Id}"] \\
\mathcal{F}(Q_2) \arrow[r, "\lambda_2"] & \mathfrak{t}_{\mathbb{Z}}/\{\pm 1\}
\end{tikzcd}$$

Similarly, $(Q_1, \lambda_{\mathbb{O}_1})$ and $(Q_2, \lambda_{\mathbb{O}_2})$ are combinatorially equivalent if the following diagram commutes:

$$\begin{tikzcd}
\mathcal{F}(Q_1) \arrow[r, "\lambda_{\mathbb{O}_1}"] \arrow[d, "f"] & \mathfrak{t}_{\mathbb{Z}} \arrow[d, "\mathrm{Id}"] \\
\mathcal{F}(Q_2) \arrow[r, "\lambda_{\mathbb{O}_2}"] & \mathfrak{t}_{\mathbb{Z}}
\end{tikzcd}$$

Note that the characteristic function λ can be obtained by ignoring signs from the omnioriented characteristic function $\lambda_{\mathbb{O}}$; we call such a λ an *induced characteristic function* from $\lambda_{\mathbb{O}}$. On the other hand, by choosing a sign for each facet, we can obtain an omnioriented characteristic function $\lambda_{\mathbb{O}}$ from the characteristic function λ; we call such a $\lambda_{\mathbb{O}}$ an *induced oriented characteristic function* from λ.

Lemma 2.6. *If $(Q_1, \lambda_{\mathbb{O}_1})$ and $(Q_2, \lambda_{\mathbb{O}_2})$ are combinatorially equivalent, then their induced (Q_1, λ_1) and (Q_2, λ_2) are also combinatorially equivalent.*

If (Q_1, λ_1) and (Q_2, λ_2) are combinatorially equivalent, then there are induced omnioriented characteristic functions $\lambda_{\mathbb{O}_1}$ and $\lambda_{\mathbb{O}_2}$ such that $(Q_1, \lambda_{\mathbb{O}_1})$ and $(Q_2, \lambda_{\mathbb{O}_2})$ are combinatorially equivalent.

We now introduce one of the key facts used to prove our main theorem (see [Wiemeler 2013, Theorems 1.3 and 6.1]).

Theorem 2.7 (Wiemeler). *Let M_1 and M_2 be 6-dimensional simply connected torus manifolds with $H^{\mathrm{odd}}(M_1) = H^{\mathrm{odd}}(M_2) = 0$, and let (Q_1, λ_1) and (Q_2, λ_2) be* *their orbit spaces with characteristic functions. Then, the following statements are equivalent*:

(1) (Q_1, λ_1) *and* (Q_2, λ_2) *are combinatorially equivalent.*

(2) M_1 *and* M_2 *are equivariantly homeomorphic.*

(3) M_1 *and* M_2 *are equivariantly diffeomorphic.*

Therefore, by Corollary 2.3 and Theorem 2.7, to classify all 6-dimensional simply connected torus manifolds with $H^{\mathrm{odd}}(M) = 0$, it is enough to classify all (Q, λ)'s up to combinatorial equivalence, where Q is a 3-dimensional disk equipped with the structure of a manifold with faces.

3. Torus graph induced from manifold with faces

Let $(M, \mathbb{O})$ be an omnioriented locally standard $2n$-dimensional torus manifold and $(Q, \lambda_{\mathbb{O}})$ be its orbit space with an omnioriented characteristic function. From the one-skeleton of $(Q, \lambda_{\mathbb{O}})$, we can define a labelled graph called a *torus graph*. One of the key steps in proving the main theorem is to classify all possible torus graphs (see Section 7). We first recall the definition of torus graph given by Maeda et al. [2007].

Let Γ be the graph of Q. Let $V(\Gamma)$ be its vertices and $E(\Gamma)$ be its oriented edges, i.e., we distinguish two edges pq and qp. For $p \in V(\Gamma)$, we denote the set of outgoing edges from p by $E_p(\Gamma)$. Because Q is an n-dimensional manifold with faces, $|E_p(\Gamma)| = n$ and each edge $e \in E(\Gamma)$ is a connected component of $\bigcap_{i=1}^{n-1} F_i$ for some $F_1, F_2, \dots, F_{n-1} \in \mathcal{F}(Q)$. Moreover, for a $p \in V(\Gamma)$ which is one of two vertices on e, there is another facet $F_n \in \mathcal{F}(Q)$ such that $\{p\}$ is a connected component of $\bigcap_{i=1}^{n} F_i$. In other words, F_n may be regarded as a normal facet of $e \in E(\Gamma)$ on $p \in V(\Gamma)$. Put $\lambda_{\mathbb{O}}(F_i) = a_i \in \mathfrak{t}_{\mathbb{Z}} \simeq \mathbb{Z}^n$. Then, there exists a unique $\alpha \in \mathfrak{t}^*_{\mathbb{Z}}$ such that

$$\langle \alpha, a_i \rangle = 0 \text{ for } i = 1, \dots, n-1 \quad \text{and} \quad \langle \alpha, a_n \rangle = +1, \tag{3-1}$$

where $\langle \cdot , \cdot \rangle$ represents the pairing of $\mathfrak{t}^*$ and $\mathfrak{t}$. Therefore, in this way, we can define a map $\mathcal{A} : E(\Gamma) \to \mathfrak{t}^*_{\mathbb{Z}}$ from the omnioriented characteristic function $\lambda_{\mathbb{O}}$. This map $\mathcal{A}$ is called an *axial function* on Γ. We call the labelled graph $(\Gamma, \mathcal{A})$ a *torus graph* induced from $(Q, \lambda_{\mathbb{O}})$ (or equivalently $(M, \mathbb{O})$). We denote such a torus graph by $\Gamma(Q, \lambda_{\mathbb{O}})$ (or $(\Gamma_M, \mathcal{A}_M)$). We can easily check the following proposition using the definition of torus graph (see also [Maeda et al. 2007]).

Proposition 3.1. *Let* $(\Gamma, \mathcal{A})$ *be a torus graph induced from* $(Q, \lambda_{\mathbb{O}})$. *Then,* Γ *is an* n*-valent regular graph, i.e.,* $|E_p(\Gamma)| = n$ *for all* $p \in V(\Gamma)$, *and* $(\Gamma, \mathcal{A})$ *satisfies the following conditions*:

(1) $\mathcal{A}(e) = \pm\mathcal{A}(\bar{e})$, *where* $\bar{e}$ *is the orientation-reversed edge of* e.

(2) $\{\mathscr{A}(e) \mid e \in E_p(\Gamma)\}$ *spans* $\mathfrak{t}^*_{\mathbb{Z}}$ *for all vertices* $p \in V(\Gamma)$.

(3) *There is a bijection* $\nabla_{pq} : E_p(\Gamma) \to E_q(\Gamma)$ *for all edges whose initial vertex is* p *and terminal vertex is* q *such that*

   (a) $\nabla_{\bar{e}} = \nabla_e^{-1}$,
   (b) $\nabla_e(e) = \bar{e}$,
   (c) $\mathscr{A}(\nabla_{pq}(e)) - \mathscr{A}(e) \equiv 0 \bmod \mathscr{A}(pq)$ *for all* $e \in E_p(\Gamma)$.

We call $\nabla = \{\nabla_e \mid e \in E(\Gamma)\}$ a *connection* on $(\Gamma, \mathscr{A})$.

Remark 3.2. The original definition of torus graph (induced from an omnioriented torus manifold) uses *tangential representations*; see [Masuda and Panov 2006; Maeda et al. 2007]. The definition of torus graph given above is essentially the same as the original definition.

In [Maeda et al. 2007], motivated by the GKM graph introduced by Guillemin and Zara [2001], an n-valent graph Γ with a label $\mathscr{A} : E(\Gamma) \to \mathfrak{t}^*_{\mathbb{Z}}$ which satisfies the three conditions in Proposition 3.1 is called an (*abstract*) *torus graph* (i.e., there might be no geometric objects which define $(\Gamma, \mathscr{A})$).

We next define the equivalence relation between two torus graphs. We call the map $f : \Gamma_1 = (V(\Gamma_1), E(\Gamma_1)) \to \Gamma_2 = (V(\Gamma_2), E(\Gamma_2))$ a *graph isomorphism* if the restricted maps $f|_V : V(\Gamma_1) \to V(\Gamma_2)$ and $f|_E : E(\Gamma_1) \to E(\Gamma_2)$ are bijective and the following diagram commutes:

$$\begin{array}{ccc} E(\Gamma_1) & \xrightarrow{f|_E} & E(\Gamma_2) \\ \downarrow{\scriptstyle \pi_{V_1}} & & \downarrow{\scriptstyle \pi_{V_2}} \\ V(\Gamma_1) & \xrightarrow{f|_V} & V(\Gamma_2) \end{array}$$

Here the map $\pi_V : E(\Gamma) \to V(\Gamma)$ is the projection onto the initial vertex, i.e., $\pi_V(pq) = p$. In other words, the bijection $f|_V$ preserves the edges. Now we may define the equivalence relation.

Definition 3.3. Let $(\Gamma_1, \mathscr{A}_1)$ and $(\Gamma_2, \mathscr{A}_2)$ be torus graphs. We say $(\Gamma_1, \mathscr{A}_1)$ and $(\Gamma_2, \mathscr{A}_2)$ are *equivalent* if there is a graph isomorphism $f : \Gamma_1 \to \Gamma_2$ such that the following diagram commutes:

$$\begin{array}{ccc} E(\Gamma_1) & \xrightarrow{\mathscr{A}_1} & \mathfrak{t}^*_{\mathbb{Z}} \\ \downarrow{\scriptstyle f|_E} & & \downarrow{\scriptstyle \mathrm{Id}} \\ E(\Gamma_2) & \xrightarrow{\mathscr{A}_2} & \mathfrak{t}^*_{\mathbb{Z}} \end{array}$$

Assume $(\Gamma, \mathscr{A}) = \Gamma(Q, \lambda_{\mathbb{O}})$. Let $\mathscr{P}_k(\Gamma, \mathscr{A})$ be the set of k-valent torus subgraphs in $(\Gamma, \mathscr{A})$, i.e., k-valent subgraphs in Γ closed under the connection ∇, where $-1 \leq k \leq n$ and we define $\mathscr{P}_{-1}(\Gamma, \mathscr{A}) = \{\varnothing\}$. Then, the set

$$\mathscr{P}(\Gamma, \mathscr{A}) = \bigcup_{k=-1}^{n} \mathscr{P}_k(\Gamma, \mathscr{A})$$

admits the structure of a simplicial poset by inclusion (see [Maeda et al. 2007]). We denote this structure by $(\mathscr{P}(\Gamma, \mathscr{A}), \preceq)$. Let $\mathscr{P}(Q)$ be the face poset of Q (see Section 2B) and $\mathscr{P}_k(Q)$ be the set of all k-dimensional faces, where $-1 \leq k \leq n$ and $\mathscr{P}_{-1}(Q) = \{\varnothing\}$. Then, each element of $\mathscr{P}_k(\Gamma, \mathscr{A})$ is nothing but the graph of an element in $\mathscr{P}_k(Q)$. This implies that the poset $(\mathscr{P}(\Gamma, \mathscr{A}), \preceq)$ is equivalent to the poset $(\mathscr{P}(Q), \preceq)$. Therefore, we have the following lemma.

Lemma 3.4. *The following two statements are equivalent*:

(1) *Two manifolds with faces with omnioriented characteristic functions* $(Q_1, \lambda_{\mathbb{O}_1})$ *and* $(Q_2, \lambda_{\mathbb{O}_2})$ *are combinatorially equivalent.*

(2) *Their induced torus graphs* $\Gamma(Q_1, \lambda_{\mathbb{O}_1})$ *and* $\Gamma(Q_2, \lambda_{\mathbb{O}_2})$ *are equivalent.*

By Lemma 2.6, Theorem 2.7 and Lemma 3.4, we have the following corollary.

Corollary 3.5. *Let* (M_1, T) *and* (M_2, T) *be* 6*-dimensional simply connected torus manifolds with vanishing odd-degree cohomology. Then, the following statements are equivalent*:

(1) (M_1, T) *and* (M_2, T) *are equivariantly diffeomorphic.*

(2) *Their orbit spaces, i.e.,* 3*-dimensional disks with the structures of manifolds with faces, with characteristic functions* $(M_1/T, \lambda_1)$ *and* $(M_2/T, \lambda_2)$ *are combinatorially equivalent.*

(3) *There are omnioriented characteristic functions* $\lambda_{\mathbb{O}_1}$ *and* $\lambda_{\mathbb{O}_2}$ *such that their induced* 3*-valent torus graphs* $\Gamma(M_1/T, \lambda_{\mathbb{O}_1})$ *and* $\Gamma(M_2/T, \lambda_{\mathbb{O}_2})$ *are equivalent.*

Therefore, to prove our main theorem (Theorem 7.1), it is enough to classify all 3-valent torus graphs $(\Gamma, \mathscr{A})$, induced from $(M, \mathbb{O})$, up to equivalence.

4. Basic 6-dimensional torus manifolds

Let (M, T) be a simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M) = 0$, and let $(\Gamma_M, \mathscr{A}_M)$ $(= (\Gamma, \mathscr{A}))$ be its torus graph induced by some omniorientation. As a preliminary to proving the main theorem (Theorem 7.1), in this section we introduce some of the basic torus graphs $(\Gamma, \mathscr{A})$ and their corresponding 6-dimensional torus manifolds (M, T).

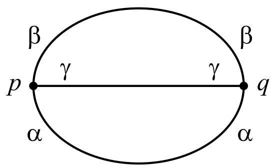


Figure 1. The torus graph $(\Gamma_{\mathrm{sp}}, \mathscr{A}_{\alpha,\beta,\gamma})$, where $\alpha, \beta, \gamma \in \mathfrak{t}^*_{\mathbb{Z}} \simeq \mathbb{Z}^3$ are a $\mathbb{Z}$-basis.

4A. 6-*sphere*. Because the induced torus graphs from (M, T) are 3-valent, if there is a 3-multiple edge, i.e., three edges that are incident to the same two vertices, then it follows from Proposition 3.1 that such a torus graph must be the torus graph in Figure 1, denoted $(\Gamma_{\mathrm{sp}}, \mathscr{A}_{\alpha,\beta,\gamma})$.

Put $\alpha = k_{11}e_1 + k_{12}e_2 + k_{13}e_3$, $\beta = k_{21}e_1 + k_{22}e_2 + k_{23}e_3$ and $\gamma = k_{31}e_1 + k_{32}e_2 + k_{33}e_3$, using the standard basis e_1, e_2, e_3 in $\mathfrak{t}^*_{\mathbb{Z}} \simeq \mathbb{Z}^3$. Then, we have

$$\det\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} = \pm 1. \tag{4-1}$$

Let $S^6 \subset \mathbb{C}^3 \oplus \mathbb{R}$ be the unit sphere, i.e., the set $(z_1, z_2, z_3, r) \in \mathbb{C}^3 \oplus \mathbb{R}$ such that $|z_1|^2 + |z_2|^2 + |z_3|^2 + r^2 = 1$. Define the T^3-action on the first three complex coordinates in S^6 by

$$(t_1, t_2, t_3)(z_1, z_2, z_3, r) \mapsto (\rho_1(t)z_1, \rho_2(t)z_2, \rho_3(t)z_3, r), \tag{4-2}$$

where $t = (t_1, t_2, t_3) \in T$ and $\rho_i : T \to S^1$, $i = 1, 2, 3$, is a 1-dimensional complex representation defined by

$$\rho_i(t_1, t_2, t_3) = t_1^{k_{i1}} t_2^{k_{i2}} t_3^{k_{i3}}.$$

Then, by choosing an appropriate omniorientation on S^6, we have that its induced torus graph is equivalent to $(\Gamma_{\mathrm{sp}}, \mathscr{A}_{\alpha,\beta,\gamma})$. Therefore, using Corollary 3.5, we have the following lemma.

Lemma 4.1. *Let $(M, \mathbb{O})$ be an omnioriented 6-dimensional simply connected torus manifold with $H^{\mathrm{odd}}(M) = 0$. If its induced torus graph is $(\Gamma_{\mathrm{sp}}, \mathscr{A}_{\alpha,\beta,\gamma})$, then (M, T) is equivariantly diffeomorphic to one of (S^6, T) defined by (4-2).*

4B. *S^4-bundles over S^2*. Assume that a 3-valent torus graph $(\Gamma, \mathscr{A})$ does not have 3-multiple edges but does have multiple edges, i.e., two edges that are incident to the same two vertices. In this section, we classify the easiest case of such torus graphs.

Because Γ is a one-skeleton of a 3-dimensional manifold with faces Q, we have $|V(\Gamma)| \geq 4$. Assume that $|V(\Gamma)| = 4$. Then, we can easily check that such a torus

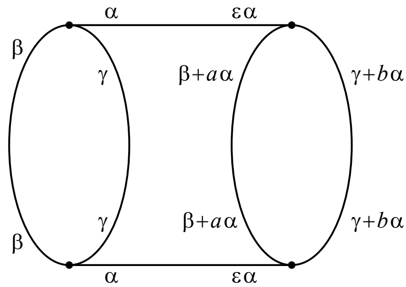


Figure 2. The torus graph $(\Gamma_S, \mathcal{A}_S) = (\Gamma_S, \mathcal{A}^{\epsilon,a,b}_{\alpha,\beta,\gamma})$, where $\epsilon = \pm 1$, $a, b \in \mathbb{Z}$, and $\alpha, \beta, \gamma \in \mathfrak{t}^*_{\mathbb{Z}}$ are a $\mathbb{Z}$-basis of $\mathfrak{t}^*_{\mathbb{Z}}$.

manifold is the one-skeleton of the 3-simplex (see Figure 3 in Section 4C) or the graph drawn in Figure 2, say Γ_S. It is well known that the torus manifold whose torus graph is the one-skeleton of the 3-simplex is equivariantly diffeomorphic to the complex projective space with some T-action (see, e.g., [Davis and Januszkiewicz 1991], and see also Figure 3 in Section 4C). So, we only study the torus manifold which induces the graph Γ_S. Because Q is homeomorphic to D^3, we may regard a Q whose one-skeleton is Γ_S as the product $D^2 \times I$, where D^2 is the 2-dimensional disk and I is the interval. By considering all functions on facets of Q which satisfy (2-2), we can classify all omnioriented characteristic functions $\lambda_{\mathcal{O}}$ on Q. Then, in the same way we induced the axial function $\mathcal{A}_S$ from $(Q, \lambda_{\mathcal{O}})$ in Section 3, we can obtain all possible axial functions on Γ_S, as shown in Figure 2.

The torus graph $(\Gamma_S, \mathcal{A}_S)$ in Figure 2 can be induced from an S^4-bundle over S^2 as follows. First, by choosing $\epsilon = \pm 1$, we may define two free T^1-actions on $S^3 \subset \mathbb{C}^2$:

$$(w, z) \mapsto (t^{-1}w, t^{\epsilon}z).$$

We denote S^3 with the above T^1-action by S^3_ϵ. Note that S^3_ϵ/T^1 is diffeomorphic to the 2-sphere S^2, and a complex line bundle over S^2 can be denoted by

$$S^3_\epsilon \times_{T^1} \mathbb{C}_k,$$

where $\mathbb{C}_k$ is the complex 1-dimensional T^1-representation space by k-times rotation for some $k \in \mathbb{Z}$. Let $S^3_\epsilon \times_{T^1} \mathbb{R}$ be the trivial real line bundle over S^2. Take the unit sphere bundle of the following Whitney sum of three vector bundles for $a, b \in \mathbb{Z}$:

$$S^3_\epsilon \times_{T^1} (\mathbb{C}_a \oplus \mathbb{C}_b \oplus \mathbb{R}).$$

Then, we obtain the S^4-bundle over S^2 denoted by

$$M(\epsilon, a, b) = S^3_\epsilon \times_{S^1} S(\mathbb{C}_a \oplus \mathbb{C}_b \oplus \mathbb{R}),$$

for $\epsilon = \pm 1$, $a, b \in \mathbb{Z}$. Namely, we can identify elements in $M(\epsilon, a, b)$ by

$$[(w, z), (x, y, r)] = [(t^{-1}w, t^{\epsilon}z), (t^{a}x, t^{b}y, r)]$$

for any $t \in T^1$ such that $|w|^2+|z|^2 = 1$ and $|x|^2+|y|^2+r^2 = 1$. Define a T^3-action on $M(\epsilon, a, b)$ by

$$[(w, z), (x, y, r)] \mapsto [(t_1 w, z), (t_2 x, t_3 y, r)],$$

where $(t_1, t_2, t_3) \in T^3$. Fix an omniorientation on $M(\epsilon, a, b)$ by the induced orientations from $S^3_{\epsilon} \times S^4 \subset \mathbb{C}^2 \times (\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{R})$. Then, considering the tangential representations around each fixed point, it is easy to check that the induced torus graph is $(\Gamma_S, \mathscr{A}^{\epsilon,a,b}_{e_1,e_2,e_3})$, where e_1, e_2, e_3 are the standard basis of $\mathfrak{t}_{\mathbb{Z}} \simeq \mathbb{Z}^3$. Therefore, by taking the appropriate automorphism of T^3, we can construct each torus graph $(\Gamma_S, \mathscr{A}_S)$ in Figure 2 from $M(\epsilon, a, b)$. Note that if $\epsilon = -1$ and $a = b$, then this is nothing but one of the torus manifolds which appeared in the classifications of torus manifolds with codimension-1 extended actions in [Kuroki 2011].

By the argument above and Corollary 3.5, we have the following lemma.

Lemma 4.2. *Let $(M, \mathbb{O})$ be an omnioriented 6-dimensional simply connected torus manifold with $H^{\mathrm{odd}}(M) = 0$. If its induced torus graph has four vertices, then (M, T) is equivariantly diffeomorphic to one of the following*:

(1) $\mathbb{C}P^3$ *with the standard T^3-action up to automorphism of T^3*;

(2) $M(\epsilon, a, b)$ *for some $\epsilon = \pm 1$ and $a, b \in \mathbb{Z}$.*

4C. *6-dimensional quasitoric manifolds.* Assume that there are no multiple edges in a 3-valent torus graph $(\Gamma, \mathscr{A})$, i.e., there are no two edges that are incident to the same two vertices. A graph Γ is called *simple* if Γ does not have both multiple edges and loops. In this section and in Section 5, we study simple torus graphs which can be realized as the one-skeleton of a manifold with faces homeomorphic to D^3.

The typical example of such torus manifolds whose torus graphs are simple is a quasitoric manifold (introduced by Davis and Januszkiewicz [1991]; see also [Buchstaber and Panov 2002]). A *quasitoric manifold* is defined by a torus manifold whose orbit space is a *simple convex polytope*, i.e., a convex polytope admitting the structure of a manifold with faces. For example, the complex projective space $\mathbb{C}P^n$ with the standard T^n-action is the quasitoric manifold whose orbit space is the n-dimensional simplex. Figure 3 shows the torus graph induced from $(\mathbb{C}P^3, \mathbb{O}_{\mathbb{C}})$, i.e., the omniorientation $\mathbb{O}_{\mathbb{C}}$ induced from the standard complex structure on $\mathbb{C}P^3$ and the standard T-action on $\mathbb{C}P^3$.

We next characterize when torus graphs are induced from simple convex polytopes, i.e., induced from quasitoric manifolds. The Steinitz theorem (see [Ziegler

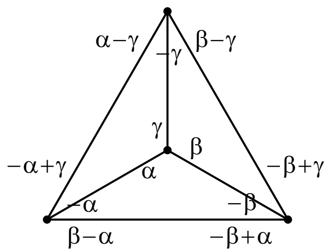


Figure 3. The torus graph induced from $(\mathbb{C}P^3, \mathbb{O}_{\mathbb{C}})$.

1995, Chapter 4]) tells us that a graph Γ is the one-skeleton of a 3-dimensional convex polytope if and only if Γ is a simple, planar and 3-connected graph, where Γ is called a 3-*connected* graph if it remains connected whenever fewer than three vertices are removed. It easily follows from the Steinitz theorem that we have the following lemma.

Lemma 4.3. *Let Q be a manifold with faces and Γ be its graph. Assume that Q is homeomorphic to the* 3*-disk D^3 and there are no multiple edges. Then, the following statements are equivalent*:

(1) *Q is combinatorially equivalent to a* 3*-dimensional simple convex polytope P.*

(2) *Γ is a* 3*-connected graph.*

Combining this result with Corollary 3.5, we have the following fact.

Lemma 4.4. *Let $(M, \mathbb{O})$ be an omnioriented* 6*-dimensional simply connected torus manifold with $H^{\mathrm{odd}}(M) = 0$. Then, the following statements are equivalent*:

(1) *(M, T) is equivariantly diffeomorphic to a quasitoric manifold.*

(2) *Its induced torus graph Γ is a* 3*-connected graph with no multiple edges.*

5. Connected sum of torus graphs and other 6-dimensional torus manifolds

By the arguments in Section 4, only the following case remains: the simply connected 6-dimensional torus manifolds with $H^{\mathrm{odd}}(M) = 0$ whose induced torus graphs are simple but not 3-connected. Such torus manifolds can be constructed using the connected sum of "oriented" torus graphs. The purpose of this section is to introduce oriented torus graphs and their connected sum (see also [Darby 2015]).

We first recall the equivariant connected sum of torus manifolds. Let M_1, M_2 be $2n$-dimensional torus manifolds and $p \in M_1^T$, $q \in M_2^T$ be fixed points. Using the slice theorem, we may take T-invariant open neighborhoods $U_1 \subset M_1$ of p and $U_2 \subset M_2$ of q. Assume that U_1 and U_2 are equivariantly diffeomorphic. Then, $U_1 \backslash \{p\}$ and $U_2 \backslash \{q\}$ are equivariantly diffeomorphic to $S^{2n-1} \times I$, where $S^{2n-1} \subset \mathbb{C}^n$ with some effective T^n-action and $I = (-\epsilon, \epsilon)$ with the trivial T^n-action for some

$\epsilon > 0$. We glue these two neighborhoods by φ defined by the identity on S^{2n-1} and the map $r \mapsto -r$ on I for $r \in I$. Namely, we can glue $M_1 \setminus \{p\}$ and $M_2 \setminus \{q\}$ by the identification

$$M_1 \setminus \{p\} \supset U_1 \setminus \{p\} \xrightarrow{\simeq} S^{2n-1} \times I \xrightarrow{\varphi} S^{2n-1} \times I \xrightarrow{\simeq} U_2 \setminus \{q\} \subset M_2 \setminus \{q\}. \tag{5-1}$$

The T^n-manifold obtained in this way is denoted by $M_1 \# M_2$ or $M_1 \#_{(p,q)} M_2$ (if we emphasize fixed points $p \in M_1^T$ and $q \in M_2^T$). Because each torus manifold has more than two fixed points, $M_1 \# M_2$ is again a torus manifold. We call this operation the *equivariant connected sum*.

Lemma 5.1. *If two torus manifolds M_1 and M_2 are simply connected and $H^{\mathrm{odd}}(M_1) = H^{\mathrm{odd}}(M_2) = 0$, then $M_1 \# M_2$ is also simply connected and $H^{\mathrm{odd}}(M_1 \# M_2) = 0$.*

Proof. It is easy to check the statement using van Kampen's theorem and the Mayer–Vietoris exact sequence. □

Assume that $(M_1, \mathcal{O}_1)$ and $(M_2, \mathcal{O}_2)$ are 6-dimensional omnioriented simply connected torus manifolds with $H^{\mathrm{odd}}(M_1) = H^{\mathrm{odd}}(M_2) = 0$. Let $(\Gamma_1, \mathcal{A}_1)$ and $(\Gamma_2, \mathcal{A}_2)$ be their induced 3-valent torus graphs. Assume that we can glue $p \in M_1^T$ and $q \in M_2^T$ by the connected sum. Then, by considering the restriction of φ in (5-1) onto $S^{2n-1} \subset \mathbb{C}^n$, i.e., the identity map, the axial functions around $p \in V(\Gamma_1)$ and $q \in V(\Gamma_2)$ must satisfy

$$\{\mathcal{A}_1(e) \mid e \in E_p(\Gamma_1)\} = \{\mathcal{A}_2(e) \mid e \in E_q(\Gamma_2)\}. \tag{5-2}$$

However, at this stage, the torus graphs $(\Gamma_1, \mathcal{A}_1)$ and $(\Gamma_2, \mathcal{A}_2)$ do not contain information about the orientations of M_1 and M_2. To do the connected sum, we need the orientations around $p \in M_1^T$ and $q \in M_2^T$. To encode the orientations around fixed points, we need the following notion.

Definition 5.2. Let $(\Gamma, \mathcal{A})$ be a torus graph. We call a triple $(\Gamma, \mathcal{A}, \sigma)$ with a map $\sigma : V(\Gamma) \to \{+1, -1\}$ an *oriented torus graph* if σ satisfies the following condition for all $e \in E(\Gamma)$:

$$\sigma(\pi_V(e))\mathcal{A}(e) = -\sigma(\pi_V(\bar{e}))\mathcal{A}(\bar{e}),$$

where $\pi_V(e) \in V(\Gamma)$ is the initial vertex of $e \in E(\Gamma)$, i.e., for $e = pq$, $\pi_V(e) = p$ and $\pi_V(\bar{e}) = q$. We call such a map σ an *orientation* of $(\Gamma, \mathcal{A})$.

Remark 5.3. Let $(M, \mathcal{O})$ be an omnioriented torus manifold. The oriented torus graph $(\Gamma, \mathcal{A}, \sigma)$ of $(M, \mathcal{O})$ is defined as follows. Let $p \in M^T$. Then, there exist exactly n characteristic submanifolds $M_1, \ldots, M_n$ such that p is a connected component of $\bigcap_{i=1}^n M_i$. Now the fixed orientations of $M_1, \ldots, M_n$ determine the decomposition of the tangential representation; i.e., $\psi_p : T_pM \xrightarrow{\simeq} V(\alpha_1) \oplus \cdots \oplus V(\alpha_n)$ is determined by fixing the orientations of $M_1, \ldots, M_n$. On the other hand, the orientation of M determines the orientation of T_pM. So, we define the map $\sigma : V(\Gamma) = M^T \to \{+1, -1\}$ by

$$\sigma(p) = \begin{cases} +1 & \text{if } \psi_p \text{ preserves the orientations,} \\ -1 & \text{if } \psi_p \text{ reverses the orientations.} \end{cases}$$

Let $(\Gamma_1, \mathscr{A}_1, \sigma_1)$ and $(\Gamma_2, \mathscr{A}_2, \sigma_2)$ be the induced oriented torus graphs from $(M_1, \mathbb{O}_1)$ and $(M_2, \mathbb{O}_2)$. If we can glue $p \in M_1^T$ and $q \in M_2^T$ by the connected sum, then both (5-2) and

$$\sigma_1(p) \neq \sigma_2(q) \tag{5-3}$$

hold ((5-3) corresponds to the fact that the orientations on T_pM_1 and T_qM_2 are different). The induced (oriented) torus graph by $M_1 \#_{(p,q)} M_2$ is nothing but the one-skeleton of the connected sum $Q_1 \#_{(p,q)} Q_2$ of manifolds with faces, where Q_i is the orbit space of M_i, $i = 1, 2$ (see [Izmest′ev 2001, Definition 3; Kuroki 2010, Section 3.1] for details about the connected sum of polytopes). Therefore, conversely, if $p \in V(\Gamma_1)$ and $q \in V(\Gamma_2)$ satisfy (5-2) and (5-3), then we can do the *connected sum of* (*oriented*) *torus graphs* between $(\Gamma_1, \mathscr{A}_1, \sigma_1)$ and $(\Gamma_2, \mathscr{A}_2, \sigma_2)$, say $(\Gamma, \mathscr{A}, \sigma) = (\Gamma_1, \mathscr{A}_1, \sigma_1) \# (\Gamma_2, \mathscr{A}_2, \sigma_2)$ or $(\Gamma_1, \mathscr{A}_1, \sigma_1) \#_{(p,q)} (\Gamma_2, \mathscr{A}_2, \sigma_2)$ (if we emphasize the vertices $p \in V(\Gamma_1)$ and $q \in V(\Gamma_2)$). More precisely, $(\Gamma, \mathscr{A}, \sigma) = (\Gamma_1, \mathscr{A}_1, \sigma_1) \# (\Gamma_2, \mathscr{A}_2, \sigma_2)$ is defined as follows (see Figure 4).

(1) $V(\Gamma) = V(\Gamma_1) \setminus \{p\} \sqcup V(\Gamma_2) \setminus \{q\}$.

(2) $E(\Gamma)$ is given by

$$(E(\Gamma_1) \setminus \{pp_1, pp_2, pp_3\}) \sqcup (E(\Gamma_2) \setminus \{qq_1, qq_2, qq_3\}) \sqcup \{p_1q_1, p_2q_2, p_3q_3\},$$

where $\mathscr{A}_1(pp_i) = \mathscr{A}_2(qq_i)$ for $i = 1, 2, 3$.

(3) $\mathscr{A} : E(\Gamma) \to (\mathfrak{t}_{\mathbb{Z}}^3)^*$ is defined by $\mathscr{A}(e) = \mathscr{A}_1(e)$ and $\mathscr{A}(f) = \mathscr{A}_2(f)$ for e in $E(\Gamma_1) \setminus \{pp_1, pp_2, pp_3\}$ and f in $E(\Gamma_2) \setminus \{qq_1, qq_2, qq_3\}$, and $\mathscr{A}(p_iq_i) = \mathscr{A}_1(p_ip)$ and $\mathscr{A}(q_ip_i) = \mathscr{A}_2(q_iq)$.

(4) $\sigma : V(\Gamma) \to \{+1, -1\}$ is defined by $\sigma(r) = \sigma_1(r)$ for $r \in V(\Gamma_1) \setminus \{p\}$ and $\sigma(r') = \sigma_2(r')$ for $r' \in V(\Gamma_2) \setminus \{q\}$.

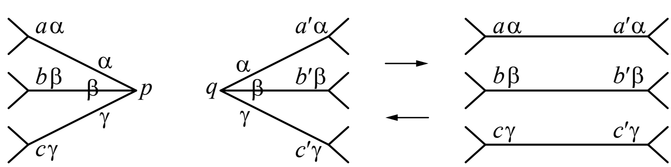


Figure 4. The local figure of the connected sum $\#_{(p,q)}$ of a torus manifold (left to right) and its inverse $\#_{(p,q)}^{-1}$ (right to left), where $\sigma_1(p) \neq \sigma_2(q)$. Here, α, β, γ are a $\mathbb{Z}$-basis of $(\mathfrak{t}_{\mathbb{Z}}^3)^*$ and $a, a', b, b', c, c' = \pm 1$.

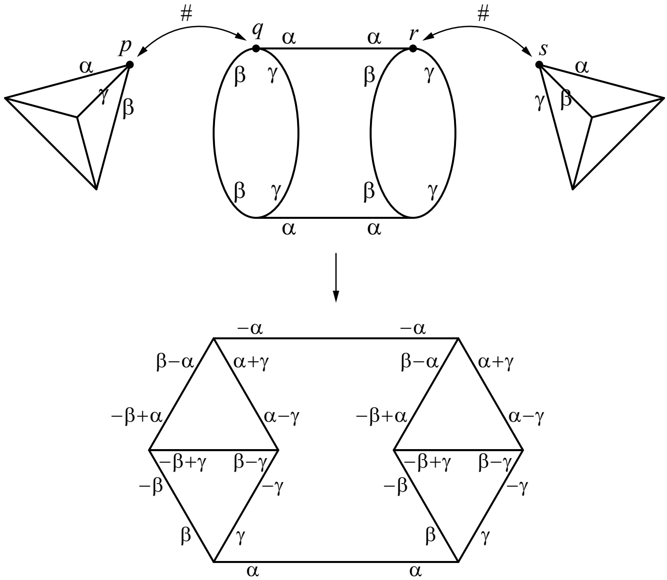


Figure 5. The torus graph (with appropriate orientations, e.g, $\sigma(p) = +1$, $\sigma(q) = -1$, $\sigma(r) = +1$, $\sigma(s) = -1$) induced from $\mathbb{C}P^3 \# (S^2 \times S^4) \# \overline{\mathbb{C}P^3}$.

Then, we can easily check that $(\Gamma, \mathscr{A}, \sigma)$ is an oriented torus graph. Using Corollary 3.5 and the arguments above, we have the following lemma.

Lemma 5.4. *Let M_1 and M_2 be 6-dimensional simply connected torus manifolds with $H^{\mathrm{odd}}(M_1) = H^{\mathrm{odd}}(M_2) = 0$, and let $(\Gamma_1, \mathscr{A}_1, \sigma_1)$ and $(\Gamma_2, \mathscr{A}_2, \sigma_2)$ be their respective induced oriented torus graphs from some omniorientations. If $(\Gamma, \mathscr{A}, \sigma) = (\Gamma_1, \mathscr{A}_1, \sigma_1) \#_{(p,q)} (\Gamma_2, \mathscr{A}_2, \sigma_2)$, then $(\Gamma, \mathscr{A}, \sigma)$ is the oriented torus graph induced from $M = M_1 \#_{(p,q)} M_2$ with some omniorientation.*

Using the connected sum, we can construct the torus manifolds which do not appear in Section 4. One such example is

$$\mathbb{C}P^3 \# (S^2 \times S^4) \# \overline{\mathbb{C}P^3},$$

where $\overline{\mathbb{C}P^3}$ is the reversed orientation of $\mathbb{C}P^3$. Figure 5 shows the torus graph induced from $\mathbb{C}P^3 \# (S^2 \times S^4) \# \overline{\mathbb{C}P^3}$ (see the axial functions in Figures 2 and 3 for details). We can easily check that this graph is 3-valent, simple and planar but not 3-connected; therefore, by Lemma 4.4, this manifold is not a quasitoric manifold.

6. Some combinatorial lemmas

To prove the main theorem (Theorem 7.1), we need the following two lemmas.

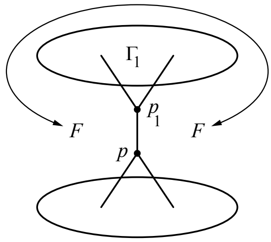


Figure 6. The figure explained in the proof of Lemma 6.1. The facet F has a self-intersection on the edge p_1p.

Lemma 6.1. *Let Q be a 3-dimensional manifold with faces which is homeomorphic to D^3 and Γ be its graph. Then, $\Gamma \setminus \{p\}$ is connected for all vertices $p \in V(\Gamma)$.*

Proof. Because Q is homeomorphic to the 3-disk D^3, Γ may be regarded as a planar graph by the stereographic projection of $\partial Q = S^2$. Assume $\Gamma \setminus \{p\}$ is not connected. Because Q is a 3-dimensional manifold with faces, there are exactly three outgoing edges from p, say pp_1, pp_2 and pp_3. Therefore, we may assume that there exists a connected component Γ_1 in $\Gamma \setminus \{p\}$ such that $p_1 \in V(\Gamma_1)$ but $p_2, p_3 \notin V(\Gamma_1)$ (see Figure 6). Since Γ_1 is also a planar 3-valent graph except on the vertex p_1 (because $p \notin V(\Gamma_1)$), there is a 2-valent subgraph in Γ_1, say $\partial\Gamma_1$, such that $\partial\Gamma_1$ splits $\partial Q = S^2$ into two connected components H_+ and H_-, where $\Gamma_1 \setminus \partial\Gamma_1 \subset H_+ \setminus \partial\Gamma_1$ but $\Gamma_1 \not\subset H_-$. This implies that there is a facet F in Q such that ∂F contains $\partial\Gamma_1$ and p_1p. However, in this case, p_1p must be a self-intersection edge of F (see Figure 6). This contradicts that Q is a manifold with faces. □

By Lemma 6.1, if Γ is not 3-connected, then there are two vertices $p, q \in V(\Gamma)$ such that $\Gamma \setminus \{p, q\}$ is not connected but both $\Gamma \setminus \{p\}$ and $\Gamma \setminus \{q\}$ are connected. More precisely, we have the following lemma.

Lemma 6.2. *Let Q be a 3-dimensional manifold with faces which is homeomorphic to D^3 and Γ be its graph. Assume that there are two vertices $p, q \in V(\Gamma)$ such that $\{p, q\} \not\subset V(F)$ for any facets F, i.e, p and q are not on the same facet F. Then, $\Gamma \setminus \{p, q\}$ is connected.*

Proof. Assume that p and q are not on the same facet of Q. Because Q is a manifold with faces, there are mutually distinct facets $F_1, \dots, F_6$ such that $\{p\}$ is a component of $F_1 \cap F_2 \cap F_3$ and $\{q\}$ is a component of $F_4 \cap F_5 \cap F_6$, and we can take vertices p_1, p_2, p_3 and q_1, q_2, q_3 such that pp_i and qq_i are all outgoing edges from p and q for $i = 1, 2, 3$. Take two vertices r and s from $\Gamma \setminus \{p, q\}$. By Lemma 6.1, $\Gamma \setminus \{q\}$ is connected. So there is a path γ from r to s in $\Gamma \setminus \{q\}$. If γ does not go through p, then r and s are connected in $\Gamma \setminus \{p, q\}$. Assume that this path γ goes through p. Then γ goes through exactly two vertices p_i, p_j (we may assume p_1 and p_2). Moreover, one of the facets F_1, F_2, F_3, say F_1, contains both p_1 and p_2.

Note that F_1 corresponds to the 2-valent subgraph in Γ. Therefore, we can take the path γ_p connecting p_1 and p_2 on F_1 which is not the path p_1pp_2. Because p and q are not on the same facet, in particular $q \notin V(F_1)$, the path γ_p does not contain q. Hence, the connected subgraph $\gamma \cup \gamma_p$ contains both r and s but does not contain both p and q. Thus, we can take the path γ' from r to s in $\gamma \cup \gamma_p \subset \Gamma \setminus \{p, q\}$. This establishes that $\Gamma \setminus \{p, q\}$ is connected. □

In summary, by Lemmas 6.1 and 6.2, we have the following fact.

Corollary 6.3. *Let Γ be a one-skeleton of a 3-dimensional manifold with faces Q. Then, for all $p \in V(\Gamma)$, $\Gamma \setminus \{p\}$ is connected. Furthermore, if $\Gamma \setminus \{p, q\}$ is not connected, then p and q are on the same facet.*

7. Proof of main theorem

The main theorem of this paper can be stated as follows:

Theorem 7.1. *Let M be a simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M) = 0$. Then, either M is equivariantly diffeomorphic to*

1. *$S^6 \subset \mathbb{C}^3 \oplus \mathbb{R}$ with a torus action induced from a (faithful) representation of T^3 on $\mathbb{C}^3$,*
2. *a 6-dimensional quasitoric manifold X, or*
3. *an S^4-bundle over S^2 which is equivariantly diffeomorphic to $M(\epsilon, a, b)$ for some $\epsilon = \pm 1$, $a, b \in \mathbb{Z}$;*

or else there are some 6-dimensional quasitoric manifolds X_h for some $h = 1, \ldots, k$, and some S^4-bundles over S^2, say $S_i = M(\epsilon_i, a_i, b_i)$ (for some $\epsilon_i = \pm 1$, $a_i, b_i \in \mathbb{Z}$ and $i = 1, \ldots, \ell$), such that M is equivariantly diffeomorphic to

$$\Big(\mathop{\#}_{h=1}^{k} X_h\Big) \# \Big(\mathop{\#}_{i=1}^{\ell} S_i\Big),$$

where # represents the equivariant connected sum around fixed points, $k + \ell \geq 2$ for $k \geq 0$, $\ell \geq 1$, and the case $k = 0$ means that there is no X_h factor.

In this final section, we prove Theorem 7.1.

Let M be a simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M) = 0$, Q be its orbit space which is homeomorphic to D^3 and $(\Gamma_M, \mathcal{A}_M)$ be its induced oriented torus graph (we omit the orientation).

Because Γ_M is a one-skeleton of a manifold with faces which is homeomorphic to D^3, it is easy to check that $|V(\Gamma_M)| \neq 1, 3$. If $|V(\Gamma_M)| = 2$, by Lemma 4.1, we have that M is equivariantly diffeomorphic to S^6, i.e., statement (1). If $|V(\Gamma_M)| = 4$, it follows from Lemma 4.2 that M is equivariantly diffeomorphic to a quasitoric manifold $\mathbb{C}P^3$ or $M(\epsilon, a, b)$ for some $\epsilon = \pm 1$, $a, b \in \mathbb{Z}$, i.e., statement (2) or (3) occurs. So we need only prove the case when $|V(\Gamma_M)| \geq 5$.

We first establish the following lemma.

Lemma 7.2. *Assume that $|V(\Gamma_M)| \geq 5$ and there is a multiple edge in Γ_M. Then, $(\Gamma_M, \mathscr{A}_M)$ can be decomposed as*

$$(\Gamma_M, \mathscr{A}_M) = (\Gamma_X, \mathscr{A}_X) \# (\Gamma_{S_1}, \mathscr{A}_{S_1}) \# \cdots \# (\Gamma_{S_{\ell'}}, \mathscr{A}_{S_{\ell'}})$$

or

$$(\Gamma_M, \mathscr{A}_M) = (\Gamma_{S_1}, \mathscr{A}_{S_1}) \# \cdots \# (\Gamma_{S_{\ell'}}, \mathscr{A}_{S_{\ell'}}),$$

where $(\Gamma_X, \mathscr{A}_X)$ is a torus graph without multiple edges and $S_i = M(\epsilon_i, a_i, b_i)$ for $i = 1, \ldots, \ell'$.

Proof. Assume two vertices p and q are connected by a multiple edge, i.e., two edges (see the bottom graph in Figure 7). Then, by the connection of the torus graph (see Proposition 3.1), it is easy to check that the axial functions around the

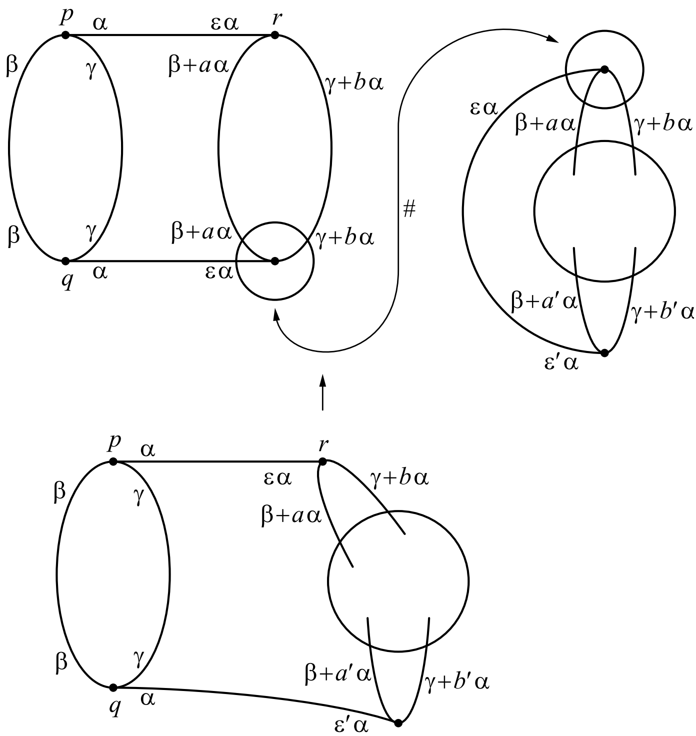


Figure 7. We may regard α, β, γ as any generators in $(\mathfrak{t}^3_{\mathbb{Z}})^*$ and $a, a', b, b' \in \mathbb{Z}$ and $\epsilon, \epsilon' = \pm 1$. The bottom graph is $(\Gamma_M, \mathscr{A}_M)$, the upper-left graph is $(\Gamma_{S_1}, \mathscr{A}_{S_1})$ and the upper-right graph is $(\Gamma_{M'}, \mathscr{A}_{M'})$. If we fix the orientation of $(\Gamma_M, \mathscr{A}_M)$ then the orientations of $(\Gamma_{S_1}, \mathscr{A}_{S_1})$ and $(\Gamma_{M'}, \mathscr{A}_{M'})$ are automatically determined.

vertex r of the bottom graph in Figure 7 satisfy the axial functions expressed in that figure, where we can take α, β, γ as a $\mathbb{Z}$-basis of $(\mathfrak{t}^3_{\mathbb{Z}})^*$. In this case, we can do an (inverse) connected sum such as the one expressed in Figure 7 (from the bottom to the top). Then, the induced torus graph $(\Gamma_M, \mathcal{A}_M)$ is decomposed into two induced torus graphs $(\Gamma_{S_1}, \mathcal{A}_{S_1})$ and $(\Gamma_{M'}, \mathcal{A}_{M'})$, where M' is some simply connected 6-dimensional torus manifold with $H^{\mathrm{odd}}(M') = 0$ by Lemma 5.1. Namely, we have

$$(\Gamma_M, \mathcal{A}_M) = (\Gamma_{M'}, \mathcal{A}_{M'}) \# (\Gamma_{S_1}, \mathcal{A}_{S_1}).$$

If there are no multiple edges in $\Gamma_{M'}$, then we may put $\Gamma_{M'} = \Gamma_X$. Assume that there is a multiple edge in $\Gamma_{M'}$. If there are only four vertices in $\Gamma_{M'}$, then we may put M' as $S_2 = M(\epsilon_2, a_2, b_2)$ by Lemma 4.2. When there are more than four vertices in $\Gamma_{M'}$, we iterate the above argument, establishing the lemma. □

Therefore, to prove Theorem 7.1, it is enough to prove the following lemma.

Lemma 7.3. *Assume that $|V(\Gamma_M)| \geq 5$ and there are no multiple edges in Γ_M. Then, $(\Gamma_M, \mathcal{A}_M)$ can be decomposed as*

$$(\Gamma_M, \mathcal{A}_M) = (\Gamma_{X_1}, \mathcal{A}_{X_1}) \# \cdots \# (\Gamma_{X_k}, \mathcal{A}_{X_k}) \# (\Gamma_{S_1}, \mathcal{A}_{S_1}) \# \cdots \# (\Gamma_{S_{\ell''}}, \mathcal{A}_{S_{\ell''}}),$$

where $(\Gamma_{X_h}, \mathcal{A}_{X_h})$ for $h = 1, \ldots, k$ is the torus graph induced from a quasitoric manifold X_h, and $S_i = M(\epsilon_i, a_i, b_i)$ for $i = 1, \ldots, \ell''$.

Proof. If Γ_M $(= \Gamma)$ is 3-connected, then it follows from Lemma 4.4 that the statement holds, i.e., $k = 1$, $\ell'' = 0$. Therefore, we may assume Γ is not 3-connected. In this case, by Corollary 6.3, there is a 2-valent torus subgraph $F \subset \Gamma$ such that $\Gamma \setminus \{p, q\}$ is not connected for some $p, q \in V(F)$.

If F is a triangle (i.e., $|V(F)| = 3$), using a method similar to that demonstrated in the proof of Lemma 6.1, we have that there is a face in Q which has a self-intersection edge. This contradicts that Q is a manifold with faces. Therefore, we may assume $|V(F)| \geq 4$. We first assume that pq is an edge of F. Then, there are two graphs Γ_1 and Γ_2 which are the connected components of $\Gamma \setminus \{p, q\}$ expressed in Figure 8. If we remove the two vertices r and q from Γ instead of p and q, where $r \in V(\Gamma_2)$ such that pr is an edge, then $\Gamma \setminus \{r, q\}$ is also not connected (see Figure 8). Therefore, we may assume that

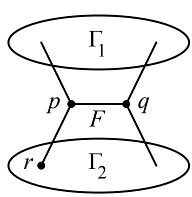


Figure 8. If we remove r and q from Γ instead of p and q, the graph is also disconnected.

(1) $p, q \in V(\Gamma)$ are such that $\Gamma \setminus \{p, q\}$ is not connected,

(2) $pq \notin E(\Gamma)$,

(3) there is a 2-valent torus subgraph (facet) F with $|V(F)| \geq 4$ in Γ such that $p, q \in V(F)$.

We call such a facet F a *singular facet*.

Let F be a singular facet. Assume $|V(F)| \geq 6$. In this case, by an argument similar to the one just before, we may take p and q to be in the position of Figure 9, i.e., p and q are on two separated edges rp and sq which are common edges of two facets F and F' in Figure 9 (note that r and s might be connected by an edge). Moreover, by considering the omnioriented characteristic functions of the facets F and F', we may take the axial functions around the facet F to be as in Figure 9.

By taking an appropriate orientation, we can do the connected sum as in Figure 10; here we denote the (oriented) torus graph containing Γ_1 by $(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1)$ and that containing Γ_2 by $(\widetilde{\Gamma}'_2, \widetilde{\mathscr{A}}'_2)$. The torus graph obtained by this connected sum is nothing but the torus graph $(\Gamma, \mathscr{A})$ in Figure 9. Note that $\widetilde{\Gamma}_1$ is simple and planar, while $\widetilde{\Gamma}'_2$ is just planar. With a method similar to that demonstrated in Figure 7, $(\widetilde{\Gamma}'_2, \widetilde{\mathscr{A}}'_2)$ can be obtained from the connected sum of $(\Gamma_S, \mathscr{A}_S)$ and the simple, planar graph $(\widetilde{\Gamma}_2, \widetilde{\mathscr{A}}_2)$ (containing Γ_2), where $(\Gamma_S, \mathscr{A}_S)$ is one of the torus graphs (by taking the appropriate axial functions) in Figure 2. Namely, the torus graph in Figure 9 can be obtained from the connected sum

$$(\Gamma, \mathscr{A}) = (\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1) \# (\Gamma_S, \mathscr{A}_S) \# (\widetilde{\Gamma}_2, \widetilde{\mathscr{A}}_2).$$

Here, it is easy to check that $\widetilde{\Gamma}_i$ consists of Γ_i and the other two facets, say $\widetilde{F}(i)$ and $\widetilde{F}'(i)$ (induced from F and F' in Γ). Because of Figure 10, the number of

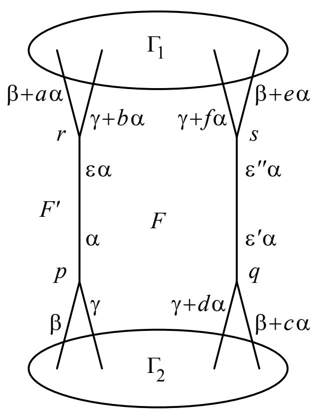


Figure 9. The axial functions around F when $|V(F)| \geq 6$, where $\epsilon, \epsilon', \epsilon'' = \pm 1$ and $a, b, c, d, e, f \in \mathbb{Z}$. Here, F' is a facet which intersects F on pr and qs.

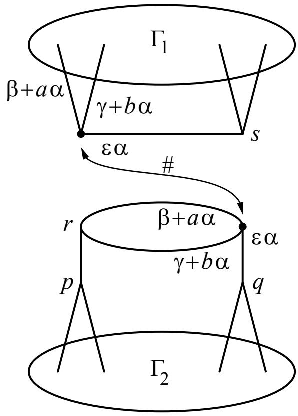


Figure 10. The torus graph $(\Gamma, \mathscr{A})$ in Figure 9 splits into two torus graphs $(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1)$ (upper) and $(\widetilde{\Gamma}'_2, \widetilde{\mathscr{A}}'_2)$ (lower). Here, we omit the axial functions around the vertices p, q, r, s because they are exactly the same as those in Figure 9.

vertices of $\widetilde{F}(i)$ and $\widetilde{F}'(i)$ is reduced; in particular, the number of vertices of the facet $\widetilde{F}(i)$ induced from the singular facet F is strictly less than 6. If both $(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1)$ and $(\widetilde{\Gamma}_2, \widetilde{\mathscr{A}}_2)$ are 3-connected, then these torus graphs are induced from quasitoric manifolds, i.e, the statements of Lemma 7.3 hold. Assume that $(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1)$ is not 3-connected. Then, by the above arguments, there is a singular facet F in $(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1)$. If $|V(F)| \geq 6$, then $(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1)$ also decomposes as

$$(\widetilde{\Gamma}_1, \widetilde{\mathscr{A}}_1) = (\widetilde{\Gamma}_3, \widetilde{\mathscr{A}}_3) \# (\Gamma_{S'}, \mathscr{A}_{S'}) \# (\widetilde{\Gamma}_4, \widetilde{\mathscr{A}}_4),$$

using arguments similar to those Figure 10. Iterating, we may reduce all singular facets with $|V(F)| \geq 6$. More precisely, we may decompose $(\Gamma, \mathscr{A})$ in Figure 9 as

$$(\Gamma, \mathscr{A}) = \mathop{\#}_{i=1}^{\ell} \{(\Gamma_i, \mathscr{A}_i) \# (\Gamma_{S_i}, \mathscr{A}_{S_i}) \# (\Gamma_{i+\ell}, \mathscr{A}_{i+\ell})\},$$

where $(\Gamma_{S_i}, \mathscr{A}_{S_i})$ for $i = 1, \ldots, \ell$ is a torus graph in Figure 2 and $(\Gamma_h, \mathscr{A}_h)$ for $h = 1, \ldots, 2\ell$ is a 3-valent simple and planar torus graph such that either

- $(\Gamma_h, \mathscr{A}_h)$ is 3-connected (in this case, induced from a quasitoric manifold), or
- all singular facets F satisfy $|V(F)| = 4$ or 5.

Assume that the number of vertices in every singular facet of the torus graph $(\Gamma, \mathscr{A})$ is less than or equal to 5. Then, such a torus graph is one of the torus graphs expressed in Figure 11. However, because Γ is the one-skeleton of a manifold with faces and is not 3-connected, it is easy to check that there exists a singular facet F' such that $F' \cap F = \{pr, qs\}$ and $|V(F')| \geq 6$. This gives a contradiction. Hence, this case does not occur. This establishes Lemma 7.3. □

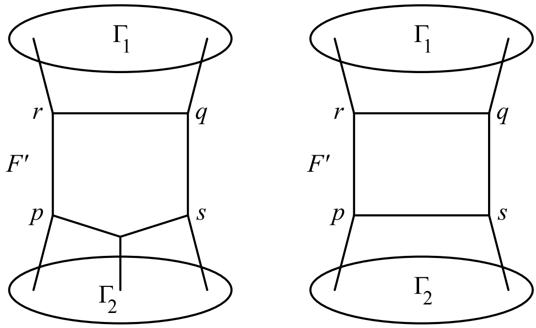


Figure 11. The singular facets F with $|V(F)| = 5$ or 4. Here, F' is a facet which intersects F on pr and qs.

Consequently, by Lemmas 5.4, 7.2 and 7.3, we have the statement of Theorem 7.1.

Finally, by Theorem 7.1 and the Mayer–Vietoris exact sequence, we also have the following well-known result.

Corollary 7.4. *Let M be a simply connected 6-dimensional torus manifold whose cohomology ring is generated by the second-degree cohomology. Then, M is a quasitoric manifold.*

Acknowledgments

The author would like to thank Professors Yael Karshon and Takashi Tsuboi for providing him excellent circumstances in which to do research.

References

[Buchstaber and Panov 2002] V. M. Buchstaber and T. E. Panov, *Torus actions and their applications in topology and combinatorics*, University Lecture Series **24**, Amer. Math. Soc., Providence, RI, 2002. MR 2003e:57039 Zbl 1012.52021

[Darby 2015] A. Darby, "Torus manifolds in equivariant complex bordism", *Topology Appl.* **189** (2015), 31–64. MR 3342571 Zbl 06435721

[Davis and Januszkiewicz 1991] M. W. Davis and T. Januszkiewicz, "Convex polytopes, Coxeter orbifolds and torus actions", *Duke Math. J.* **62**:2 (1991), 417–451. MR 92i:52012 Zbl 0733.52006

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton Univ. Press, 1993. MR 94g:14028 Zbl 0813.14039

[Guillemin and Zara 2001] V. Guillemin and C. Zara, "1-skeleta, Betti numbers, and equivariant cohomology", *Duke Math. J.* **107**:2 (2001), 283–349. MR 2002j:53110 Zbl 1020.57013

[Hattori and Masuda 2003] A. Hattori and M. Masuda, "Theory of multi-fans", *Osaka J. Math.* **40**:1 (2003), 1–68. MR 2004d:53103 Zbl 1034.57031

[Ishida et al. 2013] H. Ishida, Y. Fukukawa, and M. Masuda, "Topological toric manifolds", *Mosc. Math. J.* **13**:1 (2013), 57–98. MR 3112216 Zbl 1302.53091

[Izmest′ev 2001] I. V. Izmest′ev, "Three-dimensional manifolds defined by a coloring of the faces of a simple polytope", *Mat. Zametki* **69**:3 (2001), 375–382. In Russian; translated in *Math. Notes* **69**:3 (2001), 340–346. MR 2002g:57005 Zbl 0991.57016

[Kawakubo 1991] K. Kawakubo, *The theory of transformation groups*, Oxford Univ. Press, 1991. MR 93g:57044 Zbl 0744.57001

[Kuroki 2008] S. Kuroki, "Remarks on McGavran's paper and Nishimura's results", *Trends Math. (N.S.)* **10**:1 (2008), 77–79.

[Kuroki 2010] S. Kuroki, "Operations on 3-dimensional small covers", *Chin. Ann. Math. Ser. B* **31**:3 (2010), 393–410. MR 2011f:57029 Zbl 1217.57010

[Kuroki 2011] S. Kuroki, "Classification of torus manifolds with codimension one extended actions", *Transform. Groups* **16**:2 (2011), 481–536. MR 2012g:57059 Zbl 1246.57082

[Lee 2013] J. M. Lee, *Introduction to smooth manifolds*, 2nd ed., Graduate Texts in Mathematics **218**, Springer, New York, 2013. MR 2954043 Zbl 1258.53002

[Lü and Yu 2011] Z. Lü and L. Yu, "Topological types of 3-dimensional small covers", *Forum Math.* **23**:2 (2011), 245–284. MR 2012h:57038 Zbl 1222.52015

[Maeda et al. 2007] H. Maeda, M. Masuda, and T. Panov, "Torus graphs and simplicial posets", *Adv. Math.* **212**:2 (2007), 458–483. MR 2008e:55007 Zbl 1119.55004

[Masuda 1999] M. Masuda, "Unitary toric manifolds, multi-fans and equivariant index", *Tohoku Math. J.* (2) **51**:2 (1999), 237–265. MR 2000e:57058 Zbl 0940.57037

[Masuda 2005] M. Masuda, "h-vectors of Gorenstein* simplicial posets", *Adv. Math.* **194**:2 (2005), 332–344. MR 2006b:52009 Zbl 1063.05010

[Masuda and Panov 2006] M. Masuda and T. Panov, "On the cohomology of torus manifolds", *Osaka J. Math.* **43**:3 (2006), 711–746. MR 2007j:57039 Zbl 1111.57019

[McGavran 1976] D. McGavran, "T^3-actions on simply connected 6-manifolds, I", *Trans. Amer. Math. Soc.* **220** (1976), 59–85. MR 54 #3729 Zbl 0329.57019

[Nishimura 2012] Y. Nishimura, "Combinatorial constructions of three-dimensional small covers", *Pacific J. Math.* **256**:1 (2012), 177–199. MR 2928547 Zbl 1253.57010

[Oda 1988] T. Oda, *Convex bodies and algebraic geometry: an introduction to the theory of toric varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **15**, Springer, Berlin, 1988. MR 88m:14038 Zbl 0628.52002

[Orlik and Raymond 1970] P. Orlik and F. Raymond, "Actions of the torus on 4-manifolds, I", *Trans. Amer. Math. Soc.* **152** (1970), 531–559. MR 42 #3808 Zbl 0216.20202

[Wiemeler 2013] M. Wiemeler, "Exotic torus manifolds and equivariant smooth structures on quasi-toric manifolds", *Math. Z.* **273**:3-4 (2013), 1063–1084. MR 3030690 Zbl 1269.57014

[Ziegler 1995] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics **152**, Springer, New York, 1995. MR 96a:52011 Zbl 0823.52002

Received September 4, 2014.

SHINTARÔ KUROKI
GRADUATE SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF TOKYO
3-8-1 KOMABA MEGURO-KU
TOKYO 153-8914
JAPAN
kuroki@ms.u-tokyo.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.115

SOLUTIONS WITH LARGE NUMBER OF PEAKS FOR THE SUPERCRITICAL HÉNON EQUATION

ZHONGYUAN LIU AND SHUANGJIE PENG

This paper is concerned with the Hénon equation

$$\begin{cases} -\Delta u = |y|^{\alpha} u^{p+\varepsilon}, \ u > 0, & \text{in } B_1(0), \\ u = 0 & \text{on } \partial B_1(0), \end{cases}$$

where $B_1(0)$ is the unit ball in $\mathbb{R}^N$ $(N \geq 4)$, $p = (N+2)/(N-2)$ is the critical Sobolev exponent, $\alpha > 0$ and $\varepsilon > 0$. We show that if ε is small enough, this problem has a positive peak solution which presents a new phenomenon: the number of its peaks varies with the parameter ε at the order $\varepsilon^{-1/(N-1)}$ when $\varepsilon \to 0^+$. Moreover, all peaks of the solutions approach the boundary of $B_1(0)$ as ε goes to 0^+.

1. Introduction and main results

We study the existence of positive solutions to a type of nonlinear elliptic problem whose typical form is the supercritical problem

$$\begin{cases} -\Delta u = |y|^{\alpha} u^{p+\varepsilon}, \ u > 0, & \text{in } B_1(0), \\ u = 0 & \text{on } \partial B_1(0), \end{cases} \tag{1-1}$$

where $p = (N+2)/(N-2)$, $\alpha > 0$, $\varepsilon > 0$ and $B_1(0)$ is the unit ball in $\mathbb{R}^N$ $(N \geq 4)$.

It is well known that the problem

$$\begin{cases} -\Delta u = |y|^{\alpha} u^{q}, \ u > 0, & \text{in } B_1(0), \\ u = 0 & \text{on } \partial B_1(0), \end{cases} \tag{1-2}$$

was proposed by M. Hénon [1973] when he studied rotating stellar structures and is hence called the Hénon equation, and it has attracted a lot of interest in recent years. Ni [1982] first considered (1-2) and proved that it possesses a positive radial solution when $q \in (1, (N+2+2\alpha)/(N-2))$. Due to the appearance of the weighted term $|y|^{\alpha}$, the classical moving plane method in [Gidas et al. 1979] cannot be applied to problem (1-2). It is natural to ask whether problem (1-2) has nonradial solutions. The existence of a nonradial solution for $1 < q < p$ was obtained by

MSC2010: primary 35J60; secondary 35J65, 58E05.
Keywords: peak solutions, supercritical Hénon equation, reduction method.

Smets, Willem and Su [2002] provided α is large enough. When $q = p - \varepsilon$, Cao and Peng [2003] showed that the ground state solution is nonradial and blows up near the boundary of $B_1(0)$ as $\varepsilon \to 0$. Later on, Peng [2006] constructed multiple boundary peak solutions for problem (1-2). When $q = p$, Serra [2005] proved that problem (1-2) has a nonradial solution provided α is large enough. More recently, Wei and Yan [2013] showed that there are infinitely many nonradial solutions for problem (1-2) with $\alpha > 0$. For other results related to Hénon type problems, see [Byeon and Wang 2006; 2005; Cao et al. 2009; Hirano 2009; Pistoia and Serra 2007] and the references therein.

On the other hand, using the Pohozaev identity [1965], we know that for $q \geq \frac{N+2+2\alpha}{N-2}$ there are no solutions to problem (1-2) in star-shaped domains with respect to the origin. So it seems more interesting to consider whether there are solutions for q in the range $\left(\frac{N+2}{N-2}, \frac{N+2+2\alpha}{N-2}\right)$. However, much less is known about that case. When $q = \frac{N+2+2\alpha}{N-2} - \varepsilon$, Gladiali and Grossi [2012] showed that there exists one solution concentrating at $y = 0$ provided $0 < \alpha \leq 1$. By the results in [Gladiali et al. 2013], the same results still hold when α is not an even integer. In [Li and Peng 2009], the asymptotic behavior of the radial solutions obtained by Ni [1982] was analyzed as $\varepsilon \to 0^+$.

The purpose of this paper is to study the supercritical problem (1-1) and try to construct solutions whose number of peaks varies with ε as $\varepsilon \to 0^+$. In fact, we will consider the more general problem

$$\begin{cases} -\Delta u = K(|y|)u^{p+\varepsilon}, \ u > 0, & \text{in } B_1(0), \\ u = 0 & \text{on } \partial B_1(0), \end{cases} \tag{1-3}$$

where $K(r) \in C^1[0, 1]$ and $K(1) > 0$.

Without loss of generality, we can assume that

$$K(1) = 1.$$

The main result of this paper is as follows.

Theorem 1.1. *Assume that $N \geq 4$. If $K(r)$ satisfies $K(1) > 0$ and $K'(1) > 0$, then there exists $\varepsilon_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0)$, problem (1-3) has a solution u_ε whose number of local maximal points is of the order $\varepsilon^{-1/(N-1)}$ as $\varepsilon \to 0^+$. In particular, problem (1-1) has solutions with a large number of peaks for small $\varepsilon > 0$.*

Remark 1.2. For the case $\alpha = 0$, the well-known Pohozaev identity [1965] implies that (1-1) has no solutions for $\varepsilon > 0$. It was also shown in [Ben Ayed et al. 2003] that problem (1-1) has no single-peak solutions for ε small enough. Our results mean that the weight $|y|^\alpha$ has a great influence on the existence of peak solutions for problem (1-1).

Let us outline the main idea in the proof of Theorem 1.1. We introduce some notation first. For $x \in \mathbb{R}^N$ and $\Lambda > 0$, set

$$U_{x,\Lambda}(y) = C_N \left(\frac{\Lambda}{1+\Lambda^2|y-x|^2} \right)^{(N-2)/2}, \quad C_N = (N(N-2))^{(N-2)/4}.$$

It's well known that $U_{x,\Lambda}(y)$ are the only solutions of

$$-\Delta u = u^{(N+2)/(N-2)}, \ u > 0, \quad \text{in } \mathbb{R}^N.$$

Let

$$k = [\varepsilon^{-1/(N-1)}],$$

where $[a]$ denotes the integer part of a real number a. By the transformation $u(y) \mapsto \varepsilon^{2/(4+(N-2)\varepsilon)} u(\varepsilon^{1/(N-2)} y)$ and setting $B_* = B_{\varepsilon^{-1/(N-2)}}$, we see that (1-3) becomes

$$\begin{cases} -\Delta u = K(\varepsilon^{1/(N-2)}|y|) u^{p+\varepsilon}, \ u > 0, & \text{in } B_*(0), \\ u = 0 & \text{on } \partial B_*(0). \end{cases} \tag{1-4}$$

We denote by $PU_{x,\Lambda}$, the projection of $U_{x,\Lambda}$, the solution of the problem

$$\begin{cases} \Delta PU_{x,\Lambda} = \Delta U_{x,\Lambda} & \text{in } B_*(0), \\ PU_{x,\Lambda} = 0 & \text{on } \partial B_*(0). \end{cases} \tag{1-5}$$

Set $y = (y', y'')$, $y'' \in \mathbb{R}^{N-2}$. Define

$$\mathcal{H}_s = \Big\{ u : u \in H_0^1(B_*(0)), \ u \text{ is even in } y_h, \ h = 2, 3, \dots, N, \\ u(r\cos\theta, r\sin\theta, y'') = u\Big(r\cos\Big(\theta + \frac{2\pi j}{k}\Big), r\sin\Big(\theta + \frac{2\pi j}{k}\Big), y''\Big) \Big\}.$$

Let

$$x_j = \left(r\cos\frac{2(j-1)\pi}{k}, r\sin\frac{2(j-1)\pi}{k}, 0 \right), \quad j = 1, \dots, k,$$

where 0 is the zero vector in $\mathbb{R}^{N-2}$, and let

$$W_{r,\Lambda}(y) = \sum_{j=1}^{k} PU_{x_j,\Lambda}.$$

In what follows, we always assume that

$$r \in [\varepsilon^{-1/(N-2)}(1 - r_0\varepsilon^{1/(N-1)}), \varepsilon^{-1/(N-2)}(1 - r_1\varepsilon^{1/(N-1)})]$$

for some constants $r_1 > r_0 > 0$, and that

$$L_0 \le \Lambda \le L_1$$

for some constants $L_1 > L_0 > 0$.

We will prove Theorem 1.1 by verifying the following result.

Theorem 1.3. *Under the same assumptions as Theorem 1.1, there exists $\varepsilon_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0)$, problem (1-4) has a solution u_ε of the form*

$$u_\varepsilon = W_{r_\varepsilon,\Lambda_\varepsilon} + \phi_\varepsilon,$$

where $\phi_\varepsilon \in \mathcal{H}_s$, $\|\phi_\varepsilon\|_{L^\infty} \to 0$ as $\varepsilon \to 0^+$, $L_0 \le \Lambda_\varepsilon \le L_1$ and

$$r_\varepsilon \in [\varepsilon^{-1/(N-2)}(1 - r_0\varepsilon^{1/(N-1)}), \varepsilon^{-1/(N-2)}(1 - r_1\varepsilon^{1/(N-1)})].$$

Remark 1.4. In our result, the number of peaks k of the solution u_ε varies with the parameter ε at the order $\varepsilon^{-1/(N-1)}$ when $\varepsilon \to 0^+$. This is a new phenomenon for the Hénon equation and is in contrast to the subcritical or critical case. For example, in [Peng 2006], where $\varepsilon < 0$, it was proved that for any prescribed integer $k > 0$, there exists $\varepsilon_0 > 0$ such that for any $\varepsilon \in (-\varepsilon_0, 0)$, problem (1-4) has a solution which has exactly k peaks.

Remark 1.5. The results of this paper can be considered as a perturbation of those in [Wei and Yan 2013]. In fact the number of bubbles k can be taken to be

$$k = [\delta^{-1/(N-2)}]$$

for any $|\varepsilon| < \delta \ll 1$. When $\varepsilon = 0$, we recover Wei and Yan's result.

We use a reduction argument to prove Theorem 1.3. More precisely, we follow the method in [Wei and Yan 2010b; 2013] to construct peak solutions for problem (1-4). In those papers, where no parameter appears in the considered problem, Wei and Yan used k, the number of peaks of the solutions, as the parameter to construct infinitely many positive peak solutions. This idea is very novel and effective for obtaining infinitely many solutions to several types of problems; see [Peng and Wang 2013; Wei and Yan 2010a; 2011]. Unlike the situation in [Wei and Yan 2010b; 2013], here we deal with the supercritical case; we cannot use the variational argument. Instead, we will use the Fredholm theory of compact operators in a suitable Banach space and will employ a direct technique to eliminate the Lagrange multipliers caused from the reduction procedure. Another aspect that differs from [Wei and Yan 2010b; 2013] is that, as we mentioned before, in our proof, we use ε as the parameter in the construction of peak solutions, but in this paper the number of peaks depends on the parameter ε. As a final remark, we point out that for $\alpha = 0$, del Pino, Felmer and Musso [2003] have constructed two-peaked solutions for problem (1-1) in a special domain. Hence, we believe that the effect of the weight $|y|^\alpha$ on the existence of solutions is something like that of the domain.

This paper has the following structure. In Section 2, we carry out the finite-dimensional reduction procedure. The main results will be proved in Section 3. We put the energy expansion and some basic estimates used in Sections 2 and 3 in Appendices A and B.

2. Finite-dimensional reduction

In this section, we perform a finite-dimensional reduction. Let

$$\|u\|_* = \sup_{y\in B_*(0)} \left(\sum_{i=1}^{k} \frac{1}{(1+|y-x_i|)^{\frac{1}{2}(N-2)+\tau}} \right)^{-1} |u(y)| \tag{2-1}$$

and

$$\|v\|_{**} = \sup_{y\in B_*(0)} \left(\sum_{i=1}^{k} \frac{1}{(1+|y-x_i|)^{\frac{1}{2}(N+2)+\tau}} \right)^{-1} |v(y)|, \tag{2-2}$$

where $\tau = (N-2)/(N-1)$. We denote by L^∞_* and L^∞_{**} the function spaces defined on $B_*(0)$ with finite $\|\cdot\|_*$ and $\|\cdot\|_{**}$ norm, respectively.

Let

$$Z_{i,1} = \frac{\partial PU_{x_i,\Lambda}}{\partial r}, \qquad Z_{i,2} = \frac{\partial PU_{x_i,\Lambda}}{\partial \Lambda}.$$

First, we consider the linear problem

$$\begin{cases} -\Delta\phi-(p+\varepsilon)K(\varepsilon^{\frac{1}{N-2}}|y|)W_{r,\Lambda}^{p-1+\varepsilon}\phi = h+\sum_{j=1}^{2} c_j \sum_{i=1}^{k} U_{x_i,\Lambda}^{p-1} Z_{i,j} \text{ in } B_*(0), \\ \phi\in\mathcal{H}_s, \qquad \left\langle \sum_{i=1}^{k} U_{x_i,\Lambda}^{p-1} Z_{i,l}, \phi \right\rangle = 0, \quad l=1,2, \end{cases} \tag{2-3}$$

for some numbers c_i, where

$$\langle u, v\rangle = \int_{B_*(0)} uv.$$

Lemma 2.1. *Assume there is a sequence $\varepsilon = \varepsilon_n \to 0$ such that ϕ_ε solves (2-3) for $h = h_\varepsilon$. If $\|h_\varepsilon\|_{**}$ goes to zero as ε goes to zero, so does $\|\phi_\varepsilon\|_*$.*

Proof. The proof of this lemma is very similar to the proof of Lemma 2.1 in [Wei and Yan 2013].

We argue by contradiction. Suppose that there are $\varepsilon\to 0$, $h = h_\varepsilon$, $\Lambda_\varepsilon \in [L_0, L_1]$ and $r_\varepsilon \in [\varepsilon^{-1/(N-2)}(1-r_0\varepsilon^{1/(N-1)}), \varepsilon^{-1/(N-2)}(1-r_1\varepsilon^{1/(N-1)})]$ such that ϕ_ε solves (2-3) for $h = h_\varepsilon$, $\Lambda = \Lambda_\varepsilon$, $r = r_\varepsilon$ with $\|h_\varepsilon\|_{**} \to 0$ and $\|\phi_\varepsilon\|_* \ge c > 0$. Without loss of generality, we may assume that $\|\phi_\varepsilon\|_* = 1$.

Now rewrite (2-3) in the following integral form:

$$\phi_\varepsilon(y) = (p+\varepsilon)\int_{B_*(0)} G_\varepsilon(y,z)K(\varepsilon^{-1/(N-2)}|z|)W_{r,\Lambda}^{p-1+\varepsilon}(z)\phi_\varepsilon(z)\,dz + \int_{B_*(0)} G_\varepsilon(y,z)\Big(h_\varepsilon(z)+\sum_{j=1}^{2} c_j\sum_{i=1}^{k} Z_{i,j}(z)U_{x_i,\Lambda}^{p-1}(z)\Big)\,dz.$$

By Lemma B.3, we find

$$\begin{aligned}\left|(p+\varepsilon)\int_{B_*(0)} G_\varepsilon(y,z)K(\varepsilon^{-1/(N-2)}|z|)W_{r,\Lambda}^{p-1+\varepsilon}(z)\phi_\varepsilon(z)\,dz\right| \\ &\le (p+\varepsilon)\int_{B_*(0)} \frac{1}{|y-z|^{N-2}}K(\varepsilon^{-1/(N-2)}|z|)W_{r,\Lambda}^{p-1+\varepsilon}(z)\,|\phi_\varepsilon(z)|\,dz \\ &\le C\|\phi_\varepsilon\|_*\int_{B_*(0)} \frac{1}{|y-z|^{N-2}}W_{r,\Lambda}^{p-1+\varepsilon}\sum_{j=1}^k \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}\,dz \\ &\le C\|\phi_\varepsilon\|_*\sum_{j=1}^k \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau+\vartheta}}.\end{aligned}$$

It follows from Lemma B.2 that

$$\begin{aligned}\left|\int_{B_*(0)} G_\varepsilon(y,z)h_\varepsilon(z)\,dz\right| &\le \int_{B_*(0)} \frac{1}{|y-z|^{N-2}}|h_\varepsilon(z)|\,dz \\ &\le C\|h_\varepsilon\|_{**}\sum_{j=1}^k \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}\end{aligned}$$

and

$$\left|\int_{B_*(0)} G_\varepsilon(y,z)\sum_{j=1}^k Z_{i,l}(z)U_{x_i,\Lambda}^{p-1}(z)\,dz\right| \le C\sum_{j=1}^k \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}.$$

Next, we estimate c_ℓ, $\ell=1,2$. Multiplying (2-3) by $Z_{1,t}$ and integrating, we obtain that c_ℓ satisfies

$$\begin{aligned}\text{(2-4)}\quad \sum_{\ell=1}^2\sum_{i=1}^k \langle U_{x_i,\Lambda}^{p-1}Z_{i,\ell}, Z_{1,t}\rangle c_\ell \\ = \big\langle -\Delta\phi_\varepsilon-(p+\varepsilon)K(\varepsilon^{-1/(N-2)}|y|)W_{r,\Lambda}^{p-1+\varepsilon}\phi_\varepsilon, Z_{1,t}\big\rangle - \langle h_\varepsilon, Z_{1,t}\rangle.\end{aligned}$$

It follows from Lemma B.1 that

$$\begin{aligned}|\langle h_\varepsilon, Z_{1,\ell}\rangle| &\le C\|h_\varepsilon\|_{**}\int_{\mathbb{R}^N} \frac{1}{(1+|z-x_1|)^{N-2}}\sum_{j=1}^k \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N+2)+\tau}}\,dz \\ &\le C\|h_\varepsilon\|_{**}.\end{aligned}$$

On the other hand, using Lemma B.3, we obtain

$$\begin{aligned}&\big\langle -\Delta\phi_\varepsilon-(p+\varepsilon)K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p-1+\varepsilon}\phi_\varepsilon, Z_{1,\ell}\big\rangle \\ &\quad = \big\langle -\Delta Z_{1,\ell}-(p+\varepsilon)K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p-1+\varepsilon}Z_{1,\ell}, \phi_\varepsilon\big\rangle\end{aligned}$$

$$\begin{aligned}
&= \big\langle -\Delta Z_{1,\ell} - pK(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p-1} Z_{1,\ell}, \phi_\varepsilon \big\rangle \\
&\qquad\qquad + p\big\langle K(\varepsilon^{1/(N-2)}|y|)(W_{r,\Lambda}^{p-1} - W_{r,\Lambda}^{p-1+\varepsilon}) Z_{1,\ell}, \phi_\varepsilon \big\rangle \\
&\qquad\qquad\qquad\qquad - \varepsilon\big\langle K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p-1+\varepsilon}) Z_{1,\ell}, \phi_\varepsilon \big\rangle \\
&= o(\|\phi_\varepsilon\|_*).
\end{aligned}$$

However, there is a constant $c' > 0$ such that

$$\sum_{i=1}^{k} \langle U_{x_i,\Lambda}^{p-1} Z_{i,t}, Z_{1,\ell}\rangle = (c' + o(1))\delta_{t\ell}.$$

Hence we find from (2-4) that

$$c_\ell = o(\|\phi_\varepsilon\|_*) + O(\|h_\varepsilon\|_{**}).$$

Therefore,

$$\|\phi_\varepsilon\|_* \le o(1) + \|h_\varepsilon\|_{**} + \frac{\displaystyle\sum_{j=1}^{k} \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau+\vartheta}}}{\displaystyle\sum_{j=1}^{k} \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}}. \tag{2-5}$$

Noting that $\|\phi_\varepsilon\|_* = 1$, we obtain from (2-5) that there is $R > 0$ such that

$$\|\phi_\varepsilon(y)\|_{L^\infty(B_R(x_i))} \ge a > 0 \quad \text{for some } i. \tag{2-6}$$

Furthermore, for this particular i, the translated version $\bar{\phi}_\varepsilon(y) = \phi_\varepsilon(y - x_i)$ converges uniformly on any compact set to a solution u of

$$-\Delta u - pU_{0,\Lambda}^{p-1} u = 0 \quad \text{in } \mathbb{R}^N \tag{2-7}$$

for some $\Lambda \in [L_0, L_1]$. Since u is perpendicular to the kernel of (2-7), we have $u \equiv 0$, which contradicts $\|u(y)\|_{L^\infty(B_R(x_i))} \ge a > 0$. □

The following proposition is a direct consequence of combining Proposition 4.1 in [del Pino et al. 2003] with Lemma 2.1.

Proposition 2.2. *There exists $\varepsilon_0 > 0$ and a constant $C > 0$ such that for all $\varepsilon \le \varepsilon_0$ and all $h_\varepsilon \in L_{**}^\infty$, problem (2-3) has a unique solution $\phi_\varepsilon \equiv \mathcal{L}_\varepsilon(h_\varepsilon) \in L_*^\infty$. Moreover,*

$$\|\mathcal{L}_\varepsilon(h_\varepsilon)\|_* \le C\|h_\varepsilon\|_{**}, \qquad |c_\ell| \le C\|h_\varepsilon\|_{**}. \tag{2-8}$$

In order to prove the main theorem, we will prove that problem (1-4) admits a solution of the form $u = W_{r,\Lambda} + \phi$, where $W_{r,\Lambda} = \sum_{j=1}^{k} PU_{x_j,\Lambda}$ and $\phi \in \mathcal{H}_s$ is small and satisfies $\langle U_{x_i,\Lambda}^{p-1} Z_{i,l}, \phi\rangle = 0$, $i = 1, 2, \ldots, k$, $l = 1, 2$.

We consider the perturbation problem

$$\tag{2-9} \begin{cases} -\Delta(W_{r,\Lambda}+\phi) = K(\varepsilon^{\frac{1}{N-2}}|y|)(W_{r,\Lambda}+\phi)^{p+\varepsilon} + \sum\limits_{\ell=1}^{2} c_\ell \sum\limits_{i=1}^{k} U_{x_i,\Lambda}^{p-1} Z_{i,\ell} \text{ in } B_*(0), \\ \phi \in \mathcal{H}_s, \qquad \Big\langle \sum\limits_{i=1}^{k} U_{x_i,\Lambda}^{p-1} Z_{i,\ell}, \phi \Big\rangle = 0, \quad \ell = 1, 2. \end{cases}$$

Proposition 2.3. *There is $\varepsilon_0 > 0$ such that for any $\varepsilon \le \varepsilon_0$, any $\Lambda \in [L_0, L_1]$, and*

$$r \in \big[\varepsilon^{-\frac{1}{N-2}}(1 - r_0\varepsilon^{\frac{1}{N-1}}), \varepsilon^{-\frac{1}{N-2}}(1 - r_1\varepsilon^{\frac{1}{N-1}})\big],$$

problem (2-9) has a unique solution $\phi = \phi_{r,\Lambda}$ satisfying

$$\|\phi\|_* \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}, \qquad |c_\ell| \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)},$$

where $\sigma > 0$ is a small constant.

Let

$$N_\varepsilon(\phi) = K\big(\varepsilon^{\frac{1}{N-2}}|y|\big)\big((W_{r,\Lambda}+\phi)^{p+\varepsilon} - W_{r,\Lambda}^{p+\varepsilon} - (p+\varepsilon)W_{r,\Lambda}^{p-1+\varepsilon}\phi\big),$$

$$l_\varepsilon = K\big(\varepsilon^{\frac{1}{N-2}}|y|\big)W_{r,\Lambda}^{p+\varepsilon} - \sum_{j=1}^{k} U_{x_j,\Lambda}^{p}.$$

Then problem (2-9) can be written as

$$\tag{2-10} \begin{cases} -\Delta\phi - (p+\varepsilon)K(\varepsilon^{\frac{1}{N-2}}|y|)W_{r,\Lambda}^{p-1+\varepsilon}\phi \\ \qquad\qquad = N_\varepsilon(\phi) + l_\varepsilon + \sum\limits_{\ell=1}^{2} c_\ell \sum\limits_{i=1}^{k} U_{x_i,\Lambda}^{p-1} Z_{i,\ell} \quad \text{in } B_*(0), \\ \phi \in \mathcal{H}_s, \qquad \Big\langle \sum\limits_{i=1}^{k} U_{x_i,\Lambda}^{p-1} Z_{i,\ell}, \phi \Big\rangle = 0, \quad \ell = 1, 2. \end{cases}$$

We will use the contraction mapping theorem to prove that problem (2-9) is uniquely solvable under the condition that $\|\phi\|_*$ is small enough. So we need to estimate $N_\varepsilon(\phi)$ and l_ε.

Lemma 2.4. *If $N \ge 4$, then*

$$\|N_\varepsilon(\phi)\|_{**} \le C\|\phi\|_*^{\min\{p+\varepsilon,2\}}.$$

Proof. We have

$$|N_\varepsilon(\phi)| \le \begin{cases} C|\phi|^{p+\varepsilon}, & N \ge 7, \\ C(W_{r,\Lambda}^{p-2+\varepsilon}\phi^2 + |\phi|^{p+\varepsilon}), & N = 4, 5, 6. \end{cases}$$

Firstly, we consider $N \geq 7$. By the Hölder inequality, we have

$$
\begin{aligned}
|N_\varepsilon(\phi)| &\leq C\|\phi\|_*^{p+\varepsilon}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}\right)^{p+\varepsilon}\\
&\leq C\|\phi\|_*^{p+\varepsilon}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}\right)\\
&\qquad\times\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{p+\varepsilon}{p+\varepsilon-1}(\frac{1}{2}(N-2)+\tau)-\frac{1}{p+\varepsilon-1}(\frac{1}{2}(N+2)+\tau)}}\right)^{p+\varepsilon-1}\\
&\leq C\|\phi\|_*^{p+\varepsilon}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}\right).
\end{aligned}
$$

Thus, the result follows.

Suppose that $N = 4, 5, 6$. Using the fact that $N-2 > \frac{1}{2}(N-2)+\tau$, we get

$$
\begin{aligned}
|N_\varepsilon(\phi)| &\leq C\|\phi\|_*^{2}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}\right)^{p-2+\varepsilon}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}\right)^{2}\\
&\qquad+C\|\phi\|_*^{p+\varepsilon}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}\right)\\
&\leq C\|\phi\|_*^{2}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}\right)^{p+\varepsilon}\\
&\qquad+C\|\phi\|_*^{p+\varepsilon}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}\right)\\
&\leq C\|\phi\|_*^{2}\left(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}\right).
\end{aligned}
$$

So we have proved that for $N \geq 4$,

$$\|N_\varepsilon(\phi)\|_{**} \leq C\|\phi\|_*^{\min\{p+\varepsilon,2\}}. \qquad \square$$

Lemma 2.5. *Assume that* $r \in \left[\varepsilon^{-\frac{1}{N-2}}(1-r_0\varepsilon^{\frac{1}{N-1}}), \varepsilon^{-\frac{1}{N-2}}(1-r_1\varepsilon^{\frac{1}{N-1}})\right]$. *If* $N \geq 4$, *then*

$$\|l_\varepsilon\|_{**} \leq C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}.$$

Proof. Define

$$\Omega_j = \left\{y : y = (y', y'') \in B_{\varepsilon^{-1/(N-2)}}(0), \left\langle\frac{y'}{|y'|}, \frac{x_j}{|x_j|}\right\rangle \geq \cos\frac{\pi}{k}\right\}.$$

We have

$$
\begin{aligned}
l_\varepsilon &= K(\varepsilon^{1/(N-2)}|y|)(W_{r,\Lambda}^{p+\varepsilon}-W_{r,\Lambda}^{p})+K(\varepsilon^{1/(N-2)}|y|)\Big(W_{r,\Lambda}^{p}-\sum_{j=1}^{k}PU_{x_j,\Lambda}^{p}\Big)\\
&\quad+K(\varepsilon^{1/(N-2)}|y|)\Big(\sum_{j=1}^{k}PU_{x_j,\Lambda}^{p}-\sum_{j=1}^{k}U_{x_j,\Lambda}^{p}\Big)+\sum_{j=1}^{k}U_{x_j,\Lambda}^{p}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)\\
&=: J_0+J_1+J_2+J_3.
\end{aligned}
$$

Estimate of J_0.

$$
\begin{aligned}
|J_0| &\le C\varepsilon W_{r,\Lambda}^{p}|\ln W_{r,\Lambda}|\\
&\le C\varepsilon\Big(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}\Big)^{p}\ln\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}\\
&\le C\varepsilon\Big(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac12(N+2)+\tau}}\Big)\Big(\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{N+2}{4}\left(\frac{N-2}{2}-\frac{N-2}{N+2}\tau\right)}}\Big)^{\frac{4}{N-2}}\\
&\le C\varepsilon\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac12(N+2)+\tau}}.
\end{aligned}
$$

Estimate of J_1. From the symmetry, we can assume that $y\in\Omega_1$. Then,

$$|y-x_j|\ge|y-x_1|,\qquad y\in\Omega_1,\ j\ne1.$$

Firstly, we claim

$$\frac{1}{1+|y-x_j|}\le\frac{C}{|x_j-x_1|},\qquad y\in\Omega_1,\ j\ne1. \tag{2-11}$$

In fact, if $|y-x_1|\le\frac12|x_1-x_j|$, then $|y-x_j|\ge\frac12|x_j-x_1|$. If $|y-x_1|\ge\frac12|x_j-x_1|$, then $|y-x_j|\ge|y-x_1|\ge\frac12|x_j-x_1|$.

It's easy to verify that

$$|J_1|\le C\frac{1}{(1+|y-x_1|)^4}\sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}+C\Big(\sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}\Big)^{p}.$$

Using (2-11), taking $1<\varrho\le N-2$, we obtain for $y\in\Omega_1$,

$$\frac{1}{(1+|y-x_1|)^4}\frac{1}{(1+|y-x_j|)^{N-2}}\le C\frac{1}{(1+|y-x_1|)^{N+2-\varrho}}\frac{1}{|x_j-x_1|^{\varrho}},\qquad j\ne1.$$

Take $\varrho > \max\left\{\frac{1}{2}(N-1), 1\right\}$ satisfying $N+2-\varrho \ge \frac{1}{2}(N+2)+\tau$. Then

$$\frac{1}{(1+|y-x_1|)^4}\sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{N-2}} \le \frac{C}{(1+|y-x_1|)^{N+2-\varrho}}(k\varepsilon^{1/(N-2)})^{\varrho}$$
$$= \frac{C}{(1+|y-x_1|)^{N+2-\varrho}}\varepsilon^{\varrho/((N-1)(N-2))} \le \frac{C}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}.$$

By the Hölder inequality, we find

$$\left(\sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}\right)^{p}$$
$$\le \sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}\left(\sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{\frac{N+2}{4}\left(\frac{N-2}{2}-\tau\frac{N-2}{N+2}\right)}}\right)^{\frac{4}{N-2}}.$$

Noticing that $\dfrac{N+2}{N-2}\left(\dfrac{N-2}{2}-\tau\dfrac{N-2}{N+2}\right) > \dfrac{N-1}{2}$ if $N \ge 4$, we deduce that

$$\left(\sum_{j=2}^{k}\frac{1}{(1+|y-x_j|)^{N-2}}\right)^{p}$$
$$\le C\left(\sum_{j=2}^{k}\frac{1}{|x_j-x_1|^{\frac{N+2}{4}\left(\frac{N-2}{2}-\tau\frac{N-2}{N+2}\right)}}\right)^{\frac{4}{N-2}}\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}$$
$$\le C(k\varepsilon^{1/(N-2)})^{\frac{N+2}{N-2}\left(\frac{N-2}{2}-\tau\frac{N-2}{N+2}\right)}\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}$$
$$\le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}\sum_{j=1}^{k}\frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}.$$

Hence, we conclude that if $N \ge 4$,

$$\|J_1\|_{**} \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}.$$

Estimate of J_2. Let $H(y,x)$ be the regular part of the Green function for $-\Delta$ in $B_1(0)$ with the zero boundary condition. Let $\bar{x}_j^*$ be the reflection point of $\bar{x}_j$ with respect to $\partial B_1(0)$. Then

$$\varepsilon H(\bar{y},\bar{x}_j) = \frac{C\varepsilon}{|\bar{y}-\bar{x}_j^*|^{N-2}} \le \frac{C}{(1+|y-x_j|)^{N-2}}.$$

By direct calculation, we have

$$
\begin{aligned}
|J_2| &\le C\sum_{j=1}^{k}\frac{C\varepsilon}{(1+|y-x_j|)^4}H(\bar{y},\bar{x}_j)\\
&\le \sum_{j=1}^{k}\frac{C}{(1+|y-x_j|)^{4+(1-\gamma)(N-2)}}(\varepsilon H(\bar{y},\bar{x}_j))^{\gamma}\\
&\le C\varepsilon^{\gamma/(N-1)}\sum_{j=1}^{k}\frac{C}{(1+|y-x_j|)^{4+(1-\gamma)(N-2)}}\\
&\le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}\sum_{j=1}^{k}\frac{C}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}},
\end{aligned}
$$

where $\gamma>0$ satisfies $\gamma(N-2)/(N-1)>\frac{1}{2}$ and $4+(1-\gamma)(N-2)\ge\frac{1}{2}(N+2)+\tau$.

Estimate of J_3. For $y\in\Omega_1$ and $j>1$, using (2-11), we find

$$
U^p_{x_j,\Lambda}\le\frac{C}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}\frac{1}{|x_j-x_1|^{\frac{1}{2}(N+2)-\tau}}.
$$

Thus, we have

$$
\begin{aligned}
\left|\sum_{j=2}^{k}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)U^p_{x_j,\Lambda}\right| &\le \frac{C}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}\sum_{j=2}^{k}\frac{1}{|x_j-x_1|^{\frac{1}{2}(N+2)-\tau}}\\
&\le \frac{C}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}(k\varepsilon^{1/(N-2)})^{\frac{1}{2}(N+2)-\tau}\\
&\le \frac{C}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}.
\end{aligned}
$$

If $y\in\Omega_1$ and $\big||y|-\varepsilon^{-1/(N-2)}\big|\ge\delta\varepsilon^{-1/(N-2)}$, where $\delta>0$ is a fixed constant, then

$$
\big||y|-|x_1|\big|\ge\big||y|-\varepsilon^{-1/(N-2)}\big|-\big||x_1|-\varepsilon^{-1/(N-2)}\big|\ge\tfrac{1}{2}\delta\varepsilon^{-1/(N-2)}.
$$

So, we obtain

$$
\left|U^p_{x_1,\Lambda}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)\right|\le\frac{C}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}\varepsilon^{(\frac{1}{2}(N+2)-\tau)/(N-2)}.
$$

If $y \in \Omega_1$ and $\big||y| - \varepsilon^{-1/(N-2)}\big| \le \delta \varepsilon^{-1/(N-2)}$, then

$$\begin{aligned}\big|K(\varepsilon^{1/(N-2)}|y|) - 1\big| &\le C\big|\varepsilon^{1/(N-2)}|y| - 1\big| \\ &\le C\varepsilon^{1/(N-2)}\big(\big||y| - |x_1|\big| + \big||x_1| - \varepsilon^{-1/(N-2)}\big|\big) \\ &\le C\varepsilon^{1/(N-2)}\big||y| - |x_1|\big| + C\varepsilon^{1/(N-1)} \\ &\le C\varepsilon^{1/(N-2)}\big||y| - |x_1|\big| + C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)},\end{aligned}$$

and

$$\big||y| - |x_1|\big| \le |y|\big|-\varepsilon^{-1/(N-2)}\big| + \big||x_1| - \varepsilon^{-1/(N-2)}\big| \le 2\delta\varepsilon^{-1/(N-2)}.$$

Since

$$\begin{aligned}\frac{\varepsilon^{1/(N-2)}\big||y| - |x_1|\big|}{(1+|y-x_1|)^{N+2}} &\le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}\frac{\big||y| - |x_1|\big|^{\frac{1}{2}+\sigma}}{(1+|y-x_1|)^{N+2}} \\ &\le \frac{C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}}{(1+|y-x_1|)^{N+\frac{3}{2}-\sigma}} \le \frac{C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}},\end{aligned}$$

we get

$$\big|U^p_{x_1,\Lambda}\big(K(\varepsilon^{1/(N-2)}|y|) - 1\big)\big| \le \frac{C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}}{(1+|y-x_1|)^{\frac{1}{2}(N+2)+\tau}}.$$

As a result, we deduce

$$\|J_3\|_{**} \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}. \qquad \square$$

Proof of Proposition 2.3. Recall that

$$k = [\varepsilon^{-1/(N-1)}], \quad N \ge 4.$$

Let

$$E = \Big\{u \in \mathcal{H}_s \cap L^\infty_* : \|u\|_* \le \varepsilon^{1/(2(N-2))} \text{ and } \int_{B_*(0)} U^{p-1}_{x_i,\Lambda} Z_{i,\ell} u = 0,\ i = 1, \ldots, k,\ \ell = 1, 2\Big\}.$$

Then, (2-10) is equivalent to

$$\phi = A_\varepsilon(\phi) =: \mathcal{L}_\varepsilon(N_\varepsilon(\phi)) + \mathcal{L}_\varepsilon(l_\varepsilon),$$

where $\mathcal{L}_\varepsilon$ is defined in Proposition 2.2. We will prove that A_ε is a contraction map from E to E. First, $A_\varepsilon(E) \subset E$ because

$$\begin{aligned}\|A_\varepsilon(\phi)\|_* &\le C\|N_\varepsilon(\phi)\|_{**} + C\|l_\varepsilon\|_{**} \\ &\le C\|\phi\|_*^{\min\{p+\varepsilon,2\}} + C\|l_\varepsilon\|_{**} \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)} \le \varepsilon^{1/(2(N-2))}.\end{aligned}$$

Next we write

$$\|A_\varepsilon(\phi_1)-A_\varepsilon(\phi_2)\|_* = \|\mathcal{L}_\varepsilon(N_\varepsilon(\phi_1))-\mathcal{L}_\varepsilon(N_\varepsilon(\phi_2))\|_* \le C\|N_\varepsilon(\phi_1)-N_\varepsilon(\phi_2)\|_{**}.$$

If $N \ge 7$, then

$$|N_\varepsilon'(t)| \le C|t|^{p-1+\varepsilon}.$$

Thus, we have

$$\begin{aligned}
&|N_\varepsilon(\phi_1)-N_\varepsilon(\phi_2)|\\
&\quad\le C(|\phi_1|^{p-1+\varepsilon}+|\phi_2|^{p-1+\varepsilon})|\phi_1-\phi_2|\\
&\quad\le C(\|\phi_1\|_*^{p-1+\varepsilon}+\|\phi_2\|_*^{p-1+\varepsilon})\|\phi_1-\phi_2\|_*\left(\sum_{j=1}^k \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}\right)^{p+\varepsilon}\\
&\quad\le C(\|\phi_1\|_*^{p-1+\varepsilon}+\|\phi_2\|_*^{p-1+\varepsilon})\|\phi_1-\phi_2\|_*\sum_{j=1}^k \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}.
\end{aligned}$$

As a consequence,

$$\begin{aligned}
\|A_\varepsilon(\phi_1)-A_\varepsilon(\phi_2)\|_* &\le C\|N_\varepsilon(\phi_1)-N_\varepsilon(\phi_2)\|_{**}\\
&\le C(\|\phi_1\|_*^{p-1+\varepsilon}+\|\phi_2\|_*^{p-1+\varepsilon})\|\phi_1-\phi_2\|_* \le \tfrac{1}{2}\|\phi_1-\phi_2\|_*.
\end{aligned}$$

For $N = 4, 5, 6$,

$$|N_\varepsilon'(t)| \le CW_{r,\Lambda}^{p-2+\varepsilon}|t|+C|t|^{p-1+\varepsilon}.$$

So we have

$$\begin{aligned}
&|N_\varepsilon(\phi_1)-N_\varepsilon(\phi_2)|\\
&\quad\le C(|\phi_1|^{p-1+\varepsilon}+|\phi_2|^{p-1+\varepsilon})|\phi_1-\phi_2|+C(|\phi_1|+|\phi_2|)W_{r,\Lambda}^{p-2+\varepsilon}|\phi_1-\phi_2|\\
&\quad\le C(\|\phi_1\|_*^{p-1+\varepsilon}+\|\phi_2\|_*^{p-1+\varepsilon})\|\phi_1-\phi_2\|_*\left(\sum_{j=1}^k \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}\right)^{p+\varepsilon}\\
&\qquad+C(\|\phi_1\|_*+\|\phi_2\|_*)\|\phi_1-\phi_2\|_* W_{r,\Lambda}^{p-2+\varepsilon}\left(\sum_{j=1}^k \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N-2)+\tau}}\right)^{2}\\
&\quad\le C(\|\phi_1\|_*+\|\phi_2\|_*)\|\phi_1-\phi_2\|_*\sum_{j=1}^k \frac{1}{(1+|y-x_j|)^{\frac{1}{2}(N+2)+\tau}}.
\end{aligned}$$

In either case, we see that A_ε is a contraction map. By the contraction mapping theorem, there is a unique $\phi \in E$ such that

$$\phi = A_\varepsilon(\phi).$$

Moreover, it follows from Proposition 2.2 that

$$\|\phi\|_* \le C\|l_\varepsilon\|_{**} + C\|N_\varepsilon(\phi)\|_{**} \le C\|l_\varepsilon\|_{**} + C\|\phi\|_*^{\min\{p+\varepsilon,2\}},$$

which implies

$$\|\phi\|_* \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}, \quad |c_\ell| \le C\varepsilon^{(\frac{1}{2}+\sigma)/(N-2)}. \qquad \square$$

3. Proof of the main results

In this section, we will choose (r, Λ) such that $W_{r,\Lambda} + \phi_{r,\Lambda}$ is a solution of (1-4), where $\phi_{r,\Lambda}$ is the map obtained in Proposition 2.3.

Lemma 3.1. *If* (r, Λ) *satisfies*

(3-1)
$$\int_{B_*(0)} \left(\nabla(W_{r,\Lambda} + \phi_{r,\Lambda}) \nabla \frac{\partial W_{r,\Lambda}}{\partial r} - K(\varepsilon^{1/(N-2)}|y|)(W_{r,\Lambda} + \phi_{r,\Lambda})^{p+\varepsilon} \frac{\partial W_{r,\Lambda}}{\partial r} \right) = 0,$$

(3-2)
$$\int_{B_*(0)} \left(\nabla(W_{r,\Lambda} + \phi_{r,\Lambda}) \nabla \frac{\partial W_{r,\Lambda}}{\partial \Lambda} - K(\varepsilon^{1/(N-2)}|y|)(W_{r,\Lambda} + \phi_{r,\Lambda})^{p+\varepsilon} \frac{\partial W_{r,\Lambda}}{\partial \Lambda} \right) = 0,$$

then $W_{r,\Lambda} + \phi_{r,\Lambda}$ *is a solution of* (1-4).

Proof. It follows from Proposition 2.3 that if (3-1) and (3-2) hold, then by symmetry,

$$c_1 \left\langle U_{x_1,\Lambda}^{p-1} \frac{\partial PU_{x_1,\Lambda}}{\partial r}, \frac{\partial W_{r,\Lambda}}{\partial r} \right\rangle = 0 = c_2 \left\langle U_{x_1,\Lambda}^{p-1} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda}, \frac{\partial W_{r,\Lambda}}{\partial \Lambda} \right\rangle, \tag{3-3}$$

which implies that $c_1 = c_2 = 0$. Hence $W_{r,\Lambda} + \phi_{r,\Lambda}$ is a solution of (1-4). $\square$

In the rest of this section, we need to solve (3-1) and (3-2).

Proposition 3.2. *Equations* (3-1) *and* (3-2) *are equivalent to*

$$-\frac{\varepsilon H(\bar{x}_1, \bar{x}_1)}{\Lambda^{N-1}} + \sum_{i=2}^{k} \frac{\varepsilon G(\bar{x}_i, \bar{x}_1)}{\Lambda^{N-1}} + O(\varepsilon^{(1+\sigma)/(N-2)}) = 0 \tag{3-4}$$

and

$$\frac{B_2\varepsilon}{\Lambda^{N-2}} \frac{\partial H(\bar{x}_1, \bar{x}_1)}{\partial d} + B_3 K'(1) + \sum_{i=2}^{k} \frac{B_2\varepsilon}{\Lambda^{N-2}} \frac{\partial G(\bar{x}_i, \bar{x}_1)}{\partial d} + O(\varepsilon^{\sigma/(N-2)}) = 0, \tag{3-5}$$

respectively, where $d = 1 - \varepsilon^{1/(N-2)} r$, B_1, B_2 *and* B_3 *are the same positive constants as in Proposition A.1 and* $\sigma > 0$ *is a small constant.*

Proof. Here we prove only the first one. The second can be proved similarly by noting that $\partial/\partial d = -\varepsilon^{-1/(N-2)} \partial/\partial r$.

First, we see that

$$\int_{B_*(0)} \nabla(W_{r,\Lambda}+\phi_{r,\Lambda})\nabla\frac{\partial W_{r,\Lambda}}{\partial\Lambda}=\int_{B_*(0)}\nabla W_{r,\Lambda}\nabla\frac{\partial W_{r,\Lambda}}{\partial\Lambda},$$

and

$$\begin{aligned}\int_{B_*(0)} &K(\varepsilon^{1/(N-2)}|y|)(W_{r,\Lambda}+\phi_{r,\Lambda})^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\\ &=\int_{B_*(0)}K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\\ &\qquad+(p+\varepsilon)\int_{B_*(0)}K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p-1+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\\ &\qquad\qquad+O\Big(\int_{B_*(0)}W_{r,\Lambda}^{p-1+\varepsilon}|\phi_{r,\Lambda}|^2\Big).\end{aligned}$$

On the other hand, noticing that $\phi_{r,\Lambda}\in E$, we have

$$\begin{aligned}\int_{B_*(0)} &K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p-1+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\\ &=\int_{B_*(0)}K(\varepsilon^{1/(N-2)}|y|)\Big(W_{r,\Lambda}^{p-1+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}-\sum_{j=1}^{k}U_{x_j,\Lambda}^{p-1}\frac{\partial U_{x_j,\Lambda}}{\partial\Lambda}\Big)\phi_{r,\Lambda}\\ &\qquad+\sum_{j=1}^{k}\int_{B_*(0)}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)U_{x_j,\Lambda}^{p-1}\frac{\partial U_{x_j,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\\ &=k\int_{\Omega_1}K(\varepsilon^{1/(N-2)}|y|)\Big(W_{r,\Lambda}^{p-1+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}-\sum_{j=1}^{k}U_{x_j,\Lambda}^{p-1}\frac{\partial U_{x_j,\Lambda}}{\partial\Lambda}\Big)\phi_{r,\Lambda}\\ &\qquad+k\int_{\Omega_1}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)U_{x_1,\Lambda}^{p-1}\frac{\partial U_{x_1,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda},\end{aligned}$$

$$\begin{aligned}\Big|\int_{\Omega_1} &K(\varepsilon^{1/(N-2)}|y|)\Big(W_{r,\Lambda}^{p-1+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}-\sum_{j=1}^{k}U_{x_j,\Lambda}^{p-1}\frac{\partial U_{x_j,\Lambda}}{\partial\Lambda}\Big)\phi_{r,\Lambda}\Big|\\ &\le C\int_{\Omega_1}\Big(U_{x_1,\Lambda}^{p-1}(U_{x_1,\Lambda}-PU_{x_1,\Lambda})+U_{x_1,\Lambda}^{p-1}\sum_{j=2}^{k}U_{x_j,\Lambda}+\sum_{j=2}^{k}U_{x_j,\Lambda}^{p}\Big)|\phi_{r,\Lambda}|\\ &\qquad+O\Big(\varepsilon\int_{\Omega_1}W_{r,\Lambda}^{p-1}\ln W_{r,\Lambda}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\Big)\\ &\le C\varepsilon^{1/(N-2)(1+\sigma)},\end{aligned}$$

and

$$\begin{aligned}\left|\int_{\Omega_1}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)U_{x_1,\Lambda}^{p-1}\frac{\partial U_{x_1,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\right| &\le \left|\int_{\left||y|-\varepsilon^{-1/(N-2)}\right|\le\varepsilon^{-1/(2(N-2))}}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)U_{x_1,\Lambda}^{p-1}\frac{\partial U_{x_1,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\right|\\ &\quad+\left|\int_{\left||y|-\varepsilon^{-1/(N-2)}\right|\ge\varepsilon^{-1/(2(N-2))}}\big(K(\varepsilon^{1/(N-2)}|y|)-1\big)U_{x_1,\Lambda}^{p-1}\frac{\partial U_{x_1,\Lambda}}{\partial\Lambda}\phi_{r,\Lambda}\right|\\ &\le C\varepsilon^{1/(N-2)(1+\sigma)}.\end{aligned}$$

So, we have proved

$$\begin{aligned}&\int_{B_*(0)}\left(\nabla(W_{r,\Lambda}+\phi_{r,\Lambda})\nabla\frac{\partial W_{r,\Lambda}}{\partial\Lambda}-K(\varepsilon^{1/(N-2)}|y|)(W_{r,\Lambda}+\phi_{r,\Lambda})^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\right)\\ &\quad=\int_{B_*(0)}\left(\nabla W_{r,\Lambda}\nabla\frac{\partial W_{r,\Lambda}}{\partial\Lambda}-K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\right)+O(k\varepsilon^{(1+\sigma)/(N-2)}),\end{aligned}$$

and the result follows from Proposition A.1. □

Proof of Theorem 1.3. Note (see [Wei and Yan 2013]) that

$$H(\bar{x}_1,\bar{x}_1)=\frac{1}{2^{N-2}d^{N-2}}(1+O(d))$$

and

$$\frac{a_0}{j^N}+O\left(\frac{d}{j^{N-2}}\right)\le\frac{1}{k^{N-2}}G(\bar{x}_j,\bar{x}_1)\le\frac{a_1}{j^N}+O\left(\frac{d}{j^{N-2}}\right),$$

where $a_1\ge a_0>0$. Hence, we find that there is a constant $B_4>0$ such that

$$\sum_{j=2}^{k}G(\bar{x}_j,\bar{x}_1)=k^{N-2}\left(\frac{B_4}{|\bar{x}_1|^{N-2}}+O\left(\frac{1}{k^{N-1}}\right)+O(d)\right)=B_4k^{N-2}+O(k^{N-2}d).$$

Consequently, (3-4) and (3-5) are equivalent to

$$-\frac{A_1\varepsilon}{\Lambda^{N-1}d^{N-2}}+\frac{A_2k^{N-2}\varepsilon}{\Lambda^{N-1}}+O(\varepsilon^{(1+\sigma)/(N-2)})=0 \tag{3-6}$$

and

$$-\frac{A_3\varepsilon}{\Lambda^{N-2}d^{N-1}}+A_4+O(\varepsilon^{\sigma/(N-2)})=0, \tag{3-7}$$

respectively, for positive constants A_i, $i=1,2,3,4$. Recall that $d=1-\varepsilon^{1/(N-2)}r$. Define $\eta=dk$. Thus, (3-6) and (3-7) read

$$-\frac{A_1}{\Lambda^{N-1}\eta^{N-2}}+\frac{A_2}{\Lambda^{N-1}}+O(\varepsilon^{\sigma/(N-2)})=0 \tag{3-8}$$

and

$$-\frac{A_3}{\Lambda^{N-2}\eta^{N-1}}+A_4+O(\varepsilon^{\sigma/(N-2)})=0. \tag{3-9}$$

Let

$$f_1(\eta,\Lambda)=-\frac{A_1}{\Lambda^{N-1}\eta^{N-2}}+\frac{A_2}{\Lambda^{N-1}}$$

and

$$f_2(\eta,\Lambda)=-\frac{A_3}{\Lambda^{N-2}\eta^{N-1}}+A_4.$$

It is easy to check that $f_1=0$ and $f_2=0$ have a unique solution

$$\eta_0=\left(\frac{A_1}{A_2}\right)^{\frac{1}{N-2}},\qquad \Lambda_0=\left(\frac{A_3}{A_4\eta_0^{N-1}}\right)^{\frac{1}{N-2}}.$$

On the other hand, we have

$$\frac{\partial f_1(\eta_0,\Lambda_0)}{\partial\Lambda}=0,\qquad \frac{\partial f_2(\eta_0,\Lambda_0)}{\partial\eta}>0,$$

and

$$\frac{\partial f_1(\eta_0,\Lambda_0)}{\partial\eta}>0,\qquad \frac{\partial f_2(\eta_0,\Lambda_0)}{\partial\Lambda}>0.$$

Hence the linear operator of $f_1=0$ and $f_2=0$ at (η_0,Λ_0) is invertible. Therefore, (3-8) and (3-9) have a solution near (η_0,Λ_0). □

Appendix A: Energy expansion

Here and in Appendix B, we assume that

$$x_j=\left(r\cos\frac{2(j-1)\pi}{k},r\sin\frac{2(j-1)\pi}{k},0\right),\qquad j=1,\dots,k,$$

where $r\in\left[\varepsilon^{-\frac{1}{N-2}}(1-r_0\varepsilon^{\frac{1}{N-1}}),\varepsilon^{-\frac{1}{N-2}}(1-r_1\varepsilon^{\frac{1}{N-1}})\right]$ and 0 is the zero vector in $\mathbb{R}^{N-2}$.

Let

$$\bar{x}_j=\varepsilon^{\frac{1}{N-2}}x_j.$$

Let $G(x,y)$ be the Green function of $-\Delta$ in $B_1(0)$ with the Dirichlet boundary and let $H(x,y)$ be the regular part of the Green function.

Recall that

$$k=\left[\varepsilon^{-\frac{1}{N-1}}\right]$$

and

$$W_{r,\Lambda}(y)=\sum_{j=1}^{k}PU_{x_j,\Lambda}(y),$$

where $PU_{x,\Lambda}$ is the solution of (1-5). Moreover,

$$\phi_{x_j,\Lambda}(y) = U_{x_j,\Lambda}(y) - PU_{x_j,\Lambda}(y) = \frac{\varepsilon H(\bar{x}_j, \bar{y})}{\Lambda^{\frac{1}{2}(N-2)}} + O\left(\frac{\varepsilon^{N/(N-2)}}{d^N}\right), \tag{A-1}$$

where $d = 1 - |\bar{x}_j| = 1 - \varepsilon^{1/(N-2)}|x_j|$.

Proposition A.1. *We have*

$$\int_{B_*(0)}\left(\nabla W_{r,\Lambda}\nabla\frac{\partial W_{r,\Lambda}}{\partial\Lambda} - K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\right)$$
$$= kB_1\left(-\frac{\varepsilon H(\bar{x}_1, \bar{x}_1)}{\Lambda^{N-1}} + \sum_{i=2}^{k}\frac{\varepsilon G(\bar{x}_i, \bar{x}_1)}{\Lambda^{N-1}} + O(\varepsilon^{(1+\sigma)/(N-2)})\right),$$

and

$$\int_{B_*(0)}\left(\nabla W_{r,\Lambda}\nabla\frac{\partial W_{r,\Lambda}}{\partial r} - K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial r}\right)$$
$$= k\left(\frac{B_2\varepsilon}{\Lambda^{N-2}}\frac{\partial H(\bar{x}_1, \bar{x}_1)}{\partial r} - B_3K'(1)\varepsilon^{1/(N-2)} + \sum_{i=2}^{k}\frac{B_2\varepsilon}{\Lambda^{N-2}}\frac{\partial G(\bar{x}_i, \bar{x}_1)}{\partial r} + O(\varepsilon^{(1+\sigma)/(N-2)})\right),$$

where B_1, B_2 and B_3 are some positive constants.

Proof. The proof is quite standard now. Here we only prove the first equation. The other one can be obtained similarly.

Using symmetry, we find

$$I := \int_{B_*(0)}\left(\nabla W_{r,\Lambda}\nabla\frac{\partial W_{r,\Lambda}}{\partial\Lambda} - K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\right)$$
$$= k\left(p\sum_{i=1}^{k}\int_{B_*(0)} PU_{x_1,\Lambda}^{p-1}\frac{\partial PU_{x_1,\Lambda}}{\partial\Lambda}PU_{x_i,\Lambda} - \int_{\Omega_1}K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda}\right).$$

It is easy to check that for $y \in \Omega_1$,

$$\frac{\partial}{\partial\Lambda}W_{r,\Lambda}^{p+1} = \frac{\partial}{\partial\Lambda}PU_{x_1,\Lambda}^{p+1} + (p+1)\frac{\partial}{\partial\Lambda}\left(PU_{x_1,\Lambda}^{p}\sum_{i=2}^{k}PU_{x_i,\Lambda}\right) + O\left(U_{x_1,\Lambda}^{\frac{1}{2}(p+1)}\left(\sum_{i=2}^{k}U_{x_i,\Lambda}\right)^{\frac{1}{2}(p+1)}\right).$$

Thus, we have

$$
\begin{aligned}
(p+1)\int_{\Omega_1} & K(\varepsilon^{1/(N-2)}|y|)W_{r,\Lambda}^{p+\varepsilon}\frac{\partial W_{r,\Lambda}}{\partial\Lambda} \\
&= \int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|)\frac{\partial W_{r,\Lambda}^{p+1}}{\partial\Lambda} + O\Big(\varepsilon\int_{\Omega_1} W_{r,\Lambda}^{p+1}\ln W_{r,\Lambda}\Big) \\
&= \int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|)\frac{\partial}{\partial\Lambda}PU_{x_1,\Lambda}^{p+1} \\
&\qquad + (p+1)\int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|)\frac{\partial}{\partial\Lambda}\Big(PU_{x_1,\Lambda}^{p}\sum_{i=2}^{k} PU_{x_i,\Lambda}\Big) \\
&\qquad\quad + O\bigg(\int_{\Omega_1} U_{x_1,\Lambda}^{\frac12(p+1)}\Big(\sum_{i=2}^{k} U_{x_i,\Lambda}\Big)^{\frac12(p+1)}\bigg) + O\Big(\varepsilon\int_{\Omega_1} W_{r,\Lambda}^{p+1}\ln W_{r,\Lambda}\Big).
\end{aligned}
$$

Note that for $y\in\Omega_1$, $|y-x_i|\ge|y-x_1|$. Using (2-11), we see that for $t\in(1, N-2)$,

$$
\sum_{i=2}^{k} U_{x_i,\Lambda} \le \frac{C}{(1+|y-x_1|)^{N-2-t}}\sum_{i=2}^{k}\frac{1}{|x_i-x_1|^t}.
$$

If we take t close to $N-2$, then

$$
\int_{\Omega_1} U_{x_1,\Lambda}^{\frac12(p+1)}\Big(\sum_{i=2}^{k} U_{x_i,\Lambda}\Big)^{\frac12(p+1)} = O((k\varepsilon^{1/(N-2)})^{Nt/(N-2)}) = O(\varepsilon^{(1+\sigma)/(N-2)}).
$$

Moreover, it is easy to show that

$$
\varepsilon\int_{\Omega_1} W_{r,\Lambda}^{p+1}\ln W_{r,\Lambda} = O(\varepsilon).
$$

As a result, we obtain

$$
\begin{aligned}
I = k\bigg(&-\int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|)PU_{x_1,\Lambda}^{p}\frac{\partial PU_{x_1,\Lambda}}{\partial\Lambda} \\
&-\sum_{i=2}^{k}\int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|)PU_{x_1,\Lambda}^{p}\frac{\partial PU_{x_i,\Lambda}}{\partial\Lambda} \\
&\quad + p\sum_{i=2}^{k}\int_{\Omega_1}\big(1-K(\varepsilon^{1/(N-2)}|y|)\big)PU_{x_1,\Lambda}^{p-1}\frac{\partial PU_{x_1,\Lambda}}{\partial\Lambda}PU_{x_i,\Lambda} \\
&\qquad + p\sum_{i=2}^{k}\int_{B_*(0)\setminus\Omega_1} PU_{x_1,\Lambda}^{p-1}\frac{\partial PU_{x_1,\Lambda}}{\partial\Lambda}PU_{x_i,\Lambda} + O(\varepsilon^{(1+\sigma)/(N-2)})\bigg).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|) PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} \\
&\quad = \int_{\Omega_1} PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} + \int_{\Omega_1} \big(K(\varepsilon^{1/(N-2)}|y|) - 1\big) PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} \\
&\quad = \int_{\Omega_1} PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} \\
&\qquad\qquad + \int_{\Omega_1} \big(K(|\bar{x}_1|) - 1\big) PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} + O(\varepsilon^{(1+\sigma)/(N-2)}) \\
&\quad = \int_{\Omega_1} PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} - K'(1)d \int_{\Omega_1} PU^p_{x_1,\Lambda} \frac{\partial PU_{x_1,\Lambda}}{\partial \Lambda} + O(\varepsilon^{(1+\sigma)/(N-2)}) \\
&\quad = -\int_{\Omega_1} U^p_{x_1,\Lambda} \frac{\partial \phi_{x_1,\Lambda}}{\partial \Lambda} - p \int_{\Omega_1} U^p_{x_1,\Lambda} \frac{\partial U_{x_1,\Lambda}}{\partial \Lambda} \phi_{x_1,\Lambda} + O(\varepsilon^{(1+\sigma)/(N-2)}) \\
&\quad = \frac{B_1 \varepsilon H(\bar{x}_1, \bar{x}_1)}{\Lambda^{N-1}} + O(\varepsilon^{(1+\sigma)/(N-2)})
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_{\Omega_1} K(\varepsilon^{1/(N-2)}|y|) PU^p_{x_1,\Lambda} \frac{\partial PU_{x_i,\Lambda}}{\partial \Lambda} \\
&\qquad = \int_{\Omega_1} PU^p_{x_1,\Lambda} \frac{\partial PU_{x_i,\Lambda}}{\partial \Lambda} + \int_{\Omega_1} \big(K(\varepsilon^{1/(N-2)}|y|) - 1\big) PU^p_{x_1,\Lambda} \frac{\partial PU_{x_i,\Lambda}}{\partial \Lambda} \\
&\qquad = \int_{\Omega_1} U^p_{x_1,\Lambda} \frac{\partial U_{x_i,\Lambda}}{\partial \Lambda} - \int_{\Omega_1} U^p_{x_1,\Lambda} \frac{\partial \phi_{x_i,\Lambda}}{\partial \Lambda} + O(\varepsilon^{(1+\sigma)/(N-2)}) \\
&\qquad = -\frac{B_1 \varepsilon G(\bar{x}_i, \bar{x}_1)}{\Lambda^{N-1}} + O(\varepsilon^{(1+\sigma)/(N-2)}).
\end{aligned}
$$

Other terms can be estimated similarly. Thus, the result follows. □

Appendix B: Basic estimates

In this section, we will give some basic estimates used in the reduction procedure. We will use the same constant $C > 0$ to denote the different constants.

Lemma B.1. *Let $g_{ij} = 1/((1+|y-x_i|)^\alpha (1+|y-x_j|)^\beta)$ for each fixed i and j, $i \neq j$, where $\alpha \geq 1$ and $\beta \geq 1$ are two constants. Then for any $0 < \sigma \leq \min(\alpha, \beta)$, there is a constant $C > 0$ such that*

$$
g_{ij}(y) \leq \frac{C}{|x_i - x_j|^\sigma} \left(\frac{1}{(1+|y-x_i|)^{\alpha+\beta-\sigma}} + \frac{1}{(1+|y-x_j|)^{\alpha+\beta-\sigma}} \right).
$$

Lemma B.2. *For any constant* $0 < \sigma < N-2$, *there is a constant* $C > 0$ *such that*

$$\int_{\mathbb{R}^N} \frac{1}{|y-z|^{N-2}} \frac{1}{(1+|z|)^{2+\sigma}}\,dz \le \frac{C}{(1+|y|)^{\sigma}}.$$

The proofs of the above two lemmas can be found in [Wei and Yan 2010b].

Lemma B.3. *Suppose that* $\varepsilon > 0$ *and* $N \ge 4$. *Then there is a small* $\vartheta > 0$ *such that*

$$\int_{\mathbb{R}^N} \frac{1}{|y-z|^{N-2}} W_{r,\Lambda}^{p-1+\varepsilon}(z) \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}\,dz \le C \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau+\vartheta}}.$$

Proof. This is similar to the proof of Lemma B.3 in [Wei and Yan 2010b]. So we just sketch it. Note that

$$W_{r,\Lambda}(z) \le C \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{N-2}}.$$

As in [Wei and Yan 2010b], for $y \in \Omega_1$ we have $W_{r,\Lambda}(z) \le \frac{C}{(1+|z-x_1|)^{N-2-\tau_1}}$, where $0 < \tau_1 \le \frac{1}{2}(N-2)$. Thus,

$$W_{r,\Lambda}^{p-1+\varepsilon}(z) \le \frac{C}{(1+|z-x_1|)^{4-\frac{4\tau_1}{N-2}+(N-2-\tau_1)\varepsilon}}.$$

By virtue of Lemma B.1, for $y \in \Omega_1$ we get

$$\begin{aligned}
W_{r,\Lambda}^{p-1+\varepsilon}(z) &\sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}} \\
&\le \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+6)+\tau-\frac{4\tau_1}{N-2}+(N-2-\tau_1)\varepsilon}} \\
&\qquad + \sum_{j=2}^{k} \frac{C}{(1+|z-x_1|)^{4-\frac{4\tau_1}{N-2}+(N-2-\tau_1)\varepsilon}} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}} \\
&\le \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+6)+\tau-\frac{4\tau_1}{N-2}+(N-2-\tau_1)\varepsilon}} \\
&\qquad + \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+6)+\tau-\frac{N+2}{N-2}\tau_1+(N-2-\tau_1)\varepsilon}} \sum_{j=2}^{k} \frac{1}{|x_j-x_1|^{\tau_1}} \\
&\le \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+6)+\tau-\frac{N+2}{N-2}\tau_1+(N-2-\tau_1)\varepsilon}}.
\end{aligned}$$

Thus, we can obtain

$$\int_{\Omega_1} \frac{1}{|y-z|^{N-2}} W_{r,\Lambda}^{p-1+\varepsilon}(z) \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}\, dz$$
$$\le \int_{\Omega_1} \frac{1}{|y-z|^{N-2}} \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+6)+\tau-\frac{N+2}{N-2}\tau_1+(N-2-\tau_1)\varepsilon}}\, dz$$
$$\le \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+2)+\tau-\frac{N+2}{N-2}\tau_1+(N-2-\tau_1)\varepsilon}}.$$

As a result, for τ_1 satisfying $2-(N+2)/(N-2)\tau_1 > 0$, we find that

$$\int_{\mathbb{R}^N} \frac{1}{|y-z|^{N-2}} W_{r,\Lambda}^{p-1+\varepsilon}(z) \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}\, dz$$
$$= \sum_{i=1}^{k} \int_{\Omega_i} \frac{1}{|y-z|^{N-2}} W_{r,\Lambda}^{p-1+\varepsilon}(z) \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau}}\, dz$$
$$\le \sum_{i=1}^{k} \frac{C}{(1+|z-x_1|)^{\frac{1}{2}(N+2)+\tau-\frac{N+2}{N-2}\tau_1+(N-2-\tau_1)\varepsilon}}$$
$$\le C \sum_{j=1}^{k} \frac{1}{(1+|z-x_j|)^{\frac{1}{2}(N-2)+\tau+\vartheta}}. \qquad \square$$

Acknowledgements

The authors would like to thank the referee very much for the careful reading of the paper and valuable comments, including pointing out Remark 1.5. Z. Liu was supported by funds from the NSFC (No. 11426088, No. 11501166); S. Peng was partially supported by funds from the NSFC (No. 11125101, No. 11571130).

References

[Ben Ayed et al. 2003] M. Ben Ayed, K. El Mehdi, O. Rey, and M. Grossi, "A nonexistence result of single peaked solutions to a supercritical nonlinear problem", *Commun. Contemp. Math.* **5**:2 (2003), 179–195. MR 2004k:35140 Zbl 1066.35035

[Byeon and Wang 2005] J. Byeon and Z.-Q. Wang, "On the Hénon equation: asymptotic profile of ground states, II", *J. Differential Equations* **216**:1 (2005), 78–108. MR 2006j:35068 Zbl 1114.35070

[Byeon and Wang 2006] J. Byeon and Z.-Q. Wang, "On the Hénon equation: asymptotic profile of ground states, I", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23**:6 (2006), 803–828. MR 2007g:35046 Zbl 1114.35071

[Cao and Peng 2003] D. Cao and S. Peng, "The asymptotic behaviour of the ground state solutions for Hénon equation", *J. Math. Anal. Appl.* **278**:1 (2003), 1–17. MR 2003m:35062 Zbl 1086.35036

[Cao et al. 2009] D. Cao, S. Peng, and S. Yan, "Asymptotic behaviour of ground state solutions for the Hénon equation", *IMA J. Appl. Math.* **74**:3 (2009), 468–480. MR 2010d:35088 Zbl 1169.35377

[Gidas et al. 1979] B. Gidas, W. M. Ni, and L. Nirenberg, "Symmetry and related properties via the maximum principle", *Comm. Math. Phys.* **68**:3 (1979), 209–243. MR 80h:35043 Zbl 0425.35020

[Gladiali and Grossi 2012] F. Gladiali and M. Grossi, "Supercritical elliptic problem with non-autonomous nonlinearities", *J. Differential Equations* **253**:9 (2012), 2616–2645. MR 2959382 Zbl 1266.35105

[Gladiali et al. 2013] F. Gladiali, M. Grossi, and S. L. N. Neves, "Nonradial solutions for the Hénon equation in $\mathbb{R}^N$", *Adv. Math.* **249** (2013), 1–36. MR 3116566 Zbl 06296090

[Hénon 1973] M. Hénon, "Numerical experiments on the stability of spherical stellar systems", *Astron. Astrophys.* **24** (1973), 229–238.

[Hirano 2009] N. Hirano, "Existence of positive solutions for the Hénon equation involving critical Sobolev terms", *J. Differential Equations* **247**:5 (2009), 1311–1333. MR 2010f:35121 Zbl 1176.35083

[Li and Peng 2009] S. Li and S. Peng, "Asymptotic behavior on the Hénon equation with supercritical exponent", *Sci. China Ser. A* **52**:10 (2009), 2185–2194. MR 2011b:35159 Zbl 1184.35136

[Ni 1982] W. M. Ni, "A nonlinear Dirichlet problem on the unit ball and its applications", *Indiana Univ. Math. J.* **31**:6 (1982), 801–807. MR 84b:35051 Zbl 0515.35033

[Peng 2006] S. Peng, "Multiple boundary concentrating solutions to Dirichlet problem of Hénon equation", *Acta Math. Appl. Sin. Engl. Ser.* **22**:1 (2006), 137–162. MR 2006k:35099 Zbl 1153.35325

[Peng and Wang 2013] S. Peng and Z.-Q. Wang, "Segregated and synchronized vector solutions for nonlinear Schrödinger systems", *Arch. Ration. Mech. Anal.* **208**:1 (2013), 305–339. MR 3021550 Zbl 1260.35211

[del Pino et al. 2003] M. del Pino, P. Felmer, and M. Musso, "Two-bubble solutions in the super-critical Bahri–Coron's problem", *Calc. Var. Partial Differential Equations* **16**:2 (2003), 113–145. MR 2004a:35079 Zbl 1142.35421

[Pistoia and Serra 2007] A. Pistoia and E. Serra, "Multi-peak solutions for the Hénon equation with slightly subcritical growth", *Math. Z.* **256**:1 (2007), 75–97. MR 2008b:35091 Zbl 1134.35047

[Pohožaev 1965] S. I. Pohožaev, "On the eigenfunctions of the equation $\Delta u + \lambda f(u) = 0$", *Dokl. Akad. Nauk SSSR* **165** (1965), 36–39. In Russian; translated in *Sov. Math. Dokl.* **6** (1965), 1408–1411. MR 33 #411 Zbl 0141.30202

[Serra 2005] E. Serra, "Non radial positive solutions for the Hénon equation with critical growth", *Calc. Var. Partial Differential Equations* **23**:3 (2005), 301–326. MR 2006f:35100 Zbl 1207.35147

[Smets et al. 2002] D. Smets, M. Willem, and J. Su, "Non-radial ground states for the Hénon equation", *Commun. Contemp. Math.* **4**:3 (2002), 467–480. MR 2003g:35086 Zbl 1160.35415

[Wei and Yan 2010a] J. Wei and S. Yan, "Infinitely many positive solutions for the nonlinear Schrödinger equations in $\mathbb{R}^N$", *Calc. Var. Partial Differential Equations* **37**:3-4 (2010), 423–439. MR 2011c:35065 Zbl 1189.35106

[Wei and Yan 2010b] J. Wei and S. Yan, "Infinitely many solutions for the prescribed scalar curvature problem on $\mathbb{S}^N$", *J. Funct. Anal.* **258**:9 (2010), 3048–3081. MR 2011g:35119 Zbl 1209.53028

[Wei and Yan 2011] J. Wei and S. Yan, "Infinitely many positive solutions for an elliptic problem with critical or supercritical growth", *J. Math. Pures Appl.* (9) **96**:4 (2011), 307–333. MR 2832637 Zbl 1253.31008

[Wei and Yan 2013] J. Wei and S. Yan, "Infinitely many nonradial solutions for the Hénon equation with critical growth", *Rev. Mat. Iberoam.* **29**:3 (2013), 997–1020. MR 3090144 Zbl 1277.35330

Received November 6, 2014. Revised May 10, 2015.

ZHONGYUAN LIU
SCHOOL OF MATHEMATICS AND STATISTICS
HENAN UNIVERSITY
KAIFENG, HENAN 475004
CHINA
liuzy@henu.edu.cn

SHUANGJIE PENG
SCHOOL OF MATHEMATICS AND STATISTICS
HUBEI KEY LABORATORY OF MATHEMATICAL PHYSICS
CENTRAL CHINA NORMAL UNIVERSITY
WUHAN, HUBEI 430079
CHINA
sjpeng@mail.ccnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.141

EFFECTIVE DIVISORS ON THE PROJECTIVE LINE HAVING SMALL DIAGONALS AND SMALL HEIGHTS AND THEIR APPLICATION TO ADELIC DYNAMICS

YÛSUKE OKUYAMA

We establish a quantitative adelic equidistribution theorem for a sequence of effective divisors on the projective line over the separable closure of a product formula field having small diagonals and small g-heights with respect to an adelic normalized weight g in arbitrary characteristic and in a possibly nonseparable setting. Applying this quantitative adelic equidistribution result to adelic dynamics of f, we obtain local proximity estimates between the iterations of a rational function $f \in k(z)$ of degree > 1 and a rational function $a \in k(z)$ of degree > 0 over a product formula field k of characteristic 0.

1. Introduction

Let k be a field and denote by k_s the separable closure of k in an algebraic closure $\bar{k}$. For every $d \in \mathbb{N} \cup \{0\}$, let $k[p_0, p_1]_d$ be the set of all homogeneous polynomials in two variables over k of degree d. A *k-effective divisor* $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$ is a divisor on $\mathbb{P}^1(\bar{k})$ defined by the zeros in $\mathbb{P}^1(\bar{k})$ of some $P \in \bigcup_{d \in \mathbb{N}} k[p_0, p_1]_d$ *taking into account their multiplicities*, and is said to be on $\mathbb{P}^1(k_s)$ if $\operatorname{supp} \mathcal{Z} \subset \mathbb{P}^1(k_s)$. The defining polynomial $P(p_0, p_1)$ of $\mathcal{Z}$ is unique up to multiplication in k^* $(= k \setminus \{0\})$, and is called a *representative* of $\mathcal{Z}$. Effective divisors include Galois conjugacy classes of algebraic numbers, and are also called *Galois stable multisets* in $\mathbb{P}^1(\bar{k})$.

Our first aim in this article is to establish a *quantitative* adelic equidistribution of sequences of k-effective divisors on $\mathbb{P}^1(k_s)$, where k is a *product formula* field, having not only small g-heights (with respect to an adelic normalized weight g) but also *small diagonals* in arbitrary characteristic and in a possibly nonseparable setting. Secondly, we contribute to the study of the local *proximities* between the iterations of a rational function $f \in k(z)$ of degree > 1 and a rational function $a \in k(z)$ of degree > 0 on a chordal disk D of radius > 0 in the projective line $\mathbb{P}^1(\mathbb{C}_v)$ for each place v of k, in the setting of adelic dynamics of characteristic 0.

MSC2010: primary 37P30; secondary 11G50, 37P50, 37F10.
Keywords: product formula field, effective divisor, small diagonals, small heights, quantitative equidistribution, asymptotically Fekete configuration, local proximity sequence, adelic dynamics.

1.1. ***Arithmetic over a product formula field.*** A field k is a *product formula field* if k is equipped with

(i) a set M_k of all places of k, which are either *finite* or *infinite*,

(ii) a set $\{|\cdot|_v : v \in M_k\}$, where for each $v \in M_k$, $|\cdot|_v$ is a nontrivial absolute value of k representing v (and then by definition $|\cdot|_v$ is nonarchimedean if and only if v is finite), and

(iii) a set $\{N_v : v \in M_k\}$, where $N_v \in \mathbb{N}$ for every $v \in M_k$

such that the following *product formula* holds: if $z \in k \setminus \{0\}$ then we have $|z|_v \neq 1$ for at most finitely many $v \in M_k$ and moreover

$$\prod_{v \in M_k} |z|_v^{N_v} = 1. \tag{PF}$$

Product formula fields include number fields and function fields over curves, and a product formula field is a number field if and only if it has at least one infinite place (see, e.g., the paragraph after Definition 7.51 of [Baker and Rumely 2010]).

Let k be a product formula field. For each $v \in M_k$, let k_v be the completion of k with respect to $|\cdot|_v$ and $\mathbb{C}_v$ the completion of an algebraic closure of k_v with respect to (the extended) $|\cdot|_v$. We fix an embedding of $\bar{k}$ into $\mathbb{C}_v$ which extends that of k into k_v; by convention, the dependence on $v \in M_k$ of a local quantity induced by $|\cdot|_v$ is emphasized by adding the suffix v to it. A family $g = \{g_v : v \in M_k\}$ is an *adelic continuous weight* if

(i) for every $v \in M_k$, g_v is a continuous function on the *Berkovich* projective line $\mathsf{P}^1(\mathbb{C}_v)$ such that
$$\mu_v^g := \Delta g_v + \Omega_{\mathrm{can},v}$$
is a probability Radon measure on $\mathsf{P}^1(\mathbb{C}_v)$ (see (2-2) for the definition of the probability Radon measure $\Omega_{\mathrm{can},v}$ on $\mathsf{P}^1(\mathbb{C}_v)$, and (2-3) for the normalization of the Laplacian Δ on $\mathsf{P}^1(\mathbb{C}_v)$), and

(ii) there is a finite subset E_g in M_k such that $g_v \equiv 0$ on $\mathsf{P}^1(\mathbb{C}_v)$ for all $v \in M_k \setminus E_g$.

Moreover, g is called an *adelic normalized weight* if, in addition,

(iii) the *g_v-equilibrium energy* V_{g_v} of $\mathsf{P}^1(\mathbb{C}_v)$ vanishes for every $v \in M_k$ (see Section 2.1 for the definition of V_{g_v}).

For an adelic continuous weight $g = \{g_v : v \in M_k\}$, the family $\mu^g := \{\mu_v^g : v \in M_k\}$ is called an *adelic probability measure* (compare [Favre and Rivera-Letelier 2006, Définition 1.1]). An adelic continuous weight $g = \{g_v : v \in M_k\}$ is said to be *placewise Hölder continuous* if for every $v \in M_k$, g_v is Hölder continuous on $\mathsf{P}^1(\mathbb{C}_v)$ with respect to the small model metric d_v on $\mathsf{P}^1(\mathbb{C}_v)$ (see (3-1) for the definition of d_v).

Given $P \in \bigcup_{d\in\mathbb{N}} k[p_0, p_1]_d$ and an adelic continuous weight $g = \{g_v : v \in M_k\}$, the *g-height* of a k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$ represented by P is

$$h_g(\mathcal{Z}) := \sum_{v\in M_k} N_v \frac{M_{g_v}(P)}{\deg P}, \tag{1-1}$$

where, for every $v \in M_k$, $M_{g_v}(P)$ is the logarithmic g_v*-Mahler measure* of P (see (2-10) for the definition of $M_{g_v}(P)$ and Section 2.3 for a proof that $h_g(\mathcal{Z}) \in \mathbb{R}$); by (PF), $h_g(\mathcal{Z})$ is well defined. For every $v \in M_k$, letting $\delta_{\mathcal{S}}$ be the Dirac measure on $\mathsf{P}^1(\mathbb{C}_v)$ at a point $\mathcal{S} \in \mathsf{P}^1(\mathbb{C}_v)$, a k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$ is regarded as a positive and discrete Radon measure $\sum_{w\in\operatorname{supp}\mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})\delta_w$ on $\mathsf{P}^1(\mathbb{C}_v)$, still denoted by $\mathcal{Z}$. Then the *diagonal*

$$(\mathcal{Z}\times\mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(\bar{k})}) \;=\; \sum_{w\in\operatorname{supp}\mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2$$

of $\mathcal{Z}$ is independent of $v \in M_k$. For a sequence $(\mathcal{Z}_n)$ of k-effective divisors on $\mathbb{P}^1(\bar{k})$ satisfying $\lim_{n\to\infty}\deg\mathcal{Z}_n = \infty$, we say $(\mathcal{Z}_n)$ *has small g-heights with respect to an adelic normalized weight g* if $\limsup_{n\to\infty} h_g(\mathcal{Z}_n) \le 0$, and we say $(\mathcal{Z}_n)$ *has small diagonals* if $\lim_{n\to\infty}((\mathcal{Z}_n\times\mathcal{Z}_n)(\operatorname{diag}_{\mathbb{P}^1(\bar{k})}))/(\deg\mathcal{Z}_n)^2 = 0$.

1.2. ***Quantitative adelic equidistribution of effective divisors.*** The following is one of our main results; for the Galois conjugacy class of an algebraic number, this was due to Favre and Rivera-Letelier [2006, Théorème 7]. For the definitions of the C^1-regularity of a continuous test function ϕ on $\mathsf{P}^1(\mathbb{C}_v)$, the Lipschitz constant $\operatorname{Lip}(\phi)_v$ on $(\mathsf{P}^1(\mathbb{C}_v), \mathsf{d}_v)$, and the Dirichlet norm $\langle\phi,\phi\rangle_v$ of ϕ for each $v \in M_k$, see Section 7.

Theorem 1. *Let k be a product formula field and k_s the separable closure of k in $\bar{k}$. Let $g = \{g_v : v \in M_k\}$ be a placewise Hölder continuous adelic normalized weight. Then for every $v \in M_k$, there is $C > 0$ such that for every k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(k_s)$ and every test function $\phi \in C^1(\mathsf{P}^1(\mathbb{C}_v))$,*

$$\left|\int_{\mathsf{P}^1(\mathbb{C}_v)} \phi\, \mathrm{d}\Big(\frac{\mathcal{Z}}{\deg\mathcal{Z}} - \mu_v^g\Big)\right| \le C\cdot\max\big\{\operatorname{Lip}(\phi)_v, \langle\phi,\phi\rangle_v^{1/2}\big\}\sqrt{\max\left\{h_g(\mathcal{Z}), (\log\deg\mathcal{Z})\frac{(\mathcal{Z}\times\mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(k_s)})}{(\deg\mathcal{Z})^2}\right\}}. \tag{1-2}$$

In Theorem 1, if $v \in M_k$ is an infinite place, or equivalently, $\mathbb{C}_v \cong \mathbb{C}$, then the estimate (1-2) gives a quantitative estimate of the *Kantorovich–Wasserstein metric*

$$W\Big(\frac{\mathcal{Z}}{\deg\mathcal{Z}}, \mu_v^g\Big) = \sup_{\phi}\left|\int_{\mathbb{P}^1(\mathbb{C})} \phi\,\mathrm{d}\Big(\frac{\mathcal{Z}}{\deg\mathcal{Z}} - \mu_v^g\Big)\right|$$

between the probability Radon measures $\mathcal{Z}/\deg\mathcal{Z}$ and μ_v^g on $\mathsf{P}^1(\mathbb{C}_v) \cong \mathbb{P}^1(\mathbb{C})$, where ϕ ranges over all Lipschitz continuous functions on $\mathbb{P}^1(\mathbb{C})$ whose Lipschitz constants equal 1 with respect to the normalized chordal metric $[z, w]$ on $\mathbb{P}^1(\mathbb{C})$ (see Remark 4.2). For the details of the metric W including its role in the optimal transportation problems, see, e.g., [Villani 2009].

The next theorem is a qualitative version of Theorem 1. For a sequence of Galois conjugacy classes of algebraic numbers, this was due to Baker and Rumely [2006, Theorem 2.3], Chambert-Loir [2006, Théorème 4.2], and Favre and Rivera-Letelier [2006, Théorème 2]; see also [Szpiro, Ullmo, and Zhang 1997; Bilu 1997; Rumely 1999; Chambert-Loir 2000; Autissier 2001; Baker and Hsia 2005; Baker and Rumely 2006; Chambert-Loir 2006; Favre and Rivera-Letelier 2006], and, most recently, [Yuan 2008] on big line bundles over arithmetic varieties.

Theorem 2 (asymptotically Fekete configuration of effective divisors). *Let k be a product formula field and k_s its separable closure in $\bar{k}$. Let $g = \{g_v : v \in M_k\}$ be an adelic normalized weight. If a sequence $(\mathcal{Z}_n)$ of k-effective divisors on $\mathbb{P}^1(k_s)$ satisfying* $\lim_{n\to\infty} \deg \mathcal{Z}_n = \infty$ *has both small diagonals and small g-heights, then for every $v \in M_k$, $(\mathcal{Z}_n)$ is an asymptotically g_v-Fekete configuration on* $\mathsf{P}^1(\mathbb{C}_v)$. *In particular*, $\lim_{n\to\infty} \mathcal{Z}_n / \deg \mathcal{Z}_n = \mu_v^g$ *weakly on* $\mathsf{P}^1(\mathbb{C}_v)$.

In Theorem 2, the assertion that $(\mathcal{Z}_n)$ *is an asymptotically g_v-Fekete configuration on* $\mathsf{P}^1(\mathbb{C}_v)$ (see (2-7) for the definition), which is also called a *g_v-pseudo-equidistribution* on $\mathsf{P}^1(\mathbb{C}_v)$, is stronger than the final equidistribution assertion. For a relationship between the Kantorovich–Wasserstein metric W and (asymptotically) Fekete configurations on complex manifolds, see [Lev and Ortega-Cerdà 2012, §7]. For a recent result on the *capacity and the transfinite diameter* on complex manifolds, see [Berman and Boucksom 2010] (on $\mathbb{C}^n$, we also refer to the survey [Levenberg 2010]); for the *convergence of* (*asymptotically*) *Fekete points* on complex manifolds, see [Berman, Boucksom, and Nyström 2011].

1.3. ***Quantitative equidistribution in adelic dynamics.*** For rational functions f, a over a field k and for $n \in \mathbb{N}$, the divisor $[f^n = a]$ defined by the roots of the equation $f^n = a$ in $\mathbb{P}^1(\bar{k})$ is a k-effective divisor on $\mathbb{P}^1(\bar{k})$ if $f^n \not\equiv a$.

Let k be a product formula field. For a rational function $f \in k(z)$ of degree $d > 1$, let $\hat{g}_f := \{g_{f,v} : v \in M_k\}$ be the *adelic dynamical Green function* in the sense that for every $v \in M_k$, $g_{f,v}$ is the dynamical Green function of f on $\mathsf{P}^1(\mathbb{C}_v)$, so that $\mu_{f,v} := \mu^{g_{f,v}}$ is the f-equilibrium (or canonical) measure on $\mathsf{P}^1(\mathbb{C}_v)$ (see Section 9 for details). The family $\hat{g}_f$ is in fact an adelic normalized weight, and the $\hat{g}_f$-height function $h_{\hat{g}_f}$ coincides with the Call–Silverman *f-dynamical* (or canonical) height function. For every rational function $a \in k(z)$, the sequence $([f^n = a])$ has *strictly small* $\hat{g}_f$-heights in that $\limsup_{n\to\infty}(d^n + \deg a) \cdot h_{\hat{g}_f}([f^n = a]) < \infty$ (Lemma 9.2). Hence the following are consequences of Theorems 1 and 2, respectively.

Theorem 3. *Let k be a product formula field and k_s its separable closure in $\bar{k}$. Let $f \in k(z)$ be a rational function of degree $d > 1$ and $a \in k(z)$ a rational function.*

Then for every $v \in M_k$, there exists a constant $C > 0$ such that for every test function $\phi \in C^1(\mathsf{P}^1(\mathbb{C}_v))$ and every $n \in \mathbb{N}$,

$$\text{(1-3)} \quad \left| \int_{\mathsf{P}^1(\mathbb{C}_v)} \phi \, \mathrm{d}\Big(\frac{[f^n = a]}{d^n + \deg a} - \mu_{f,v} \Big) \right| \\ \le C \cdot \max\big\{ \mathrm{Lip}(\phi)_v, \langle \phi, \phi \rangle_v^{1/2} \big\} \sqrt{\frac{n \cdot ([f^n = a] \times [f^n = a])(\mathrm{diag}_{\mathbb{P}^1(k_s)})}{(d^n + \deg a)^2}}$$

if $f^n \not\equiv a$ and the divisor $[f^n = a]$ on $\mathbb{P}^1(\bar{k})$ is on $\mathbb{P}^1(k_s)$.

Theorem 4. *Let k be a product formula field and k_s its separable closure in $\bar{k}$. Let $f \in k(z)$ be a rational function of degree $d > 1$ and $a \in k(z)$ a rational function. If the sequence $([f^n = a])$ has small diagonals and the divisor $[f^n = a]$ is on $\mathbb{P}^1(k_s)$ for every sufficiently large $n \in \mathbb{N}$, then for every $v \in M_k$, $([f^n = a])$ is an asymptotically $g_{f,v}$-Fekete configuration on $\mathsf{P}^1(\mathbb{C}_v)$. In particular,*

$$\lim_{n \to \infty} \frac{[f^n = a]}{d^n + \deg a} = \mu_{f,v}$$

weakly on $\mathsf{P}^1(\mathbb{C}_v)$.

The final equidistribution assertion in Theorem 4 has been established in [Brolin 1965; Ljubich 1983; Freire, Lopes, and Mañé 1983] in complex dynamics, and in [Favre and Rivera-Letelier 2010] in (not necessarily adelic) nonarchimedean dynamics (of characteristic 0 when $\deg a > 0$). For every constant $a \in \mathbb{P}^1(k)$, the estimate (1-3) in Theorem 3 has been obtained in [Okuyama 2013b, Theorems 4 and 5] in complex and (not necessarily adelic) nonarchimedean dynamics of characteristic 0. In complex dynamics, for every $f \in \mathbb{C}(z)$ of degree $d > 1$, every constant $a \in \mathbb{P}^1(\mathbb{C})$, and every $\phi \in C^2(\mathbb{P}^1(\mathbb{C}))$, a finer estimate than (1-3) has been obtained in [Drasin and Okuyama 2007, Theorem 2 and (4.2)].

1.4. ***Application to a motivating question.*** Let K be an algebraically closed field that is complete with respect to a nontrivial absolute value $|\cdot|$, and $[z, w]$ be the normalized *chordal metric* on $\mathbb{P}^1 = \mathbb{P}^1(K)$ (see (2-1)). A subset D in $\mathbb{P}^1$ is called a *chordal disk* (in $\mathbb{P}^1$) if $D = \{z \in \mathbb{P}^1 : [z, w] \le r\}$ for some $w \in \mathbb{P}^1$ and some *radius* $r \ge 0$. Even in the specific case $a = \mathrm{Id}$ (see, e.g., [Cremer 1928; Siegel 1942; Brjuno 1971; 1972; Herman and Yoccoz 1983; Yoccoz 1988; 1995; Pérez-Marco 1993; 2001]), which is one of the most interesting cases and is related to *the difficulty of small denominators* in nonarchimedean and complex dynamics, the following question has not been completely understood.

Question. How uniformly close on a chordal disk D of radius > 0 can the sequence (f^n) of the iterations of a rational function $f \in K(z)$ of degree > 1 be to a rational function $a \in K(z)$ of degree > 0?

For a study of this question on the projective space $\mathbb{P}^N(K)$, see [Okuyama 2010]. The following estimate of the *local proximity sequence* $(\sup_D[f^n, a]_v)$ is an application of Theorem 3 to this question in the setting of adelic dynamics.

Theorem 5. *Let k be a product formula field of characteristic 0. Let $f \in k(z)$ be a rational function of degree > 1 and $a \in k(z)$ a rational function of degree > 0. Then for every $v \in M_k$ and every chordal disk D in $\mathbb{P}^1(\mathbb{C}_v)$ of radius > 0, as $n \to \infty$,*

$$\log \sup_D [f^n, a]_v = O\Big(\sqrt{n \cdot \big([f^n = a] \times [f^n = a]\big)\big(\mathrm{diag}_{\mathbb{P}^1(\bar{k})}\big)}\Big). \tag{1-4}$$

Here, the implicit constant in $O(\,\cdot\,)$ possibly depends on f and a.

In the case that $a = \mathrm{Id}$, we will see that $\big([f^n = \mathrm{Id}] \times [f^n = \mathrm{Id}]\big)\big(\mathrm{diag}_{\mathbb{P}^1(\bar{k})}\big) = O(d^n)$ as $n \to \infty$ in Section 10. Hence Theorem 5 concludes the following.

Theorem 6. *Let k be a product formula field of characteristic 0. Let $f \in k(z)$ be a rational function of degree $d > 1$. Then for every $v \in M_k$ and every chordal disk D in $\mathbb{P}^1(\mathbb{C}_v)$ of radius > 0,*

$$\log \sup_D [f^n, \mathrm{Id}]_v = O\big(\sqrt{nd^n}\big) \quad \textit{as } n \to \infty. \tag{1-5}$$

1.5. *The unit $D^*(p)$.* The next result generalizes the obvious fact that the discriminant of a polynomial in one variable over a field k is in k. The unit $D^*(p)$ plays an important role in the nonseparable case and might have been studied before, but we could find no relevant literature.

Theorem 7. *Let k be a field and k_s the separable closure of k in an algebraic closure $\bar{k}$ of k. For every $p(z) \in k[z]$ of degree > 0, let $\{z_1, \dots, z_m\}$ be the set of all distinct zeros of $p(z)$ in $\bar{k}$ so that $p(z) = a \cdot \prod_{j=1}^m (z - z_j)^{d_j}$ in $\bar{k}[z]$ for some $a \in k \setminus \{0\}$ and some sequence $(d_j)_{j=1}^m$ in $\mathbb{N}$. If $\{z_1, \dots, z_m\} \subset k_s$, then*

$$D^*(p) := \prod_{j=1}^m \prod_{i: i \neq j} (z_j - z_i)^{d_i d_j} \in k \setminus \{0\},$$

where, a priori, this $D^*(p)$ is always in $\bar{k} \setminus \{0\}$.

1.6. *Organization of this article.* In Section 2, we recall background from potential theory and arithmetic on the Berkovich projective line. In Section 3, we extend Favre and Rivera-Letelier's regularization $[\,\cdot\,]_\epsilon$ of discrete Radon measures and establish required estimates on them, and in Section 4 we see the negativity of regularized Fekete sums and a Cauchy–Schwarz inequality. In Sections 5 and 6, we compute the g-Fekete sums $(\mathcal{Z}, \mathcal{Z})_g$ and estimate the regularized g-Fekete sums $(\mathcal{Z}_\epsilon, \mathcal{Z}_\epsilon)_g$ with respect to a k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$. In Section 7, we prove Theorems 1 and 2; the arguments are more or less adaptions of those in the proofs of [Favre and Rivera-Letelier 2006, Théorème 7] and [Baker and Rumely 2010, Theorem 10.24], respectively. In Section 8, we review background from nonarchimedean and complex dynamics. Finally, we prove Theorems 3 and 4 in Section 9, Theorems 5 and 6 in Section 10, and Theorem 7 in Section 11.

2. Background from potential theory and arithmetic

Notation 2.1. For a field k, the origin of k^2 is also denoted by $0 = 0_k$, and we write $\pi = \pi_k : k^2 \setminus \{0\} \to \mathbb{P}^1 = \mathbb{P}^1(k)$ for the canonical projection, so that $\pi(0, 1) = \infty$ and $\pi(p_0, p_1) = p_1/p_0$ if $p_0 \neq 0$. Set the wedge product $(z_0, z_1) \wedge (w_0, w_1) := z_0 w_1 - z_1 w_0$ on k^2.

Let K be an algebraically closed field that is complete with respect to a nontrivial absolute value $|\cdot|$, which is said to be *nonarchimedean* if the strong triangle inequality $|z+w| \leq \max\{|z|, |w|\}$ holds, and *archimedean* otherwise. On K^2, let $\|(p_0, p_1)\|$ be either the maximal norm $\max\{|p_0|, |p_1|\}$ (for nonarchimedean K) or the euclidean norm $\sqrt{|p_0|^2 + |p_1|^2}$ (for archimedean K). The *normalized chordal metric* $[z, w]$ on $\mathbb{P}^1 = \mathbb{P}^1(K)$ is the function

$$(z, w) \mapsto [z, w] = |p \wedge q| / \big(\|p\| \cdot \|q\|\big) \leq 1 \tag{2-1}$$

on $\mathbb{P}^1 \times \mathbb{P}^1$, where $p \in \pi^{-1}(z)$, $q \in \pi^{-1}(w)$. The metric topology on $\mathbb{P}^1$ with respect to $[z, w]$ agrees with the relative topology on $\mathbb{P}^1$ from the *Berkovich projective line* $\mathsf{P}^1 = \mathsf{P}^1(K)$, which is a compact augmentation of $\mathbb{P}^1$ containing $\mathbb{P}^1$ as a dense subset, and is isomorphic to $\mathbb{P}^1$ if and only if K is archimedean (see Section 3.2 for more details when K is nonarchimedean). Letting $\delta_{\mathcal{S}}$ be the Dirac measure on P^1 at a point $\mathcal{S} \in \mathsf{P}^1$, set

$$\Omega_{\mathrm{can}} := \begin{cases} \delta_{\mathcal{S}_{\mathrm{can}}} & \text{for nonarchimedean } K, \\ \omega & \text{for archimedean } K, \end{cases} \tag{2-2}$$

where $\mathcal{S}_{\mathrm{can}}$ is the canonical (or Gauss) point in P^1 for nonarchimedean K (see Section 3.2 for the definition), and ω is the Fubini–Study area element on $\mathbb{P}^1$ normalized as $\omega(\mathbb{P}^1) = 1$ for archimedean K. For nonarchimedean K, the *generalized Hsia kernel* $[\mathcal{S}, \mathcal{S}']_{\mathrm{can}}$ on P^1 with respect to $\mathcal{S}_{\mathrm{can}}$ is the unique (jointly) upper semicontinuous and separately continuous extension of the normalized chordal metric $[z, w]$ on $\mathbb{P}^1 (\times \mathbb{P}^1)$ to $\mathsf{P}^1 \times \mathsf{P}^1$ (see (3-4) for a more concrete description). By convention, for archimedean K, the kernel function $[\mathcal{S}, \mathcal{S}']_{\mathrm{can}}$ is defined by $[z, w]$ itself. Let $\Delta = \Delta_{\mathsf{P}^1}$ be the distributional Laplacian on P^1 normalized so that for each $\mathcal{S}' \in \mathsf{P}^1$,

$$\Delta \log [\,\cdot\,, \mathcal{S}']_{\mathrm{can}} = \delta_{\mathcal{S}'} - \Omega_{\mathrm{can}} \quad \text{on } \mathsf{P}^1. \tag{2-3}$$

For the construction of the Laplacian Δ in the nonarchimedean case, see [Baker and Rumely 2010, §5; Favre and Jonsson 2004, §7.7; Thuillier 2005, §3] and also [Jonsson 2015, §2.5]. In [Baker and Rumely 2010], the opposite sign convention for Δ is adopted.

2.1. *Potential theory on* P^1 *with external fields.* For the foundation of the potential theory on the (Berkovich) projective line, see [Baker and Rumely 2010; Favre and Rivera-Letelier 2010; Thuillier 2005], and also [Jonsson 2015; Tsuji 1959, III §11] ([Thuillier 2005] is on more general curves than lines and [Tsuji 1959, III §11] is on $\mathbb{P}^1(\mathbb{C})$). We also refer to [Saff and Totik 1997] for the generalities of *weighted* potential theory, i.e., logarithmic potential theory *with external fields*.

A *continuous weight* g *on* P^1 is a continuous function on P^1 such that

$$\mu^g := \Delta g + \Omega_{\mathrm{can}}$$

is a probability Radon measure on P^1. For a continuous weight g on P^1, the *g-potential kernel* on P^1 (or the negative of an Arakelov Green kernel function on P^1 relative to μ^g [Baker and Rumely 2010, §8.10]) is the function

$$\Phi_g(\mathcal{S}, \mathcal{S}') := \log\,[\mathcal{S}, \mathcal{S}']_{\mathrm{can}} - g(\mathcal{S}) - g(\mathcal{S}') \quad \text{on } \mathsf{P}^1 \times \mathsf{P}^1, \tag{2-4}$$

and the *g-potential* of a Radon measure ν on P^1 is the function

$$U_{g,\nu}(\,\cdot\,) := \int_{\mathsf{P}^1} \Phi_g(\,\cdot\,, \mathcal{S}')\,\mathrm{d}\nu(\mathcal{S}') \quad \text{on } \mathsf{P}^1. \tag{2-5}$$

By Fubini's theorem, $\Delta U_{g,\nu} = \nu - \nu(\mathsf{P}^1)\mu^g$ on P^1. The *g-equilibrium energy* $V_g \in (-\infty, +\infty)$ *of* P^1 is the supremum of the *g-energy* functional

$$\nu \mapsto \int_{\mathsf{P}^1 \times \mathsf{P}^1} \Phi_g\,\mathrm{d}(\nu \times \nu) = \int_{\mathsf{P}^1} U_{g,\nu}\,\mathrm{d}\nu \tag{2-6}$$

on the space of all probability Radon measures ν on P^1; indeed, $V_g > -\infty$ since $V_g \ge \int_{\mathsf{P}^1 \times \mathsf{P}^1} \Phi_g\,\mathrm{d}(\Omega_{\mathrm{can}} \times \Omega_{\mathrm{can}}) > -\infty$. A probability Radon measure μ on P^1 at which the g-energy functional (2-6) attains the supremum V_g is called a *g-equilibrium mass distribution on* P^1; in fact the unique g-equilibrium mass distribution on P^1 is μ^g, and moreover, $U_{g,\mu^g} \equiv V_g$ on P^1 (for nonarchimedean K, see [Baker and Rumely 2010, Theorem 8.67, Proposition 8.70]). For a discussion on such a Gauss variational problem, see [Saff and Totik 1997, Chapter 1].

A *normalized weight* g *on* P^1 is a continuous weight on P^1 satisfying $V_g = 0$; for every continuous weight g on P^1, $\bar{g} := g + V_g/2$ is the unique normalized weight on P^1 such that $\mu^{\bar{g}} = \mu^g$.

For a continuous weight g on P^1 and a Radon measure ν on P^1, the *g-Fekete* *sum* with respect to ν is

$$(\nu,\nu)_g := \int_{\mathsf{P}^1\times\mathsf{P}^1\setminus\mathrm{diag}_{\mathbb{P}^1(K)}} \Phi_g \,\mathrm{d}(\nu\times\nu),$$

which generalizes the classical *Fekete sum* associated with a finite subset in $\mathbb{C}$ (see [Fekete 1930a; 1930b; 1933]). If $\operatorname{supp}\nu$ is a discrete (so finite) subset in P^1, i.e., if ν is a *discrete* measure on P^1, then $(\nu,\nu)_g$ is always finite (even if $\operatorname{supp}\nu\subset\mathbb{P}^1$).

For a continuous weight g on P^1, a sequence (ν_n) of positive and discrete Radon measures on P^1 satisfying $\lim_{n\to\infty}\nu_n(\mathsf{P}^1)=\infty$ is called an *asymptotically g-Fekete configuration on* P^1 if the sequence (ν_n) not only has *small diagonals* in that $(\nu_n\times\nu_n)(\mathrm{diag}_{\mathbb{P}^1(K)})=o(\nu_n(\mathsf{P}^1)^2)$ as $n\to\infty$ but also satisfies $\lim_{n\to\infty}(\nu_n,\nu_n)_g/(\nu_n(\mathsf{P}^1))^2=V_g$; under the former small diagonals condition, the latter one is equivalent to the weaker

$$\liminf_{n\to\infty}\frac{(\nu_n,\nu_n)_g}{(\nu_n(\mathsf{P}^1))^2}\ge V_g, \tag{2-7}$$

since we always have

$$\limsup_{n\to\infty}\frac{(\nu_n,\nu_n)_g}{(\nu_n(\mathsf{P}^1))^2}\le V_g \tag{2-8}$$

(see, e.g., [Baker and Rumely 2010, Lemma 7.54]). By a classical argument (see [Saff and Totik 1997, Theorem 1.3 in Chapter III]), if (ν_n) is an asymptotically g-Fekete configuration on P^1, then $\lim_{n\to\infty}\nu_n/\nu_n(\mathsf{P}^1)=\mu^g$ weakly on P^1.

2.2. *Local arithmetic on* P^1.

Let k be a field.

Definition 2.2. A field extension K/k is an *algebraic and metric augmentation* of k if K is algebraically closed and (topologically) complete with respect to a nontrivial absolute value $|\cdot|$ (e.g., $\mathbb{C}_v$ is an algebraic and metric augmentation of a product formula field k for every $v\in M_k$).

For every $P\in\bigcup_{d\in\mathbb{N}}k[p_0,p_1]_d$, there is a sequence $(q_j^P)_{j=1}^{\deg P}$ in $\bar{k}^2\setminus\{0\}$ giving a factorization

$$P(p_0,p_1)=\prod_{j=1}^{\deg P}\big((p_0,p_1)\wedge q_j^P\big) \tag{2-9}$$

of P in $\bar{k}[p_0,p_1]$. Set $z_j^P:=\pi(q_j^P)\in\mathbb{P}^1(\bar{k})$ for each $j\in\{1,2,\dots,\deg P\}$. Although the sequence $(q_j^P)_{j=1}^{\deg P}$ is not unique, the sequence $(z_j^P)_{j=1}^{\deg P}$ in $\mathbb{P}^1(\bar{k})$ is independent of the choice of $(q_j^P)_{j=1}^{\deg P}$ up to permutations. Let in addition K be an algebraic and metric completion of k. Then the sum $M^\#(P):=\sum_{j=1}^{\deg P}\log\|q_j^P\|$ is also independent of the choice of $(q_j^P)_{j=1}^{\deg P}$, and for every continuous weight g on $\mathsf{P}^1=\mathsf{P}^1(K)$, the *logarithmic g-Mahler measure* of P is

$$M_g(P) := \sum_{j=1}^{\deg P} g(z_j^P) + M^\#(P). \tag{2-10}$$

The function $S_P := |P(\,\cdot\,/\|\cdot\|)|$ on $K^2 \setminus \{0\}$ descends to $\mathbb{P}^1(K)$ and in turn extends continuously to P^1 so that $\log S_P = \sum_{j=1}^{\deg P} \log[\,\cdot\,, z_j^P]_{\mathrm{can}} + M^\#(P)$ on P^1, which can be rewritten as $\log S_P - (\deg P)g = \sum_{j=1}^{\deg P} \Phi_g(\,\cdot\,, z_j^P) + M_g(P)$ on P^1. Integrating both sides against $\mathrm{d}\mu^g$ over P^1, by $U_{g,\mu^g} \equiv V_g$ on P^1, we have the *Jensen-type* formula

$$M_g(P) = \int_{\mathsf{P}^1} \bigl(\log S_P - (\deg P)g\bigr)\,\mathrm{d}\mu^g - (\deg P)V_g. \tag{2-11}$$

2.3. ***A lemma on global arithmetic.*** Let k be a product formula field. The proof of the next result is not based on a field extension of k.

Lemma 2.3. *For every* $P \in \bigcup_{d\in\mathbb{N}} k[p_0, p_1]_d$, *we have* $\sum_{v\in M_k} N_v \cdot M^\#(P)_v \in \mathbb{R}_{\ge 0}$.

Proof. Let $(q_j^P)_{j=1}^{\deg P}$ be a sequence in $\bar{k}^2 \setminus \{0\}$ giving a factorization (2-9) of P, and let $L(P(1,\cdot)) \in k \setminus \{0\}$ be the coefficient of the maximal degree term of $P(1,z) \in k[z]$. Setting $q_j^P = ((q_j^P)_0, (q_j^P)_1)$, for each $j \in \{1, 2, \dots, \deg P\}$, we have

$$L(P(1,\cdot)) = (-1)^{\deg P - \deg_\infty P}\Bigl(\prod_{j:\pi(q_j^P)=\infty} (q_j^P)_1\Bigr)\Bigl(\prod_{j:\pi(q_j^P)\neq\infty} (q_j^P)_0\Bigr)$$

since for each $j \in \{1, 2, \dots, \deg P\}$,

$$q_j^P = \begin{cases} (q_j^P)_0 \cdot (1, \pi(q_j^P)) & \text{if } \pi(q_j^P) \neq \infty, \\ (q_j^P)_1 \cdot (0, 1) & \text{if } \pi(q_j^P) = \infty. \end{cases}$$

Thus we have $\sum_{v\in M_k} N_v \cdot M^\#(P)_v \ge \sum_{v\in M_k} N_v \log|L(P(1,\cdot))|_v = 0$, where the final equality is by (PF).

For each $i, j \in \mathbb{N}\cup\{0\}$ satisfying $i + j = \deg P$, if the coefficient $a_{i,j} \in k$ of the expansion $P(p_0, p_1) = \sum_{i+j=\deg P} a_{i,j} p_0^i p_1^j$ in $k[p_0, p_1]_{\deg P}$ does not vanish, then by (PF), there is a finite subset $E_{i,j}$ in M_k such that $|a_{i,j}|_v = 1$ for every $v \in M_k \setminus E_{i,j}$. Set $E_P := \{\text{infinite places of } k\} \cup \bigcup_{i,j\in\mathbb{N}\cup\{0\}:a_{i,j}\neq 0} E_{i,j}$. For every $v \in M_k \setminus E_P$, by the strong triangle inequality, $|P(p_0, p_1)|_v$ is bounded above by

$$\max\{\max\{|p_0|_v, |p_1|_v\}^{i+j} : i, j \in \mathbb{N}\cup\{0\}, i + j = \deg P\} = \|(p_0, p_1)\|_v^{\deg P}$$

on $\mathbb{C}_v^2$, so that $\log S_{P,v} \le 0$ on $\mathbb{P}^1(\mathbb{C}_v)$ and in turn on $\mathsf{P}^1(\mathbb{C}_v)$. Set $g^0 := \{g_v^0 : v \in M_k\}$ with $g_v^0 \equiv 0$ on $\mathsf{P}^1(\mathbb{C}_v)$ for every $v \in M_k$; then g^0 is an adelic continuous weight. For every finite $v \in M_k$, we have $\mu_v^{g^0} = \delta_{\mathcal{S}_{\mathrm{can},v}}$ on $\mathsf{P}^1(\mathbb{C}_v)$ and moreover $V_{g_v^0} = \log[\mathcal{S}_{\mathrm{can},v}, \mathcal{S}_{\mathrm{can},v}]_{\mathrm{can},v} = 0$, so that by the Jensen-type formula (2-11), we have $M^\#(P)_v = M_{g_v^0}(P) = \log S_{P,v}(\mathcal{S}_{\mathrm{can},v})$. Hence, $M^\#(P)_v \le 0$ for every $v \in M_k \setminus E_P$, and we conclude that $\sum_{v\in M_k} N_v \cdot M^\#(P)_v < \infty$ since $\#E_P < \infty$. $\square$

3. Regularization of discrete Radon measures whose supports are in $\mathbb{P}^1$

Let K be an algebraically closed field complete with respect to a nontrivial absolute value $|\cdot|$.

3.1. ***The small model metric*** d ***and the Hsia kernel*** $|\mathcal{S}-\mathcal{S}'|_\infty$**.** The kernel function $[\mathcal{S},\mathcal{S}']_{\mathrm{can}}$ is not necessarily a metric on $\mathsf{P}^1=\mathsf{P}^1(K)$; indeed, for every $\mathcal{S}\in\mathsf{P}^1$, $[\mathcal{S},\mathcal{S}]_{\mathrm{can}}$ vanishes if and only if $\mathcal{S}\in\mathbb{P}^1=\mathbb{P}^1(K)$. The *small model metric* d on P^1 is the function

$$\mathsf{d}(\mathcal{S},\mathcal{S}') := [\mathcal{S},\mathcal{S}']_{\mathrm{can}} - \frac{[\mathcal{S},\mathcal{S}]_{\mathrm{can}}+[\mathcal{S}',\mathcal{S}']_{\mathrm{can}}}{2} \quad \text{on } \mathsf{P}^1\times\mathsf{P}^1, \tag{3-1}$$

which extends the normalized chordal metric $[z,w]$ on $\mathbb{P}^1$ (but this d does not induce the topology of P^1; see [Baker and Rumely 2010, §2.7; Favre and Rivera-Letelier 2006, §4.7] for details). On the other hand, the *Hsia kernel* $|\mathcal{S}-\mathcal{S}'|_\infty$ on the *Berkovich affine line* $\mathsf{A}^1=\mathsf{A}^1(K)=\mathsf{P}^1\setminus\{\infty\}$ is the function

$$|\mathcal{S}-\mathcal{S}'|_\infty := [\mathcal{S},\mathcal{S}']_{\mathrm{can}}\cdot[\mathcal{S},\infty]_{\mathrm{can}}^{-1}\cdot[\mathcal{S}',\infty]_{\mathrm{can}}^{-1} \quad \text{on } \mathsf{A}^1\times\mathsf{A}^1, \tag{3-2}$$

although the difference $\mathcal{S}-\mathcal{S}'$ itself is not defined unless both $\mathcal{S},\mathcal{S}'\in K$ (for details, see [Baker and Rumely 2010, Chapter 4]). The kernel $|\mathcal{S}-\mathcal{S}'|_\infty$ is the unique (jointly) upper semicontinuous and separately continuous extension of the function $|z-w|$ on $K\times K$ to $\mathsf{A}^1\times\mathsf{A}^1$.

3.2. ***A short description of*** P^1 ***for nonarchimedean*** K**.** Suppose that K is nonarchimedean. A subset B in K is called a (K-closed) *disk* in K if it has the form $B=\{z\in K:|z-a|\le r\}$ for some $a\in K$ and some *radius* $r\ge 0$. By the strong triangle inequality, *two disks in* K *either nest or are disjoint*. This alternative extends to any two decreasing infinite sequences of disks in K such that they either *infinitely nest* or *are eventually disjoint*, and so induces a *cofinal equivalence relation* among them.

Example 3.1. Instead of giving a formal definition of the cofinal equivalence class $\mathcal{S}$ of a decreasing infinite sequence (B_n) of disks in K, let us be practical: each $z\in K$ is regarded as the cofinal equivalence class of the constant sequence (B_n) of the disks $B_n\equiv\{z\}$ in K (of radii $\equiv 0$). More generally, for every cofinal equivalence class $\mathcal{S}$ of a decreasing infinite sequence (B_n) of disks in K, the intersection $B_{\mathcal{S}}:=\bigcap_{n\in\mathbb{N}}B_n$ is independent of the choice of the representatives (B_n) of $\mathcal{S}$, and if $B_{\mathcal{S}}\neq\varnothing$, then $B_{\mathcal{S}}$ is still a disk in K and the $\mathcal{S}$ is represented by the constant sequence $(\tilde{B}_n)$ of the disks $\tilde{B}_n\equiv B_{\mathcal{S}}$ in K.

As a set, the set of all cofinal equivalence classes $\mathcal{S}$ of decreasing infinite sequences (B_n) of disks in K and in addition $\infty\in\mathbb{P}^1$ is nothing but P^1 ([Berkovich 1990, p. 17]; see also [Baker and Rumely 2010, §2; Favre and Rivera-Letelier 2006, §3; Benedetto 2010, §6.1]): for example, the *canonical* (or *Gauss*) point $\mathcal{S}_{\mathrm{can}}$ in

P^1 is represented by the ring of K-integers $\mathcal{O}_K := \{z \in K : |z| \le 1\}$, which is a disk in K. The above alternative induces a partial ordering $\succeq$ on P^1 such that for every $\mathcal{S}, \mathcal{S}' \in \mathsf{P}^1$ satisfying $B_{\mathcal{S}}, B_{\mathcal{S}'} \neq \varnothing$, we have $\mathcal{S} \succeq \mathcal{S}'$ if and only if $B_{\mathcal{S}} \supset B_{\mathcal{S}'}$ (the description is a little complicated when one of $B_{\mathcal{S}}, B_{\mathcal{S}'}$ equals $\varnothing$). For every $\mathcal{S}, \mathcal{S}' \in \mathsf{P}^1$ satisfying $\mathcal{S} \succeq \mathcal{S}'$, the *segment* between $\mathcal{S}$ and $\mathcal{S}'$ in P^1 is the set of all points $\mathcal{S}'' \in \mathsf{P}^1$ satisfying $\mathcal{S} \succeq \mathcal{S}'' \succeq \mathcal{S}'$, which can be equipped with either the ordering induced by $\succeq$ on P^1 or its opposite. All those (oriented) segments make P^1 a *tree* in the sense of Jonsson [2015, §2, Definition 2.2]. The (Gelfand) topology of P^1 coincides with the (weak) topology of P^1 as a tree.

For each $\mathcal{S} \in \mathsf{P}^1 \setminus \{\infty\}$ represented by (B_n), set

$$\operatorname{diam} \mathcal{S} := \lim_{n\to\infty} \operatorname{diam} B_n \quad (= \operatorname{diam} B_{\mathcal{S}} \text{ if } B_{\mathcal{S}} \neq \varnothing),$$

where $\operatorname{diam} B$ denotes the diameter of a disk B in K with respect to $|\cdot|$; by convention, for $\mathcal{S} = \infty$, we set $B_\infty := K$ and $\operatorname{diam} \infty := +\infty$. The *hyperbolic space* is $\mathsf{H}^1 = \mathsf{H}^1(K) := \mathsf{P}^1 \setminus \mathbb{P}^1 = \{\mathcal{S} \in \mathsf{P}^1 : \operatorname{diam} \mathcal{S} \in (0, +\infty)\}$. The *big model* (or *hyperbolic*) *metric* ρ on H^1 is a path metric on H^1 (but does not induce the relative topology of H^1 induced by P^1) so that for every $\mathcal{S}, \mathcal{S}' \in \mathsf{H}^1$ satisfying $\mathcal{S} \succeq \mathcal{S}'$,

$$\rho(\mathcal{S}, \mathcal{S}') = \log(\operatorname{diam} \mathcal{S} / \operatorname{diam} \mathcal{S}') \tag{3-3}$$

(see, e.g., [Baker and Rumely 2010, §2.7]). In terms of ρ, the generalized Hsia kernel $[\mathcal{S}, \mathcal{S}']_{\mathrm{can}}$ with respect to $\mathcal{S}_{\mathrm{can}}$ is interpreted as a Gromov product

$$\log\, [\mathcal{S}, \mathcal{S}']_{\mathrm{can}} = -\rho(\mathcal{S}'', \mathcal{S}_{\mathrm{can}}) \quad \text{on } \mathsf{H}^1 \times \mathsf{H}^1, \tag{3-4}$$

where $\mathcal{S}''$ is the unique point in H^1 lying between $\mathcal{S}$ and $\mathcal{S}'$, between $\mathcal{S}'$ and $\mathcal{S}_{\mathrm{can}}$, and between $\mathcal{S}_{\mathrm{can}}$ and $\mathcal{S}$ (see [Favre and Rivera-Letelier 2006, §3.4]). Similarly, for every $\mathcal{S}, \mathcal{S}' \in \mathsf{A}^1$,

$$|\mathcal{S} - \mathcal{S}'|_\infty = \operatorname{diam} \mathcal{S}'', \tag{3-5}$$

where $\mathcal{S}''$ is the smallest point in A^1 satisfying both $\mathcal{S}'' \succeq \mathcal{S}$ and $\mathcal{S}'' \succeq \mathcal{S}'$ with respect to the partial ordering $\succeq$ on P^1.

For every $\epsilon > 0$, a continuous mapping

$$\pi_\epsilon : \mathsf{A}^1 \to \mathsf{A}^1$$

is defined by $\pi_\epsilon(\mathcal{S}) := \mathcal{S}''$ for every $\mathcal{S} \in \mathsf{A}^1$, where $\mathcal{S}'' \in \{\mathcal{S} \in \mathsf{P}^1 : \operatorname{diam} \mathcal{S} \in [\epsilon, +\infty)\}$ is the unique point between ∞ and $\mathcal{S}$ satisfying $\operatorname{diam} \mathcal{S}'' = \max\{\epsilon, \operatorname{diam} \mathcal{S}\}$ (see [Favre and Rivera-Letelier 2006, §4.6] for details).

3.3. ***Regularization on*** P^1**.** When K is archimedean, fix a nonnegative smooth decreasing function $\xi : [0, \infty) \to [0, 1]$ such that $\operatorname{supp} \xi \subset [0, 1]$ and $\int_0^\infty \xi(x)\, \mathrm{d}x = 1$, and set $\xi_\epsilon(x) := \xi(x/\epsilon)/\epsilon$ on $[0, +\infty)$ for each $\epsilon > 0$. For every $z \in K$ and every

$\epsilon > 0$, the *ϵ-regularization* $[z]_\epsilon$ of δ_z is the convolution $\xi_\epsilon * \delta_z$ on $\mathbb{P}^1$, i.e., for any continuous test function ϕ on $\mathbb{P}^1$,

$$(\xi_\epsilon * \delta_z)(\phi) = \int_0^\epsilon \xi_\epsilon(r)\,\mathrm{d}r \int_0^{2\pi} \phi(z + re^{i\theta})\,\frac{\mathrm{d}\theta}{2\pi}.$$

When K is nonarchimedean, for every $z \in K$ and every $\epsilon > 0$, the *ϵ-regularization* $[z]_\epsilon$ of δ_z is defined by $[z]_\epsilon := (\pi_\epsilon)_*\delta_z = \delta_{\pi_\epsilon(z)}$ on P^1 [Favre and Rivera-Letelier 2006, p. 343]. In both cases, $[z]_\epsilon$ is a probability Radon measure on P^1, the *chordal* potential $\mathsf{P}^1 \ni \mathcal{S} \mapsto \int_{\mathsf{P}^1} \log[\mathcal{S}, \mathcal{S}']_{\mathrm{can}}\,\mathrm{d}[z]_\epsilon(\mathcal{S}')$ of $[z]_\epsilon$ is a continuous function on P^1, and for every $z, w \in K$ and every $\epsilon > 0$, the estimate

$$\int_{\mathsf{A}^1\times\mathsf{A}^1} \log|\mathcal{S} - \mathcal{S}'|_\infty\,\mathrm{d}([z]_\epsilon \times [w]_\epsilon)(\mathcal{S}, \mathcal{S}') \geq \begin{cases} \log|z - w| & \text{if } z \neq w, \\ C_{\mathrm{abs}} + \log\epsilon & \text{if } z = w \end{cases} \tag{3-6}$$

holds, where $C_{\mathrm{abs}} \leq 0$ is an absolute constant and in fact $C_{\mathrm{abs}} = 0$ for nonarchimedean K [Favre and Rivera-Letelier 2006, Lemmes 2.10, 4.11, and their proofs].

Let us extend the ϵ-regularization $[\,\cdot\,]_\epsilon$ and the estimate (3-6) to P^1. Set $\iota(z) := 1/z \in \mathrm{PGL}(2, K)$, which extends to an automorphism on P^1 (see Fact 8.2), so that $\iota^2 = \mathrm{Id}$ on P^1 and $[\iota(\mathcal{S}), \iota(\mathcal{S}')]_{\mathrm{can}} = [\mathcal{S}, \mathcal{S}']_{\mathrm{can}}$ (so $\mathsf{d}(\iota(\mathcal{S}), \iota(\mathcal{S}')) = \mathsf{d}(\mathcal{S}, \mathcal{S}')$) on $\mathsf{P}^1 \times \mathsf{P}^1$. For every $\epsilon > 0$, set $[\infty]_\epsilon := \iota_*[0]_\epsilon$.

For every $z \in \mathbb{P}^1$ and every $\epsilon > 0$, we have

$$\operatorname{supp}[z]_\epsilon \subset \{\mathcal{S} \in \mathsf{P}^1 : \mathsf{d}(\mathcal{S}, z) \leq \epsilon\}, \tag{3-7}$$

as follows immediately from the definitions of $|\mathcal{S} - \mathcal{S}'|_\infty$ (and (3-5)), d, and $[z]_\epsilon$ when $z \in K$, and from (3-7) applied to $z = 0$ and the invariance of d under ι when $z = \infty$. Moreover, for every $z \in K$ and every $\epsilon > 0$,

$$\sup_{\mathcal{S} \in \operatorname{supp}[z]_\epsilon} |\log[\mathcal{S}, \infty]_{\mathrm{can}} - \log[z, \infty]| \leq \epsilon \tag{3-8}$$

by a direct computation of $\log[\,\cdot\,, \infty]_{\mathrm{can}} - \log[z, \infty]$ on K, using that $\operatorname{supp}[z]_\epsilon \subset \{\mathcal{S} \in \mathbb{P}^1 : |\mathcal{S} - z|_\infty \leq \epsilon\}$ and the density of K in A^1.

Lemma 3.2. *Let g be a continuous weight on P^1 having a modulus of continuity η on $(\mathsf{P}^1, \mathsf{d})$. Then for every $\epsilon > 0$ and every $z, w \in \mathbb{P}^1$,*

$$\begin{aligned} &\int_{\mathsf{P}^1\times\mathsf{P}^1} \Phi_g\,\mathrm{d}([z]_\epsilon \times [w]_\epsilon) \\ &\qquad \geq \begin{cases} \Phi_g(z, w) - 2\epsilon - 2\eta(\epsilon) & \textit{if } z \neq w, \\ C_{\mathrm{abs}} + \log\epsilon - 2\epsilon + 2\log[z, \infty] - 2\eta(\epsilon) - 2g(z) & \textit{if } z = w \in K, \\ C_{\mathrm{abs}} + \log\epsilon - 2\epsilon - 2\eta(\epsilon) - 2g(\infty) & \textit{if } z = w = \infty. \end{cases} \end{aligned} \tag{3-9}$$

Proof. Since $\Phi_g(\mathcal{S}, \mathcal{S}') = \log[\mathcal{S}, \mathcal{S}']_{\mathrm{can}} - g(\mathcal{S}) - g(\mathcal{S}')$ on $\mathsf{P}^1 \times \mathsf{P}^1$, by (3-7), we can assume $g \equiv 0$ (and $\eta \equiv 0$) on P^1 without loss of generality. For every $z, w \in K$, by the definition (3-2) of $|\mathcal{S}-\mathcal{S}'|_\infty$ and (3-8),

$$\int_{\mathsf{P}^1\times\mathsf{P}^1}\log\,[\mathcal{S},\mathcal{S}']_{\mathrm{can}}\,\mathrm{d}([z]_\epsilon\times[w]_\epsilon)(\mathcal{S},\mathcal{S}')$$
$$\geq\int_{\mathsf{A}^1\times\mathsf{A}^1}\log|\mathcal{S}-\mathcal{S}'|_\infty\,\mathrm{d}([z]_\epsilon\times[w]_\epsilon)(\mathcal{S},\mathcal{S}')-2\epsilon+\log\,[z,\infty]+\log\,[w,\infty],$$

which with the estimate (3-6) yields (3-9) (for $g\equiv\eta\equiv 0$) in this case. The estimate (3-9) (for $g\equiv\eta\equiv 0$) in the case $z=w=\infty$ follows from $[\infty]_\epsilon=\iota_*[0]_\epsilon$, $[\iota(\mathcal{S}),\iota(\mathcal{S}')]_{\mathrm{can}}=[\mathcal{S},\mathcal{S}']_{\mathrm{can}}$, and the estimate (3-9) for $z=w=0$.

There remains the case that $z=\infty$ and $w\in K$ (so $z\neq w$). If K is nonarchimedean, then for every $w\in K$ and $\epsilon>0$, the equalities $[\infty]_\epsilon=\iota_*[0]_\epsilon$ and $[\iota(\mathcal{S}),\iota(\mathcal{S}')]_{\mathrm{can}}=[\mathcal{S},\mathcal{S}']_{\mathrm{can}}$, together with the interpretation (3-4) of $[\mathcal{S},\mathcal{S}']_{\mathrm{can}}$, yield

$$\int_{\mathsf{P}^1\times\mathsf{P}^1}\log\,[\mathcal{S},\mathcal{S}']_{\mathrm{can}}\,\mathrm{d}([\infty]_\epsilon\times[w]_\epsilon)(\mathcal{S},\mathcal{S}')$$
$$=\int_{\mathsf{P}^1\times\mathsf{P}^1}\log\,[\mathcal{S},\mathcal{S}']_{\mathrm{can}}\,\mathrm{d}([0]_\epsilon\times\iota_*[w]_\epsilon)(\mathcal{S},\mathcal{S}')=\log\,[\pi_\epsilon(0),\iota(\pi_\epsilon(w))]_{\mathrm{can}}$$
$$\geq\log\,[0,\iota(w)]=\log\,[\infty,w]\geq\log\,[\infty,w]-2\epsilon,$$

which implies the estimate (3-9) (for $g\equiv\eta\equiv 0$) in the case $z=\infty$ and $w\in K$ when K is nonarchimedean. If K is archimedean, then for every $w\in K$ and every $r,r'>0$, we have

$$\int_0^{2\pi}\frac{\mathrm{d}\phi}{2\pi}\int_0^{2\pi}\log\left|(0+re^{i\theta})-\frac{1}{w+r'e^{i\phi}}\right|\frac{\mathrm{d}\theta}{2\pi}$$
$$=\int_0^{2\pi}\max\{-\log|w+r'e^{i\phi}|,\log r\}\frac{\mathrm{d}\phi}{2\pi}\geq-\int_0^{2\pi}\log\left|(w+r'e^{i\phi})-0\right|\frac{\mathrm{d}\phi}{2\pi},$$

so that for every $w\in K\cong\mathsf{A}^1$ and every $\epsilon>0$,

$$\int_{\mathsf{A}^1\times\mathsf{A}^1}\log|\mathcal{S}-\mathcal{S}'|_\infty\,\mathrm{d}([0]_\epsilon\times\iota_*[w]_\epsilon)(\mathcal{S},\mathcal{S}')$$
$$=\int_{\mathsf{A}^1\times\mathsf{A}^1}\log|\mathcal{S}-\iota(\mathcal{S}')|_\infty\,\mathrm{d}([0]_\epsilon\times[w]_\epsilon)(\mathcal{S},\mathcal{S}')\geq-\int_{\mathsf{A}^1}\log|\mathcal{S}'-0|_\infty\,\mathrm{d}[w]_\epsilon(\mathcal{S}').$$

On the other hand, for every $w\in K$ and every $\epsilon>0$, by the definition (2-1) of the chordal metric $[z,w]$ on $\mathbb{P}^1\cong\mathsf{P}^1$ (and $[0,\infty]=1$),

$$\int_{\mathsf{P}^1}\log\,[\mathcal{S}',\infty]_{\mathrm{can}}\,\mathrm{d}(\iota_*[w]_\epsilon)(\mathcal{S}')=\int_{\mathsf{P}^1}\log\,[\mathcal{S}',0]_{\mathrm{can}}\,\mathrm{d}[w]_\epsilon(\mathcal{S}')$$
$$=\int_{\mathsf{A}^1}\log|\mathcal{S}'-0|_\infty\,\mathrm{d}[w]_\epsilon(\mathcal{S}')+\int_{\mathsf{P}^1}\log\,[\mathcal{S}',\infty]_{\mathrm{can}}\,\mathrm{d}[w]_\epsilon(\mathcal{S}').$$

From these computations and (3-8), for every $w\in K$ and every $\epsilon>0$, we get

$$\begin{aligned}\int_{\mathsf{P}^1\times\mathsf{P}^1} &\log{[\mathcal{S},\mathcal{S}']_{\mathrm{can}}}\,\mathrm{d}([\infty]_\epsilon\times[w]_\epsilon)(\mathcal{S},\mathcal{S}')\\ &=\int_{\mathsf{P}^1\times\mathsf{P}^1}\log{[\mathcal{S},\mathcal{S}']_{\mathrm{can}}}\,\mathrm{d}([0]_\epsilon\times\iota_*[w]_\epsilon)(\mathcal{S},\mathcal{S}')\\ &\ge\int_{\mathsf{P}^1}\log{[\mathcal{S},\infty]_{\mathrm{can}}}\,\mathrm{d}[0]_\epsilon(\mathcal{S})+\int_{\mathsf{P}^1}\log{[\mathcal{S}',\infty]_{\mathrm{can}}}\,\mathrm{d}[w]_\epsilon(\mathcal{S}')\\ &\ge\log{[0,\infty]}+\log{[w,\infty]}-2\epsilon=\log{[w,\infty]}-2\epsilon,\end{aligned}$$

which implies the estimate (3-9) (for $g\equiv\eta\equiv 0$) in the case $z=\infty$ and $w\in K$ when K is archimedean. □

4. The negativity of regularized Fekete sums and a Cauchy–Schwarz inequality

Let K be an algebraically closed field that is complete with respect to a nontrivial absolute value $|\cdot|$. For every $\epsilon>0$ and every discrete measure ν on $\mathsf{P}^1=\mathsf{P}^1(K)$ whose support is in $\mathbb{P}^1=\mathbb{P}^1(K)$, the *$\epsilon$-regularization* of ν is

$$\nu_\epsilon:=\sum_{w\in\operatorname{supp}\nu}\nu(\{w\})[w]_\epsilon\quad\text{on }\mathsf{P}^1.$$

For every continuous weight g on P^1, let us call $(\nu_\epsilon,\nu_\epsilon)_g$ the *ϵ-regularized g-Fekete sum* with respect to this ν.

4.1. *C^1-regularity and the Dirichlet norm.*

Recall the description of P^1 given in Section 3.2. For nonarchimedean K, a function ϕ on $\mathsf{P}^1=\mathsf{P}^1(K)$ is *in* $C^1(\mathsf{P}^1)$ if

(i) ϕ is continuous on P^1 and locally constant except for a union $\mathcal{T}$ of at most finitely many segments in $\mathsf{H}^1=\mathsf{H}^1(K)$, which are oriented by the partial ordering $\succeq$ on P^1, and

(ii) the derivative ϕ' with respect to the length parameter induced by the hyperbolic metric ρ on each segment in $\mathcal{T}$ exists and is continuous on $\mathcal{T}$.

The *Dirichlet norm* of $\phi\in C^1(\mathsf{P}^1)$ is defined by $\langle\phi,\phi\rangle^{1/2}:=\left(\int_{\mathcal{T}}(\phi')^2\,\mathrm{d}\rho\right)^{1/2}$, where $\mathrm{d}\rho$ is the 1-dimensional Hausdorff measure on H^1 with respect to ρ (for details, see [Favre and Rivera-Letelier 2006, §5.5]). When K is archimedean, the C^1-regularity and the Dirichlet norm of a function ϕ on $\mathsf{P}^1\cong\mathbb{P}^1$ is defined with respect to the complex (or differentiable) structure of $\mathbb{P}^1$. For completeness, we include a proof of the following.

Proposition 4.1. *Every ϕ in $C^1(\mathsf{P}^1)$ is Lipschitz continuous on* $(\mathsf{P}^1,\mathsf{d})$.

Proof. When K is archimedean, this is obvious. Suppose that K is nonarchimedean and let $\phi\in C^1(\mathsf{P}^1)$. By definition, ϕ is locally constant on P^1 except for a union

$\mathcal{T}$ of at most finitely many segments in H^1, and is Lipschitz continuous on $\mathcal{T}$ with respect to ρ. The set $\mathcal{T}$ is compact in (H^1, ρ), and for every $\mathcal{S}, \mathcal{S}' \in \mathsf{H}^1$, by the definition (3-1) of d, (3-4), and (3-3), if $\mathcal{S}_{\text{can}} \succeq \mathcal{S} \succeq \mathcal{S}'$, then

$$\mathsf{d}(\mathcal{S}, \mathcal{S}') = \operatorname{diam} \mathcal{S} - \frac{\operatorname{diam} \mathcal{S} + \operatorname{diam} \mathcal{S}'}{2} = \frac{\operatorname{diam} \mathcal{S} - \operatorname{diam} \mathcal{S}'}{2} \geq \frac{\operatorname{diam} \mathcal{S}'}{2} \rho(\mathcal{S}, \mathcal{S}'),$$

and similarly, if $\mathcal{S}_{\text{can}} \preceq \mathcal{S} \preceq \mathcal{S}'$, then $\mathsf{d}(\mathcal{S}, \mathcal{S}') \geq \rho(\mathcal{S}, \mathcal{S}')/(2 \operatorname{diam} \mathcal{S}')$. Hence we conclude that ϕ is also Lipschitz continuous on $\mathcal{T}$ with respect to d, and in turn on the whole P^1 with respect to d. □

The Lipschitz constant of a Lipschitz continuous function ϕ on $(\mathsf{P}^1, \mathsf{d})$ is denoted by $\operatorname{Lip}(\phi)$.

Remark 4.2. When K is archimedean (so $\mathsf{P}^1 \cong \mathbb{P}^1$), we have $\langle \phi, \phi \rangle^{1/2} \leq \operatorname{Lip}(\phi)$ for every $\phi \in C^1(\mathbb{P}^1)$. Moreover, every Lipschitz continuous function ϕ on $(\mathbb{P}^1, [z, w])$ is approximated by functions in $C^1(\mathbb{P}^1)$ in the Lipschitz norm.

4.2. ***The negativity of $(\nu_\epsilon, \nu_\epsilon)_g$ and a Cauchy–Schwarz inequality.*** For every Radon measure μ on P^1 satisfying $\mu(\mathsf{P}^1) = 0$, if the chordal potential of μ, which is defined by $\mathcal{S} \mapsto \int_{\mathsf{P}^1} \log [\mathcal{S}, \mathcal{S}']_{\text{can}} \, d\mu(\mathcal{S}')$, is continuous on P^1, then we have the *positivity* property $\int_{\mathsf{P}^1 \times \mathsf{P}^1} (-\log |\mathcal{S} - \mathcal{S}'|_\infty) \, d(\mu \times \mu)(\mathcal{S}, \mathcal{S}') \geq 0$ (see [Favre and Rivera-Letelier 2006, §2.5 and §4.5]) and in fact the *Cauchy–Schwarz inequality*

$$\left| \int_{\mathsf{P}^1} \phi \, d\mu \right|^2 \leq \langle \phi, \phi \rangle \cdot \int_{\mathsf{P}^1 \times \mathsf{P}^1} (-\log |\mathcal{S} - \mathcal{S}'|_\infty) \, d(\mu \times \mu)(\mathcal{S}, \mathcal{S}') \tag{4-1}$$

for every test function $\phi \in C^1(\mathsf{P}^1)$ (see [Favre and Rivera-Letelier 2006, (32) and (33)]).

In particular, for every $\epsilon > 0$, every normalized weight g on P^1, every test function $\phi \in C^1(\mathsf{P}^1)$, and every discrete measure ν on P^1 whose support is in $\mathbb{P}^1$, the computation

$$\begin{aligned} 0 &\leq \int_{\mathsf{P}^1 \times \mathsf{P}^1} (-\log |\mathcal{S} - \mathcal{S}'|_\infty) \, d((\nu_\epsilon - (\nu(\mathsf{P}^1))\mu^g) \times (\nu_\epsilon - (\nu(\mathsf{P}^1))\mu^g))(\mathcal{S}, \mathcal{S}') \\ &= \int_{\mathsf{P}^1 \times \mathsf{P}^1} (-\Phi_g) \, d((\nu_\epsilon - (\nu(\mathsf{P}^1))\mu^g) \times (\nu_\epsilon - (\nu(\mathsf{P}^1))\mu^g)) = -(\nu_\epsilon, \nu_\epsilon)_g \end{aligned}$$

(recalling $U_{g,\mu^g} \equiv 0$ on P^1) yields not only the *negativity* $(\nu_\epsilon, \nu_\epsilon)_g \leq 0$ but, with the Cauchy–Schwarz inequality (4-1) and the triangle inequality, also the estimate

$$\begin{aligned} \left| \int_{\mathsf{P}^1} \phi \, d\big(\nu - \nu(\mathsf{P}^1)\mu^g\big) \right| &= \left| \int_{\mathsf{P}^1} \phi \, d\big((\nu - \nu_\epsilon) + (\nu_\epsilon - (\deg \nu)\mu^g)\big) \right| \\ &\leq (\deg \nu) \operatorname{Lip}(\phi)\epsilon + \langle \phi, \phi \rangle^{1/2} \cdot (-(\nu_\epsilon, \nu_\epsilon)_g)^{1/2}. \end{aligned} \tag{4-2}$$

5. Computations of Fekete sums $(\mathcal{Z}, \mathcal{Z})_g$

Let k be a field. For a k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$, set

$$D^*(\mathcal{Z}|\bar{k}) := \prod_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} \prod_{w' \in \operatorname{supp} \mathcal{Z} \setminus \{w, \infty\}} (w - w')^{(\operatorname{ord}_w \mathcal{Z})(\operatorname{ord}_{w'} \mathcal{Z})} \in \bar{k} \setminus \{0\},$$

which is in fact in $k \setminus \{0\}$ by Theorem 7 if $\mathcal{Z}$ is on $\mathbb{P}^1(k_s)$. For every $P \in \bigcup_{d \in \mathbb{N}} k[p_0, p_1]_d$, let $L(P(1, \cdot)) \in k \setminus \{0\}$ be the coefficient of the maximal degree term of $P(1, z) \in k[z]$ (appearing in Section 2.3).

Lemma 5.1. *Let k be a field. Let $\mathcal{Z}$ be a k-effective divisor on $\mathbb{P}^1(\bar{k})$ represented by $P \in \bigcup_{d \in \mathbb{N}} k[p_0, p_1]_d$, and let $(q_j^P)_{j=1}^{\deg P}$ be a sequence in $\bar{k}^2 \setminus \{0\}$ giving a factorization (2-9) of P. For each $j \in \{1, 2, \dots, \deg P\}$, set $q_j^P = ((q_j^P)_0, (q_j^P)_0)$ and $z_j := \pi(q_j^P) \in \mathbb{P}^1(\bar{k})$. Suppose $(q_j^P)_{j=1}^{\deg P}$ is normalized with respect to a distinguished zero $w_0 \in \mathbb{P}^1(\bar{k})$ of P so that for each $j \in \{1, 2, \dots, \deg P\}$,*

$$\begin{cases} (q_j^P)_0 = 1 & \text{if } z_j \notin \{w_0, \infty\}, \\ (q_j^P)_1 = 1 & \text{if } w_0 \neq z_j = \infty. \end{cases} \tag{5-1}$$

Then

$$L(P(1, \cdot)) = (-1)^{\deg P - \deg_\infty P} \cdot \begin{cases} \prod_{j: z_j = w_0} (q_j^P)_0 & \text{if } w_0 \neq \infty, \\ \prod_{j: z_j = w_0} (q_j^P)_1 & \text{if } w_0 = \infty, \end{cases} \tag{5-2}$$

and

$$\begin{aligned} &\prod_{j=1}^{\deg P} \prod_{i: z_i \neq z_j} (q_i^P \wedge q_j^P) \\ &\qquad = (-1)^{\deg_\infty P (\deg P - \deg_\infty P)} \cdot L(P(1, \cdot))^{2(\deg P - \deg_{w_0} P)} \cdot D^*(\mathcal{Z}|\bar{k}). \end{aligned} \tag{5-3}$$

Proof. Without normalizing the sequence $(q_j^P)_{j=1}^{\deg P}$ we have, by direct computation,

$$\begin{aligned} &\prod_{j=1}^{\deg P} \prod_{i: z_i \neq z_j} (q_i^P \wedge q_j^P) \\ &= \prod_{\substack{j: z_j = \infty \\ i: z_i \neq \infty}} \big((q_i^P)_0 (q_j^P)_1\big) \cdot \prod_{\substack{j: z_j \neq \infty \\ i: z_i = \infty}} \big(-(q_i^P)_1 (q_j^P)_0\big)\big) \cdot \prod_{\substack{j: z_j \neq \infty \\ i: z_i \notin \{z_j, \infty\}}} \big((q_i^P)_0 (q_j^P)_0 (z_j - z_i)\big) \\ &= (-1)^{\deg_\infty P (\deg P - \deg_\infty P)} \cdot \Bigg(\prod_{j: z_j = \infty} \Bigg((q_j^P)_1^{\deg P - \deg_\infty P} \cdot \prod_{i: z_i \neq \infty} (q_i^P)_0 \Bigg) \Bigg)^2 \\ &\qquad \cdot \Bigg(\prod_{j: z_j \neq \infty} \Bigg((q_j^P)_0^{\deg P - \deg_\infty P - \deg_{z_j} P} \cdot \prod_{i: z_i \notin \{z_j, \infty\}} (q_i^P)_0 \Bigg) \Bigg) \cdot D^*(\mathcal{Z}|\bar{k}). \end{aligned} \tag{5-4}$$

Let us normalize (q_j^P) so that the normalization (5-1) holds with respect to a distinguished zero $w_0 \in \mathbb{P}^1(\bar{k})$ of P. Then (5-2) follows from

$$L(P(1,\cdot)) = (-1)^{\deg P - \deg_\infty P} \cdot \Bigl(\prod_{j:z_j=\infty} (q_j^P)_1\Bigr)\Bigl(\prod_{j:z_j\neq\infty} (q_j^P)_0\Bigr)$$

and the normalization (5-1).

Let us show (5-3). If $w_0 = \infty$, then under the normalization (5-1), the equality (5-4) yields

$$\prod_{j=1}^{\deg P} \prod_{i:z_i\neq z_j} (q_i^P \wedge q_j^P) = (-1)^{\deg_\infty P(\deg P - \deg_\infty P)} \cdot \Bigl(\prod_{j:z_j=\infty} (q_j^P)_1\Bigr)^{2(\deg P - \deg_\infty P)} \cdot 1 \cdot D^*(\mathcal{Z}|\bar{k}),$$

which with (5-2) implies (5-3) when $w_0 = \infty$. If $w_0 \neq \infty$, then under the normalization (5-1), the equality (5-4) yields

$$\begin{aligned}\prod_{j=1}^{\deg P} \prod_{i:z_i\neq z_j} (q_i^P \wedge q_j^P) &= (-1)^{\deg_\infty P(\deg P - \deg_\infty P)} \cdot \Bigl(\prod_{i:z_i=w_0} (q_i^P)_0\Bigr)^{2\deg_\infty P} \\ &\quad \cdot \Bigl(\prod_{j:z_j=w_0} \bigl((q_j^P)_0^{\deg P - \deg_\infty P - \deg_{z_j} P} \cdot 1\bigr)\Bigr) \\ &\quad \cdot \Bigl(\prod_{j:z_j\notin\{w_0,\infty\}} \Bigl(1 \cdot \prod_{i:z_i=w_0} (q_i^P)_0\Bigr)\Bigr) \cdot D^*(\mathcal{Z}|\bar{k}) \\ &= (-1)^{\deg_\infty P(\deg P - \deg_\infty P)} \cdot \Bigl(\prod_{i:z_i=w_0} (q_i^P)_0\Bigr)^{2\deg_\infty P + 2(\deg P - \deg_\infty P - \deg_{w_0} P)} \cdot D^*(\mathcal{Z}|\bar{k}),\end{aligned}$$

which with (5-2) implies (5-3) when $w_0 \neq \infty$. $\square$

Lemma 5.2 (local computation). *Let k be a field and K an algebraic and metric augmentation of k* (see Section 2.2). *For every continuous weight g on $\mathsf{P}^1 = \mathsf{P}^1(K)$ and every k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$ represented by a homogeneous polynomial $P \in \bigcup_{d\in\mathbb{N}} k[p_0, p_1]_d$, we have*

$$(\mathcal{Z},\mathcal{Z})_g + 2 \cdot \sum_{w\in \operatorname{supp} \mathcal{Z}\setminus\{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log[w,\infty] - 2 \cdot \sum_{w\in\operatorname{supp}\mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 g(w) = 2(\deg\mathcal{Z}) \log|L(P(1,\cdot))| + \log|D^*(\mathcal{Z}|\bar{k})| - 2(\deg\mathcal{Z}) M_g(P). \tag{5-5}$$

Proof. Let $\mathcal{Z}$ and P be as in the statement and let $(q_j^P)_{j=1}^{\deg P}$ be a sequence in $\bar{k}^2 \setminus \{0\}$ giving a factorization (2-9) of P and satisfying the normalization (5-1) with respect to a distinguished zero $w_0 \in \mathbb{P}^1(\bar{k})$ of P. Set $z_j := \pi(q_j^P) \in \mathbb{P}^1(\bar{k})$ for each $j \in \{1, 2, \dots, \deg P\}$. Since by definition

$$\Phi_g(z, z') = \log [z, z'] - g(z) - g(z')$$

on $\mathbb{P}^1(K) \times \mathbb{P}^1(K)$, we have

$$(\mathcal{Z}, \mathcal{Z})_g = \log\biggl(\prod_{j=1}^{\deg P} \prod_{i: z_i \neq z_j} |q_i^P \wedge q_j^P| \biggr) - 2 \cdot \sum_{j=1}^{\deg P} \sum_{i: z_i \neq z_j} (g(z_i) + \log \|q_i^P\|);$$

by (5-3),

$$\log\biggl(\prod_{j=1}^{\deg P} \prod_{i: z_i \neq z_j} |q_i^P \wedge q_j^P| \biggr) = 2(\deg P - \deg_{w_0} P) \log \bigl| L(P(1, \cdot)) \bigr| + \log \bigl| D^*(\mathcal{Z}|\bar{k}) \bigr|,$$

and we also have

$$\begin{aligned}
\sum_{j=1}^{\deg P} \sum_{i: z_i \neq z_j} \bigl(g(z_i) + \log \|q_i^P\|\bigr)
&= \sum_{j=1}^{\deg P} \sum_{i=1}^{\deg P} \bigl(g(z_i) + \log \|q_i^P\|\bigr) - \sum_{j=1}^{\deg P} \sum_{i: z_i = z_j} \bigl(g(z_i) + \log \|q_i^P\|\bigr) \\
&= (\deg P) M_g(P) - \sum_{j=1}^{\deg P} (\deg_{z_j} P) g(z_j) - \sum_{j=1}^{\deg P} \sum_{i: z_i = z_j} \log \|q_i^P\|,
\end{aligned}$$

where the final equality is by the definition (2-10) of $M_g(P)$. Hence

$$\begin{aligned}
(\mathcal{Z}, \mathcal{Z})_g &= 2(\deg P) \log \bigl| L(P(1, \cdot)) \bigr| + \log \bigl| D^*(\mathcal{Z}|\bar{k}) \bigr| - 2(\deg P) M_g(P) \\
&\quad + 2 \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 g(w) - 2\biggl((\deg_{w_0} P) \log |L(P(1, \cdot))| - \sum_{j=1}^{\deg P} \sum_{i: z_i = z_j} \log \|q_i^P\| \biggr).
\end{aligned}$$

For each $j \in \{1, 2, \dots, \deg P\}$, also set $q_j^P = ((q_j^P)_0, (q_j^P)_0)$. If $\infty \notin \operatorname{supp} \mathcal{Z}$, then $w_0 \neq \infty$, and by the normalization (5-1) and the equality (5-2),

$$\begin{aligned}
(\deg_{w_0} P) \log \bigl| L(P(1, \cdot)) \bigr| - \sum_{j=1}^{\deg P} \sum_{i: z_i = z_j} \log \|q_i^P\|
&= -\sum_{j=1}^{\deg P} \sum_{i: z_i = z_j} \bigl(\log \|q_i^P\| - \log |(q_i^P)_0| \bigr) = \sum_{j=1}^{\deg P} \sum_{i: z_i = z_j} \log [z_i, \infty] \\
&= \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty] = \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty].
\end{aligned}$$

If $\infty \in \operatorname{supp} \mathcal{Z}$, then we can set $w_0 = \infty$, and by the normalization (5-1) and the equality (5-2) (and $q_i^P = (q_i^P)_1 \cdot (0, 1)$ when $z_i = \infty$),

$$(\deg_{w_0} P) \log \left| L(P(1, \cdot)) \right| - \sum_{j=1}^{\deg P} \sum_{i : z_i = z_j} \log \| q_i^P \|$$
$$= - \sum_{j : z_j = \infty} \sum_{i : z_i = z_j} \left(\log \| q_i^P \| - \log |(q_i^P)_1| \right) - \sum_{j : z_j \neq \infty} \sum_{i : z_i = z_j} \left(\log \| q_i^P \| - \log |(q_i^P)_0| \right)$$
$$= \sum_{j : z_j \neq \infty} \sum_{i : z_i = z_j} \log [z_i, \infty] = \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty].$$

This completes the proof. □

Lemma 5.3 (global computation). *Let k be a product formula field and k_s the separable closure of k in $\bar{k}$. Then for every adelic continuous weight $g = \{g_v : v \in M_k\}$ and every k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(k_s)$,*

$$\sum_{v \in M_k} N_v \left((\mathcal{Z}, \mathcal{Z})_{g_v} + 2 \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty]_v \right) = -2(\deg \mathcal{Z})^2 h_g(\mathcal{Z}) + 2 \sum_{v \in M_k} N_v \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 g_v(w). \tag{5-6}$$

Proof. Let $P \in \bigcup_{d \in \mathbb{N}} k[p_0, p_1]_d$ be a representative of $\mathcal{Z}$. Summing up the product of N_v and (5-5) (for this P) over all $v \in M_k$, we have

$$\sum_{v \in M_k} N_v \left((\mathcal{Z}, \mathcal{Z})_{g_v} + 2 \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty]_v - 2 \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 g_v(w) \right) = -2(\deg \mathcal{Z})^2 h_g(\mathcal{Z})$$

by the product formula (PF) (since $L(P(1, \cdot)) \in k \setminus \{0\}$ and, under the assumption that $\mathcal{Z}$ is on $\mathbb{P}^1(k_s)$, $D^*(\mathcal{Z}|\bar{k}) \in k \setminus \{0\}$) and the definition (1-1) of $h_g(\mathcal{Z})$. □

6. Estimates of regularized Fekete sums $(\mathcal{Z}_\epsilon, \mathcal{Z}_\epsilon)_g$

6.1. ***Local estimate.*** Let k be a field and K an algebraic and metric augmentation of k. Let $\mathcal{Z}$ be a k-effective divisor on $\mathbb{P}^1(\bar{k})$, which we regard as the Radon measure

$$\sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z}) \delta_w$$

on $\mathsf{P}^1 = \mathsf{P}^1(K)$, and let g be a continuous weight on P^1 such that g is a $1/\kappa$-Hölder continuous function on $(\mathsf{P}^1, \mathsf{d})$ for some $\kappa \geq 1$ having the $1/\kappa$-Hölder constant $C(g) \geq 0$.

Lemma 6.1. *For every* $\epsilon > 0$,

$$(\mathcal{Z}_\epsilon, \mathcal{Z}_\epsilon)_g \geq (\mathcal{Z}, \mathcal{Z})_g + 2 \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty] - 2 \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 g(w) \\ + (C_{\mathrm{abs}} + \log \epsilon) \cdot (\mathcal{Z} \times \mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(\bar{k})}) - 2(\deg \mathcal{Z})^2 (\epsilon + C(g)\epsilon^{1/\kappa}).$$

Proof. Set $\eta(\epsilon) = C(g)\epsilon^{1/\kappa}$. For every $\epsilon > 0$, using (3-9),

$$\begin{aligned}
&(\mathcal{Z}_\epsilon, \mathcal{Z}_\epsilon)_g - (\mathcal{Z}, \mathcal{Z})_g \\
&\quad = \int_{\mathsf{P}^1 \times \mathsf{P}^1} \Phi_g \, \mathrm{d}(\mathcal{Z}_\epsilon \times \mathcal{Z}_\epsilon) - \int_{\mathsf{P}^1 \times \mathsf{P}^1 \setminus \operatorname{diag}_{\mathbb{P}^1(K)}} \Phi_g \, \mathrm{d}(\mathcal{Z} \times \mathcal{Z}) \\
&\quad = \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 \int_{\mathsf{P}^1 \times \mathsf{P}^1} \Phi_g \, \mathrm{d}([w]_\epsilon \times [w]_\epsilon) \\
&\qquad + \sum_{(z,w) \in \mathbb{P}^1 \times \mathbb{P}^1 \setminus \operatorname{diag}_{\mathbb{P}^1}} \left(\int_{\mathsf{P}^1 \times \mathsf{P}^1} \Phi_g(\mathcal{S}, \mathcal{S}') \, \mathrm{d}([z]_\epsilon \times [w]_\epsilon)(\mathcal{S}, \mathcal{S}') - \Phi_g(z, w) \right) \\
&\quad \geq \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \big(C_{\mathrm{abs}} + \log \epsilon - 2\epsilon + 2 \log [w, \infty] - 2\eta(\epsilon) - 2g(w) \big) \\
&\qquad\qquad + \big(\mathcal{Z}(\{\infty\})\big)^2 (C_{\mathrm{abs}} + \log \epsilon - 2\epsilon - 2\eta(\epsilon) - 2g(\infty)) \\
&\qquad\qquad + \big((\deg \mathcal{Z})^2 - (\mathcal{Z} \times \mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(\bar{k})})\big)(-2\epsilon - 2\eta(\epsilon)) \\
&\quad = \big((\mathcal{Z} \times \mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(\bar{k})})\big)\big(C_{\mathrm{abs}} + \log \epsilon - 2\epsilon - 2\eta(\epsilon)\big) \\
&\qquad\qquad + 2 \sum_{w \in \operatorname{supp} \mathcal{Z} \setminus \{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log [w, \infty] - 2 \sum_{w \in \operatorname{supp} \mathcal{Z}} (\operatorname{ord}_w \mathcal{Z})^2 g(w) \\
&\qquad\qquad + \big((\deg \mathcal{Z})^2 - (\mathcal{Z} \times \mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(\bar{k})})\big)(-2\epsilon - 2\eta(\epsilon)),
\end{aligned}$$

which completes the proof. □

6.2. ***Global estimate.*** Let k be a product formula field, and $\mathcal{Z}$ a k-effective divisor on $\mathbb{P}^1(k_s)$. Let $g = \{g_v : v \in M_k\}$ be a placewise Hölder continuous adelic normalized weight, so for every $v \in M_k$, g_v is a normalized weight on $\mathsf{P}^1(\mathbb{C}_v)$ and is a $1/\kappa_v$-Hölder continuous function on $(\mathsf{P}^1(\mathbb{C}_v), \mathsf{d}_v)$ for some $\kappa_v \geq 1$ having the $1/\kappa_v$-Hölder constant $C(g_v) \geq 0$.

Lemma 6.2. *For every* $v_0 \in M_k$ *and every* $\epsilon > 0$,

$$N_{v_0}(\mathcal{Z}_\epsilon, \mathcal{Z}_\epsilon)_{g_{v_0}} \geq -2(\deg \mathcal{Z})^2 h_g(\mathcal{Z}) + (C_{\mathrm{abs}} + \log \epsilon) \cdot (\mathcal{Z} \times \mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(k_s)}) \cdot \sum_{v \in E_g \cup \{v_0\}} N_v \\ - 2 (\deg \mathcal{Z})^2 \sum_{v \in E_g \cup \{v_0\}} N_v (\epsilon + C(g_v)\epsilon^{1/\kappa_{v_0}}).$$

Proof. Fix $v_0 \in M_k$. We use, for every $v \in M_k$, the notation

$$W_v := (\mathcal{Z},\mathcal{Z})_{g_v} + 2\sum_{w\in\operatorname{supp}\mathcal{Z}\setminus\{\infty\}} (\operatorname{ord}_w \mathcal{Z})^2 \log[w,\infty]_v - 2\sum_{w\in\operatorname{supp}\mathcal{Z}} (\operatorname{ord}_w\mathcal{Z})^2 g_v(w).$$

Since $(\mathcal{Z}_\epsilon,\mathcal{Z}_\epsilon)_{g_v}\le 0$ for every $\epsilon>0$ and every $v\in M_k$ (see Section 4.2), using also Lemma 6.1, we have

$$\begin{aligned}N_{v_0}(\mathcal{Z}_\epsilon,\mathcal{Z}_\epsilon)_{g_{v_0}} &\ge \sum_{v\in E_g\cup\{v_0\}} N_v(\mathcal{Z}_\epsilon,\mathcal{Z}_\epsilon)_{g_{v_0}}\\ &\ge \sum_{v\in E_g\cup\{v_0\}} N_v W_v + (C_{\mathrm{abs}}+\log\epsilon)\cdot(\mathcal{Z}\times\mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(k_s)})\cdot\sum_{v\in E_g\cup\{v_0\}} N_v\\ &\qquad - 2(\deg\mathcal{Z})^2\sum_{v\in E_g\cup\{v_0\}} N_v(\epsilon + C(g_v)\epsilon^{1/\kappa_{v_0}}).\end{aligned}$$

Moreover, since for every $v\in M_k\setminus E_g$, $g_v\equiv 0$ on $\mathsf{P}^1(\mathbb{C}_v)$ and $(\mathcal{Z},\mathcal{Z})_{g_v}\le 0$, using also (5-6), we have

$$\sum_{v\in E_g\cup\{v_0\}} N_v W_v \ge \sum_{v\in M_k} N_v W_v = -2(\deg\mathcal{Z})^2 h_g(\mathcal{Z}),$$

which completes the proof. $\square$

7. Proofs of Theorems 1 and 2

Proof of Theorem 1. Fix $v_0\in M_k$. For every $v\in M_k$, g_v is a $1/\kappa_v$-Hölder continuous function on $(\mathsf{P}^1(\mathbb{C}_v),\mathsf{d}_v)$ for some $\kappa_v\ge 1$ having the $1/\kappa_v$-Hölder constant $C(g_v)\ge 0$. Set $\epsilon = 1/(\deg\mathcal{Z})^{2\kappa_{v_0}}$. For every test function $\phi\in C^1(\mathsf{P}^1(\mathbb{C}_{v_0}))$, by (4-2) and Lemma 6.2,

$$\begin{aligned}\left|\int_{\mathsf{P}^1(\mathbb{C}_{v_0})}\phi\,\mathrm{d}\Big(\frac{\mathcal{Z}}{\deg\mathcal{Z}}-\mu^g_{v_0}\Big)\right| &\le \frac{\operatorname{Lip}(\phi)_{v_0}}{(\deg\mathcal{Z})^{2\kappa_0}} + \frac{\langle\phi,\phi\rangle_{v_0}^{1/2}}{N_{v_0}^{1/2}}\\ &\cdot\Bigg(2\cdot h_g(\mathcal{Z}) + (-C_{\mathrm{abs}}+2\kappa_{v_0}\log\deg\mathcal{Z})\cdot\frac{(\mathcal{Z}\times\mathcal{Z})(\operatorname{diag}_{\mathbb{P}^1(k_s)})}{(\deg\mathcal{Z})^2}\cdot\sum_{v\in E_g\cup\{v_0\}} N_v\\ &\qquad + 2\sum_{v\in E_g\cup\{v_0\}} N_v\Big(\frac{1}{(\deg\mathcal{Z})^{2\kappa_0}}+\frac{C(g_v)}{(\deg\mathcal{Z})^2}\Big)\Bigg)^{1/2},\end{aligned}$$

which completes the proof. $\square$

Proof of Theorem 2. Fix $v_0\in M_k$. For every $n\in\mathbb{N}$, we have $(\mathcal{Z}_n,\mathcal{Z}_n)_{g_v}\le 0$ if $v\in M_k\setminus E_g$. Hence by (2-8), (5-6), and the assumption that $V_{g_v}=0$ for every

$v \in M_k$, we obtain

$$N_{v_0}\frac{(\mathcal{Z}_n,\mathcal{Z}_n)_{g_{v_0}}}{(\deg\mathcal{Z}_n)^2}+\#E_g\cdot o(1)\geq\sum_{v\in M_k}N_v\frac{(\mathcal{Z}_n,\mathcal{Z}_n)_{g_v}}{(\deg\mathcal{Z}_n)^2}$$
$$\geq -2\cdot h_g(\mathcal{Z}_n)-2\frac{(\mathcal{Z}_n\times\mathcal{Z}_n)(\mathrm{diag}_{\mathbb{P}^1(k_s)})}{(\deg\mathcal{Z}_n)^2}\sum_{v\in E_g}N_v\sup_{\mathsf{P}^1(\mathbb{C}_v)}|g_v|\quad\text{as } n\to\infty;$$

thus, under the assumption that $(\mathcal{Z}_n)$ has both small diagonals and small g-heights, we have $\liminf_{n\to\infty}(\mathcal{Z}_n,\mathcal{Z}_n)_{g_{v_0}}/(\deg\mathcal{Z}_n)^2\geq 0=V_{g_{v_0}}$. Hence (2-7) holds for g_{v_0} and $(\mathcal{Z}_n)$, and the proof is complete. □

8. Nonarchimedean and complex dynamics

Fact 8.1. Let k be a field. For a rational function $\phi\in k(z)$, we call

$$F_\phi=((F_\phi)_0,(F_\phi)_1)\ \in\bigcup_{d\in\mathbb{N}\cup\{0\}}(k[p_0,p_1]_d\times k[p_0,p_1]_d)$$

a *lift* of ϕ if $\pi\circ F_\phi=\phi\circ\pi$ on $k^2\setminus\{0\}$ and, in addition, $F_\phi^{-1}(0)=\{0\}$ when $\deg\phi>0$. The latter nondegeneracy condition is equivalent to the nonvanishing of $\mathrm{Res}(F_\phi):=\mathrm{Res}((F_\phi)_0,(F_\phi)_1)$; for the definition of the homogeneous resultant $\mathrm{Res}(P,Q)\in k$ for $P,Q\in\bigcup_{d\in\mathbb{N}\cup\{0\}}k[p_0,p_1]_d$, see, e.g., [Silverman 2007, §2.4]. Such a lift F_ϕ of ϕ is unique up to multiplication in k^*, and is in fact in $k[p_0,p_1]_{\deg\phi}\times k[p_0,p_1]_{\deg\phi}$.

Let K be an algebraically closed field that is complete with respect to a nontrivial absolute value $|\cdot|$.

8.1. ***The dynamical Green function g_f on P^1.*** For the foundation of a potential-theoretical study of dynamics on the Berkovich projective line, see [Baker and Rumely 2010; Favre and Rivera-Letelier 2010] for nonarchimedean K and, e.g., [Berteloot and Mayer 2001, §VIII] for archimedean K ($\cong\mathbb{C}$).

Fact 8.2. Let $\phi\in K(z)$ be a rational function of degree $d_0\in\mathbb{N}\cup\{0\}$. The action of ϕ on $\mathbb{P}^1=\mathbb{P}^1(K)$ uniquely extends to a continuous endomorphism on $\mathsf{P}^1=\mathsf{P}^1(K)$. When $d_0>0$, the extended ϕ is surjective, open, and discrete and preserves $\mathbb{P}^1$ and $\mathsf{H}^1=\mathsf{H}^1(K)$, the local degree function $z\mapsto\deg_z\phi$ on $\mathbb{P}^1$ also canonically extends to P^1, and the (mapping) degree of the extended $\phi:\mathsf{P}^1\to\mathsf{P}^1$ still equals d_0 (see [Baker and Rumely 2010, §2.3, §9; Benedetto 2010, §6.3]): in particular, the extended action of ϕ on P^1 induces a push-forward ϕ_* and a pullback ϕ^* on the spaces of continuous functions and of Radon measures on P^1. When $d_0=0$, the extended ϕ is still constant, and we set $\phi^*\mu:=0$ on P^1 for every Radon measure μ on P^1 by convention. Let $F_\phi\in K[p_0,p_1]_{\deg\phi}\times K[p_0,p_1]_{\deg\phi}$ be a lift of ϕ. The function

$$T_{F_\phi} := \log\big\| F_\phi(\,\cdot\,/\|\cdot\|)\big\| = \log\|F_\phi\| - (\deg\phi)\log\|\cdot\| \tag{8-1}$$

on $K^2 \setminus \{0\}$ descends to $\mathbb{P}^1$ and in turn extends continuously to P^1, satisfying $\Delta T_{F_\phi} = \phi^*\Omega_{\mathrm{can}} - (\deg\phi)\Omega_{\mathrm{can}}$ on P^1 (see, e.g., [Okuyama 2013a, Definition 2.8]). Moreover, ϕ is a Lipschitz continuous endomorphism on $(\mathsf{P}^1, \mathsf{d})$ and T_{F_ϕ} is a Lipschitz continuous function on $(\mathsf{P}^1, \mathsf{d})$ (for nonarchimedean K, see [Baker and Rumely 2010, Proposition 9.37]). For every $n \in \mathbb{N}$, the homogeneous polynomial $F_\phi^n \in K[p_0, p_1]_{\deg\phi^n} \times K[p_0, p_1]_{\deg\phi^n}$ is a lift of ϕ^n.

Let $f \in K(z)$ be a rational function of degree $d > 1$, and consider a lift $F \in K[p_0, p_1]_d \times K[p_0, p_1]_d$ of f. The uniform limit $g_F := \lim_{n\to\infty} T_{F^n}/d^n$ on P^1 exists, and more precisely, for every $n \in \mathbb{N}$,

$$\sup_{\mathsf{P}^1}\left| g_F - \frac{T_{F^n}}{d^n}\right| \le \frac{\sup_{\mathsf{P}^1}|T_F|}{d^n(d-1)}. \tag{8-2}$$

The limit g_F is called the *dynamical Green function of F* on P^1 and is a continuous weight on P^1. The probability Radon measure

$$\mu_f := \mu^{g_F} = \Delta g_F + \Omega_{\mathrm{can}} = \lim_{n\to\infty}\frac{(f^n)^*\Omega_{\mathrm{can}}}{d^n} \quad \text{weakly on } \mathsf{P}^1$$

is independent of the choice of F and satisfies $f^*\mu_f = d\cdot\mu_f$ on P^1. It is called the *f-equilibrium* (*or canonical*) *measure* on P^1. Moreover, g_F is a Hölder continuous function on $(\mathsf{P}^1, \mathsf{d})$ (for nonarchimedean K, see [Favre and Rivera-Letelier 2006, §6.6]). The remarkable *energy formula*

$$V_{g_F} = -\frac{\log|\mathrm{Res}\, F|}{d(d-1)} \tag{8-3}$$

was first established by DeMarco [2003] for archimedean K and was generalized to rational functions defined over a number field by Baker and Rumely [2006] (for a simple proof of (8-3) which also works for general K, see [Baker 2009, Appendix A] or [Okuyama and Stawiska 2011, Appendix]). The *dynamical Green function g_f of f on* P^1 is the unique normalized weight on P^1 such that $\mu^{g_f} = \mu_f$, i.e., for any lift F of f, $g_f \equiv g_F + V_{g_F}/2$ on P^1.

8.2. ***A Berkovich space version of the quasiperiodicity region $\mathcal{E}_f$.*** For nonarchimedean dynamics, see [Baker and Rumely 2010, §10; Favre and Rivera-Letelier 2010, §2.3; Benedetto 2010, §6.4]. For complex dynamics, see, e.g., [Milnor 2006].

Let $f \in K(z)$ be a rational function of degree > 1. The *Berkovich Julia set* of f is

$$\mathsf{J}(f) := \left\{\mathcal{S} \in \mathsf{P}^1 : \bigcap_{U \text{ open in } \mathsf{P}^1 \text{ containing } \mathcal{S}} \Big(\bigcup_{n\in\mathbb{N}} f^n(U)\Big) = \mathsf{P}^1 \setminus E(f)\right\},$$

where $E(f) := \left\{a \in \mathbb{P}^1 : \#\bigcup_{n\in\mathbb{N}} f^{-n}(a) < \infty\right\}$ is the *exceptional set* of f. The

Berkovich Fatou set is $\mathsf{F}(f) := \mathsf{P}^1 \setminus \mathsf{J}(f)$. By definition, $\mathsf{J}(f)$ is closed and $\mathsf{F}(f)$ is open in P^1, both $\mathsf{J}(f)$ and $\mathsf{F}(f)$ are totally invariant under f, and $\mathsf{J}(f)$ has no interior point unless $\mathsf{J}(f) = \mathsf{P}^1$. The *classical Julia set* $\mathsf{J}(f) \cap \mathbb{P}^1$ (resp. the *classical Fatou set* $\mathsf{F}(f) \cap \mathbb{P}^1$) coincides with the set of all nonequicontinuity points (resp. the region of local equicontinuity) of the family $\{f^n : n \in \mathbb{N}\}$ as a family of endomorphisms on $(\mathbb{P}^1, [z, w])$.

A component U of $\mathsf{F}(f)$ is called a *Berkovich Fatou component* of f, and is said to be *cyclic* under f if $f^n(U) = U$ for some $n \in \mathbb{N}$, which is called a *period* of U under f. Following [Fatou 1920, §28], a cyclic Berkovich Fatou component U of f having a period $n \in \mathbb{N}$ is called a *singular domain* of f if $f^n : U \to U$ is injective. Let $\mathcal{E}_f$ be the set of all points $\mathcal{S} \in \mathsf{P}^1$ having an open neighborhood V in P^1 such that $\liminf_{n\to\infty} \sup_{V\cap\mathbb{P}^1}[f^n, \mathrm{Id}] = 0$, which is a Berkovich space version of Rivera-Letelier's *quasiperiodicity region* of f. When K is archimedean, $\mathcal{E}_f$ *coincides with* the union of all singular domains of f, and when K is nonarchimedean, $\mathcal{E}_f$ is still open and forward invariant under f and *is contained in* the union of all singular domains of f (see [Okuyama 2013a, Lemma 4.4]).

The following function T_* is Rivera-Letelier's *iterative logarithm* of f on $\mathcal{E}_f \cap \mathbb{P}^1$, which is a nonarchimedean counterpart of the uniformization of a Siegel disk or a Herman ring of f.

Theorem 8.3 ([Rivera-Letelier 2003, §3.2, §4.2]. See also [Favre and Rivera-Letelier 2010, Théorème 2.15]). *Suppose that K is nonarchimedean and has characteristic* 0 *and residual characteristic p. Let $f \in K(z)$ be a rational function on $\mathbb{P}^1$ of degree > 1 and suppose that $\mathcal{E}_f \neq \emptyset$, which implies $p > 0$ by* [Favre and Rivera-Letelier 2010, Lemme 2.14]. *Then for every component Y of $\mathcal{E}_f$ not containing ∞, there are $k_0 \in \mathbb{N}$, a continuous action $T : \mathbb{Z}_p \times (Y \cap K) \ni (\omega, y) \mapsto T^\omega(y) \in Y \cap K$, and a nonconstant K-valued holomorphic function T_* on $Y \cap K$ such that for every $m \in \mathbb{Z}$, $(f^{k_0})^m = T^m$ on $Y \cap K$, that for every $\omega \in \mathbb{Z}_p$, T^ω is a biholomorphism on $Y \cap K$, and that for every $\omega_0 \in \mathbb{Z}_p$,*

$$\lim_{\mathbb{Z}_p \ni \omega \to \omega_0} \frac{T^\omega - T^{\omega_0}}{\omega - \omega_0} = T_* \circ T^{\omega_0} \quad \textit{locally uniformly on } Y \cap K. \tag{8-4}$$

8.3. ***The fundamental relationship between μ_f and $\mathsf{J}(f)$.*** If K is archimedean, the inclusion $\operatorname{supp}\mu_f \subset \mathsf{J}(f)$ is classical, but it is not trivial from the definition of $\mathsf{J}(f)$ when K is nonarchimedean. For an elementary proof, see [Okuyama 2013a, proof of Theorem 2.18]. Actually the equality $\operatorname{supp}\mu_f = \mathsf{J}(f)$ holds, but we will dispense with the reverse (and easier) inclusion $\mathsf{J}(f) \subset \operatorname{supp}\mu_f$.

9. Proofs of Theorems 3 and 4

Let k be a product formula field. The proof of the following is based not only on (PF) but also on elimination theory (and the strong triangle inequality).

Theorem 9.1 [Baker and Rumely 2006, Lemma 3.1]. *Let k be a product formula field. For every $\phi \in k(z)$ and every lift $F_\phi \in k[p_0, p_1]_{\deg \phi} \times k[p_0, p_1]_{\deg \phi}$ of ϕ, there exists a finite subset E_{F_ϕ} in M_k containing all the infinite places of k such that for every $v \in M_k \setminus E_{F_\phi}$, we have $|\mathrm{Res}\, F_\phi|_v = 1$ and $\|F_\phi(\cdot)\|_v = \|\cdot\|_v^{\deg\phi}$ on $\mathbb{C}_v^2$.*

Let $f \in k(z)$ be a rational function of degree > 1 and $F \in k[p_0, p_1]_d \times k[p_0, p_1]_d$ a lift of f. Then the family $\hat{g}_f = \{g_{f,v} : v \in M_k\}$ is an adelic normalized weight, where $g_{f,v}$ is the dynamical Green function of f on $\mathsf{P}^1(\mathbb{C}_v)$ for every $v \in M_k$. Indeed, letting $g_{F,v}$ be the dynamical Green function of F on $\mathsf{P}^1(\mathbb{C}_v)$ for each $v \in M_k$ and E_F be a finite subset in M_k obtained by Theorem 9.1 applied to F, for every $v \in M_k \setminus E_F$ we have $T_{F^n,v} \equiv 0$ on $\mathsf{P}^1(\mathbb{C}_v)$ for every $n \in \mathbb{N}$, giving $g_{f,v} \equiv g_{F,v} \equiv 0$ on $\mathsf{P}^1(\mathbb{C}_v)$. We call the adelic normalized weight $\hat{g}_f = \{g_{f,v} : v \in M_k\}$ and the adelic probability measure $\hat{\mu}_f := \mu^{\hat{g}_f}$ the *adelic dynamical Green function* of f and the *adelic f-equilibrium (or canonical) measure*, respectively. Here, for every $v \in M_k$, $\mu_{f,v} := \mu^{g_{f,v}} = \mu_v^{\hat{g}_f}$ (as in Section 1) is the f-equilibrium (or canonical) measure on $\mathsf{P}^1(\mathbb{C}_v)$.

Lemma 9.2. *Let k be a product formula field. Let $f, a \in k(z)$ be rational functions and suppose $d := \deg f > 1$. Then the sequence $([f^n = a])$ of k-effective divisors on $\mathbb{P}^1(\bar{k})$ has strictly small $\hat{g}_f$-heights in that*

$$\limsup_{n\to\infty} (d^n + \deg a) \cdot h_{\hat{g}_f}([f^n = a]) < \infty.$$

Proof. Let $F \in k[p_0, p_1]_d \times k[p_0, p_1]_d$ and $A \in k[p_0, p_1]_{\deg a} \times k[p_0, p_1]_{\deg a}$ be lifts of f and a, respectively. Then $F^n \wedge A \in k[p_0, p_1]_{d^n+\deg a} \times k[p_0, p_1]_{d^n+\deg a}$ is a representative of $[f^n = a]$ for every $n \in \mathbb{N}$ such that $f^n \not\equiv a$. Let E_F, E_A be finite subsets in M_k obtained by applying Theorem 9.1 to F, A, respectively, so that for every $v \in M_k \setminus (E_F \cup E_A)$ and every $n \in \mathbb{N}$, we have $T_{F^n,v} \equiv T_{A,v} \equiv 0$ and $g_{F,v} \equiv 0$ on $\mathsf{P}^1(\mathbb{C}_v)$. For every $v \in M_k$ and every sufficiently large $n \in \mathbb{N}$, since $|F^n \wedge A|_v \le \|F^n\|_v \|A\|_v$ on $\mathbb{C}_v^2 \setminus \{0\}$, we have $\log S_{F^n \wedge A,v} \le T_{F^n,v} + T_{A,v}$ on $\mathbb{P}^1(\mathbb{C}_v)$ and in turn on $\mathsf{P}^1(\mathbb{C}_v)$ (recalling that $S_{F^n\wedge A,v} = |(F^n \wedge A)(\cdot/\|\cdot\|_v)|_v$ on $\mathbb{P}^1(\mathbb{C}_v)$), so using also $g_{f,v} \equiv g_{F,v} + V_{g_{F,v}}/2$ on $\mathsf{P}^1(\mathbb{C}_v)$, we obtain

$$\frac{\log S_{F^n\wedge A,v}}{d^n + \deg a} - g_{f,v} \le \frac{T_{F^n,v} + T_{A,v}}{d^n + \deg a} - \Big(g_{F,v} + \frac{1}{2} V_{g_{F,v}}\Big) \quad \text{on } \mathsf{P}^1(\mathbb{C}_v).$$

Hence, by the definition (1-1) of $h_{\hat{g}_f}$, the Jensen-type formula (2-11), the energy formula (8-3) (with $\mathrm{Res}\, F \in k \setminus \{0\}$), and (PF), we have

$$\begin{aligned} h_{\hat{g}_f}([f^n = a]) &\le \sum_{v \in M_k} N_v \int_{\mathsf{P}^1(\mathbb{C}_v)} \Big(\frac{T_{F^n,v} + T_{A,v}}{d^n + \deg a} - g_{F,v}\Big)\, d\mu_{f,v} - \frac{3}{2} \sum_{v\in M_k} N_v \cdot V_{g_{F,v}} \\ &= \sum_{v \in E_F \cup E_A} N_v \int_{\mathsf{P}^1(\mathbb{C}_v)} \Big(\frac{T_{F^n,v} + T_{A,v}}{d^n + \deg a} - g_{F,v}\Big)\, d\mu_{f,v} \\ &= O(d^{-n}) \quad \text{as } n \to \infty, \end{aligned}$$

where the final order estimate is by (8-2) and $\#(E_F \cup E_A) < \infty$. □

With the help of Lemma 9.2, Theorems 3 and 4 follow from Theorems 1 and 2, respectively.

We omit the proof of the following characterization of $h_{\hat{g}_f}$, which we will dispense with in this article.

Lemma 9.3. *Let k be a product formula field. Then for every rational function $f \in k(z)$ of degree $d > 1$, the $\hat{g}_f$-height function $h_{\hat{g}_f}$ coincides with the* Call–Silverman f-dynamical (or canonical) height function *in that for every k-effective divisor $\mathcal{Z}$ on $\mathbb{P}^1(\bar{k})$, ($f_*\mathcal{Z}$ is also a k-effective divisor on $\mathbb{P}^1(\bar{k})$, and) the equality $(h_{\hat{g}_f} \circ f_*)(\mathcal{Z}) = (d \cdot h_{\hat{g}_f})(\mathcal{Z})$ holds.*

10. Proofs of Theorems 5 and 6

Let K be an algebraically closed field that is complete with respect to a nontrivial absolute value $|\cdot|$. For subsets $A, B \subset \mathbb{P}^1$, set $[A, B] := \inf_{z \in A, z' \in B}[z, z']$.

Let $f, a \in K(z)$ be rational functions and suppose that $d := \deg f > 1$. Let $N \in \mathbb{N}$ be so large that $f^n \not\equiv a$ if $n > N$. Then $\big(\overline{\bigcup_{n>N} \operatorname{supp}[f^n = a]} \cup \mathsf{J}(f)\big) \cap \mathbb{P}^1$ is closed in $\mathbb{P}^1$.

Lemma 10.1. *Suppose that K has characteristic* 0. *Let D be a chordal disk in $\mathbb{P}^1$ of radius > 0 satisfying $\liminf_{n\to\infty} \sup_D [f^n, a] = 0$. Then*:

(i) $a(D) \subset \mathcal{E}_f$.

(ii) $D \setminus \big(\overline{\bigcup_{n>N} \operatorname{supp}[f^n = a]} \cup \mathsf{J}(f)\big) \neq \varnothing$.

(iii) *There is a chordal disk D' in $\mathbb{P}^1 \setminus \mathsf{J}(f)$ of radius > 0 such that*

$$\liminf_{n\to\infty} [f^n(D'), a(D')] > 0.$$

Proof of (i). Since $\liminf_{n\to\infty} \sup_D [f^n, a] = 0$, there is a sequence (n_j) in $\mathbb{N}$ such that $\lim_{j\to\infty} \sup_D [f^{n_j}, a] = 0$ and $\lim_{j\to\infty}(n_{j+1} - n_j) = \infty$. For every $z \in D$, set $D'' := \{w \in \mathbb{P}^1 : [w, a(z)] \le r\}$ in $a(D)$ for $r > 0$ small enough. Then $\liminf_{j\to\infty} \sup_{D''}[f^{n_{j+1}-n_j}, \mathrm{Id}] \le \limsup_{j\to\infty} \sup_D [f^{n_{j+1}}, f^{n_j}] = 0$, so that $a(z) \in \mathcal{E}_f$. Hence $a(D) \subset \mathcal{E}_f$. □

Proof of (ii). When K is archimedean, let Y be the component of $\mathcal{E}_f$ containing $a(D)$, which is by the first assertion either a Siegel disk or a Herman ring of f. Setting $k_0 := \min\{n \in \mathbb{N} : f^n(Y) = Y\}$, there are a sequence (n_j) and an N in $\mathbb{N}$ with the properties that $f^{n_N}(D) \subset Y$, that $k_0 \mid (n_j - n_N)$ for every $j \ge N$, and that $a = \lim_{j\to\infty} (f^{k_0})^{(n_j - n_N)/k_0} \circ f^{n_N}$ uniformly on D. Then $D \cap \mathsf{J}(f) = \varnothing$. Let $\lambda \in \mathbb{C}$ be the rotation number of Y, so that there exists a holomorphic injection $h : Y \to \mathbb{C}$ such that $h \circ f^{k_0} = \lambda \cdot h$ on Y. Then $|\lambda| = 1$ but λ is not a root of unity (by $d > 1$). Choosing a subsequence of (n_j) if necessary, $\lambda_a := \lim_{j\to\infty} \lambda^{(n_j - n_N)/k_0} \in \mathbb{C}$

exists. For every $n \ge n_N$, if $k_0 \nmid (n-n_N)$, then $D \cap \operatorname{supp}[f^n = a] = \varnothing$, whereas if $k_0 \mid (n-n_N)$, then $h \circ f^n - h \circ a = (\lambda^{(n-n_N)/k_0} - \lambda_a) \cdot (h \circ f^{n_N})$ on D, so $\big(D \setminus (h \circ f^{n_N})^{-1}(0)\big) \cap \operatorname{supp}[f^n = a] = \varnothing$ if n is large enough.

When K is nonarchimedean, let Y be the component of $\mathcal{E}_f$ containing $a(D)$. Without loss of generality, we assume that $\infty \notin Y$, and then applying Theorem 8.3 to this Y, we obtain $p \in \mathbb{N}$, $k_0 \in \mathbb{N}$, T, and T_* as in the theorem. There are a sequence (n_j) and an N in $\mathbb{N}$ such that $f^{n_N}(D) \subset Y$, $k_0 \mid (n_j - n_N)$ for every $j \ge N$, and $a = \lim_{j\to\infty} (f^{k_0})^{(n_j-n_N)/k_0} \circ f^{n_N}$ uniformly on D. Then $D \cap \mathsf{J}(f) = \varnothing$. Choosing a subsequence of (n_j) if necessary, $\omega_a := \lim_{j\to\infty}(n_j - n_N)/k_0 \in \mathbb{Z}_p$ exists. For every $n \ge n_N$, if $k_0 \nmid (n-n_N)$, then $D \cap \operatorname{supp}[f^n = a] = \varnothing$, whereas if $k_0 \mid (n-n_N)$, then

$$f^n - a = (T^{(n-n_N)/k_0} - T^{\omega_a}) \circ f^{n_N} \tag{10-1}$$

on D. Choose $b \in D \setminus \{\infty\}$ and $r \in |K^*|$ small enough that the (K-closed) disk $B = \{z \in K : |z-b| \le r\}$ is contained in D, and fix $\epsilon \in |K^*|$ so small that for $Z_\epsilon := \bigcup_{w \in B \cap (T_* \circ T^{\omega_a} \circ f^{n_N})^{-1}(0)} \{z \in B : |z-w| < \epsilon\}$, we have $B \setminus Z_\epsilon \ne \varnothing$. The maximum modulus principle from rigid analysis (see [Bosch, Güntzer, and Remmert 1984, §6.2.1, §7.3.4]) gives $\min_{z \in f^{n_N}(B \setminus Z_\epsilon)} |T_* \circ T^{\omega_a}(z)| > 0$, so that by the uniform convergence (8-4) and the equality (10-1), $(B \setminus Z_\epsilon) \cap \operatorname{supp}[f^n = a] = \varnothing$ if n is large enough. □

Proof of (iii). By the first assertion, there is a unique singular domain U of f containing $a(D)$. Fix $n_0 \in \mathbb{N}$ such that $f^{n_0}(U) = U$, and set $\mathcal{C} := \bigcup_{j=0}^{n_0-1} f^j(U)$. Then there is a component V of $f^{-1}(\mathcal{C}) \setminus \mathcal{C}$ since $f : \mathcal{C} \to \mathcal{C}$ is injective and $d > 1$. Fix a chordal disk D'' of radius > 0 in $a^{-1}(V) \cap (\mathbb{P}^1 \setminus \mathsf{J}(f))$, so that $a(D'') \subset V \subset f^{-1}(\mathcal{C}) \setminus \mathcal{C}$. If $a(D'') \cap \bigcup_{n \in \mathbb{N}\cup\{0\}} f^n(D'') = \varnothing$, then we are done by setting $D' = \{z \in \mathbb{P}^1 : [z,b] \le r\}$ for some $b \in D''$ and $r > 0$ small enough. But if there is $N \in \mathbb{N} \cup \{0\}$ such that $a(D'') \cap f^N(D'') \ne \varnothing$, then by setting $D' := \{z \in \mathbb{P}^1 : [z,b] \le r\}$ for some $b \in D'' \cap f^{-N}(a(D''))$ and $r > 0$ small enough, we get $\liminf_{n\to\infty} [a(D'), f^n(D')] > 0$ from

$$a(D') \cap \bigcup_{n \ge N+1} f^n(D') \subset a(D'') \cap \bigcup_{n \in \mathbb{N}} f^n(a(D'')) \subset V \cap \mathcal{C} = \varnothing. \qquad \square$$

Lemma 10.2. *For every $w_0 \in \mathbb{P}^1 \setminus \big(\overline{\bigcup_{n>N} \operatorname{supp}[f^n = a]} \cup \mathsf{J}(f)\big)$, there is a function $\phi_0 \in C^1(\mathsf{P}^1)$ such that $\phi_0 \equiv \log[w_0, \cdot]_{\mathrm{can}}$ on $\bigcup_{n>N} \operatorname{supp}[f^n = a] \cup \mathsf{J}(f)$.*

Proof. Fix $w_0 \in \mathbb{P}^1 \setminus \big(\overline{\bigcup_{n>N} \operatorname{supp}[f^n = a]} \cup \mathsf{J}(f)\big)$. Without loss of generality, we can assume that $w_0 \ne \infty$, and fix $\epsilon > 0$ so small that

$$\big\{\mathcal{S} \in \mathsf{P}^1 : |\mathcal{S} - w_0|_\infty \le \epsilon\big\} \subset \mathsf{P}^1 \setminus \Big(\overline{\bigcup_{n>N} \operatorname{supp}[f^n = a]} \cup \mathsf{J}(f)\Big)$$

(recall Sections 3.1 and 3.2 here).

When K is nonarchimedean, by the definition of the map $\pi_\epsilon : \mathsf{A}^1 \to \mathsf{A}^1$, we have $\{\mathcal{S} \in \mathsf{P}^1 : \mathcal{S} \preceq \pi_\epsilon(w_0)\} = \{\mathcal{S} \in \mathsf{P}^1 : |\mathcal{S} - w_0|_\infty \le \epsilon\}$. The function

$$\mathcal{S} \mapsto \phi_0(\mathcal{S}) := \begin{cases} \log\,[w_0, \pi_\epsilon(w_0)]_{\mathrm{can}} & \text{if } \mathcal{S} \preceq \pi_\epsilon(w_0), \\ \log\,[w_0, \mathcal{S}]_{\mathrm{can}} & \text{otherwise} \end{cases} \quad \text{on } \mathsf{P}^1$$

is in $C^1(\mathsf{P}^1)$ since it is continuous on P^1, locally constant on P^1 except for the segment $\mathcal{I}$ in H^1 joining $\pi_\epsilon(w_0)$ and $\mathcal{S}_{\mathrm{can}}$, and linear on $\mathcal{I}$ with respect to the length parameter induced by the hyperbolic metric ρ on H^1. When K is archimedean (so $\mathsf{P}^1 \cong \mathbb{P}^1$), there is a function $\phi_0 \in C^1(\mathbb{P}^1)$ satisfying

$$z \mapsto \phi_0(z) = \begin{cases} \int_{\mathbb{P}^1} \log\,[w_0, w]\, \mathrm{d}[z]_{\epsilon/2}(w) & \text{if } |z - w_0| \le \epsilon/2, \\ \log\,[w_0, z] & \text{if } |z - w_0| \ge \epsilon \text{ or } z = \infty. \end{cases}$$

In both cases, the given $\phi_0 \in C^1(\mathsf{P}^1)$ satisfies the desired property. $\square$

Fact 10.3. For rational functions $\phi, \psi \in K(z)$, the *chordal proximity function*

$$\mathcal{S} \mapsto [\phi, \psi]_{\mathrm{can}}(\mathcal{S}) \quad \text{on } \mathsf{P}^1$$

between ϕ and ψ is the unique continuous extension of the function $z \mapsto [\phi(z), \psi(z)]$ on $\mathbb{P}^1$ to P^1 (see [Okuyama 2013a, Proposition 2.9] for its construction, as well as Remark 2.10 of the same paper), and for every continuous weight g on P^1, we also define its weighted version by $\Phi(\phi, \psi)_g := \log\,[\phi, \psi]_{\mathrm{can}} - g \circ \phi - g \circ \psi$ on P^1.

For every $n \in \mathbb{N}$ such that $f^n \not\equiv a$, recall the *Riesz decomposition*

$$\text{(10-2)} \quad \Phi(f^n, a)_{g_f} = U_{g_f, [f^n = a] - (d^n + \deg a)\mu_f} - U_{g_f, a^*\mu_f} + \int_{\mathsf{P}^1} \Phi(f^n, a)_{g_f}\, \mathrm{d}\mu_f$$

on P^1, and also $U_{g_f, a^*\mu_f} = g_f \circ a + U_{g_f, a^*\Omega_{\mathrm{can}}} - \int_{\mathsf{P}^1} (g_f \circ a)\, \mathrm{d}\mu_f$ on P^1 [Okuyama 2013a, Lemma 2.19].

Proof of Theorem 5. Let k be a product formula field of characteristic 0. Let $f \in k(z)$ be a rational function of degree $d > 1$ and $a \in k(z)$ a rational function of degree > 0. Let $N \in \mathbb{N}$ be so large that $f^n \not\equiv a$ if $n > N$. Fix $v \in M_k$. Let D be a chordal disk in $\mathbb{P}^1(\mathbb{C}_v)$ of radius > 0, and assume that $\liminf_{n\to\infty} \sup_D [f^n, a]_v = 0$; otherwise we are done. By Lemma 10.1, there are not only a point $w_0 \in D \setminus \big(\overline{\bigcup_{n>N}[f^n = a]} \cup \mathsf{J}(f)_v\big)$ but also a chordal disk D' in $\mathbb{P}^1(\mathbb{C}_v) \setminus \mathsf{J}(f)_v$ of radius > 0 such that $\liminf_{n\to\infty} [f^n(D'), a(D')]_v > 0$. Fix a point $w_1 \in D'$. Then also $w_1 \in \mathbb{P}^1 \setminus \big(\overline{\bigcup_{n>N}[f^n = a]} \cup \mathsf{J}(f)_v\big)$.

For every $n \in \mathbb{N}$ large enough and every $j \in \{0, 1\}$, by (10-2),

$$\text{(10-3)} \quad \log\,[f^n(w_j), a(w_j)]_v - g_{f,v}(f^n(w_j)) - g_{f,v}(a(w_j))$$
$$= U_{g_{f,v}, [f^n = a] - (d^n + \deg a)\mu_{f,v}}(w_j) - U_{g_{f,v}, a^*\mu_{f,v}}(w_j) + \int_{\mathsf{P}^1(\mathbb{C}_v)} \Phi(f^n, a)_{g_{f,v}}\, \mathrm{d}\mu_{f,v},$$

so that taking the difference of both sides in (10-3) for each $j \in \{0,1\}$ and noting that $g_{f,v}$ and $U_{g_{f,v},a^*\mu_{f,v}}$ are bounded on $\mathsf{P}^1(\mathbb{C}_v)$, we have

$$\begin{aligned}\log[f^n(w_0),a(w_0)]_v - \log[f^n(w_1),a(w_1)]_v &\\ = \int_{\mathsf{P}^1(\mathbb{C}_v)} \log[w_0,\mathcal{S}']_{\mathrm{can},v}\, \mathrm{d}([f^n=a]-(d^n+\deg a)\mu_f)(\mathcal{S}') &\\ - \int_{\mathsf{P}^1(\mathbb{C}_v)} \log[w_1,\mathcal{S}']_{\mathrm{can},v}\, \mathrm{d}([f^n=a]-(d^n+\deg a)\mu_f)(\mathcal{S}') + O(1)&\end{aligned}$$

as $n\to\infty$. In the left hand side, by the choice of w_0 and w_1, we have

$$\log\sup_D [f^n,a]_v \ge \log[f^n(w_0),a(w_0)]_v$$

and

$$\liminf_{n\to\infty}\log[f^n(w_1),a(w_1)]_v \ge \liminf_{n\to\infty}\log[f^n(D'),a(D')]_v > -\infty,$$

so that as $n\to\infty$,

$$\log\sup_D[f^n,a]_v + O(1) \ge \log[f^n(w_0),a(w_0)]_v - \log[f^n(w_1),a(w_1)]_v.$$

In the right hand side, for each $j\in\{0,1\}$, by Lemma 10.2 applied to w_j, the inclusion $\operatorname{supp}\mu_f \subset \mathsf{J}(f)$, and Theorem 3 (and $k_s=\bar{k}$ in the characteristic 0 case), we have

$$\begin{aligned}\int_{\mathsf{P}^1(\mathbb{C}_v)} \log[w_j,\mathcal{S}']_{\mathrm{can},v}\, \mathrm{d}\big([f^n=a]-(d^n+\deg a)\mu_f\big)(\mathcal{S}') &\\ = O\Big(\sqrt{n\cdot\big([f^n=a]\times[f^n=a]\big)\big(\mathrm{diag}_{\mathbb{P}^1(\bar{k})}\big)}\Big) \quad \text{as } n\to\infty.&\end{aligned}$$

These estimates complete the proof of (1-4) for this $v\in M_k$. □

Fact 10.4. For a rational function $f(z)\in k(z)$ over a field k, a point $w\in\mathbb{P}^1(\bar{k})$ is called a *multiple* periodic point of f if $[f^n=\mathrm{Id}](\{w\})>1$ for some $n\in\mathbb{N}$. For a rational function $f(z)\in k(z)$ over a field k of characteristic 0, there are *at most finitely many* multiple periodic points of f in $\mathbb{P}^1(\bar{k})$; this is well known in the case that $k=\mathbb{C}$ (see, e.g., [Milnor 2006, §13]), and holds in general *by the Lefschetz principle* (see, e.g., [Eklof 1973]).

Proof of Theorem 6. As noted above, f has at most finitely many multiple periodic points in $\mathbb{P}^1(\bar{k})$, and for every multiple periodic point w of f, setting $p=p_w:=\min\{n\in\mathbb{N}:[f^n=\mathrm{Id}](\{w\})>1\}$, by the (formal) power series expansion $f^p(z)=w+(z-w)+C(z-w)^{[f^p=\mathrm{Id}](\{w\})}+\cdots$ of f^p around w, we also have $\sup_{n\in\mathbb{N}}[f^n=\mathrm{Id}](\{w\})\le[f^p=\mathrm{Id}](\{w\})$ under the characteristic 0 assumption.

Hence $\sup_{n\in\mathbb{N}}\big(\sup_{w\in\operatorname{supp}[f^n=\mathrm{Id}]}[f^n=\mathrm{Id}](\{w\})\big)<\infty$, so that

$$\big([f^n=\mathrm{Id}]\times[f^n=\mathrm{Id}]\big)\big(\mathrm{diag}_{\mathbb{P}^1(\bar{k})}\big)\le(d^n+1)\cdot\sup_{w\in\operatorname{supp}[f^n=\mathrm{Id}]}[f^n=\mathrm{Id}](\{w\})=O(d^n)$$

as $n\to\infty$. Now (1-5) follows from (1-4). □

11. Proof of Theorem 7

Let k be a field and k_s the separable closure of k in $\bar{k}$. Let $p(z)\in k[z]$ be a polynomial of degree >0 and $\{z_1,\dots,z_m\}$ the set of all distinct zeros of $p(z)$ in $\bar{k}$ so that $p(z)=a\cdot\prod_{j=1}^m(z-z_j)^{d_j}$ in $\bar{k}[z]$ for some $a\in k\setminus\{0\}$ and some sequence $(d_j)_{j=1}^m$ in $\mathbb{N}$. For a while, we do not assume $\{z_1,\dots,z_m\}\subset k_s$. Let $\{p_1(z),p_2(z),\dots,p_N(z)\}$ be the set of all mutually distinct, nonconstant, irreducible, and monic factors of $p(z)$ in $k[z]$, so that $p(z)=a\cdot\prod_{\ell=1}^N p_\ell(z)^{s_\ell}$ in $k[z]$ for some sequence $(s_\ell)_{\ell=1}^N$ in $\mathbb{N}$. For every $\ell\in\{1,2,\dots,N\}$, by the irreducibility of $p_\ell(z)$ in $k[z]$, $p_\ell(z)$ is the unique monic minimal polynomial in $k[z]$ of each zero of $p_\ell(z)$ in $\bar{k}$, so $p_\ell(z)$ and $p_n(z)$ have no common zeros in $\bar{k}$ if $\ell\neq n$. Hence for each $j\in\{1,2,\dots,m\}$, there is a unique $\ell=:\ell(j)\in\{1,2,\dots,N\}$ such that $p_\ell(z_j)=0$.

Now suppose that $\{z_1,z_2,\dots,z_m\}\subset k_s$. Then for every $\ell\in\{1,2,\dots,N\}$, $p_\ell(z)=\prod_{i:\ell(i)=\ell}(z-z_i)$ in $\bar{k}[z]$, so that

$$d_i=s_{\ell(i)} \tag{11-1}$$

for every $i\in\{1,2,\dots,m\}$. For every distinct $\ell,n\in\{1,2,\dots,N\}$,

$$\prod_{j:\ell(j)=\ell}\ \prod_{i:\ell(i)=n}(z_j-z_i)=\prod_{j:\ell(j)=\ell}p_n(z_j)=R(p_\ell,p_n), \tag{11-2}$$

where $R(p,q)\in k$ is the (usual) resultant of $p(z),q(z)\in k[z]$. The derivation $p_\ell'(z)$ of $p_\ell(z)$ in $k[z]$ satisfies

$$p_\ell'(z)=\sum_{h:\ell(h)=\ell}\Bigg(\prod_{\substack{i:i\neq h,\\ \ell(i)=\ell}}(z-z_i)\Bigg)$$

in $\bar{k}[z]$. Hence for every $\ell\in\{1,2,\dots,N\}$,

$$\prod_{j:\ell(j)=\ell}\ \prod_{\substack{i:i\neq j,\\ \ell(i)=\ell}}(z_j-z_i)=\prod_{j:\ell(j)=\ell}p_\ell'(z_j)=R(p_\ell,p_\ell'). \tag{11-3}$$

By (11-1), (11-3), and (11-2), we have

$$\begin{aligned}D^*(p) &:= \prod_{j=1}^{m}\prod_{i:i\neq j}(z_j-z_i)^{d_i d_j} = \prod_{j=1}^{m}\prod_{i:i\neq j}(z_j-z_i)^{s_{\ell(i)}s_{\ell(j)}}\\ &= \prod_{\ell=1}^{N}\Biggl(\prod_{j:\ell(j)=\ell}\Biggl(\Biggl(\prod_{\substack{i:i\neq j,\\ \ell(i)=\ell}}(z_j-z_i)^{s_\ell^2}\Biggr)\Biggl(\prod_{n:n\neq\ell}\prod_{i:\ell(i)=n}(z_j-z_i)^{s_n s_\ell}\Biggr)\Biggr)\Biggr)\\ &= \prod_{\ell=1}^{N}\Biggl(R(p_\ell,p'_\ell)^{s_\ell^2}\cdot\prod_{n:n\neq\ell}R(p_\ell,p_n)^{s_n s_\ell}\Biggr),\end{aligned}$$

which is in $k\setminus\{0\}$. Now the proof is complete. □

Acknowledgements

The author thanks the referee for their very careful scrutiny and invaluable comments, and also thanks Professors Mamoru Asada and Joe Silverman for discussions on Theorem 7 and comments. This work was partly done during the author's visits to National Taiwan Normal University and Academia Sinica, and the author thanks the institutes and Professor Liang-Chung Hsia for their hospitality. This work was partially supported by JSPS Grant-in-Aid for Young Scientists (B), 24740087.

References

[Autissier 2001] P. Autissier, "Points entiers sur les surfaces arithmétiques", *J. Reine Angew. Math.* **531** (2001), 201–235. MR 2002a:11066 Zbl 1007.11041

[Baker 2009] M. H. Baker, "A finiteness theorem for canonical heights attached to rational maps over function fields", *J. Reine Angew. Math.* **626** (2009), 205–233. MR 2011c:14075 Zbl 1187.37133

[Baker and Hsia 2005] M. H. Baker and L.-C. Hsia, "Canonical heights, transfinite diameters, and polynomial dynamics", *J. Reine Angew. Math.* **585** (2005), 61–92. MR 2006i:11071 Zbl 1071.11040

[Baker and Rumely 2006] M. H. Baker and R. Rumely, "Equidistribution of small points, rational dynamics, and potential theory", *Ann. Inst. Fourier* (*Grenoble*) **56**:3 (2006), 625–688. MR 2007m:11082 Zbl 1234.11082

[Baker and Rumely 2010] M. H. Baker and R. Rumely, *Potential theory and dynamics on the Berkovich projective line*, Mathematical Surveys and Monographs **159**, American Mathematical Society, Providence, RI, 2010. MR 2012d:37213 Zbl 1196.14002

[Benedetto 2010] R. Benedetto, "Non-Archimedean dynamics in dimension one: lecture notes", lecture notes, 2010, http://math.arizona.edu/~swc/aws/2010/2010BenedettoNotes-09Mar.pdf.

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI, 1990. MR 91k:32038 Zbl 0715.14013

[Berman and Boucksom 2010] R. Berman and S. Boucksom, "Growth of balls of holomorphic sections and energy at equilibrium", *Invent. Math.* **181**:2 (2010), 337–394. MR 2011h:32021 Zbl 1208.32020

[Berman, Boucksom, and Nyström 2011] R. Berman, S. Boucksom, and D. Witt Nyström, "Fekete points and convergence towards equilibrium measures on complex manifolds", *Acta Math.* **207**:1 (2011), 1–27. MR 2012j:32036 Zbl 1241.32030

[Berteloot and Mayer 2001] F. Berteloot and V. Mayer, *Rudiments de dynamique holomorphe*, Cours Spécialisés **7**, Société Mathématique de France, Paris, 2001. MR 2005b:37087 Zbl 1051.37019

[Bilu 1997] Y. Bilu, "Limit distribution of small points on algebraic tori", *Duke Math. J.* **89**:3 (1997), 465–476. MR 98m:11067 Zbl 0918.11035

[Bosch, Güntzer, and Remmert 1984] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis: a systematic approach to rigid analytic geometry*, Grundlehren der Mathematischen Wissenschaften **261**, Springer, Berlin, 1984. MR 86b:32031 Zbl 0539.14017

[Brjuno 1971] A. D. Brjuno, "Аналитическая форма дифференциальных уравнений, 1", *Trudy Moskov. Mat. Obšč.* **25** (1971), 119–262. Translated as "Analytic form of differential equations, 1" in *Trans. Mosc. Math. Soc.* **25** (1971), 131–288. MR 51 #13365 Zbl 0263.34003

[Brjuno 1972] A. D. Brjuno, "Аналитическая форма дифференциальных уравнений, 2", *Trudy Moskov. Mat. Obšč.* **26** (1972), 199–239. Translated as "Analytic form of differential equations, 2" in *Trans. Mosc. Math. Soc.* **26** (1972), 199–239. MR 51 #13365 Zbl 0269.34006

[Brolin 1965] H. Brolin, "Invariant sets under iteration of rational functions", *Ark. Mat.* **6** (1965), 103–144. MR 33 #2805 Zbl 0127.03401

[Chambert-Loir 2000] A. Chambert-Loir, "Points de petite hauteur sur les variétés semi-abéliennes", *Ann. Sci. École Norm. Sup.* (4) **33**:6 (2000), 789–821. MR 2002e:14037 Zbl 1018.11034

[Chambert-Loir 2006] A. Chambert-Loir, "Mesures et équidistribution sur les espaces de Berkovich", *J. Reine Angew. Math.* **595** (2006), 215–235. MR 2008b:14040 Zbl 1112.14022

[Cremer 1928] H. Cremer, "Zum Zentrumproblem", *Math. Ann.* **98**:1 (1928), 151–163. MR 1512397 JFM 53.0303.04

[DeMarco 2003] L. DeMarco, "Dynamics of rational maps: Lyapunov exponents, bifurcations, and capacity", *Math. Ann.* **326**:1 (2003), 43–73. MR 2004f:32044 Zbl 1032.37029

[Drasin and Okuyama 2007] D. Drasin and Y. Okuyama, "Equidistribution and Nevanlinna theory", *Bull. Lond. Math. Soc.* **39**:4 (2007), 603–613. MR 2008f:37099 Zbl 1123.37018

[Eklof 1973] P. C. Eklof, "Lefschetz's principle and local functors", *Proc. Amer. Math. Soc.* **37** (1973), 333–339. MR 48 #3736 Zbl 0254.14004

[Fatou 1920] P. Fatou, "Sur les équations fonctionnelles", *Bull. Soc. Math. France* **48** (1920), 208–314. MR 1504797 JFM 47.0921.02

[Favre and Jonsson 2004] C. Favre and M. Jonsson, *The valuative tree*, Lecture Notes in Mathematics **1853**, Springer, Berlin, 2004. MR 2006a:13008 Zbl 1064.14024

[Favre and Rivera-Letelier 2006] C. Favre and J. Rivera-Letelier, "Équidistribution quantitative des points de petite hauteur sur la droite projective", *Math. Ann.* **335**:2 (2006), 311–361. Correction in **339**:4 (2007), 799–801. MR 2007g:11074 Zbl 1175.11029

[Favre and Rivera-Letelier 2010] C. Favre and J. Rivera-Letelier, "Théorie ergodique des fractions rationnelles sur un corps ultramétrique", *Proc. Lond. Math. Soc.* (3) **100**:1 (2010), 116–154. MR 2011b:37190 Zbl 1254.37064

[Fekete 1930a] M. Fekete, "Über den transfiniten Durchmesser ebener Punktmengen, 1", *Math. Z.* **32**:1 (1930), 108–114. MR 1545154 JFM 56.0090.01

[Fekete 1930b] M. Fekete, "Über den transfiniten Durchmesser ebener Punktmengen, 2", *Math. Z.* **32**:1 (1930), 215–221. MR 1545162 JFM 56.0112.02

[Fekete 1933] M. Fekete, "Über den transfiniten Durchmesser ebener Punktmengen, 3", *Math. Z.* **37**:1 (1933), 635–646. MR 1545425 Zbl 0007.40204

[Freire, Lopes, and Mañé 1983] A. Freire, A. Lopes, and R. Mañé, "An invariant measure for rational maps", *Bol. Soc. Brasil. Mat.* **14**:1 (1983), 45–62. MR 85m:58110b Zbl 0568.58027

[Herman and Yoccoz 1983] M. Herman and J.-C. Yoccoz, "Generalizations of some theorems of small divisors to non-Archimedean fields", pp. 408–447 in *Geometric dynamics* (Rio de Janeiro, 1981), edited by J. Palis, Jr., Lecture Notes in Math. **1007**, Springer, Berlin, 1983. MR 85i:12012 Zbl 0528.58031

[Jonsson 2015] M. Jonsson, "Dynamics of Berkovich spaces in low dimensions", pp. 205–366 in *Berkovich spaces and applications* (Santiago de Chile/Paris, 2008/2010), edited by A. Ducros et al., Lecture Notes in Math. **2119**, Springer, Cham, 2015. MR 3330767 Zbl 06463429

[Lev and Ortega-Cerdà 2012] N. Lev and J. Ortega-Cerdà, "Equidistribution estimates for Fekete points on complex manifolds", preprint, 2012. arXiv 1210.8059v1

[Levenberg 2010] N. Levenberg, "Weighted pluripotential theory results of Berman–Boucksom", preprint, 2010. arXiv 1010.4035

[Ljubich 1983] M. J. Ljubich, "Entropy properties of rational endomorphisms of the Riemann sphere", *Ergodic Theory Dynam. Systems* **3**:3 (1983), 351–385. MR 85k:58049 Zbl 0537.58035

[Milnor 2006] J. Milnor, *Dynamics in one complex variable*, 3rd ed., Annals of Mathematics Studies **160**, Princeton University Press, 2006. MR 2006g:37070 Zbl 1085.30002

[Okuyama 2010] Y. Okuyama, "Nonlinearity of morphisms in non-Archimedean and complex dynamics", *Michigan Math. J.* **59**:3 (2010), 505–515. MR 2012d:37208 Zbl 1242.37063

[Okuyama 2013a] Y. Okuyama, "Adelic equidistribution, characterization of equidistribution, and a general equidistribution theorem in non-Archimedean dynamics", *Acta Arith.* **161**:2 (2013), 101–125. MR 3141914 Zbl 1302.37070

[Okuyama 2013b] Y. Okuyama, "Fekete configuration, quantitative equidistribution and wandering critical orbits in non-Archimedean dynamics", *Math. Z.* **273**:3-4 (2013), 811–837. MR 3030679 Zbl 06149057

[Okuyama and Stawiska 2011] Y. Okuyama and M. Stawiska, "Potential theory and a characterization of polynomials in complex dynamics", *Conform. Geom. Dyn.* **15** (2011), 152–159. MR 2846305 Zbl 1252.37036

[Pérez-Marco 1993] R. Pérez-Marco, "Sur les dynamiques holomorphes non linéarisables et une conjecture de V. I. Arnold", *Ann. Sci. École Norm. Sup.* (4) **26**:5 (1993), 565–644. MR 95a:58103 Zbl 0812.58051

[Pérez-Marco 2001] R. Pérez-Marco, "Total convergence or general divergence in small divisors", *Comm. Math. Phys.* **223**:3 (2001), 451–464. MR 2003d:37063 Zbl 1161.37331

[Rivera-Letelier 2003] J. Rivera-Letelier, "Dynamique des fonctions rationnelles sur des corps locaux", pp. 147–230 in *Geometric methods in dynamics, II* (Rio de Janeiro, 2000), edited by W. de Melo et al., Astérisque **287**, Société Mathématique de France, Paris, 2003. MR 2005f:37100 Zbl 1140.37336

[Rumely 1999] R. Rumely, "On Bilu's equidistribution theorem", pp. 159–166 in *Spectral problems in geometry and arithmetic* (Iowa City, IA, 1997), edited by T. Branson, Contemp. Math. **237**, American Mathematical Society, Providence, RI, 1999. MR 2000g:11060 Zbl 1029.11030

[Saff and Totik 1997] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, Grundlehren der Mathematischen Wissenschaften **316**, Springer, Berlin, 1997. MR 99h:31001 Zbl 0881.31001

[Siegel 1942] C. L. Siegel, "Iteration of analytic functions", *Ann. of Math.* (2) **43** (1942), 607–612. MR 4,76c Zbl 0061.14904

[Silverman 2007] J. H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics **241**, Springer, New York, 2007. MR 2008c:11002 Zbl 1130.37001

[Szpiro, Ullmo, and Zhang 1997] L. Szpiro, E. Ullmo, and S. Zhang, "Équirépartition des petits points", *Invent. Math.* **127**:2 (1997), 337–347. MR 98i:14027 Zbl 0991.11035

[Thuillier 2005] A. Thuillier, *Théorie du potentiel sur les courbes en géométrie analytique non Archimédienne: applications à la théorie d'Arakelov*, thesis, Université Rennes 1, 2005, https://tel.archives-ouvertes.fr/tel-00010990.

[Tsuji 1959] M. Tsuji, *Potential theory in modern function theory*, Maruzen, Tokyo, 1959. Reprinted by Chelsea, New York, 1975. MR 22 #5712 Zbl 0087.28401

[Villani 2009] C. Villani, *Optimal transport: old and new*, Grundlehren der Mathematischen Wissenschaften **338**, Springer, Berlin, 2009. MR 2010f:49001 Zbl 1156.53003

[Yoccoz 1988] J.-C. Yoccoz, "Linéarisation des germes de difféomorphismes holomorphes de $(\mathbb{C}, 0)$", *C. R. Acad. Sci. Paris Sér. I Math.* **306**:1 (1988), 55–58. MR 89i:58123 Zbl 0668.58010

[Yoccoz 1995] J.-C. Yoccoz, "Théorème de Siegel, nombres de Bruno et polynômes quadratiques", pp. 3–88 in *Petits diviseurs en dimension* 1, Astérisque **231**, Société Mathématique de France, Paris, 1995. MR 96m:58214 Zbl 0836.30001

[Yuan 2008] X. Yuan, "Big line bundles over arithmetic varieties", *Invent. Math.* **173**:3 (2008), 603–649. MR 2010b:14049 Zbl 1146.14016

Received July 29, 2014. Revised April 5, 2015.

YÛSUKE OKUYAMA
DIVISION OF MATHEMATICS
KYOTO INSTITUTE OF TECHNOLOGY
SAKYO-KU, KYOTO 606-8585
JAPAN
okuyama@kit.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.177

COMPUTING HIGHER FROBENIUS–SCHUR INDICATORS IN FUSION CATEGORIES CONSTRUCTED FROM INCLUSIONS OF FINITE GROUPS

PETER SCHAUENBURG

We consider a subclass of the class of group-theoretical fusion categories: To every finite group G and subgroup H one can associate the category of G-graded vector spaces with a two-sided H-action compatible with the grading. We derive a formula that computes higher Frobenius-Schur indicators for the objects in such a category using the combinatorics and representation theory of the groups involved in their construction. We calculate some explicit examples for inclusions of symmetric groups.

1. Introduction

Higher Frobenius–Schur indicators are invariants of an object in a pivotal fusion category (and hence also invariants of that category). They generalize the degree two Frobenius-Schur indicator — which was originally defined for a representation of a finite group by its namesakes in 1906 — to higher degrees and more general objects. Categorical versions of degree two indicators were studied by Bantay [1997], as well as Fuchs, Ganchev, Szlachányi, and Vescernyés [Fuchs et al. 1999]; indicators for modules over semisimple Hopf algebras were introduced by Linchenko and Montgomery [2000] and studied in depth by Kashina, Sommerhäuser, and Zhu [2006]. The degree two indicators for modules over semisimple quasi-Hopf algebras were treated by Mason and Ng [2005]. The higher indicators for pivotal fusion categories that we deal with in the present paper were introduced in [Ng and Schauenburg 2008; 2007b; 2007a].

Frobenius–Schur indicators have become a tool for the structure theory and classification of fusion categories. The problem we deal with here, however, is simply how to calculate them in very specific examples. More concretely we will deal with a specific class of group-theoretical fusion categories [Ostrik 2003; Etingof, Nikshych and Ostrik 2005]. Degree two indicators for Hopf algebras associated

Research partially supported through a FABER Grant by the *Conseil régional de Bourgogne*.
MSC2010: primary 18D10, 16T05; secondary 20C15.
Keywords: fusion category, Frobenius-Schur indicator.

with such categories have been studied in [Kashina, Mason and Montgomery 2002; Jedwab and Montgomery 2009]. In [Kashina, Sommerhäuser and Zhu 2006] formulas for higher indicators of smash product Hopf algebras associated to a group acting by automorphisms on another group were given. This class of examples includes the Drinfeld double of a finite group. For such doubles, the explicit formulas were used to study the question of integrality of the indicators in [Iovanov, Mason and Montgomery 2014]. Extensive computer calculations, in particular with a view towards the question of whether the indicators of the doubles of symmetric groups are positive, were conducted in [Courter 2012]; examples for certain other groups can be found in [Keilberg 2014; 2012].

Natale [2005] has derived formulas for the degree two Frobenius–Schur indicators of the objects in general group-theoretical fusion categories. Her approach is based on the fact that a group-theoretical fusion category can be written as the module category over a quasi-Hopf algebra which is known explicitly. Then the explicit definition of degree two indicators of modules over quasi-Hopf algebras in [Mason and Ng 2005] can be applied.

In principle the same approach, now using the higher indicator formula for quasi-Hopf algebras from [Ng and Schauenburg 2008], could be used to obtain higher indicator formulas for group-theoretical categories. However, those formulas involve iterated applications of the associator elements of the relevant quasi-Hopf algebra dealing with the parentheses of iterated tensor products in the category. Applying them with the explicit quasi-Hopf structure deriving from the data of a group-theoretical fusion category seems a formidable task.

We will take an entirely different approach. The formula from [Ng and Schauenburg 2007a, Theorem 4.1], generalizing the "third formula" from [Kashina, Sommerhäuser and Zhu 2006], links higher Frobenius–Schur indicators in a spherical fusion category $\mathcal{C}$ to the ribbon structure of the Drinfeld center $\mathcal{Z}(\mathcal{C})$ and the functor from $\mathcal{C}$ to $\mathcal{Z}(\mathcal{C})$ adjoint to the underlying functor. The "third formula" was used in [Shimizu 2011] to calculate indicators in Tambara–Yamagami categories; in our context the approach is aided by the fact that the centers of group-theoretical fusion categories are easy to determine: a group-theoretical fusion category is the monoidal category of bimodules over the (twisted) group algebra of a subgroup H of a finite group G inside the category Vect_G of G-graded vector spaces (twisted by a three-cocycle on G). By [Schauenburg 2001], the Drinfeld center of such a bimodule category is equivalent to the Drinfeld center of the "ambient" category. In different language this means that group-theoretical fusion categories are Morita equivalent to the category of graded vector spaces with twisted associativity; see the survey [Nikshych 2013]. We will treat the case of a group-theoretical fusion category defined without cocycles. Thus $\mathcal{C} = {}^G_H\mathcal{M}_H$, the center is $\mathcal{Z}({}^G_H\mathcal{M}_H) = \mathcal{Z}(\mathrm{Vect}_G)$, equivalent to the category of modules over the Drinfeld double of G.

In a sense, the underlying functor $\mathcal{Z}(\mathrm{Vect}_G) \to {}^G_H\mathcal{M}_H$ is already known explicitly from [Schauenburg 2001], but we need to do more. Simple objects in ${}^G_H\mathcal{M}_H$ are parametrized by group-theoretical data, namely (equivalence classes of) pairs consisting of an element of G and an irreducible representation of a certain stabilizer subgroup of H. Simple objects of $\mathcal{Z}(\mathrm{Vect}_G)$ are also classified by group-theoretical data, (equivalence classes of) pairs consisting of an element of G and an irreducible representation of its centralizer. In Section 3, we will describe the underlying functor $\mathcal{Z}(\mathrm{Vect}_G) \to {}^G_H\mathcal{M}_H$ on the level of simple objects by a formula involving only the combinatorics and representation theory of subgroups of G. Given this description, one can turn things around and describe the adjoint functor ${}^G_H\mathcal{M}_H \to \mathcal{Z}(\mathrm{Vect}_G)$ equally explicitly. Admittedly the resulting description, while completely explicit and entirely on the level of groups, subgroups, and group representations, is quite unwieldy — this is perhaps natural, since one has to deal with how conjugacy classes and centralizers (involved in the description of modules over the Drinfeld double) relate to double cosets of a chosen subgroup, and stabilizers of one-sided cosets under the regular action (involved in the description of ${}^G_H\mathcal{M}_H$).

In Section 4, we will use the description of the adjoint functor and the "third formula" to obtain a formula for the higher indicators of the simple objects of ${}^G_H\mathcal{M}_H$. Luckily we do not need complete information about the adjoint, but only the traces of the ribbon structure on the images under the adjoint. This allows us to dramatically simplify the immediate result based on the complicated description of the adjoint to obtain a surprisingly simple-looking formula for the higher indicators. It is in fact even simpler than Natale's formula for second indicators, and uses only group characters and the combinatorics of group elements and subgroups, without mentioning the associated quasi-Hopf algebra and its characters at all. One should admit, though, that characters of the associated quasi-Hopf algebra are in turn described in more "basic" terms in [Natale 2005]. Also, our results are marred by the obvious limitation that they do not treat general group-theoretical categories, but only those in whose definition the relevant group cocycles are trivial — we have amended this limitation in [Schauenburg 2015].

We also treat variants of the indicator formula that are more complicated, involving passing to orbits under the action of auxiliary subgroups, but computationally advantageous for the same reason that they pass from sums over the entire group H to sums over certain orbits.

In Section 5, we will explicitly calculate indicators in several examples of fusion categories associated to an inclusion of symmetric groups $S_{n-2} \subset S_n$. We use the "simple" version of our indicator formula for the cases $n = 4, 5$. The cases $n = 6, 7$ illustrate how the more complicated versions reduce the size of the calculations needed down to a manageable size.

2. Preliminaries

Throughout the paper, G is a finite group, and $H \subset G$ a subgroup. We denote the adjoint action of G on itself by $x \triangleright g = xgx^{-1}$. If V is a representation of a subgroup $K \subset G$, and $x \in G$, we denote by $x \triangleright V$ the twisted representation of $x \triangleright K$ with the same underlying vector space V on which $y \in x \triangleright K$ acts like $x^{-1} \triangleright y \in K$.

We work over the field $\mathbb{C}$ of complex numbers; representations are complex representations; and characters are ordinary characters.

The category ${}^G_H\mathcal{M}_H := {}^{\mathbb{C}G}_{\mathbb{C}H}\mathcal{M}_{\mathbb{C}H}$ is defined as the category of $\mathbb{C}H$-bimodules over the group algebra of H, considered as an algebra in the category of $\mathbb{C}G$-comodules, that is, of G-graded vector spaces. Thus, an object of ${}^G_H\mathcal{M}_H$ is a G-graded vector space $M \in \mathrm{Vect}_G$ with a two-sided H-action compatible with the grading in the sense that $|hmk| = h|m|k$ for $h, k \in H$ and $m \in M$.

The category ${}^G_H\mathcal{M}_H$ is a fusion category. The tensor product is the tensor product of $\mathbb{C}H$-bimodules. Simple objects are parametrized by irreducible representations of the stabilizers of right cosets of H in G. More precisely, let $D \in H\backslash G/H$ be a double coset of H in G, let $d \in D$, and let $S = \mathrm{Stab}_H(dH) = H \cap (d \triangleright H)$ be the stabilizer in H of the right coset dH under the action of H on its right cosets in G. Then the subcategory ${}^D_H\mathcal{M}_H \subset {}^G_H\mathcal{M}_H$, defined to contain those objects the degrees of all of whose homogeneous elements lie in D, is equivalent to the category $\mathrm{Rep}(S)$ of representations of S. The equivalence ${}^D_H\mathcal{M}_H \to \mathrm{Rep}(S)$ takes M to $(M_{dH})/H \cong (M/H)_{dH/H}$, the space of those vectors in the quotient of M by the right action of H whose degrees lie in the right coset of d. Details are in [Zhu 2001; Schauenburg 2002a]. We will denote the inverse equivalence by $\mathcal{F}_d : \mathrm{Rep}(\mathrm{Stab}_H(dH)) \to {}^{HdH}_{H}\mathcal{M}_H$, so that we have a category equivalence

$$\bigoplus_d \mathrm{Rep}(\mathrm{Stab}_H(dH)) \xrightarrow{(\mathcal{F}_d)_d} {}^G_H\mathcal{M}_H$$

in which the sum runs over a set of representatives of the double cosets of H in G. Of course ${}^D_H\mathcal{M}_H$ can be described by choosing a different representative of D. If $h \in H$, then dh has the same right coset as d, and $\mathcal{F}_{dh} = \mathcal{F}_d$, while $\mathrm{Stab}_H(hdH) = h \triangleright \mathrm{Stab}_H(dH)$ and $\mathcal{F}_d(W) = \mathcal{F}_{hd}(h \triangleright W)$ for $W \in \mathrm{Rep}(\mathrm{Stab}_H(dh))$.

In the special case $H = G$, the above description, with the neutral element representing the sole class of G in G, amounts to the (well-known) equivalence $\mathrm{Rep}(G) \cong {}^G_G\mathcal{M}_G$ sending $V \in \mathrm{Rep}(G)$ to $V \otimes \mathbb{C}G$ with the regular right G-action and the diagonal left G-action. This is a monoidal category equivalence.

The category ${}^G_G\mathcal{YD} = {}^{\mathbb{C}G}_{\mathbb{C}G}\mathcal{YD}$ of (left-left) Yetter–Drinfeld modules over $\mathbb{C}G$ has objects the G-graded vector spaces with a left G-action compatible with the grading in the sense that $|gv| = g|v|g^{-1}$ for $g \in G$ and $v \in V \in {}^G_G\mathcal{YD}$. The category ${}^G_G\mathcal{YD}$ is the (right) center of the category ${}^G\mathcal{M}$ of G-graded vector spaces: the half-braiding

$c : U \otimes V \to V \otimes U$ between a graded vector space U and a Yetter–Drinfeld module V is given by $u \otimes v \mapsto |u|v \otimes u$. To calculate indicators using the "third formula" we also need the fact that the canonical pivotal structure of ${}^G_G\mathcal{YD}$ is given by the ordinary vector space isomorphism $V \to V^{**}$, so that pivotal trace and ordinary trace coincide. Finally, the ribbon automorphism θ of an object $V \in {}^G_G\mathcal{YD}$ is given by $\theta(v) = |v|v$.

Simple objects of ${}^G_G\mathcal{YD}$ are parametrized by irreducible representations of the centralizers in G of elements of G. (In fact, this can be viewed as a special case of the description of graded bimodules above, as we shall review in Example 4.7 below). More precisely, let $g \in G$ and $C_G(g)$, the centralizer of g in G. Then, a functor

$$\mathcal{G}_g : \mathrm{Rep}(C_G(g)) \to {}^G_G\mathcal{YD}$$

can be defined by sending $V \in \mathrm{Rep}(C_G(g))$ to the $\mathbb{C}G$-module

$$\mathrm{Ind}^G_{C_G(g)} V = \mathbb{C}G \otimes_{\mathbb{C}C_G(g)} V,$$

endowed with the grading given by $|x \otimes v| = xgx^{-1}$ for $x \in G$ and $v \in V$. We note the special case $g = 1$ which recovers the canonical (monoidal) inclusion functor $\mathrm{Rep}(G) \to {}^G_G\mathcal{YD}$. Summing over different elements, we obtain a category equivalence

$$\bigoplus_g \mathrm{Rep}(C_G(g)) \xrightarrow{(\mathcal{G}_g)_g} {}^G_G\mathcal{YD}.$$

The sum runs over a set of representatives of the conjugacy classes of G, and the image of the functor $\mathcal{G}_g$ consists of those Yetter–Drinfeld modules, the degrees of whose homogeneous elements lie in the conjugacy class of g. We note for later use that the ribbon automorphism of $\mathcal{G}_g(V)$ is $\theta(x\otimes v) = (x \triangleright g)(x \otimes v) = xg \otimes v = x \otimes gv$; the trace of θ^m is therefore $[G : C_G(g)]\,\chi(g^m)$, where χ denotes the character of V.

As a final piece of notation, we will write $\langle M, N\rangle := \dim_{\mathbb{C}}(\mathrm{Hom}_{\mathcal{C}}(M, N))$ for objects M, N in a semisimple category.

3. The center and the adjoint

By a result of Müger [2003], the Drinfeld center $\mathcal{Z}(\mathcal{C})$ of a pivotal fusion category $\mathcal{C}$ is a modular category, and the underlying functor $\mathcal{Z}(\mathcal{C}) \to \mathcal{C}$ has a two-sided adjoint $\mathcal{K}$. To handle the center of ${}^G_H\mathcal{M}_H$ and the adjoint functor $\mathcal{K}$ we use the fact [Schauenburg 2001] that the center of a category of bimodules in a tensor category $\mathcal{C}$ coincides, in many cases including the present one, with the center of $\mathcal{C}$ itself.

To be precise, we will use the "right center" $\overline{\mathcal{Z}}(\mathcal{C})$ whose objects are pairs (V, c) in which $c : X \otimes V \to V \otimes X$ is a half-braiding defined for any $X \in \mathcal{C}$, and we denote by $\overline{\mathcal{K}}$ the adjoint functor of the underlying functor $\overline{\mathcal{Z}}(\mathcal{C}) \to \mathcal{C}$.

Then, writing $\mathcal{C} = {}^G\mathcal{M} = \mathrm{Vect}_G$ for the category of G-graded vector spaces, we have a category equivalence

$$ {}^G_G\mathcal{YD} \cong \overline{\mathcal{Z}}(\mathcal{C}) \longrightarrow \overline{\mathcal{Z}}({}_{\mathbb{C}H}\mathcal{C}_{\mathbb{C}H}) = \overline{\mathcal{Z}}\big({}^G_H\mathcal{M}_H\big) $$

which sends $(N, c) \in \overline{\mathcal{Z}}(\mathcal{C})$ to an object of $\overline{\mathcal{Z}}({}_{\mathbb{C}H}\mathcal{C}_{\mathbb{C}H})$ whose underlying right $\mathbb{C}H$-module is $N \otimes \mathbb{C}H$, whose left $\mathbb{C}H$-module structure is given by

$$ \mathbb{C}H \otimes N \otimes \mathbb{C}H \xrightarrow{c\otimes\mathbb{C}H} N \otimes \mathbb{C}H \otimes \mathbb{C}H \xrightarrow{N\otimes\nabla} N \otimes \mathbb{C}H, $$

and whose half-braiding (which we do not need) is induced by the half-braiding of N.

Thus, we identify $\overline{\mathcal{Z}}({}^G_H\mathcal{M}_H) = {}^G_G\mathcal{YD}$, and we identify the underlying functor $\overline{\mathcal{Z}}({}^G_H\mathcal{M}_H) \to {}^G_H\mathcal{M}_H$ with the functor

$$ \mathcal{U} : {}^G_G\mathcal{YD} \ni N \longrightarrow N \otimes \mathbb{C}H \in {}^G_H\mathcal{M}_H, $$

where the obvious right $\mathbb{C}H$-module $N \otimes \mathbb{C}H$ has left module structure given by $a(n \otimes b) = an \otimes ab$ and grading given by $|n \otimes b| = |n|b$.

Next, let $g \in G$, set $C := C_G(g)$, and let $V \in \mathrm{Rep}(C)$. We consider

$$ \mathcal{U}\mathcal{G}_g(V) = \mathbb{C}G \underset{\mathbb{C}C}{\otimes} V \otimes \mathbb{C}H \in {}^G_H\mathcal{M}_H. $$

Let $\mathfrak{X}_g$ be a set of representatives of the double cosets in $H\backslash G/C$, giving the decomposition $G = \bigsqcup_{x\in\mathfrak{X}_g} HxC$. Then each $\mathbb{C}HxC \underset{\mathbb{C}C}{\otimes} V \otimes \mathbb{C}H \subset \mathbb{C}G \underset{\mathbb{C}C}{\otimes} V \otimes \mathbb{C}H$ is a subobject in ${}^G_H\mathcal{M}_H$, and we have

$$ \mathcal{U}\mathcal{G}_g(V) = \bigoplus_{x\in\mathfrak{X}_g} \mathbb{C}HxC \underset{\mathbb{C}C}{\otimes} V \otimes \mathbb{C}H. $$

Note that the degrees of the homogeneous elements of $\mathbb{C}HxC \otimes_{\mathbb{C}C} V \otimes H$ lie in the double coset $H(x \triangleright g)H$, so that $\mathbb{C}HxC \otimes_{\mathbb{C}C} V \otimes \mathbb{C}H$ is in the image of the functor $\mathcal{F}_{x\triangleright g}$. To calculate the preimage, observe first that the degree of $hxc \otimes v \otimes h' \in \mathbb{C}HxC \otimes_{\mathbb{C}C} V \otimes \mathbb{C}H$ is $(hx \triangleright g)h'$, and thus is in $(x \triangleright g)H$ if and only if $h \in \mathrm{Stab}_H((x \triangleright g)H) =: J$. Hence

$$ \mathbb{C}HxC \underset{\mathbb{C}C}{\otimes} V \otimes \mathbb{C}H = \mathcal{F}_{x\triangleright g}(\mathbb{C}JxC \underset{\mathbb{C}C}{\otimes} V). $$

Next, observe that for $j, \tilde{j} \in J$ and $c, \tilde{c} \in C$, we have $jxc = \tilde{j}x\tilde{c}$ if and only if $\tilde{j}^{-1}j = x \triangleright (\tilde{c}c^{-1})$, which implies that we have an isomorphism

$$ \mathbb{C}JxC \underset{\mathbb{C}C}{\otimes} V \ni jxc \otimes v \mapsto j \otimes cv \in \mathbb{C}J \underset{\mathbb{C}[J\cap(x\triangleright C)]}{\otimes} (x \triangleright V). $$

Note that $J \cap (x \triangleright C) = \mathrm{Stab}_H((x \triangleright g)H) \cap C_G(x \triangleright g) = H \cap C_G(x \triangleright g) = H \cap x \triangleright C$.

We have shown:

$$\mathbb{C}HxC \underset{\mathbb{C}C}{\otimes} V \otimes \mathbb{C}H = \mathcal{F}_{x\triangleright g}\big(\operatorname{Ind}_{H\cap(x\triangleright C)}^{\operatorname{Stab}_H((x\triangleright g)H)} \operatorname{Res}^{x\triangleright C}_{H\cap(x\triangleright C))}(x\triangleright V)\big),$$

whence

$$\mathcal{UG}_g(V) = \bigoplus_{x\in\mathfrak{X}_g} \mathcal{F}_{x\triangleright g}\big(\operatorname{Ind}_{H\cap(x\triangleright C)}^{\operatorname{Stab}_H((x\triangleright g)H)} \operatorname{Res}^{x\triangleright C}_{H\cap(x\triangleright C)}(x\triangleright V)\big).$$

Let $d\in G$ and $S = \operatorname{Stab}_H(dH)$. Let $\mathfrak{H}_d$ be a set of representatives of H/S. Thus the double coset HdH is the disjoint union $HdH = \bigsqcup_{h\in\mathfrak{H}_d} hdH$; that is, $\mathfrak{H}_d d$ is a set of representatives of the right cosets contained in HdH.

If $x\triangleright g\in HdH$, then there is a unique $h\in\mathfrak{H}_d$ such that $(x\triangleright g)H = hdH$; thus, $\operatorname{Stab}_H((x\triangleright g)H) = \operatorname{Stab}_H(hdH) = h\triangleright S$, and for a representation N of $\operatorname{Stab}_H((x\triangleright g)H)$ we have $\mathcal{F}_{x\triangleright g}N = \mathcal{F}_{hd}N = \mathcal{F}_d(h^{-1}\triangleright N)$. Again $H\cap(x\triangleright C) = (h\triangleright S)\cap(x\triangleright C)$. Therefore,

$$\begin{aligned}(\mathcal{UG}_g(V))_{HdH} &= \bigoplus \mathcal{F}_{hd}\big(\operatorname{Ind}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C)} \operatorname{Res}^{x\triangleright C}_{(h\triangleright S)\cap(x\triangleright C)}(x\triangleright V)\big)\\ &= \bigoplus \mathcal{F}_d\big(h^{-1}\triangleright\big(\operatorname{Ind}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C)} \operatorname{Res}^{x\triangleright C}_{(h\triangleright S)\cap(x\triangleright C)}(x\triangleright V)\big)\big),\end{aligned}$$

where the sum is over all $x\in\mathfrak{X}_g$ and $h\in\mathfrak{H}_d$ such that $x\triangleright g\in hdH$, and if $W\in\operatorname{Irr}(S)$, then

$$\begin{aligned}\big\langle \mathcal{UG}_g(V), \mathcal{F}_d(W)\big\rangle &= \sum\big\langle h^{-1}\triangleright\big(\operatorname{Ind}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C)} \operatorname{Res}^{x\triangleright C}_{(h\triangleright S)\cap(x\triangleright C)}(x\triangleright V)\big), W\big\rangle\\ &= \sum\big\langle \operatorname{Ind}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C)} \operatorname{Res}^{x\triangleright C}_{(h\triangleright S)\cap(x\triangleright C)}(x\triangleright V), h\triangleright W\big\rangle.\end{aligned}$$

For the adjoint $\overline{\mathcal{K}}$ of $\mathcal{U}$, this implies, by Frobenius reciprocity:

$$\big\langle \overline{\mathcal{K}}\mathcal{F}_d(W), \mathcal{G}_g(V)\big\rangle = \sum\big\langle x^{-1}\triangleright\big(\operatorname{Ind}^{x\triangleright C}_{(h\triangleright S)\cap(x\triangleright C)} \operatorname{Res}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C)}(h\triangleright W)\big), V\big\rangle.$$

This means that we have calculated a formula for the adjoint $\overline{\mathcal{K}}$: denoting by $\mathfrak{C}$ a system of representatives for the conjugacy classes of G, we have

$$\begin{aligned}\overline{\mathcal{K}}\mathcal{F}_d(W) &= \sum \mathcal{G}_g\big(x^{-1}\triangleright\big(\operatorname{Ind}^{x\triangleright C_G(g)}_{(h\triangleright S)\cap(x\triangleright C_G(g))} \operatorname{Res}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C_G(g))}(h\triangleright W)\big)\big)\\ &= \sum \mathcal{G}_{x\triangleright g}\big(\operatorname{Ind}^{x\triangleright C_G(g)}_{(h\triangleright S)\cap(x\triangleright C_G(g))} \operatorname{Res}^{h\triangleright S}_{(h\triangleright S)\cap(x\triangleright C_G(g))}(h\triangleright W)\big),\end{aligned}$$

where the sum is over all $g\in\mathfrak{C}$, $x\in\mathfrak{X}_g$, and $h\in\mathfrak{H}_d$ such that $x\triangleright g\in hdH$. While this is clearly not a particularly pleasant or practical formula, we can say something in its favor: It expresses the functor $\overline{\mathcal{K}}$ entirely in terms of the groups involved and their representations, using, of course, the translation of group representations to objects in the two categories involved via the functors $\mathcal{F}$ and $\mathcal{G}$.

4. Indicator formulas for group inclusions

We retain the notations of the previous section, and proceed to calculate the higher Frobenius–Schur indicators of objects in ${}^G_H\mathcal{M}_H$. This is based on the categorical version of the "third formula" in [Kashina, Sommerhäuser and Zhu 2006, §6.4] that calculates indicators in a fusion category $\mathcal{C}$ through the adjoint $\overline{\mathcal{K}}$.

The formula obtained above for the adjoint $\overline{\mathcal{K}} : {}^G_H\mathcal{M}_H \to {}^G_G\mathcal{YD}$ yields, via [Ng and Schauenburg 2007a, Theorem 4.1], a formula for the higher indicators of the simple objects of ${}^G_H\mathcal{M}_H$. Since we are dealing with the right center, the relevant formula [op. cit., Remark 4.3] is

$$\nu_m(X) = \frac{1}{|G|}\operatorname{Tr}\big(\theta^{-m}_{\overline{\mathcal{K}}(X)}\big).$$

We proceed to use the information available on $\overline{\mathcal{K}}$ to apply it.

First, let η' be a character of $(h \triangleright S) \cap (x \triangleright C)$, and $\chi = \operatorname{Ind}^{x\triangleright C}_{(h\triangleright S)\cap(x\triangleright C)}(\eta')$. Then by a standard formula for induced characters,

$$\begin{aligned}\chi(x \triangleright g^m) &= \frac{1}{|(h \triangleright S) \cap (x \triangleright C)|} \sum_{\substack{y\in x\triangleright C\\ y\triangleright x\triangleright g^m \in h\triangleright S}} \eta'(y \triangleright x \triangleright g^m)\\ &= \begin{cases}[x \triangleright C : (h \triangleright S) \cap (x \triangleright C)]\,\eta'(x \triangleright g^m) & \text{if } x \triangleright g^m \in h \triangleright S,\\ 0 & \text{otherwise,}\end{cases}\end{aligned}$$

as elements in $x \triangleright C$ commute with $x \triangleright g^m$.

Let η be the character of $W \in \operatorname{Rep}(S)$, and let χ be the character of $V := \operatorname{Ind}^{x\triangleright C}_{h\triangleright S\cap x\triangleright C}\operatorname{Res}^{h\triangleright S}_{h\triangleright S\cap x\triangleright C}(h \triangleright \eta)$. Then

$$\begin{aligned}\operatorname{Tr}\big(\theta^m_{\mathcal{G}_{x\triangleright g}(V)}\big) &= [G : x \triangleright C]\,\chi(x \triangleright g^m)\\ &= \begin{cases}[G : (h \triangleright S) \cap (x \triangleright C)]\,\eta(h^{-1}x \triangleright g^m) & \text{if } x \triangleright g^m \in h \triangleright S,\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

By the formula for $\overline{\mathcal{K}}(\mathcal{F}_d(W))$ obtained in the previous section, this finally implies (using $|(h \triangleright S) \cap (x \triangleright C_G(g))| = |S \cap (h^{-1}x \triangleright C_G(g)| = |S \cap C_G(h^{-1}x \triangleright g)|$) that

$$\nu_m(\mathcal{F}_d(W)) = \sum \frac{1}{|S \cap C_G(h^{-1}x \triangleright g)|}\,\overline{\eta}(h^{-1}x \triangleright g^m), \tag{1}$$

where the sum is over $g \in \mathfrak{C}$, $x \in \mathfrak{X}_g$, and $h \in \mathfrak{H}_d$ such that $x \triangleright g \in hdH$ and $x \triangleright g^m \in h \triangleright S$. Surely this sum is not pleasant to work with; it involves summing over all conjugacy classes of the group and all representatives of certain double cosets, as well as over the coset representatives in $\mathfrak{H}_d$, albeit that last sum involves either no summand (for many combinations of g and x we might have $x \triangleright g \notin HdH$), or just one summand (the representative of the unique right coset containing $x \triangleright g$).

We shall process it further using the observation

$$HdH = \bigsqcup_{\substack{g\in\mathfrak{C},\ x\in\mathfrak{X}_g\\ x\triangleright g\in HdH}} H\triangleright(x\triangleright g) = \bigsqcup_{\substack{g\in\mathfrak{C},\ x\in\mathfrak{X}_g\\ h\in\mathfrak{H}_d\\ x\triangleright g\in hdH}} H\triangleright(h^{-1}x\triangleright g). \tag{2}$$

For the first equality, one has to check when $x\triangleright g$ and $y\triangleright g$, for $x, y\in G$, are in the same orbit of the action of H on G by conjugation:

$$\begin{aligned} \exists h\in H : h\triangleright(x\triangleright g) = y\triangleright g \quad &\iff \quad \exists h\in H : hxgx^{-1}h^{-1} = ygy^{-1}\\ &\iff \quad \exists h\in H : y^{-1}hx\in C_G(g)\\ &\iff \quad x\in HyC_G(g), \end{aligned}$$

while the second is an obvious reparametrization.

Thus, the set

$$\mathfrak{R}_d = \{h^{-1}x\triangleright g \mid g\in\mathfrak{C}, x\in\mathfrak{X}_g, h\in\mathfrak{H}_d, x\triangleright g\in hdH\} \tag{3}$$

is a set of representatives of the orbits of the action of H on HdH by conjugation. Moreover, $\mathfrak{R}_d\subset dH$. Thus, $\mathfrak{R}_d$ is a set of representatives of the orbits of the action of S on dH by conjugation. We have very nearly proved the main result of the paper:

Theorem 4.1. *Let G be a finite group, $H\subset G$ a subgroup, $d\in G$, $S=\mathrm{Stab}_H(dH)$, $W\in\mathrm{Rep}(S)$ with character η, and $\mathcal{F}_d(W)$ the object of ${}^G_H\mathcal{M}_H$ corresponding to W. Then*

$$\nu_m(\mathcal{F}_d(W)) = \frac{1}{|S|}\sum_{\substack{r\in dH\\ r^m\in S}}\overline{\eta}(r^m) = \frac{1}{|S|}\sum_{\substack{h\in H\\ (dh)^m\in S}}\overline{\eta}((dh)^m). \tag{4}$$

Proof. Substituting (3) in the indicator formula (1) yields

$$\nu_m(\mathcal{F}_d(W)) = \sum_{\substack{r\in\mathfrak{R}_d\\ r^m\in S}}\frac{1}{|S\cap C_G(r)|}\overline{\eta}(r^m). \tag{5}$$

But for $s\in S$ we have $(s\triangleright r)^m\in S\iff r^m\in S$, and $\eta((s\triangleright r)^m)=\eta(r^m)$ whenever $r^m\in S$. Since $S\cap C_G(r)$ is the stabilizer of r under the adjoint action of S, the first equality in (4) follows. The second equality is a trivial reparametrization. □

In the following we keep the notations of Theorem 4.1.

Remark 4.2. Note that for $r\in dH$ we have $r^m\in S\iff r^m\in H$. Thus we could modify the conditions in the sums (4) and subsequent similar sums, but in the examples that we treated, it seemed easier to check whether an element is in S than to check whether it is in H.

Remark 4.3. For $m \in \mathbb{N}$, the elements

$$\mu_m(d) := \frac{1}{|S|} \sum_{\substack{r \in dH \\ r^m \in S}} r^m = \frac{1}{|S|} \sum_{\substack{h \in H \\ (dh)^m \in S}} (dh)^m \in \mathbb{C}S \tag{6}$$

are central in the group algebra $\mathbb{C}S$, and $\nu_m(\mathcal{F}_d(W)) = \eta(\mu_m(d))$.

Remark 4.4. If $d \in C_G(H)$, then $S = H$, and for $h \in H$ we have $(dh)^m = d^m h^m \in H$ if and only if $d^m \in H$, so that

$$\mu_m(d) = \begin{cases} d^m \dfrac{1}{|H|} \displaystyle\sum_{h \in H} h^m & \text{if } d^m \in H, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

and therefore, since $d^m \in H$ is in the center of H,

$$\nu_m(\mathcal{F}_d(W)) = \begin{cases} \dfrac{\bar{\eta}(d^m)}{\eta(1)} \nu_m(W) & \text{if } d^m \in H, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The most obvious case of this is when $d = 1$; the image of $\mathcal{F}_1$ is the monoidal subcategory ${}^H_H\mathcal{M}_H \subset {}^G_H\mathcal{M}_H$, which is monoidally equivalent to $\operatorname{Rep}(H)$. The formula (8) can also be used to easily obtain examples where the higher indicators are not real: the cyclic group G of order 9, its generator d, its subgroup H of order 3, and a nontrivial irreducible character of the latter will do to obtain $\nu_3(\mathcal{F}_d(W))$, a nontrivial third root of unity.

Lemma 4.5. *Let $y \in S$. Then*

$$\sum_{\chi \in \operatorname{Irr}(S)} \nu_m(\mathcal{F}_d(\chi))\, \chi(y) = |\{h \in H \mid (dh)^m = y\}|. \tag{9}$$

In fact the function $\zeta_m(y) = |\{h \in H \mid (dh)^m = y\}|$ is easily seen to be a class function on S, so one can verify (9) by taking its scalar product with an irreducible character η. The left hand side gives the m-th indicator by the orthogonality relations, the right hand side by (4).

Remark 4.6. Assume that $H \subset G$ is part of an exact factorization, i.e., there exists a subgroup $L \subset G$ such that $LH = G$ and $L \cap G = \{1\}$. As pointed out in [Schauenburg 2002b], the category ${}^G_H\mathcal{M}_H$ is then equivalent to the category of modules over a bismash product Hopf algebra $\mathbb{C}^L \# \mathbb{C}H$. Thus, our results comprise a method to calculate indicators for bismash product Hopf algebras (of which the double below is a special case).

Example 4.7. Let Γ be a finite group, $G = \Gamma \times \Gamma$ with diagonal embedding $\Delta : \Gamma \to \Gamma \times \Gamma$, and $H = \Delta(\Gamma)$. It is well known that the category ${}^G_H\mathcal{M}_H \cong {}^\Gamma_\Gamma\mathcal{M}^\Gamma_\Gamma$ is equivalent to the module category of the Drinfeld double of Γ (in fact this is a special case of [Schauenburg 1994]).

Let $\mathfrak{G}$ be a cross section of the conjugacy classes of Γ. Then $\{(\gamma, 1) \mid \gamma \in \mathfrak{G}\}$ is a cross section of the double cosets of H in G. Let $d = (\gamma, 1)$. Then $S = \mathrm{Stab}_H(dH) = \Delta(C_\Gamma(\gamma))$. Let $h = \Delta(\theta) \in H$ and $m \in \mathbb{N}$. Then $(dh)^m = (\gamma\theta, \theta)^m = ((\gamma\theta)^m, \theta^m)$, thus $(dh)^m \in S$ if and only if $(\gamma\theta)^m = \theta^m$. Therefore, our indicator formula yields

$$\nu_m(\mathcal{F}_d(W)) = \frac{1}{|C_\Gamma(\gamma)|} \sum_{\substack{\theta\in\Gamma\\ (\gamma\theta)^m=\theta^m}} \bar{\eta}(\theta^m). \tag{10}$$

This formula was obtained in [Kashina, Sommerhäuser and Zhu 2006]; see also [Iovanov, Mason and Montgomery 2014], where the corresponding special case of (9) can be found. Note that we can replace $\bar{\eta}$ by η since the indicators in this case are known to be real.

In the proof of Theorem 4.1 we have obtained the simple looking indicator formula (4) via the more complicated formula (5). But in fact the latter is, in some respects, better than the former: it involves a sum over fewer terms, namely orbits of the adjoint action of S instead of individual elements of dH. Of course, for this simplification we could have taken any section of the orbits on dH instead of $\mathfrak{R}_d$. In fact, we can also pass to orbits over a group different from S; also, it may be convenient to take orbits in H of the action on H corresponding to the adjoint action on dH:

Proposition 4.8. *In the notation of Theorem 4.1, set $E = C_G(d) \cap SC_G(S) \cap N_G(H)$. Then, $SE = ES$ is a subgroup of G. Let $S' \subset SE$ be a subgroup, and let $\mathfrak{R}'_d$ be a section of the orbits of dH under the adjoint action of S' on dH. Then,*

$$\nu_m(\mathcal{F}_d(W)) = \frac{1}{|S|} \sum_{\substack{r\in\mathfrak{R}'_d\\ r^m\in S}} \frac{|S'|}{|S'\cap C_G(r)|} \bar{\eta}(r^m). \tag{11}$$

Alternatively, let S' act on H by "twisted conjugation" defined by the formula $s \mathbin{\tilde{\triangleright}} h = (d^{-1} \triangleright s)hs^{-1}$. Let $\mathfrak{T}'_d$ be a system of representatives of the orbits. Then,

$$\nu_m(\mathcal{F}_d(W)) = \frac{1}{|S|} \sum_{\substack{h\in\mathfrak{T}'_d\\ (dh)^m\in S}} \frac{|S'|}{|S'\cap C_G(dh)|} \bar{\eta}((dh)^m). \tag{12}$$

Proof. Let $x \in E$ and $u \in S = H \cap (d \triangleright H)$. Then $x \triangleright u \in (x \triangleright H) \cap (xd \triangleright H) = H \cap d \triangleright H = S$ since $x \triangleright H = H$ and $xd = dx$ by hypothesis. Thus E normalizes S, and $SE = ES$ is a subgroup of G. Now let $x \in E$ and $h \in H$. Since $x \in SC_G(S)$, we have $(dh)^m \in S$ if and only if $x \triangleright (dh)^m \in S$; in fact, these two elements are then conjugate in S. The condition $x \in C_G(d)$ implies $x \triangleright (dh)^m = (d(x \triangleright h))^m$, and $x \in N_G(H)$ implies $x \triangleright h \in H$. Thus the action of S' on dH is well defined, and the condition $r^m \in S$ is invariant along the orbits, as well as the values $\eta(r^m)$

along those orbits where $r^m \in S$. This implies (11), since $S' \cap C_G(r)$ is the stabilizer of r. Since $s \triangleright (dh) = d(s \,\tilde{\triangleright}\, h)$ for $s \in S$ and $h \in H$, we obtain (12) by a simple reparametrization. □

Remark 4.9. The previous result is perhaps the most useful if $S' \subset C_G(d)$, so that the twisted adjoint action coincides with the adjoint action. At any rate, it allows us to replace H by a set of orbit representatives before passing to the nastier part of the calculations involved in applying the indicator formula to concrete examples.

To set notation for subsequent calculations, let $\overline{G}$ be the set of orbits of G under the adjoint action of S', and $\overline{S}$ the image of S in $\overline{G}$. We do not distinguish notationally elements of $\overline{G}$ from those of G. We also let $\tilde{H}$ be the set of orbits of the twisted adjoint action of S' on H, and $Q(d) := \sum_{h \in H} h \in \mathbb{C}\tilde{H}$. Set

$$T(d) := \sum_{h \in \mathfrak{T}'_d} [S' : S' \cap C_G(dh)]\, dh = d\, Q(d) \in \mathbb{C}\overline{G}. \tag{13}$$

Let $\mathbb{C}\overline{G} \ni x \mapsto x^{[m]} \in \mathbb{C}\overline{G}$ be the linear map induced by taking m-th powers of group elements. Let $\pi : \mathbb{C}\overline{G} \to \mathbb{C}\overline{S}$ be the linear projection annihilating $\overline{G} \setminus \overline{S}$. Then

$$\nu_m(\mathcal{F}_d(W)) = \overline{\eta}(\bar{\mu}_m(d)) \quad \text{with} \quad \bar{\mu}_m(d) = \frac{1}{|S|} \pi(T(d)^{[m]}). \tag{14}$$

Of course $\bar{\mu}_m(d)$ is just the image of $\mu_m(d)$ in $\mathbb{C}\overline{S}$.

5. Example calculations

Consider the symmetric group S_n and the subgroup $S_m \subset S_n$ for $m < n$. For $d \in S_n$ the stabilizer $\mathrm{Stab}_{S_m}(dS_n) = S_m \cap d \triangleright S_m$ consists of those permutations $\sigma \in S_m$ for which $d^{-1} \triangleright \sigma \in S_m$. For $d^{-1} \triangleright \sigma$ to fix every element greater than m it is necessary and sufficient that σ fix every element k with $d^{-1}(k) \notin \{1, \dots, m\}$. Thus $\mathrm{Stab}_{S_m}(dS_m) = S_{\{1,\dots,m\} \cap \{d(1),\dots,d(m)\}}$ is a symmetric group. We have seen that in general higher indicators for the objects of ${}^{G}_{H}\mathcal{M}_H$ are nonnegative rational linear combinations of character values of the stabilizers $\mathrm{Stab}_H\, dH$. Moreover, higher indicators for any pivotal fusion category are cyclotomic integers.

Proposition 5.1. *Let $m < n$. Then all values of the higher Frobenius–Schur indicators for the objects of ${}^{S_n}_{S_m}\mathcal{M}_{S_m}$ are integers.*

The following example shows that this can fail if we embed S_m into S_n in a different fashion.

Example 5.2. Consider

$$G = S_9 \supset H = \{\sigma \in S_9 \mid i \equiv j \pmod 3 \Rightarrow \sigma(i) \equiv \sigma(j) \pmod 3\},$$

so H is the subgroup of those permutations in S_9 that preserve conjugacy modulo 3. Thus $H \cong S_3$ is generated by $t = (123)(456)(789)$ and $s = (12)(45)(78)$.

The element $d = (147258369) \in S_9$ satisfies $d^3 = t$, so in particular $d^{-1} \triangleright t \in H$. On the other hand $d^{-1} \triangleright s = (12)(45)(79)$, so $d^{-1} \triangleright s \notin H$ because $1 \equiv 7 \pmod 3$ while $2 \not\equiv 9 \pmod 3$. It follows that $S = \mathrm{Stab}_H(dH) = \langle t \rangle$.

To compute $\mu_3(d)$, first observe that $d^3 = (dt)^3 = (dt^2)^3 = t$. The computation $ds = (157369)(248)$ and $(ds)^3 = (13)(56)(79) \notin S$ shows that $(dh)^3 \notin S$ for $h \in H \setminus \{1, t, t^2\}$, since such h are conjugate to s by powers of t, which commute with d. Thus $\mu_3(d) = t$.

In particular $\nu_3(\mathcal{F}_d(\eta)) = \zeta^{-1}$ is not real when $\eta(t) = \zeta$ is a nontrivial third root of unity.

We will now compute some of the indicator values for the canonically embedded subgroups $S_{n-2} \subset S_n$ (as we shall see, this contains, in a sense, the case $S_{n-1} \subset S_n$, or rather $S_{n-2} \subset S_{n-1}$). We note already that all the indicator values we will find are nonnegative.

For $n \geq 4$, it is easy to check that S_{n-2} has seven double cosets in S_n:

$$\begin{aligned}
&\{\sigma \in S_n \mid \sigma(n-1) = n-1, \sigma(n) = n)\} = S_{n-2},\\
&\{\sigma \in S_n \mid \sigma(n-1) \neq n-1, \sigma(n) = n)\},\\
&\{\sigma \in S_n \mid \sigma(n-1) = n-1, \sigma(n) \neq n)\},\\
&\{\sigma \in S_n \mid \sigma(n-1) = n, \sigma(n) = n-1)\},\\
&\{\sigma \in S_n \mid \sigma(n-1) = n, \sigma(n) \neq n-1)\},\\
&\{\sigma \in S_n \mid \sigma(n-1) \neq n, \sigma(n) = n-1)\},\\
&\{\sigma \in S_n \mid \{\sigma(n-1), \sigma(n)\} \cap \{n-1, n\} = \varnothing\}.
\end{aligned}$$

A convenient set of double coset representatives is $d_1 = (\,)$, $d_2 = (n-2, n-1)$, $d_3 = (n-2, n)$, $d_4 = (n-1, n)$, $d_5 = (n-2, n-1, n)$, $d_6 = (n-2, n, n-1)$, and $d_7 = (n-3, n-1)(n-2, n)$.

Note that d_2 and d_3 are conjugate by $(n-1, n)$. The same holds for d_5 and d_6. We have $\mathrm{Stab}_{S_{n-2}}(d_2 S_{n-2}) = \mathrm{Stab}_{S_{n-2}}(d_5 S_{n-2}) = S_{n-3}$, $\mathrm{Stab}_{S_{n-2}}(d_7 S_{n-2}) = S_{n-4}$, and $\mathrm{Stab}_{S_{n-2}}(d_4 S_{n-2}) = S_{n-2}$.

Note that every d_i commutes with the elements in $\mathrm{Stab}_{S_{n-2}}(d_i S_{n-2})$; this is particular to our choice of representatives. It implies that the twisted conjugation action of the stabilizers on the group S_{n-2} from Proposition 4.8 is the ordinary adjoint action.

Note further that d_4 commutes with the elements of S_{n-2}. By Remark 4.4 it follows that

$$\nu_m(\mathcal{F}_{(n-1,n)}(W)) = \begin{cases} \nu_m(W) & \text{if } m \text{ is even,} \\ 0 & \text{if } m \text{ is odd,} \end{cases} \tag{15}$$

for any $W \in \mathrm{Rep}(S_{n-2})$, while $\nu_m(\mathcal{F}_{()}(W)) = \nu_m(W)$.

Note also that $d_2 \in S_{n-1}$. Thus, the indicators for objects in $\mathcal{F}_{d_2}(\mathrm{Rep}(S_{n-2}))$ can also be viewed as indicators in the subcategory ${}^{S_{n-1}}_{S_{n-2}}\mathcal{M}_{S_{n-2}}$. The subgroup $S_{n-2} \subset S_{n-1}$ is part of an exact factorization, $S_{n-1} = C_{n-1} \cdot S_{n-2}$, where C_{n-1} denotes the cyclic group generated by the $(n-1)$-cycle $(1, 2, \dots, n-1)$. As remarked already, these indicators are indicators for modules over a bismash product Hopf algebra $\mathbb{C}^{C_{n-2}} \# \mathbb{C}S_{n-1}$. Observe that the exact factorization suggests a different choice of coset representative, namely the $(n-1)$-cycle instead of d_2. We have the feeling that d_2 is the better choice since the $(n-1)$-cycle does not commute with elements in the corresponding stabilizer.

Since the images of $\mathcal{F}_{d_2}$ and $\mathcal{F}_{d_3}$ are mapped to each other by an autoequivalence, as well as the images of $\mathcal{F}_{d_5}$ and $\mathcal{F}_{d_6}$, we can concentrate on the indicators of the objects in the images of $\mathcal{F}_{d_i}$ for $i = 2, 5, 7$. We will treat some of them below for small values of n.

$S_2 \subset S_4$. Consider $H = \langle (1\ 2) \rangle \subset G = S_4$. We have the following double coset representatives, with their right cosets and double cosets:

i	d_i	$d_i H \setminus \{d_i\}$	$H d_i H \setminus d_i H$	$\mathrm{Stab}_H(d_i H)$
1	$()$	$(1\ 2)$		H
2	$(2\ 3)$	$(1\ 2\ 3)$	$(1\ 3), (1\ 3\ 2)$	$\{()\}$
3	$(2\ 4)$	$(1\ 2\ 4)$	$(1\ 4), (1\ 4\ 2)$	$\{()\}$
4	$(3\ 4)$	$(1\ 2)(3\ 4)$		H
5	$(2\ 3\ 4)$	$(1\ 2\ 3\ 4)$	$(1\ 3\ 4), (1\ 3\ 4\ 2)$	$\{()\}$
6	$(2\ 4\ 3)$	$(1\ 2\ 4\ 3)$	$(1\ 4\ 3), (1\ 4\ 3\ 2)$	$\{()\}$
7	$(2\ 3)(1\ 4)$	$(1\ 4\ 2\ 3)$	$(1\ 4\ 2\ 3), (1\ 3\ 2\ 4)$	$\{()\}$

We proceed to list the sequences of the higher Frobenius–Schur indicators for all the simple objects of ${}^G_H\mathcal{M}_H$ in the images of the functors $\mathcal{F}_{d_i}$. These sequences are periodic, and we list them for one complete period.

For d_1, they are the sequences of the higher Frobenius–Schur indicators of the representations of H, namely $(1, \dots)$ with period one for the trivial representation, and $(1, 0, \dots)$ with period two for the nontrivial representation.

In all other cases, the only powers of the elements of $d_i H$ that lie in the stabilizer $\mathrm{Stab}_H(d_i H)$ are identity elements. (This requires only a glance for d_4, as the stabilizer itself is trivial in the other cases.) Thus, regardless of the choice of representation also in the d_4 case, the indicator ν_m counts how many of the two m-th powers of the two elements of $d_i H$ are trivial; the count is then divided by two in the d_4 case. Thus the indicator sequences, up to a full period, are

$$(\nu_m(\mathcal{F}_{d_i}(W)))_m = \begin{cases} (0, 1, 1, 1, 0, 2, \dots) & \text{for } i = 2, 3, \\ (0, 1, \dots) & \text{for } i = 4, \\ (0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 2, \dots) & \text{for } i = 5, 6, \\ (0, 1, 0, 2, \dots) & \text{for } i = 7. \end{cases}$$

(Note that the case d_4 was already treated above using Remark 4.4.)

$S_3 \subset S_5$. In this case we have the right cosets

i	d_i	$d_i S_3 \setminus \{d_i\}$
1	()	(12), (13), (23), (123), (132)
2	(34)	(12)(34), (143), (243), (1243), (1432)
3	conjugate preceding row by (45)	
4	(45)	(12)(45), (13)(45), (23)(45), (123)(45), (132)(45)
5	(345)	(12)(345), (1453), (2453), (12453), (14532)
6	conjugate preceding row by (45)	
7	(2435)	(14352), (15243), (25)(34), (143)(25), (152)(34)

We have

$$\mathrm{Stab}_{S_3}(d_i S_3) = \begin{cases} S_3 & \text{for } i = 1, 4, \\ S_2 = \langle (1\ 2) \rangle & \text{for } i = 2, 3, 5, 6, \\ \{()\} & \text{for } i = 7. \end{cases}$$

As indicated above, we will only treat the indicators for d_2, d_5, and d_7.

One sees that for $i = 2$, the only possibility for a power of an element of $d_i S_3$ to be in $\mathrm{Stab}_{S_3}(d_i S_3)$ is if that power is trivial. The same is of course true for $i = 7$. So the m-th indicators for the simple objects in the images of $\mathcal{F}_{d_i}$ for $i = 2, 7$ do not "see" the representations of $\mathrm{Stab}_{S_3}(d_i S_3)$, but only count the number of elements whose orders divide m; the count has to be divided by 2 if $i = 2$. We have

$$\nu_m(\mathcal{F}_{d_2}(W)) = \begin{cases} 0, & (m, 12) = 1, \\ 1, & (m, 12) = 2, 3, \\ 2, & (m, 12) = 4, 6, \\ 3, & (m, 12) = 12; \end{cases} \qquad \nu_m(\mathcal{F}_{d_7}(W)) = \begin{cases} 0, & (m, 60) = 1, 3, \\ 1, & (m, 60) = 2, \\ 2, & (m, 60) = 4, 5, 15, \\ 3, & (m, 60) = 6, 10, \\ 4, & (m, 60) = 12, 20, \\ 5, & (m, 60) = 30, \\ 6, & (m, 60) = 60. \end{cases}$$

Finally $d_5 S_3$ contains one element, $(1\ 2)(3\ 4\ 5)$, whose third power is in $\mathrm{Stab}_{S_3}(d_5 S_3) \setminus \{()\}$. Powers of the other elements are only in the stabilizer when they are trivial. Thus, we obtain

$$\mu_m(d_5) = \mu_m(d_6) = \begin{cases} 0 & \text{when } (m, 60) = 1, 2, \\ \frac{1}{2}(() + (1\,2)) & \text{when } (m, 60) = 3, \\ () & \text{when } (m, 60) = 4, 5, 6, 10, \\ 2() & \text{when } (m, 60) = 12, 20, 30, \\ \frac{1}{2}(3() + (1\,2)) & \text{when } (m, 60) = 15, \\ 3() & \text{when } (m, 60) = 60. \end{cases}$$

For the trivial representation W_0 of $\langle(1\,2)\rangle$, this yields

$$\nu_m(\mathcal{F}_{d_5}(W_0)) = \nu_m(\mathcal{F}_{d_6}(W_0)) = \begin{cases} 0 & \text{when } (m, 60) = 1, 2, \\ 1 & \text{when } (m, 60) = 3, 4, 5, 6, 10, \\ 2 & \text{when } (m, 60) = 12, 15, 20, 30, \\ 3 & \text{when } (m, 60) = 60. \end{cases}$$

For the nontrivial irreducible representation W_1 of $\langle(1\,2)\rangle$, we obtain

$$\nu_m(\mathcal{F}_{d_5}(W_1)) = \nu_m(\mathcal{F}_{d_6}(W_1)) = \begin{cases} 0 & \text{when } (m, 60) = 1, 2, 3, \\ 1 & \text{when } (m, 60) = 4, 5, 6, 10, 15, \\ 2 & \text{when } (m, 60) = 12, 20, 30, \\ 3 & \text{when } (m, 60) = 60. \end{cases}$$

$\boldsymbol{S_4 \subset S_6}$**.** Since $|S_4| = 24$, it seems worth reducing the size of calculations in this case by considering orbits of S_4 as outlined in Proposition 4.8. We will use $S' = \mathrm{Stab}_{S_4}(d_i S_4)$.

For $i = 2, 5$ the stabilizer is S_3. The orbits of S_4 under the adjoint action of S_3 are obtained by subdividing the well-known conjugacy classes of S_4 according to the placement of the letter 4 in the respective cycle structure. Trusting details to the reader, we state:

$$\begin{aligned} Q(d_i) = {} & () + 3(12) + 3(14) \\ & + 2(123) + 6(124) \\ & + 3(12)(34) \\ & + 6(1234). \end{aligned}$$

From this we obtain

$$\begin{aligned} T((45)) = (45)Q((45)) = {} & (45) + 3(12)(45) + 3(154) \\ & + 2(123)(45) + 6(1254) \\ & + 3(12)(354) \\ & + 6(12354) \end{aligned}$$

and

$$\begin{aligned} T((456)) = (456)\,Q((456)) = (456) &+ 3(12)(456) + 3(1564) \\ &+ 2(123)(456) + 6(12564) \\ &+ 3(12)(3564) \\ &+ 6(123564). \end{aligned}$$

Thus (omitting the neutral element and writing $3 := 3() \in \mathbb{C}\bar{S}$, etc.),

$$\begin{aligned}
\bar{\mu}_2((45)) &= \tfrac{1}{6}(1+3+2(123)) = \tfrac{1}{3}(2+(123)),\\
\bar{\mu}_3((45)) &= \tfrac{1}{6}(3+3(12)) = \tfrac{1}{2}(1+(12)),\\
\bar{\mu}_4((45)) &= \tfrac{1}{6}(1+3+2(123)+6) = \tfrac{1}{3}(5+(123)),\\
\bar{\mu}_5((45)) &= 1,\\
\bar{\mu}_6((45)) &= \tfrac{1}{6}(1+3+3+2+3) = 2,\\
\bar{\mu}_{10}((45)) &= \tfrac{1}{6}(1+3+2(123)+6) = \tfrac{1}{3}(5+(123)) = \bar{\mu}_4((45)),\\
\bar{\mu}_{12}((45)) &= \tfrac{1}{6}(1+3+3+2+6+3) = 3,\\
\bar{\mu}_{15}((45)) &= \tfrac{1}{6}(3+3(12)+6) = \tfrac{1}{2}(3+(12)),\\
\bar{\mu}_{30}((45)) &= \tfrac{1}{6}(1+3+3+2+3+6) = 3 = \bar{\mu}_{12}((45)),\\
\bar{\mu}_{20}((45)) &= \tfrac{1}{6}(1+3+2(123)+6+6) = \tfrac{1}{3}(8+(123)),\\
\bar{\mu}_{60}((45)) &= 4,\\
\bar{\mu}_2((456)) &= 0,\\
\bar{\mu}_3((456)) &= \tfrac{1}{6}(1+3(12)+2) = \tfrac{1}{2}(1+(12)),\\
\bar{\mu}_4((456)) &= \tfrac{1}{6}(3+3) = 1,\\
\bar{\mu}_5((456)) &= 1,\\
\bar{\mu}_6((456)) &= \tfrac{1}{6}(1+3+2+6) = 2,\\
\bar{\mu}_{10}((456)) &= 1,\\
\bar{\mu}_{12}((456)) &= \tfrac{1}{6}(1+3+3+2+3+6) = 3,\\
\bar{\mu}_{15}((456)) &= \tfrac{1}{6}(1+3(12)+2+6) = \tfrac{1}{2}(3+(12)),\\
\bar{\mu}_{20}((456)) &= \tfrac{1}{6}(3+6+3) = 2\\
\bar{\mu}_{30}((456)) &= \tfrac{1}{6}(1+3+2+6+6) = 3,\\
\bar{\mu}_{60}((456)) &= 4.
\end{aligned}$$

For d_7, the calculations are even more tedious; we now need the S_2-orbits of S_4, that is, the subdivision of the conjugacy classes of S_4 according to the placement of the letters 3, 4 in the cycle structure. Thus,

$$\begin{aligned}Q((35)(46)) = (\,) + (12) + 2(13) + 2(14) + (34)\\ + 2(123) + 2(124) + 2(134) + 2(143)\\ + (12)(34) + 2(13)(24)\\ + 2(1234) + 2(1243) + 2(1324)\end{aligned}$$

and

$$\begin{aligned}T((35)(46)) = (35)(46) + (12)(35)(46) + 2(153)(46) + 2(164)(35) + (3645)\\ + 2(1253)(46) + 2(1264)(35) + 2(15364) + 2(16453)\\ + (12)(3645) + 2(153)(264)\\ + 2(125364) + 2(126453) + 2(153264),\end{aligned}$$

giving

$$\begin{aligned}
&\bar{\mu}_2((35)(46)) = \bar{\mu}_3((35)(46)) = 1, && \bar{\mu}_{12}((35)(46)) = 10,\\
&\bar{\mu}_4((35)(46)) = 4, && \bar{\mu}_{15}((35)(46)) = 3,\\
&\bar{\mu}_5((35)(46)) = 2, && \bar{\mu}_{20}((35)(46)) = 6,\\
&\bar{\mu}_6((35)(46)) = 7, && \bar{\mu}_{30}((35)(46)) = 9,\\
&\bar{\mu}_{10}((35)(46)) = 3, && \bar{\mu}_{60}((35)(46)) = 12.
\end{aligned}$$

In particular, the indicators of the two simples in the image of $\mathcal{F}_{(35)(46)}$ are identical; while for the other cases, we have to distinguish between the three irreducible representations of S_3, to wit, the trivial representation W_0, the sign representation W_1, and the two-dimensional irreducible W_2. We obtain:

object		m (for ν_m)										
d_i	W_j	2	3	4	5	6	10	12	15	20	30	60
(45)	W_0	1	1	2	1	2	2	3	2	3	3	4
	W_1	1	0	2	1	2	2	3	1	3	3	4
	W_2	1	1	3	2	4	3	6	3	5	6	8
(456)	W_0	0	1	1	1	2	1	3	2	2	3	4
	W_1	0	0	1	1	2	1	3	1	2	3	4
	W_2	0	1	2	2	4	2	6	3	4	6	8
(35)(46)	any	1	1	4	2	7	3	10	3	6	9	12

$S_5 \subset S_7$. If we want to deal with the representations associated to $d_7 = (46)(57)$ as in the preceding example, we calculate with a sum $Q((46)(57))$ with as many terms as there are orbits in S_5 of the adjoint action of S_3. One can check that there are 28 orbits. But we can reduce the task considerably (if not quite by half) by extending the stabilizer to a larger group S' as indicated in Proposition 4.8. As the element $(45)(67)$ commutes with d_7 and $\mathrm{Stab}_{S_5}(d_7S_5)$, and normalizes S_5, we can choose $S' = S_3 \cdot \langle(45)(67)\rangle$. Thus, we get

$$\begin{aligned} Q((46)(57)) = {} & () + 3(12) + 6(14) + (45) \\ & + 2(123) + 12(124) + 6(145) \\ & + 6(12)(34) + 3(12)(45) + 6(14)(25) \\ & + 12(1234) + 12(1245) + 6(1425) \\ & + 6(12)(345) + 12(14)(235) + 2(45)(123) \\ & + 12(12345) + 12(12435) \end{aligned}$$

with “only” 18 terms. We calculate

$$\begin{aligned} T((46)(57)) = {} & (46)(57) + 3(12)(46)(57) + 6(164)(57) + (4756) \\ & + 2(123)(46)(57) + 12(1264)(57) + 6(16475) \\ & + 6(12)(364)(57) + 3(12)(4756) + 6(164)(275) \\ & + 12(12364)(57) + 12(126475) + 6(164275) \\ & + 6(12)(36475) + 12(164)(2375) + 2(4756)(123) \\ & + 12(1236475) + 12(1264375). \end{aligned}$$

From here, we can go through all the divisors m of the exponent 420 of S_7 to obtain the elements $\bar{\mu}_m$ and the indicators for the three irreducible representations of S_3. The Table 1 calculates $\bar{\mu}_m$ in two stages, giving first an “unsimplified” version of $\pi(T^{[m]})$ in an attempt to hint at how this intermediate result can really be read off quite directly from the expression for T obtained above.

For good measure, we shall also finish the calculations for $d_2 = (56)$ and $d_5 = (567)$. In each case $\mathrm{Stab}_{S_5}(d_iS_5) = S_4$, and

$$\begin{aligned} Q(d_i) = {} & () + 6(12) + 4(15) + 8(123) + 12(125) \\ & + 3(12)(34) + 12(12)(35) + 6(1234) + 24(1235) \\ & + 8(123)(45) + 12(125)(34) + 24(12345). \end{aligned}$$

Thus,

$$\begin{aligned} T((56)) = {} & (56) + 6(12)(56) + 4(165) + 8(123)(56) + 12(1265) \\ & + 3(12)(34)(56) + 12(12)(365) + 6(1234)(56) + 24(12365) \\ & + 8(123)(465) + 12(1265)(34) + 24(123465), \end{aligned}$$

m	$\pi(T((46)(57))^{[m]})$	$\bar{\mu}_m((46)(57))$	$\nu_m(\mathcal{F}_{(46)(57)}(W_j))$ W_0	W_1	W_2
2	$1+3+2(123)$	$\frac{1}{3}(2+(123))$	1	1	1
3	6	1	1	1	2
4	$1+3+1+2(123)+12+3+2(123)$	$\frac{1}{3}(10+2(123))$	4	4	6
5	$6+6(12)$	$1+(12)$	2	0	2
6	$1+3+6+2+6+6+12+6$	7	7	7	14
7	$12+12$	4	4	4	8
10	$1+3+2(123)+6+12+6$	$\frac{1}{3}(14+(123))$	5	5	9
12	$1+3+6+1+2+12+6+3+6+12+6+12+2$	12	12	12	24
14	$1+3+2(123)+12+12$	$\frac{1}{3}(14+(123))$	5	5	9
15	$6+6+6(12)$	$2+(12)$	3	1	4
20	$1+3+1+2(123)+12+6+3+12+6+2(123)$	$\frac{1}{3}(22+2(123))$	8	8	14
21	$6+12+12$	5	5	5	10
28	$1+3+1+12+3+2(123)+12+12$	$\frac{1}{3}(22+2(123))$	8	8	14
30	$1+3+6+2+6+6+6+12+12+6+6$	11	11	11	22
35	$6+6(12)+12+12$	$5+(12)$	6	4	10
42	$1+3+6+2+6+6+12+6+12+12$	11	11	11	22
60	$1+3+6+1+2+12+6+6+3+6+12+12+6+6+12+2$	16	16	16	32
70	$1+3+2(123)+6+12+6+12+12$	$\frac{1}{3}(26+(123))$	9	9	17
84	$1+3+6+1+2+12+6+3+6+12+6+12+2+12+12$	16	16	16	32
105	$6+6+6(12)+12+12$	$6+(12)$	7	5	12
140	$1+3+1+2(123)+12+6+3+12+6+2(123)+12+12$	$\frac{1}{3}(34+2(123))$	12	12	22
210	$1+3+6+2+6+6+6+12+12+6+6+12+12$	15	15	15	30
420		20	20	20	40

Table 1. Indicator calculations on $\mathrm{Im}(\mathcal{F}_{(46)(57)}) \subset {}^{S_7}_{S_5}\mathcal{M}_{S_5}$.

and

$$\begin{aligned}T((567)) = {} & (567)+6(12)(567)+4(1675)+8(123)(567)+12(12675)\\ & +3(12)(34)(567)+12(12)(3675)+6(1234)(567)+24(123675)\\ & +8(123)(4675)+12(12675)(34)+24(1234675).\end{aligned}$$

Thus, we obtain the elements $\bar{\mu}_m((56))$ and $\bar{\mu}_m((567))$ listed in Table 2.

From this information, together with the character table of S_4 given in Table 3, one can then calculate all the indicator values for the simples in the images of $\mathcal{F}_{(56)}$ and $\mathcal{F}_{(567)}$; see Table 4.

m	$\bar{\mu}_m((56))$	$\bar{\mu}_m((567))$
2	$\frac{1}{14}(5+4(123)+3(12)(34))$	0
3	$\frac{1}{2}(1+(12))$	$\frac{1}{8}(3+2(12)+(12)(34)+2(1234))$
4	$\frac{1}{3}(5+(123))$	$\frac{1}{6}(4+2(123))$
5	1	$\frac{1}{2}(1+(12))$
6	$\frac{1}{4}(11+(12)(34))$	$\frac{1}{4}(7+(12)(34))$
7	0	1
10	$\frac{1}{12}(17+4(123)+3(12)(34))$	1
12	4	3
14	$\frac{1}{12}(5+4(123)+3(12)(34))$	1
15	$\frac{1}{2}(3+(12))$	$\frac{1}{8}(7+6(12)+(12)(34)+2(1234))$
20	$\frac{1}{3}(8+(123))$	$\frac{1}{6}(10+2(123))$
21	$\frac{1}{2}(1+(12))$	$\frac{1}{8}(11+2(12)+(12)(34)+2(1234))$
28	$\frac{1}{6}(10+2(123))$	$\frac{1}{3}(5+(123))$
30	$\frac{1}{4}(15+(12)(34))$	$\frac{1}{4}(11+(12)(34))$
35	1	$\frac{1}{2}(3+(12))$
42	$\frac{1}{4}(11+(12)(34))$	$\frac{1}{4}(11+(12)(34))$
60	5	4
70	$\frac{1}{12}(17+4(123)+3(12)(34))$	2
84	4	4
105	$\frac{1}{2}(3+(12))$	$\frac{1}{8}(15+6(12)+(12)(34)+2(1234))$
210	$\frac{1}{4}(15+(12)(34))$	$\frac{1}{4}(15+(12)(34))$
420	5	5

Table 2. $\bar{\mu}_m((56)), \bar{\mu}_m((567)) \in \mathbb{C}S_4$ for indicators in ${}_{S_5}^{S_7}\mathcal{M}_{S_5}$.

	()	(12)	(123)	(12)(34)	(1234)
η_0	1	1	1	1	1
η_1	1	−1	1	1	−1
η_2	2	0	−1	2	0
η_3	3	1	0	−1	−1
η_4	3	−1	0	−1	1

Table 3. Character table of S_4.

	$\nu_m(\mathcal{F}_{(56)}(W_i))$					$\nu_m(\mathcal{F}_{(567)}(W_i))$				
m	W_0	W_1	W_2	W_3	W_4	W_0	W_1	W_2	W_3	W_4
2	1	1	1	1	1	0	0	0	0	0
3	1	0	1	2	1	1	0	1	1	1
4	2	2	3	5	5	1	1	1	2	2
5	1	1	2	3	3	1	0	1	2	1
6	3	3	6	8	8	2	2	4	5	5
7	0	0	0	0	0	1	1	2	3	3
10	2	2	3	4	4	1	1	2	3	3
12	4	4	8	12	12	3	3	6	9	9
14	1	1	1	1	1	1	1	2	3	3
15	2	1	3	5	4	2	0	2	3	2
20	3	3	5	8	8	2	2	3	5	5
21	1	0	1	2	1	2	1	3	4	4
28	2	2	3	5	5	2	2	3	5	5
30	4	4	8	11	11	3	3	6	8	8
35	1	1	2	3	3	2	1	3	5	4
42	3	3	6	8	8	3	3	6	8	8
60	5	5	10	15	14	4	4	8	12	12
70	2	2	3	4	4	2	2	4	6	6
84	4	4	8	12	12	4	4	8	12	12
105	2	1	3	5	4	3	1	4	6	5
210	4	4	8	11	11	4	4	8	11	11
420	5	5	10	15	15	5	5	10	15	15

Table 4. Indicators on $\operatorname{Im}(\mathcal{F}_{(56)})$, $\operatorname{Im}(\mathcal{F}_{(567)}) \subset {}^{S_7}_{S_5}\mathcal{M}_{S_5}$.

The GAP [2014] code on the next page can be used to calculate the higher indicators for objects in ${}^{G}_{H}\mathcal{M}_{H}$ for any finite group G and subgroup H available to GAP. It uses the simple but inefficient formula (4). Moreover it is written in the most

```
IndicatorForOneRep:=function(m,G,H,d,S,eta)
  local h,sum;
  sum:=0;
  for h in H do
    if (d*h)^m in S
      then sum:=sum+((d*h)^(-m))^eta;
    fi;
  od;
  return(sum/Size(S));
end;

IndicatorsForDoubleCoset:=function(G,H,d)
  local S,eta,irreps,m;
  S:=Intersection(H,H^(d^(-1)));
  irreps:=Irr(S);
  for m in DivisorsInt(Exponent(G)) do
    Print(m,":");
    for eta in irreps do
      Print(IndicatorForOneRep(m,G,H,d,S,eta),",");
    od;
    Print("\n");
  od;
end;
```

GAP code to compute indicators in ${}^G_H\mathcal{M}_H$.

straightforward manner, makes hardly any attempt to reduce the load of calculations, and blindly repeats the same steps several times instead. For the moment, we do not pursue the quest to write better code (storing intermediate results such as the elements μ_m instead of recalculating them for each representation), nor the task to make use of the improved formula in Proposition 4.8 to speed up matters. The clumsy code is sufficient to do any of the calculations done above "by hand" again in seconds. Thus it *could* have been used to verify these results *if* the author *had* had any reason to mistrust his capability to perform flawless computations. Also, *if* the original calculations *had* contained errors, the GAP code *could* have been used to track those down and possibly correct them.

As it stands, the code was also sufficiently efficient to check that the inclusions $S_6 \subset S_8$ as well as $S_7 \subset S_9$ continue to produce only nonnegative indicator values.

References

[Bantay 1997] P. Bantay, "The Frobenius–Schur indicator in conformal field theory", *Phys. Lett. B* **394**:1-2 (1997), 87–88. MR 98c:81195 Zbl 0925.81331

[Courter 2012] R. Courter, *Computing higher indicators for the double of a symmetric group*, Ph.D. thesis, University of Southern California, Los Angeles, 2012, Available at http://search.proquest.com/docview/1151507400. MR 3103664

[Etingof, Nikshych and Ostrik 2005] P. Etingof, D. Nikshych, and V. Ostrik, "On fusion categories", *Ann. of Math.* (2) **162**:2 (2005), 581–642. MR 2006m:16051 Zbl 1125.16025

[Fuchs et al. 1999] J. Fuchs, A. C. Ganchev, K. Szlachányi, and P. Vecsernyés, "S_4 symmetry of 6j symbols and Frobenius–Schur indicators in rigid monoidal C^* categories", *J. Math. Phys.* **40**:1 (1999), 408–426. MR 99k:81111 Zbl 0986.81044

[GAP 2014] The GAP Group, *GAP: groups, algorithms, and programming*, Version 4.7.4, 2014, Available at http://www.gap-system.org.

[Iovanov, Mason and Montgomery 2014] M. Iovanov, G. Mason, and S. Montgomery, "FSZ-groups and Frobenius–Schur indicators of quantum doubles", *Math. Res. Lett.* **21**:4 (2014), 757–779. MR 3275646 Zbl 06379829

[Jedwab and Montgomery 2009] A. Jedwab and S. Montgomery, "Representations of some Hopf algebras associated to the symmetric group S_n", *Algebr. Represent. Theory* **12**:1 (2009), 1–17. MR 2010d:16045 Zbl 1200.16048

[Kashina, Mason and Montgomery 2002] Y. Kashina, G. Mason, and S. Montgomery, "Computing the Frobenius–Schur indicator for abelian extensions of Hopf algebras", *J. Algebra* **251**:2 (2002), 888–913. MR 2003f:16061 Zbl 1012.16040

[Kashina, Sommerhäuser and Zhu 2006] Y. Kashina, Y. Sommerhäuser, and Y. Zhu, *On higher Frobenius–Schur indicators*, Memoirs of the American Mathematical Society **181**:855, American Mathematical Society, Providence, RI, 2006. MR 2007k:16071 Zbl 1163.16029

[Keilberg 2012] M. Keilberg, "Higher indicators for some groups and their doubles", *J. Algebra Appl.* **11**:2 (2012), Article ID #1250030. MR 2925444 Zbl 1255.16035

[Keilberg 2014] M. Keilberg, "Higher indicators for the doubles of some totally orthogonal groups", *Comm. Algebra* **42**:7 (2014), 2969–2998. MR 3178056 Zbl 06298117

[Linchenko and Montgomery 2000] V. Linchenko and S. Montgomery, "A Frobenius–Schur theorem for Hopf algebras", *Algebr. Represent. Theory* **3**:4 (2000), 347–355. MR 2001k:16073 Zbl 0971.16018

[Mason and Ng 2005] G. Mason and S.-H. Ng, "Central invariants and Frobenius–Schur indicators for semisimple quasi-Hopf algebras", *Adv. Math.* **190**:1 (2005), 161–195. MR 2005h:16066 Zbl 1100.16033

[Müger 2003] M. Müger, "From subfactors to categories and topology, II: The quantum double of tensor categories and subfactors", *J. Pure Appl. Algebra* **180**:1-2 (2003), 159–219. MR 2004f:18014 Zbl 1033.18003

[Natale 2005] S. Natale, "Frobenius–Schur indicators for a class of fusion categories", *Pacific J. Math.* **221**:2 (2005), 353–377. MR 2007j:16070 Zbl 1108.16035

[Ng and Schauenburg 2007a] S.-H. Ng and P. Schauenburg, "Frobenius–Schur indicators and exponents of spherical categories", *Adv. Math.* **211**:1 (2007), 34–71. MR 2008b:16067 Zbl 1138.16017

[Ng and Schauenburg 2007b] S.-H. Ng and P. Schauenburg, "Higher Frobenius–Schur indicators for pivotal categories", pp. 63–90 in *Hopf algebras and generalizations* (Evanston, IL, 2004), edited by L. H. Kauffman et al., Contemp. Math. **441**, American Mathematical Society, Providence, RI, 2007. MR 2008m:18015 Zbl 1153.18008

[Ng and Schauenburg 2008] S.-H. Ng and P. Schauenburg, "Central invariants and higher indicators for semisimple quasi-Hopf algebras", *Trans. Amer. Math. Soc.* **360**:4 (2008), 1839–1860. MR 2009d:16065 Zbl 1141.16028

[Nikshych 2013] D. Nikshych, "Morita equivalence methods in classification of fusion categories", pp. 289–325 in *Hopf algebras and tensor categories*, edited by N. Andruskiewitsch et al., Contemp. Math. **585**, American Mathematical Society, Providence, RI, 2013. MR 3077244 Zbl 06342030

[Ostrik 2003] V. Ostrik, "Module categories over the Drinfeld double of a finite group", *Int. Math. Res. Not.* **2003**:27 (2003), 1507–1520. MR 2004h:18005 Zbl 1044.18005

[Schauenburg 1994] P. Schauenburg, "Hopf modules and Yetter–Drinfel'd modules", *J. Algebra* **169**:3 (1994), 874–890. MR 95j:16047 Zbl 0810.16037

[Schauenburg 2001] P. Schauenburg, "The monoidal center construction and bimodules", *J. Pure Appl. Algebra* **158**:2-3 (2001), 325–346. MR 2002f:18013 Zbl 0984.18006

[Schauenburg 2002a] P. Schauenburg, "Hopf algebra extensions and monoidal categories", pp. 321–381 in *New directions in Hopf algebras*, edited by S. Montgomery and H.-J. Schneider, Math. Sci. Res. Inst. Publ. **43**, Cambridge University Press, 2002. MR 2003k:16055 Zbl 1014.16037

[Schauenburg 2002b] P. Schauenburg, "Hopf bimodules, coquasibialgebras, and an exact sequence of Kac", *Adv. Math.* **165**:2 (2002), 194–263. MR 2003e:16052 Zbl 1006.16054

[Schauenburg 2015] P. Schauenburg, "A higher Frobenius–Schur indicator formula for group-theoretical fusion categories", *Comm. Math. Phys.* **340** (2015), 833–849. MR 3397032 Zbl 06490964

[Shimizu 2011] K. Shimizu, "Frobenius–Schur indicators in Tambara–Yamagami categories", *J. Algebra* **332** (2011), 543–564. MR 2012b:18014 Zbl 1236.18012

[Zhu 2001] Y. Zhu, "Hecke algebras and representation ring of Hopf algebras", pp. 219–227 in *First International Congress of Chinese Mathematicians* (Beijing, 1998), edited by L. Yang and S.-T. Yau, AMS/IP Stud. Adv. Math. **20**, American Mathematical Society, Providence, RI, 2001. MR 2002c:20011 Zbl 1064.20011

Received August 13, 2014. Revised January 23, 2015.

PETER SCHAUENBURG
INSTITUT DE MATHÉMATIQUES DE BOURGOGNE
UMR 5584 CNRS
UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ
F-21000 DIJON
FRANCE
peter.schauenburg@u-bourgogne.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.203

CHORDAL GENERATORS AND THE HYDRODYNAMIC NORMALIZATION FOR THE UNIT BALL

SEBASTIAN SCHLEISSINGER

Let $c \geq 0$ and denote by $\mathcal{K}(\mathbb{H}, c)$ the set of all infinitesimal generators $G : \mathbb{H} \to \mathbb{C}$ on the upper half-plane $\mathbb{H}$ such that $\limsup_{y\to\infty} y \cdot |G(iy)| \leq c$. This class is related to univalent functions $f : \mathbb{H} \to \mathbb{H}$ with hydrodynamic normalization and appears in the so-called chordal Loewner equation.

In this paper, we generalize the class $\mathcal{K}(\mathbb{H}, c)$ and the hydrodynamic normalization to the Euclidean unit ball in $\mathbb{C}^n$. The generalization is based on the observation that $G \in \mathcal{K}(\mathbb{H}, c)$ can be characterized by an inequality for the hyperbolic length of $G(z)$.

1. Introduction

One-parameter semigroups. Let $\mathbb{B}_n = \{z \in \mathbb{C}^n \mid \|z\| < 1\}$ be the Euclidean unit ball in $\mathbb{C}^n$. In one dimension we write $\mathbb{D} := \mathbb{B}_1$ for the unit disc.

Definition 1.1. A continuous one-real-parameter semigroup of holomorphic functions on $\mathbb{B}_n$ is a map $[0, \infty) \ni t \mapsto \Phi_t \in \mathcal{H}(\mathbb{B}_n, \mathbb{B}_n)$ satisfying the following conditions:

(1) Φ_0 is the identity.

(2) $\Phi_{t+s} = \Phi_t \circ \Phi_s$ for all $t, s \geq 0$.

(3) Φ_t tends to the identity locally uniformly in $\mathbb{B}_n$, when t tends to 0.

Supported by the ERC grant "HEVO - Holomorphic Evolution Equations" n. 277691.
MSC2010: 20M20, 32A40, 37L05.
Keywords: semigroups of holomorphic mappings, infinitesimal generators, hydrodynamic normalization, chordal Loewner equation.

Given such a semigroup $\{\Phi_t\}_{t\geq 0}$ and a point $z \in \mathbb{B}_n$, the limit

$$G(z) := \lim_{t\to 0} \frac{\Phi_t(z) - z}{t}$$

exists and the vector field $G : \mathbb{B}_n \to \mathbb{C}^n$, called the *infinitesimal generator*[1] of Φ_t, is a holomorphic function (see, e.g., [Abate 1992]). We denote by $\mathrm{Inf}(\mathbb{B}_n)$ the set of all infinitesimal generators of semigroups in $\mathbb{B}_n$. For any $z \in \mathbb{B}_n$, the map $w(t) := \Phi_t(z)$ is the solution of the initial value problem

$$\frac{dw(t)}{dt} = G(w(t)), \quad w(0) = z. \tag{1-1}$$

There are various characterizations of holomorphic functions $G : \mathbb{B}_n \to \mathbb{C}^n$ that are infinitesimal generators; see [Reich and Shoikhet 2005, Section 7.3], [Bracci et al. 2010, Theorem 0.2], [Bracci et al. 2014, p. 193].

The set $\mathrm{Inf}(\mathbb{D})$, i.e., all infinitesimal generators in the unit disc, can be characterized completely by the Berkson–Porta representation formula [1978]

$$\mathrm{Inf}(\mathbb{D}) = \{z \mapsto (\tau - z)(1 - \bar{\tau}z)p(z) \mid \tau \in \overline{\mathbb{D}},\ p \in \mathcal{H}(\mathbb{D}, \mathbb{C}) \\ \text{with } \mathrm{Re}(p(z)) \geq 0 \text{ for all } z \in \mathbb{D}\}. \tag{1-2}$$

Remark 1.2. Let $F : \mathbb{D} \to \mathbb{D}$ be a holomorphic self-map. Recall the Denjoy–Wolff theorem (see, e.g., [Reich and Shoikhet 2005, Theorem 5.1]): If F is not an elliptic automorphism (i.e., an automorphism with exactly one fixed point in $\mathbb{D}$), then there exists one point $\tau \in \overline{\mathbb{D}}$ (the Denjoy–Wolff point of F) such that the iterates F^n converge locally uniformly in $\mathbb{D}$ to the constant map τ.

If $\{\Phi_t\}_{t\geq 0}$ is a semigroup on $\mathbb{D}$, then we call $\tau \in \overline{\mathbb{D}}$ the Denjoy–Wolff point of $\{\Phi_t\}_{t\geq 0}$ if τ is the Denjoy–Wolff point of Φ_1, which is equivalent to $\lim_{t\to\infty} \Phi_t = \tau$ locally uniformly.

If an infinitesimal generator in the unit disc does not generate a semigroup of elliptic automorphisms of $\mathbb{D}$, then the point $\tau \in \overline{\mathbb{D}}$ from formula (1-2) is exactly the Denjoy–Wolff point of the semigroup.

There are two special cases of infinitesimal generators in $\mathbb{D}$ that have been studied intensively and turned out to be quite useful in Loewner theory and its applications. The two different cases arise from certain normalizations of the Berkson–Porta data τ and p from formula (1-2). In the *radial* case, one considers those elements $G \in \mathrm{Inf}(\mathbb{D})$ whose Berkson–Porta data τ and p satisfy

$$\tau = 0 \quad \text{and} \quad p(0) = 1,$$

i.e., $G(z) = -zp(z)$.

[1] There is no standard convention in the literature and often $-G$ is called the infinitesimal generator of the semigroup.

This class plays a central role in studying the class S of all univalent functions $f : \mathbb{D} \to \mathbb{C}$ with $f(0) = 0$, $f'(0) = 1$, via the powerful tools of Loewner's theory, which considers a nonautonomous version of (1-1); see, e.g., [Pommerenke 1975, Chapter 6]. The class of radial generators as well as the class S have been generalized in this context to the polydisc $\mathbb{D}^n$ (see [Poreda 1987a; 1987b]), and to the unit ball $\mathbb{B}_n$ (see [Graham and Kohr 2003] for a collection of several results and references).

The second class, the set of all *chordal* generators[2], consists of all $G \in \mathrm{Inf}(\mathbb{D})$ whose Berkson–Porta data τ and p satisfy

$$\tau = 1 \quad \text{and} \quad \angle \lim_{z \to 1} \frac{p(z)}{z-1} \text{ is finite.}$$

The aim of this paper is to introduce a generalization of the chordal class for the unit ball $\mathbb{B}_n$.

The hydrodynamic normalization in one dimension. Instead of fixing an interior point, like in the class S, it can be of interest to investigate univalent self-mappings of $\mathbb{D}$ that fix a boundary point. In this case, one usually passes from $\mathbb{D}$ to the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$.

A class of such mappings that is easy to describe and that appears in several applications is the set of all univalent mappings $f : \mathbb{H} \to \mathbb{H}$ that fix the boundary point ∞ and have the so-called *hydrodynamic normalization*. Basic properties of this class can be found in [Goryaĭnov and Ba 1992]; see also [Bauer 2005; Contreras et al. 2010]. One of its main applications is the chordal Loewner equation; see [Abate et al. 2010, Section 4] for further references.

A univalent function $f : \mathbb{H} \to \mathbb{H}$ has *hydrodynamic normalization* (at ∞) if f has the expansion

$$f(z) = z - \frac{c}{z} + \gamma(z),$$

where $c \geq 0$, which is usually called *half-plane capacity*, and γ satisfies

$$\angle \lim_{z \to \infty} z \cdot \gamma(z) = 0.$$

We denote by $\mathfrak{P}$ the set of all these functions. Then $\mathfrak{P}$ is a semigroup and the functional $l : \mathfrak{P} \to [0, \infty)$, $l(f) = c$, is additive: if $f_1, f_2 \in \mathfrak{P}$, then $f_1 \circ f_2 \in \mathfrak{P}$ and $l(f_1 \circ f_2) = l(f_1) + l(f_2)$.

Remark 1.3. Let $f \in \mathfrak{P}$ with $l(f) = c$. If we transfer f to the unit disc by conjugation by the Cayley transform, then we obtain a function $\tilde{f} : \mathbb{D} \to \mathbb{D}$ having

[2] Note that there is no standard use of the words "radial" and "chordal" in the literature. In [Contreras et al. 2010], e.g., an element $G \in \mathrm{Inf}(\mathbb{D})$ is called *radial* if $\tau \in \mathbb{D}$ and chordal if $\tau \in \partial\mathbb{D}$.

the expansion

$$\tilde{f}(z) = z - \frac{c}{4}(z-1)^3 + \tilde{\gamma}(z),$$

where $\angle \lim_{z\to 1} \tilde{\gamma}(z)/(z-1)^3 = 0$.

If $\{\Phi_t\}_{t\geq 0}$ is a one-real-parameter semigroup contained in $\mathfrak{P}$ with $l(\Phi_1) = a$, then it is easy to see that $l(\Phi_t) = a\cdot t$. If H is the generator of this semigroup, then we also define $l(H) := a$.

We will be interested in the following set of chordal generators.

Definition 1.4. By $\mathcal{K}(\mathbb{H}, c)$ we denote the set of all infinitesimal generators H of one-real-parameter semigroups $\{\Phi_t\}_{t\geq 0}$ contained in $\mathfrak{P}$ with $l(H) \leq c$.

Remark 1.5. The set $\mathcal{K}(\mathbb{H}, c)$ can be characterized in various ways; see [Goryaĭnov and Ba 1992, Section 1] and [Maassen 1992, Proposition 2.2].

It is known that $H \in \mathcal{K}(\mathbb{H}, c)$ for some $c \geq 0$ if and only if H maps $\mathbb{H}$ into $\overline{\mathbb{H}}$ and

$$\limsup_{y\to\infty} y|H(iy)| \leq c. \tag{1-3}$$

In fact, $l(H) = \limsup_{y\to\infty} y|H(iy)|$.

Furthermore, this is equivalent to H maps $\mathbb{H}$ into $\overline{\mathbb{H}}$ and

$$|H(z)| \leq \frac{c}{\operatorname{Im}(z)} \tag{1-4}$$

for all $z \in \mathbb{H}$. The number $l(H)$ is the smallest constant such that this inequality holds.

Finally, it is known that this property is equivalent to the fact that $-G$ is the Cauchy transform of a finite, nonnegative Borel measure μ on $\mathbb{R}$, i.e.,

$$H(z) = \int_{\mathbb{R}} \frac{\mu(du)}{u-z}. \tag{1-5}$$

The number $l(H)$ can be calculated by $l(H) = \mu(\mathbb{R})$.

Remark 1.6. It is easy to see that the following holds: if $f \in \mathfrak{P}$ with $c = l(f)$, then $H := f - \mathrm{id} \in \mathcal{K}(\mathbb{H}, c)$ with $l(H) = c$.

Let $C : \mathbb{H} \to \mathbb{D}$, $C(z) = (z-i)/(z+i)$, be the Cayley map. We define $\mathcal{K}(\mathbb{D}, c)$ by

$$\mathcal{K}(\mathbb{D}, c) = \{C'(C^{-1}) \cdot (H \circ C^{-1}) \mid H \in \mathcal{K}(\mathbb{H}, c)\}.$$ [3]

The rest of this paper is organized as follows: In Section 2 we look for an invariant characterization of chordal generators, i.e., of the sets $\mathcal{K}(\mathbb{H}, c)$ and $\mathcal{K}(\mathbb{D}, c)$, and we introduce the class $\mathcal{K}(\mathbb{B}_n, c)$ for the higher-dimensional unit ball. It will turn out to be quite useful to study "slices" of this class, which is done in Section 3. In Section 4 we introduce and study the class $\mathfrak{P}_n$, a higher-dimensional analog of the class $\mathfrak{P}$.

[3] If $\{\Phi_t\}_{t\geq 0}$ is a semigroup in $\mathbb{H}$ with generator H, then $\{C \circ \Phi_t \circ C^{-1}\}_{t\geq 0}$ is a semigroup in $\mathbb{D}$ and its generator is given by $C'(C^{-1}) \cdot (H \circ C^{-1})$.

2. Chordal generators in higher dimensions

Invariant formulation for $\mathcal{K}(\mathbb{D}, c)$ and $\mathcal{K}(\mathbb{H}, c)$. For $R > 0$, we let $E_{\mathbb{D}}(1, R)$ be the horodisc in $\mathbb{D}$ with center 1 and radius R, i.e.,

$$E_{\mathbb{D}}(1, R) = \left\{ z \in \mathbb{D} \;\middle|\; \frac{1}{|u_{\mathbb{D}}(z)|} < R \right\},$$

where $u_{\mathbb{D}}(z) = -(1 - |z|^2)/|1 - z|^2$ is the Poisson kernel in $\mathbb{D}$ with respect to 1.

By using the Cayley map, we define analogously

$$E_{\mathbb{H}}(\infty, R) = C^{-1}(E_{\mathbb{D}}(1, R)) = \left\{ z \in \mathbb{H} \;\middle|\; \frac{1}{\operatorname{Im}(z)} < R \right\}.$$

For $z \in \mathbb{D}$ and a tangent vector $v \in \mathbb{C}$, we denote by $|v|_{\mathbb{D},z}$ the hyperbolic length of v, i.e.,

$$|v|_{\mathbb{D},z} := \frac{2|v|}{1 - |z|^2}.$$

Furthermore, we let $R_{\mathbb{D}}(z)$ be the radius R of the horodisc $E_{\mathbb{D}}(1, R)$ that satisfies $z \in \partial E(1, R)$; in short, $R_{\mathbb{D}}(z) = 1/|u_{\mathbb{D}}(z)|$. Analogously, for $z \in \mathbb{H}$ and $v \in \mathbb{C}$, we define $R_{\mathbb{H}}(z) := 1/\operatorname{Im}(z)$ and the hyperbolic length $|v|_{\mathbb{H},z} := |v|/\operatorname{Im}(z)$.

According to (1-4), we know that $H \in \mathcal{K}(\mathbb{H}, c)$ if and only if H maps $\mathbb{H}$ into $\overline{\mathbb{H}}$ and $|H(z)| \le c/\operatorname{Im}(z)$ for all $z \in \mathbb{H}$. By using the Berkson–Porta formula, it is easy to see that we can rephrase this to: $H \in \mathcal{K}(\mathbb{H}, c)$ if and only if $H \in \operatorname{Inf}(\mathbb{H})$ and $|H(z)| \le c/\operatorname{Im}(z)$ for all $z \in \mathbb{H}$.

The last inequality is equivalent to $|H(z)|/\operatorname{Im}(z) \le c/\operatorname{Im}(z)^2$ or

$$|H(z)|_{\mathbb{H},z} \le \frac{c}{\operatorname{Im}(z)^2} = c \cdot R_{\mathbb{H}}(z)^2.$$

If we pass from $\mathbb{H}$ to $\mathbb{D}$ and transform H into $G = C'(C^{-1}) \cdot (H \circ C^{-1})$, then G satisfies $|G(C(z))|_{\mathbb{D},C(z)} = |H(z)|_{\mathbb{H},z}$ and we immediately get the following characterization.

Proposition 2.1. *Let $G \in \operatorname{Inf}(\mathbb{D})$. Then*

$$G \in \mathcal{K}(\mathbb{D}, c) \quad \Longleftrightarrow \quad |G(z)|_{\mathbb{D},z} \le c \cdot R_{\mathbb{D}}(z)^2 \text{ for all } z \in \mathbb{D}.$$

Let $H \in \operatorname{Inf}(\mathbb{H})$. Then

$$H \in \mathcal{K}(\mathbb{H}, c) \quad \Longleftrightarrow \quad |H(z)|_{\mathbb{H},z} \le c \cdot R_{\mathbb{H}}(z)^2 \text{ for all } z \in \mathbb{H}.$$

Chordal generators in the unit ball. For $n \in \mathbb{N}$, let u_n be the pluricomplex Poisson kernel in $\mathbb{B}_n$ with pole at $e_1 := (1, 0, \ldots, 0)$, i.e.,

$$u_{\mathbb{B}_n,p} = -\frac{1 - \|z\|^2}{|1 - z_1|^2}.$$

The level sets of $u_{\mathbb{B}_n}$ are exactly the boundaries of horospheres with center e_1; more precisely, the set

$$E_{\mathbb{B}_n}(e_1, R) := \{z \in \mathbb{B}_n \mid |u_{\mathbb{B}_n}(z)|^{-1} < R\}, \quad R > 0,$$

is the horosphere with center e_1 and radius R.

Furthermore, for $z \in \mathbb{B}_n$ and $v \in \mathbb{C}^n$, we denote by $\|v\|_{\mathbb{B}_n,z}$ the Kobayashi-hyperbolic length of the vector v with respect to z.

Motivated by Proposition 2.1, we make the following definition.

Definition 2.2. Let $c \geq 0$. We define the class $\mathcal{K}(\mathbb{B}_n, c)$ to be the set of all infinitesimal generators G on $\mathbb{B}_n$ such that, for all $z \in \mathbb{B}_n$,

$$\|G(z)\|_{\mathbb{B}_n,z} \leq \frac{c}{u_{\mathbb{B}_n}(z)^2}. \tag{2-1}$$

Remark 2.3. $\mathcal{K}(\mathbb{B}_n, c)$ is a compact family: Montel's theorem and the definition of $\mathcal{K}(\mathbb{B}_n, c)$ immediately imply that it is a normal family. If a sequence $(G_n) \subset \mathcal{K}(\mathbb{B}_n, c)$ converges locally uniformly to $G : \mathbb{B}_n \to \mathbb{C}^n$, then G is holomorphic and also an infinitesimal generator, which can be seen by using the characterization given in [Bracci et al. 2010, Theorem 0.2]. Of course, G also satisfies (2-1) and we conclude $G \in \mathcal{K}(\mathbb{B}_n, c)$.

Just as we passed from $\mathbb{D}$ to $\mathbb{H}$ in one dimension, we can pass from the unit ball $\mathbb{B}_n$ to the Siegel upper half-space $\mathbb{H}_n = \{(z_1, \tilde{z}) \in \mathbb{C}^n \mid \operatorname{Im}(z_1) > \|\tilde{z}\|^2\}$ in order to get simpler formulas:

The Cayley map

$$C : \mathbb{H}_n \to \mathbb{B}_n, \quad C(z) = (C_1(z), \dots, C_n(z)) = \left(\frac{z_1 - i}{z_1 + i}, \frac{2z_2}{z_1 + i}, \dots, \frac{2z_n}{z_1 + i}\right),$$

maps $\mathbb{H}_n$ biholomorphically onto $\mathbb{B}_n$. It extends to a homeomorphism from the one-point compactification $\widehat{\mathbb{H}}_n = \mathbb{H}_n \cup \partial\mathbb{H}_n \cup \{\infty\}$ of $\mathbb{H}_n \cup \partial\mathbb{H}_n$ to the closure of $\mathbb{B}^n$.

The pluricomplex Poisson kernel transforms as follows:

$$u_{\mathbb{H}_n}(z) := u_{\mathbb{B}_n}(C(z)) = -\operatorname{Im}(z_1) + \|\tilde{z}\|^2.$$

Thus, we define the horosphere $E_{\mathbb{H}_n}(\infty, R)$ with center ∞ and radius $R > 0$ by

$$E_{\mathbb{H}_n}(\infty, R) := \left\{z \in \mathbb{H}_n \;\middle|\; \operatorname{Im}(z_1) - \|\tilde{z}\|^2 > \frac{1}{R}\right\}.$$

For $v \in \mathbb{C}^n$ and $z \in \mathbb{H}_n$, we let $\|v\|_{\mathbb{H}_n,z}$ be the Kobayashi hyperbolic length of v.

Let $c \geq 0$. We define the class $\mathcal{K}(\mathbb{H}_n, c)$ to be the set of all infinitesimal generators H on $\mathbb{H}_n$ satisfying the inequality

$$\|H(z)\|_{\mathbb{H}_n,z} \leq \frac{c}{u_{\mathbb{H}_n}(z)^2}$$

for all $z \in \mathbb{H}_n$. Then we have

$$\mathcal{K}(\mathbb{B}_n, c) = \big\{C'(C^{-1}) \cdot (H \circ C^{-1}) \mid H \in \mathcal{K}(\mathbb{H}_n, c)\big\}.$$

From now on we will stay in the upper half-space $\mathbb{H}_n$, where most of the computations we need take a simpler form.

3. Slices

Normalized geodesics and slices. For any $H \in \mathrm{Inf}(\mathbb{H}_n)$, one can consider one-dimensional slices by using the so-called *Lempert projection devices*; see [Bracci and Shoikhet 2014, Section 3].

If $w \in \mathbb{H}_n$, then there exists a unique complex geodesic passing through w and ∞. Let us choose a parametrization $\varphi : \mathbb{H} \to \mathbb{H}_n$ of this geodesic. There exists a unique holomorphic map $P : \mathbb{H}_n \to \mathbb{H}_n$ with $P^2 = P$ and $P \circ \varphi = \varphi$. Define $\widetilde{P} = \varphi^{-1} \circ P$. Then

$$h_\varphi : \mathbb{H} \to \mathbb{C}, \quad h_\varphi(\zeta) = d\widetilde{P}(\varphi(\zeta)) \cdot H(\varphi(\zeta)),$$

is an infinitesimal generator on $\mathbb{H}$; see [Bracci and Shoikhet 2014, p. 6].

We will need special parametrizations of these geodesics: In [Bracci and Patrizio 2005, p. 516], it is shown that for any complex geodesic $\varphi : \mathbb{H} \to \mathbb{H}_n$ with $\varphi(\infty) = \infty$, there exists $a_\varphi > 0$ such that

$$u_{\mathbb{H}_n}(\varphi(\zeta)) = a_\varphi \cdot u_{\mathbb{H}}(\zeta)$$

for all $\zeta \in \mathbb{H}$. Call a geodesic $\varphi : \mathbb{H} \to \mathbb{H}_n$ *normalized* if $\varphi(\infty) = \infty$ and $a_\varphi = 1$.

Lemma 3.1. *Let $a \in \mathbb{C}$ and $\gamma \in \mathbb{C}^{n-1}$ such that $(a, \gamma) \in \mathbb{H}_n$. Then the map*

$$\varphi_\gamma : \mathbb{H} \to \mathbb{H}_n, \quad \varphi_\gamma(\zeta) := (\zeta + i\|\gamma\|^2, \gamma),$$

is a normalized geodesic through (a, γ). Furthermore, if $H = (H_1, \widetilde{H}) \in \mathrm{Inf}(\mathbb{H}_n)$, then the slice $h_\gamma := h_{\varphi_\gamma}$ of H with respect to φ_γ is given by

$$h_\gamma(\zeta) = H_1(\varphi_\gamma(\zeta)) - 2i\bar{\gamma}^T \cdot \widetilde{H}(\varphi_\gamma(\zeta)). \tag{3-1}$$

Proof. Let $\psi : \mathbb{D} \to \mathbb{B}_n$ be a complex geodesic with $\psi(1) = e_1$. As a parametrization for ψ, one can choose (see [Bracci and Shoikhet 2014, Section 3])

$$\psi(\zeta) = (\alpha^2(\zeta - 1) + 1, \alpha(\zeta - 1)\beta),$$

where $\alpha > 0$ and $\beta \in \mathbb{C}^{n-1}$ such that $\|\beta\|^2 = 1 - \alpha^2$. Then

$$C^{-1}(\psi(\zeta)) = \left(i\frac{2 + \alpha^2(\zeta - 1)}{\alpha^2(1 - \zeta)}, i\beta/\alpha\right)$$

and

$$\zeta \mapsto C^{-1}\big(\psi(C_1(\zeta))\big) = \left(-i + \frac{\zeta+i}{\alpha^2}, i\beta/\alpha\right)$$
$$= \left(\frac{\zeta}{\alpha^2} + i\,\frac{1-\alpha^2}{\alpha^2}, i\beta/\alpha\right) = \left(\frac{\zeta}{\alpha^2} + i\left\|\frac{\beta}{\alpha}\right\|^2, i\beta/\alpha\right)$$

is a complex geodesic from $\mathbb{H}$ to $\mathbb{H}_n$. A reparametrization (ζ/α^2 to ζ) and setting $\gamma = i\beta/\alpha$ gives the geodesic

$$\varphi_\gamma(\zeta) = (\zeta + i\,\|\gamma\|^2, \gamma). \tag{3-2}$$

This complex geodesic is normalized because it satisfies $\varphi_\gamma(\infty) = \infty$ and

$$u_{\mathbb{H}_n}(\varphi_\gamma(\zeta)) = \operatorname{Im}(\zeta + i\,\|\gamma\|^2) - \|\gamma\|^2 = \operatorname{Im}(\zeta) = u_{\mathbb{H}}(\zeta).$$

The projection onto $\varphi_\gamma(\mathbb{H})$ is given by

$$P(z_1, \tilde{z}) = (z_1 - 2i\,\bar{\gamma}^T \cdot \tilde{z} + 2i\,\|\gamma\|^2, \gamma). \tag{3-3}$$

Clearly, P is holomorphic and maps $\mathbb{H}_n$ onto $\varphi_\gamma(\mathbb{H})$ because

$$\begin{aligned}\operatorname{Im}(z_1 - 2i\,\bar{\gamma}^T \cdot \tilde{z} + 2i\,\|\gamma\|^2) &= \operatorname{Im}(z_1) - 2\operatorname{Im}(i\,\bar{\gamma}^T \cdot \tilde{z}) + 2\|\gamma\|^2 \\ &\geq \|\tilde{z}\|^2 - 2\|\gamma\|\,\|\tilde{z}\| + \|\gamma\|^2 + \|\gamma\|^2 \\ &= (\|\gamma\| - \|\tilde{z}\|)^2 + \|\gamma\|^2 \geq \|\gamma\|^2.\end{aligned}$$

Furthermore,

$$\begin{aligned}(P \circ P)(z_1, \tilde{z}) &= (z_1 - 2i\,\bar{\gamma}^T \tilde{z} + 2i\,\|\gamma\|^2 - 2i\,\bar{\gamma}^T \gamma + 2i\,\|\gamma\|^2, \gamma) \\ &= (z_1 - 2i\,\bar{\gamma}^T \tilde{z} + 2i\,\|\gamma\|^2, \gamma) = P(z_1, \tilde{z}).\end{aligned}$$

Thus, the inverse $\widetilde{P} : \mathbb{H}_2 \to \mathbb{H}$, $\widetilde{P} = \varphi_\gamma^{-1} \circ P$, is given by

$$\widetilde{P}(z_1, \tilde{z}) = (z_1 - 2i\,\bar{\gamma}^T \tilde{z} + i\,\|\gamma\|^2).$$

If $H(z) = (H_1(z), \widetilde{H}(z))$ is a generator on $\mathbb{H}_n$, we get the slice reduction

$$\begin{aligned}h_{\varphi_\gamma}(\zeta) &= d\widetilde{P}(\varphi_\gamma(\zeta)) \cdot H(\varphi_\gamma(\zeta)) \\ &= H_1(\varphi_\gamma(\zeta)) - 2i\,\bar{\gamma}^T \cdot \widetilde{H}(\varphi_\gamma(\zeta)). \qquad \square\end{aligned}$$

Some explicit formulas. Later on we will need explicit formulas of the Kobayashi norms of $dP(z)H(z)$ and $H(z) - dP(z) \cdot H(z)$. The following lemma is proven in the Appendix.

Lemma 3.2. *Let $a \in \mathbb{C}$, $p, v \in \mathbb{C}^{n-1}$ and $z = (z_1, \tilde{z}) \in \mathbb{H}_n$. Then the following formulas hold*:

$$\left\| \begin{pmatrix} a \\ 0 \end{pmatrix} \right\|_{\mathbb{H}_n, z} = \frac{|a|}{|u_{\mathbb{H}_n}(z)|}, \tag{3-4}$$

$$\left\| \begin{pmatrix} 2i\,\bar{p}^T v \\ v \end{pmatrix} \right\|_{\mathbb{H}_n, z} = 2\frac{\sqrt{\|v\|^2 \, |u_{\mathbb{H}_n}(z)| + |\overline{(p-\tilde{z})}^T v|^2}}{|u_{\mathbb{H}_n}(z)|}, \tag{3-5}$$

$$\left\| \begin{pmatrix} a - 2i\,\bar{\tilde{z}}^T v \\ 0 \end{pmatrix} + \begin{pmatrix} 2i\,\bar{\tilde{z}}^T v \\ v \end{pmatrix} \right\|^2_{\mathbb{H}_n, z} = \left\| \begin{pmatrix} a - 2i\,\bar{\tilde{z}}^T v \\ 0 \end{pmatrix} \right\|^2_{\mathbb{H}_n, z} + \left\| \begin{pmatrix} 2i\,\bar{\tilde{z}}^T v \\ v \end{pmatrix} \right\|^2_{\mathbb{H}_n, z}. \tag{3-6}$$

By using Lemma 3.2 we obtain the following explicit expressions.

Lemma 3.3. *Let $H = (H_1, \widetilde{H}) \in \operatorname{Inf}(\mathbb{H}_n)$ and fix $z \in \mathbb{H}_n$. Denote by P the projection onto the complex geodesic through z and ∞. Then the following formulas hold*:

$$\begin{aligned} dP(z) \cdot H(z) &= (H_1(z) - 2i\,\bar{\tilde{z}}^T \widetilde{H}(z), 0), \\ H(z) - dP(z) \cdot H(z) &= (2i\,\bar{\tilde{z}}^T \widetilde{H}(z), \widetilde{H}(z)). \end{aligned} \tag{3-7}$$

Furthermore,

$$\|H(z)\|^2_{\mathbb{H}_n, z} = \|dP(z) \cdot H(z)\|^2_{\mathbb{H}_n, z} + \|H(z) - dP(z) \cdot H(z)\|^2_{\mathbb{H}_n, z}, \tag{3-8}$$

$$\|dP(z) H(z)\|_{\mathbb{H}_n, z} = \frac{|H_1(z) - 2i\,\bar{\tilde{z}}^T \widetilde{H}(z)|}{|u_{\mathbb{H}_n}(z)|}, \tag{3-9}$$

$$\|H(z) - dP(z) \cdot H(z)\|_{\mathbb{H}_n, z} = 2\frac{\|\widetilde{H}(z)\|}{\sqrt{|u_{\mathbb{H}_n}(z)|}}. \tag{3-10}$$

Proof. The formulas for $dP(z)H(z)$ and $H(z) - dP(z)H(z)$ follow from the explicit form (3-3).

Equation (3-8) follows from (3-6) with $a = H_1(z)$ and $v = \widetilde{H}(z)$.

Furthermore, (3-9) follows directly from (3-4) with $a = H_1(z) - 2i\,\bar{\tilde{z}}^T \widetilde{H}(z)$ and (3-10) from (3-5) by setting $p = \tilde{z}$ and $v = \widetilde{H}$. □

Slices of generators in $\mathcal{K}(\mathbb{H}_n, c)$ and examples.

Proposition 3.4. *Let $c \geq 0$ and $H \in \mathcal{K}(\mathbb{H}_n, c)$. Then every normalized slice h_γ of H belongs to $\mathcal{K}(\mathbb{H}, c)$.*

Proof. Fix $\gamma \in \mathbb{C}^{n-1}$ and $\zeta \in \mathbb{H}$ and let $z = \varphi_\gamma(\zeta)$.

Furthermore, let P be the projection onto $\varphi_\gamma(\mathbb{H})$. Now we write $H(z)$ as

$$H(z) = dP(z) \cdot H(z) + (H(z) - dP(z)H(z)).$$

As $H \in \mathcal{K}(\mathbb{H}_n, c)$, equation (3-8) implies

$$\|H(z)\|^2_{\mathbb{H}_n,z} = \|dP(z)\cdot H(z)\|^2_{\mathbb{H}_n,z} + \|H(z) - dP(z)H(z)\|^2_{\mathbb{H}_n,z} \le \frac{c^2}{u_{\mathbb{H}_n}(z)^4}.$$

In particular,

$$\|dP(z)\cdot H(z)\|_{\mathbb{H}_n,z} \le \frac{c}{u_{\mathbb{H}_n}(z)^2}. \tag{3-11}$$

By the definition of the slice h_γ, we have

$$dP(\varphi_\gamma(\zeta))\cdot H(\varphi_\gamma(\zeta)) = (d\varphi_\gamma)(\zeta)\cdot h_\gamma(\zeta),$$

and consequently

$$\|dP(\varphi_\gamma(\zeta))\cdot H(\varphi_\gamma(z))\|_{\mathbb{H}_n,\varphi_\gamma(\zeta)} = \|(d\varphi_\gamma)(\zeta)\cdot h_\gamma(\zeta)\|_{\mathbb{H}_n,\varphi_\gamma(\zeta)} = |h_\gamma(\zeta)|_{\mathbb{H},\zeta}.$$

The last equality holds as φ_γ is a complex geodesic. Equation (3-11) implies

$$|h_\gamma(\zeta)|_{\mathbb{H},\zeta} \le \frac{c}{u_{\mathbb{H}_n}(\varphi_\gamma(\zeta))^2} = \frac{c}{u_{\mathbb{H}}(\zeta)^2},$$

where the last equality holds as φ_γ is normalized. Hence, $h_\gamma \in \mathcal{K}(\mathbb{H}, c)$. □

Remark 3.5. If two holomorphic functions $H_1, H_2 : \mathbb{H}_n \to \mathbb{C}^n$ have the same slices, i.e., $dP(z)H_1(z) = dP(z)H_2(z)$ for all $z \in \mathbb{H}_n$, then $H_1 = H_2$; see the proof of Theorem 3.2 in [Casavecchia 2010].

Example 3.6. The family $\{\Phi_t(z) = (z_1, e^{-it/z_1} z_2)\}_{t\ge 0}$ is a semigroup on $\mathbb{H}_2$. Its generator H is given by

$$H(z_1, z_2) = \left(0, -i\frac{z_2}{z_1}\right).$$

Thus, for $\gamma \in \mathbb{C}$, the slice h_γ has the form

$$h_\gamma(z) = -2i\bar{\gamma}\cdot -i\frac{\gamma}{z + i|\gamma|^2} = \frac{-2|\gamma|^2}{z + i|\gamma|^2}.$$

Consequently, the limit $\lim_{y\to\infty} y\cdot|h(iy)| = 2|\gamma|^2$ exists, but does not have an upper bound that is independent of γ. Proposition 3.4 implies that for any $c \ge 0$, $H \notin \mathcal{K}(\mathbb{H}_2, c)$.

Example 3.7. Let

$$H : \mathbb{H}_2 \to \mathbb{C}^2, \qquad H(z_1, z_2) = \begin{pmatrix} -1/z_1 \\ z_2/2z_1^2 \end{pmatrix}.$$

For $\gamma \in \mathbb{C}$, the slice h_γ is given by

$$\begin{aligned} h_\gamma(\zeta) &= \frac{-1}{\zeta + i|\gamma|^2} - 2i\bar{\gamma} \cdot \frac{\gamma}{2(\zeta + i|\gamma|^2)^2} \\ &= \frac{-\zeta - 2i|\gamma|^2}{(\zeta + i|\gamma|^2)^2} = \frac{(-\zeta - 2i|\gamma|^2)(\bar{\zeta}^2 - 2i|\gamma|^2\bar{\zeta} - |\gamma|^4)}{\left|\zeta + i|\gamma|^2\right|^4}. \end{aligned}$$

Let us write $\zeta = x + iy$, $x \in \mathbb{R}$, $y \in (0, \infty)$. Then a small calculation gives

$$\operatorname{Im}(h_\gamma(\zeta)) = \frac{y(x^2+y^2) + 4y^2|\gamma|^2 + 5y|\gamma|^4 + 2|\gamma|^6}{\left|\zeta + i|\gamma|^2\right|^4} > 0.$$

Furthermore,

$$\limsup_{y\to\infty} y|h_\gamma(iy)| = 1.$$

Hence, $h_\gamma \in \mathcal{K}(\mathbb{H}, 1)$. So each slice is an infinitesimal generator in $\mathbb{H}$ and by [Bracci and Shoikhet 2014, Proposition 3.8], the function H is an infinitesimal generator in $\mathbb{H}_2$.

Now let $(z_1, z_2) \in \mathbb{H}_2$ and write $z_1 = x + iy$, $x, y \in \mathbb{R}$. Then we get

$$\begin{aligned} u_{\mathbb{H}_2}(z)^4 \cdot \|H(z)\|^2_{\mathbb{H}_2,z} &= (y - |z_2|^2)^2 \cdot \frac{x^2 + y^2 + 3|z_2|^2 y}{(x^2+y^2)^2} \\ &\underset{y \ge |z_2|^2}{\le} y^2 \cdot \frac{x^2+y^2+3y^2}{(x^2+y^2)^2} \le \frac{x^2+4y^2}{x^2+y^2} \le 4 \end{aligned}$$

(an explicit formula of the Kobayashi metric is given in the Appendix). Consequently, $H \in \mathcal{K}(\mathbb{H}_2, 2)$.

Question 3.8. Let $H : \mathbb{H}_n \to \mathbb{C}^n$ be an infinitesimal generator. Assume there exists $c \ge 0$ such that $h_\gamma \in \mathcal{K}(\mathbb{H}, c)$ for every $\gamma \in \mathbb{C}^{n-1}$. Does this imply that $H \in \mathcal{K}(\mathbb{H}_n, C)$ for some $C \ge c$?

4. Univalent functions with hydrodynamic normalization

Motivated by Remark 1.6, we define the following generalization of the class $\mathfrak{P}$, where id stands for the identity mapping on $\mathbb{H}_n$.

Definition 4.1.

$$\mathfrak{P}_n := \left\{ f : \mathbb{H}_n \to \mathbb{H}_n \mid f \text{ is univalent and } f - \mathrm{id} \in \mathcal{K}(\mathbb{H}_n, c) \text{ for some } c \ge 0 \right\}.$$

Remark 4.2. It is important to note that if $f : \mathbb{H}_n \to \mathbb{H}_n$ is a holomorphic self-mapping, then the map $f - \mathrm{id}$ is automatically an infinitesimal generator; see [Reich and Shoikhet 2005, p. 207].

Basic properties of $\mathfrak{P}_n$. The following proposition summarizes some basic properties of $\mathfrak{P}_n$.

Proposition 4.3. (a) *$\mathfrak{P}_n$ contains no automorphism of $\mathbb{H}_n$ except the identity.*

(b) *Let $\alpha : \mathbb{H}_n \to \mathbb{H}_n$ be an automorphism of $\mathbb{H}_n$ with $\alpha(\infty) = \infty$. If $f \in \mathfrak{P}_n$, then $\alpha^{-1} \circ f \circ \alpha \in \mathfrak{P}_n$.*

(c) *Let $f \in \mathfrak{P}_n$. Then $f(E_{\mathbb{H}_n}(\infty, R)) \subset E_{\mathbb{H}_n}(\infty, R)$ for every $R > 0$.*

(d) *Let $f \in \mathfrak{P}_n$ and write $f(z) = z + H(z)$ with $H = (H_1, \tilde{H}) \in \mathcal{K}(\mathbb{H}_n, c)$. Then*

$$\|\tilde{H}(z)\|^2 \le |H_1(z) - 2i\bar{\tilde{z}}^T \tilde{H}| \quad \textit{for all } z = (z_1, \tilde{z}) \in \mathbb{H}_n. \tag{4-1}$$

(e) *Let $f \in \mathfrak{P}_n$. Then there exists $R > 0$ such that $E_{\mathbb{H}_n}(\infty, R) \subset f(\mathbb{H}_n)$.*

Proof. The statements (a) and (b) can easily be shown by using the explicit form of automorphisms of $\mathbb{H}_n$; see [Abate 1989, Proposition 2.2.4].

The statement (c) is just Julia's lemma: Write $f(z) = z + H(z)$ and let us pass to the unit ball and define $\tilde{f} : \mathbb{B}_n \to \mathbb{B}_n$, $\tilde{f} = C \circ f \circ C^{-1}$. Then

$$\tilde{f} = \frac{1}{2i + H_1(C^{-1}(z)) - z_1 H_1(C^{-1}(z))}\left(\begin{pmatrix}(1 - z_1)H_1(C^{-1}(z))\\ 2(1 - z_1)\tilde{H}(C^{-1}(z))\end{pmatrix} + 2iz\right).$$

By taking the sequence $z_n = (1 - 1/n, 0)$, it is easy to see that

$$\lim_{n\to\infty} \tilde{f}(z_n) = e_1 \quad \text{and} \quad \lim_{n\to\infty} \frac{1 - \|\tilde{f}(z_n)\|}{1 - \|z_n\|} = 1,$$

i.e., e_1 is a boundary regular fixed point of $\tilde{f}$ with boundary dilatation coefficient ≤ 1. Julia's lemma (see [Abate 1989, Theorem 2.2.21]) implies that $\tilde{f}(E_{\mathbb{B}_n}(e_1, R)) \subset E_{\mathbb{B}_n}(e_1, R)$ for any $R > 0$.

Inequality (d) follows directly from (c): Let $z = (z_1, \tilde{z}) \in \mathbb{H}_n$. Another formulation of (c) is $-u_{\mathbb{H}_n}(z + H(z)) \ge -u_{\mathbb{H}_n}(z)$, or more explicitly

$$\begin{aligned}\mathrm{Im}(z_1) + \mathrm{Im}(H_1(z)) - \|\tilde{z} + \tilde{H}(z)\|^2 &\ge \mathrm{Im}(z_1) - \|\tilde{z}\|^2\\ \iff \mathrm{Im}(H_1(z)) &\ge \|\tilde{z} + \tilde{H}(z)\|^2 - \|\tilde{z}\|^2 = 2\mathrm{Re}(\bar{\tilde{z}}^T \tilde{H}(z)) + \|\tilde{H}(z)\|^2\\ \iff \mathrm{Im}(H_1(z) - 2i\bar{\tilde{z}}^T \tilde{H}(z)) &\ge \|\tilde{H}(z)\|^2.\end{aligned}$$

From this inequality it follows that $\|\tilde{H}(z)\|^2 \le |H_1(z) - 2i\bar{\tilde{z}}^T \tilde{H}|$ for all $z \in \mathbb{H}_n$.

Finally we prove (e):

Let $f \in \mathfrak{P}_n$ and write $f(z) = z + H(z)$ with $H \in \mathcal{K}(\mathbb{H}_n, c)$. Because of (c), f maps the horosphere $E_{\mathbb{H}_n}(\infty, 1)$ into itself. Hence the statement is proven if we can show that $u_{\mathbb{H}_n}$ is bounded on $f(\partial E_{\mathbb{H}_n}(\infty, 1))$.

Let $z \in \mathbb{H}_n$ with $z \in \partial E_{\mathbb{H}_n}(\infty, 1)$, i.e., $|u_{\mathbb{H}_n}(z)| = 1$. Furthermore, we choose $\zeta \in \mathbb{H}$ and $\gamma \in \mathbb{C}$ such that $\varphi_\gamma(\zeta) = z$. Note that this implies $|u_{\mathbb{H}}(\zeta)| = \mathrm{Im}(\zeta) = 1$.

Let P be the projection onto $\varphi_\gamma(\mathbb{H})$.

Then we have

$$|u_{\mathbb{H}_n}(f(z))| = |u_{\mathbb{H}_n}(z+H(z))| = \Big|u_{\mathbb{H}_n}(\underbrace{z+dP(z)H(z)}_{=:\,w}+\underbrace{H(z)-dP(z)H(z)}_{=:\,v})\Big|.$$

As $dP(z)\cdot dP(z) = dP(z)$, we have $dP(z)\cdot v = 0$. A small calculation (see also [Casavecchia 2010, Lemma 3.1]) gives $v \in T_z^{\mathbb{C}}\partial E_{\mathbb{H}_n}(\infty,1)$. Furthermore, also $w \in \varphi_\gamma(\mathbb{H})$ and $dP(z) = dP(w)$ and we get $v \in T_w^{\mathbb{C}}\partial E_{\mathbb{H}_n}(\infty, |u_{\mathbb{H}_n}(w)|^{-1})$. As $E_{\mathbb{H}_n}(\infty, |u_{\mathbb{H}_n}(w)|^{-1}) = \big\{z\in\mathbb{H}_n \mid |u_{\mathbb{H}_n}(z)| > |u_{\mathbb{H}_n}(w)|\big\}$ is convex, this implies

$$\begin{aligned}|u_{\mathbb{H}_n}(w+v)| \le |u_{\mathbb{H}_n}(w)| = |u_{\mathbb{H}_n}(z+dP(z)H(z))| &\underset{\text{Lemma 3.3}}{=} |u_{\mathbb{H}_n}(z+(h_\gamma(\zeta),0))|\\ &= \operatorname{Im}(z_1)-\|\tilde z\|^2+\operatorname{Im}(h_\gamma(\zeta)) \le \operatorname{Im}(z_1)-\|\tilde z\|^2+|h_\gamma(\zeta)|\\ &= |u_{\mathbb{H}_n}(z)|+|h_\gamma(\zeta)| = 1+|h_\gamma(\zeta)| \le 1+\frac{c}{\operatorname{Im}(\zeta)} = 1+c.\end{aligned}$$

Consequently, $f(\mathbb{H}_n)\supset f(E_{\mathbb{H}_n}(\infty,1))\supset E_{\mathbb{H}_n}(\infty,1+c)$. □

Theorem 4.4. *$\mathfrak{P}_n$ is a semigroup: if $f,g\in\mathfrak{P}_n$, then $f\circ g\in\mathfrak{P}_n$.*

Proof. Let $f,g\in\mathfrak{P}_n$ with $F=(F_1,\widetilde F) := f-\mathrm{id}$, $G=(G_1,\widetilde G) := g-\mathrm{id}$ and

$$\|F(z)\|_{\mathbb{H}_n,z} \le \frac{c}{u_{\mathbb{H}_n}(z)^2},\qquad \|G(z)\|_{\mathbb{H}_n,z}\le\frac{d}{u_{\mathbb{H}_n}(z)^2}$$

for all $z\in\mathbb{H}_n$. Let $z=(z_1,\tilde z)\in\mathbb{H}_n$ and $p=(p_1,\tilde p) := z+G(z)$.

From Remark 4.2, we know that $f\circ g-\mathrm{id}$ is an infinitesimal generator on $\mathbb{H}_n$. It remains to estimate the hyperbolic metric of this generator. We have

$$\begin{aligned}\|(f\circ g)(z)-z\|_{\mathbb{H}_n,z} &= \|G(z)+F(z+G(z))\|_{\mathbb{H}_n,z}\\ &\le \|G(z)\|_{\mathbb{H}_n,z}+\|F(z+G(z))\|_{\mathbb{H}_n,z} \le \frac{d}{u_{\mathbb{H}_n}(z)^2}+\|F(p)\|_{\mathbb{H}_n,z}\\ &\le \frac{d}{u_{\mathbb{H}_n}(z)^2}+\|(F_1(p)-2i\,\bar{\tilde p}^T\,\widetilde F(p),0)\|_{\mathbb{H}_n,z}+\|(2i\,\bar{\tilde p}^T\,\widetilde F(p),\widetilde F(p))\|_{\mathbb{H}_n,z}.\end{aligned}$$

Note that $F_1(p)-2i\,\bar{\tilde p}^T\,\widetilde F(p)$ corresponds to the slice of F with respect to the geodesic through p and infinity. Because of Proposition 3.4, we know that

$$|F_1(p)-2i\,\bar{\tilde p}^T\,\widetilde F(p)| \le \frac{c}{|u_{\mathbb{H}_n}(p)|}\le\frac{c}{|u_{\mathbb{H}_n}(z)|},$$

where the second inequality follows from Proposition 4.3 (c). Together with (3-4), this implies

$$\|(F_1(p)-2i\,\bar{\tilde p}^T\,\widetilde F(p),0)\|_{\mathbb{H}_n,z} = \frac{|(F_1(p)-2i\,\bar{\tilde p}^T\,\widetilde F(p)|}{|u_{\mathbb{H}_n}(z)|}\le\frac{c}{u_{\mathbb{H}_n}(z)^2}. \tag{4-2}$$

It remains to show that there exists a constant $C > 0$ such that

$$\|(2i\,\bar{\tilde{p}}^T\,\widetilde{F}(p),\widetilde{F}(p))\|_{\mathbb{H}_n,z} \le \frac{C}{u_{\mathbb{H}_n}(z)^2}.$$

First, (3-5) gives

(4-3)

$$\begin{aligned}\|(2i\,\bar{\tilde{p}}^T\,\widetilde{F}(p),\widetilde{F}(p))\|_{\mathbb{H}_n,z} &= 2\frac{\sqrt{\|\widetilde{F}(p)\|^2\,|u_{\mathbb{H}_n}(z)| + |\overline{(\tilde{p}-\tilde{z})}^T\,\widetilde{F}(p)|^2}}{|u_{\mathbb{H}_n}(z)|} \\ &\le 2\frac{\sqrt{\|\widetilde{F}(p)\|^2\,|u_{\mathbb{H}_n}(z)| + \|(\tilde{p}-\tilde{z})\|^2\cdot\|\widetilde{F}(p)\|^2}}{|u_{\mathbb{H}_n}(z)|} \\ &= 2\frac{\|\widetilde{F}(p)\|}{|u_{\mathbb{H}_n}(z)|}\sqrt{|u_{\mathbb{H}_n}(z)| + \|\widetilde{G}(z)\|^2}.\end{aligned}$$

Now we differentiate between two cases.

Case 1: $|u_{\mathbb{H}_n}(z)| \ge 1$. The equations (3-8) and (3-10) imply

$$2\frac{\|\widetilde{F}(p)\|}{\sqrt{|u_{\mathbb{H}_n(p)}|}} \le \|\widetilde{F}(p)\|_{\mathbb{H}_n,p} \le \frac{c}{u_{\mathbb{H}_n}(p)^2};$$

thus

$$\|\widetilde{F}(p)\| \le \frac{c}{2|u_{\mathbb{H}_n}(p)|^{3/2}} \le \frac{c}{2|u_{\mathbb{H}_n}(z)|^{3/2}}. \tag{4-4}$$

In the same way, we get

$$\|\widetilde{G}(z)\| \le \frac{d}{2|u_{\mathbb{H}_n}(z)|^{3/2}}. \tag{4-5}$$

Combining (4-4) with (4-3) gives

$$\begin{aligned}\|(2i\,\bar{\tilde{p}}^T\,\widetilde{F}(p),\widetilde{F}(p))\|_{\mathbb{H}_n,z} &\le \frac{c}{|u_{\mathbb{H}_n}(z)|\,|u_{\mathbb{H}_n}(z)|^{3/2}}\sqrt{|u_{\mathbb{H}_n}(z)| + \|\widetilde{G}(z)\|^2} \\ &= \frac{c}{|u_{\mathbb{H}_n}(z)|^2}\sqrt{1+\frac{\|\widetilde{G}(z)\|^2}{|u_{\mathbb{H}_n}(z)|}} \\ &\underset{(4\text{-}5)}{\le} \frac{c}{|u_{\mathbb{H}_n}(z)|^2}\sqrt{1+\frac{d^2}{4|u_{\mathbb{H}_n}(z)|^4}} \\ &\le \frac{c\sqrt{1+d^2/4}}{|u_{\mathbb{H}_n}(z)|^2}.\end{aligned}$$

Case 2: $|u_{\mathbb{H}_n}(z)| \le 1$. From (4-2) we know that $|F_1(p) - 2i\,\bar{\tilde{p}}^T \widetilde{F}(p)| \le c/|u_{\mathbb{H}_n}(z)|$, and (4-1) implies

$$\|\widetilde{F}(p)\| \le \frac{\sqrt{c}}{\sqrt{|u_{\mathbb{H}_n}(z)|}}.$$

Similarly we get

$$\|\widetilde{G}(z)\| \le \frac{\sqrt{d}}{\sqrt{|u_{\mathbb{H}_n}(z)|}}.$$

Hence, with (4-3) we obtain

$$\begin{aligned}
\|(2i\,\bar{\tilde{p}}^T \widetilde{F}(p), \widetilde{F}(p))\|_{\mathbb{H}_n,z} &\le 2\frac{\sqrt{c}}{|u_{\mathbb{H}_n}(z)|^{3/2}}\sqrt{|u_{\mathbb{H}_n}(z)| + \|\widetilde{G}(z)\|^2} \\
&\le 2\frac{\sqrt{c}}{|u_{\mathbb{H}_n}(z)|^{3/2}}\sqrt{|u_{\mathbb{H}_n}(z)| + \frac{d}{|u_{\mathbb{H}_n}(z)|}} \\
&= 2\frac{\sqrt{c}}{|u_{\mathbb{H}_n}(z)|^{2}}\sqrt{u_{\mathbb{H}_n}(z)^2 + d} \\
&\le 2\frac{\sqrt{c}}{|u_{\mathbb{H}_n}(z)|^{2}}\sqrt{1 + d}.
\end{aligned}$$

□

On the Loewner equation with a $\mathcal{K}(\mathbb{H}_n, c)$-Herglotz vector field. Let $\{\Phi_t\}_{t\ge 0}$ be a semigroup on $\mathbb{H}_n$ with generator $H \in \mathcal{K}(\mathbb{H}_n, c)$. Next we will show that this implies $\Phi_t \in \mathfrak{P}_n$ for every $t \ge 0$.

In fact we can prove a little more by considering a nonautonomous version of (1-1). To this end, let $\{H_t : \mathbb{H}_n \to \mathbb{C}^n\}_{t\ge 0}$ be a $\mathcal{K}(\mathbb{H}_n, c)$-Herglotz vector field, i.e., $H_t \in \mathcal{K}(\mathbb{H}_n, c)$ for almost every $t \ge 0$ and the map $t \mapsto H_t(z)$ is measurable for every $z \in \mathbb{H}_n$; see [Arosio and Bracci 2011, Definition 1.2]. In this case, one can solve the nonautonomous version of (1-1), namely the Loewner equation

$$\frac{\partial \varphi_t(z)}{\partial t} = H_t(\varphi_t(z)), \quad \varphi_0(z) = z \in \mathbb{H}_n, \tag{4-6}$$

which gives a family $\{\varphi_t\}_{t\ge 0}$ of univalent self-mappings of $\mathbb{H}_n$; see [Arosio and Bracci 2011, Theorem 1.4].

Theorem 4.5. *If $\{H_t\}_{t\ge 0}$ is a $\mathcal{K}(\mathbb{H}_n, c)$-Herglotz vector field and $\{\varphi_t\}_{t\ge 0}$ the solution to (4-6), then $\varphi_t \in \mathfrak{P}_n$ for every $t \ge 0$.*

Proof. Firstly, for every $t \ge 0$ and $R > 0$, the map φ_t maps the horosphere $E_{\mathbb{H}_n}(\infty, R)$ into itself, i.e.,

$$|u_{\mathbb{H}_n}(\varphi_t(z))| \ge |u_{\mathbb{H}_n}(z)| \tag{4-7}$$

for every $z \in \mathbb{H}_n$. This can be seen as follows:

First, consider the autonomous case $H_t(z) = J(z)$ for every $t \geq 0$ and some $J \in \mathcal{K}(\mathbb{H}_n, c)$. Let G be the corresponding generator in the unit ball, i.e., $G = C'(C^{-1}) \cdot (J \circ C^{-1})$. Then G satisfies the inequality

$$\|G(z)\| \leq \|G(z)\|_{\mathbb{B}_n,z} \leq \frac{c}{u_{\mathbb{B}_n}(z)^2} = \frac{c|1-z_1|^4}{(1-\|z\|^2)^2}.$$

Putting $z = r \cdot e_1$ gives

$$\|G(re_1)\| \leq \frac{c(1-r)^4}{(1-r^2)^2} = \frac{c(1-r)^2}{(1+r)^2}.$$

From this it follows immediately that

$$\lim_{(0,1)\ni r\to 1} G(re_1) = 0 \quad \text{and} \quad \lim_{(0,1)\ni r\to 1} \frac{G_1(re_1)}{r-1} = 0.$$

Theorem 0.3 in [Bracci et al. 2010] implies that e_1 is a boundary regular fixed point for the generated semigroup with boundary dilatation coefficient 1. Hence we can apply Julia's lemma and obtain (4-7).

Now assume that $H_t(z)$ is piecewise constant with respect to time. By using the previous case, we see that (4-7) also holds in this case.

Finally, for a general $\mathcal{K}(\mathbb{H}_n, c)$-Herglotz vector field $H_t(z)$, we can approximate the solution φ_t by a sequence $\varphi_{t,n}$ such that for each n, the family $\{\varphi_{t,n}\}_{t\geq 0}$ solves (4-6) with a piecewise constant $\mathcal{K}(\mathbb{H}_n, c)$-Herglotz vector field. By using the continuity of $u_{\mathbb{H}_n}(z)$, we see that (4-7) also holds for φ_t.

Let $z = (z_1, z_2) \in \mathbb{H}_n$ and write $\varphi_t = (\varphi_{1,t}, \tilde{\varphi}_t)$, $H_t = (H_{1,t}, \widetilde{H}_t)$. The mapping φ_t satisfies the integral equation

$$\varphi_t(z) = z + \int_0^t H_s(\varphi_s(z))\, ds.$$

Similarly to the proof of Theorem 4.4, (4-4), we deduce from the fact that $H_t \in \mathcal{K}(\mathbb{H}_n, c)$ for almost every $t \geq 0$ and equations (3-8) and (3-10) that

$$\|\widetilde{H}_t(\varphi_t(z))\| \leq \frac{c}{2|u_{\mathbb{H}_n}(z)|^{3/2}} \tag{4-8}$$

for every $z \in \mathbb{H}_n$ and almost every $t \geq 0$, and similarly to (4-2), we deduce that

$$\left\|(H_{1,t}(\varphi_t(z)) - 2i\bar{\tilde{\varphi}}_t^{\,T}\, \widetilde{H}_t(\varphi_t(z)), 0)\right\|_{\mathbb{H}_n,z} \leq \frac{c}{u_{\mathbb{H}_n}(z)^2} \tag{4-9}$$

for every $z \in \mathbb{H}_n$ and almost every $t \geq 0$.

First we get

$$\|\tilde{\varphi}_s - \tilde{z}\| \leq \int_0^s \|\widetilde{H}_\tau(\varphi_\tau(z))\|\, d\tau \leq \int_0^s \frac{c}{2|u_{\mathbb{H}_n}(z)|^{3/2}}\, d\tau = \frac{cs}{2|u_{\mathbb{H}_n}(z)|^{3/2}}. \tag{4-10}$$

Suppose $|u_{\mathbb{H}_n}(z)| \geq 1$. Then we have

$$\begin{aligned}
\|\varphi_t(z)-z\|_{\mathbb{H}_n,z} &\leq \int_0^t \|H_s(\varphi_s(z))\|_{\mathbb{H}_n,z}\,ds \\
&\leq \int_0^t \left\| \begin{pmatrix} H_{1,s}(\varphi_s(z))-2i\,\bar{\tilde{\varphi}}_s^{\,T}\,\widetilde{H}_s(\varphi_s(z)) \\ 0 \end{pmatrix} \right\|_{\mathbb{H}_n,z} ds \\
&\qquad\qquad + \int_0^t \left\| \begin{pmatrix} 2i\,\bar{\tilde{\varphi}}_s^{\,T}\,\widetilde{H}_s(\varphi_s(z)) \\ \widetilde{H}_s(\varphi_s(z)) \end{pmatrix} \right\|_{\mathbb{H}_n,z} ds \\
&\underset{(4\text{-}9),(3\text{-}5)}{\leq} \int_0^t \frac{c}{u_{\mathbb{H}_n}(z)^2}\,ds + \int_0^t 2\frac{\|\widetilde{H}_s(\varphi_s(z))\|}{|u_{\mathbb{H}_n}(z)|}\sqrt{|u_{\mathbb{H}_n}(z)|+\|\tilde{\varphi}_s-\tilde{z}\|^2}\,ds \\
&\underset{(4\text{-}8),(4\text{-}10)}{\leq} \int_0^t \frac{c}{u_{\mathbb{H}_n}(z)^2}\,ds + \int_0^t \frac{c}{|u_{\mathbb{H}_n}(z)|^{5/2}}\sqrt{|u_{\mathbb{H}_n}(z)|+\frac{c^2s^2}{4|u_{\mathbb{H}_n}(z)|^3}}\,ds \\
&= \frac{ct}{u_{\mathbb{H}_n}(z)^2} + \int_0^t \frac{c}{|u_{\mathbb{H}_n}(z)|^2}\sqrt{1+\frac{c^2s^2}{4|u_{\mathbb{H}_n}(z)|^4}}\,ds \\
&\leq \frac{ct}{u_{\mathbb{H}_n}(z)^2} + \int_0^t \frac{c}{|u_{\mathbb{H}_n}(z)|^2}\sqrt{1+c^2s^2}\,ds \\
&= c\cdot\frac{t+\int_0^t \sqrt{1+c^2s^2}\,ds}{u_{\mathbb{H}_n}(z)^2}.
\end{aligned}$$

The case $|u_{\mathbb{H}_n}(z)| \leq 1$ is treated similarly, compare with the proof of Theorem 4.4, and we conclude that for every $t \geq 0$, there exists $C > 0$ such that $\|\varphi_t(z)-z\|_{\mathbb{H}_n} \leq C/u_{\mathbb{H}_n}(z)^2$ for all $z \in \mathbb{H}_n$. Together with Remark 4.2, this implies that $\varphi_t \in \mathfrak{P}_n$. □

Question 4.6. Let $f \in \mathfrak{P}_1$. In [Goryaĭnov and Ba 1992, Section 4], it is shown that there exists a $\mathcal{K}(\mathbb{H}, c)$-Herglotz vector field H_t and a time $T \geq 0$ such that $f = \varphi_T$, where $\{\varphi_t\}_{t\geq 0}$ is the solution of (4-6). What can be said in the higher-dimensional case?

On the behavior of iterates. Let $F : \mathbb{B}_n \to \mathbb{B}_n$ be holomorphic. We say that $p \in \overline{\mathbb{B}}_n$ is the Denjoy–Wolff point of F if $F^n \to p$ for $n \to \infty$ locally uniformly. The basic results about the behavior of the iterates F^n for $n \to \infty$ can be found in [Abate 1989, Chapter 2.2]. In particular we have (Theorem 2.2.31)

(4-11)
$$F \text{ has a Denjoy–Wolff point on the boundary } \partial\mathbb{B}_n \iff F \text{ has no fixed points.}$$

Now let $f \in \mathfrak{P}_n$. For $n = 1$, f has the Denjoy–Wolff point ∞ if f is not the identity: As f is not an elliptic automorphism, the classical Denjoy–Wolff theorem implies that f has a Denjoy–Wolff point. This point has to be ∞, e.g., because of Proposition 4.3 (c).

Next we will show that this is also true in higher dimensions, provided that f extends smoothly to the boundary point ∞. There are different possible definitions of smoothness of f near ∞. We will use the following one: Let $H(z) = f(z) - z$, and denote by $G : \mathbb{B}_n \to \mathbb{C}^n$ the corresponding generator on $\mathbb{B}_n$; i.e., we have

$$H(z) = (C^{-1})'(C(z)) \cdot G(C(z))$$

and a small computation shows

$$H_1(z) = -\frac{i}{2}(z_1 + i)^2 \cdot G_1(C(z)).$$

Our smoothness condition will be that G_1 has a C^3-extension to e_1; i.e., we can write

$$G_1(z) = \sum_{\substack{k_1+\cdots+k_n \le 3 \\ k_1,\dots,k_n \ge 0}} a_{k_1,\dots,k_n} (z_1 - 1)^{k_1} \cdot z_2^{k_2} \cdot \cdots \cdot z_n^{k_n} + \mathcal{O}(\|z - e_1\|^3),$$

which translates to

$$H_1(z) = -\frac{i}{2}(z_1 + i)^2 \cdot \sum_{k_1+\cdots+k_n \le 3} a_{k_1,\dots,k_n} \left(\frac{-2i}{z_1 + i}\right)^{k_1} \cdot \left(\frac{2z_2}{z_1 + i}\right)^{k_2} \cdot \cdots \cdot \left(\frac{2z_n}{z_1 + i}\right)^{k_n} + \mathcal{O}(\|C(z) - e_1\|^3),$$

or

(4-12)
$$\begin{aligned} H_1(z) = b_{0,\dots,0} \cdot (z_1 + i)^2 + (z_1 + i) \cdot \sum_{k_1+\cdots+k_n=1} b_{k_1,\dots,k_n} z_2^{k_2} \cdot \cdots \cdot z_n^{k_n} \\ + \sum_{k_1+\cdots+k_n=2} b_{k_1,\dots,k_n} z_2^{k_2} \cdot \cdots \cdot z_n^{k_n} + (z_1 + i)^{-1} \cdot \sum_{k_1+\cdots+k_n=3} b_{k_1,\dots,k_n} z_2^{k_2} \cdot \cdots \cdot z_n^{k_n} \\ + \mathcal{O}\big(|z_1 + i|^{-1} \cdot \|(1, z_2, \dots, z_n)\|^3\big) \end{aligned}$$

for some coefficients $b_{k_1,\dots,k_n} \in \mathbb{C}$.

Theorem 4.7. *Let $f \in \mathfrak{P}_n$, $f \neq \mathrm{id}$, and assume that (4-12) is satisfied. Then ∞ is the Denjoy–Wolff point of f.*

Proof. Write $f(z) = z + H(z)$, where $H \in \mathcal{K}(\mathbb{H}_n, c)$ and $H = (H_1, \widetilde{H})$. Let $\gamma \in \mathbb{C}^{n-1}$. If we can show that the slice $h_\gamma(\zeta) = H_1(\varphi(\zeta)) - 2i\bar{\gamma}^T \widetilde{H}(\varphi_\gamma(\zeta))$ has no zeros, then we are done:

This implies that H has no zeros because of (3-7) and (3-8). Hence, f has no fixed points and (4-11) implies that f has a Denjoy–Wolff point. This point has to be ∞ because of Proposition 4.3 (c).

Similarly to the proof of Theorem 4.4, (4-4), we have

$$\|\widetilde{H}(z)\| \le \frac{c}{2|u_{\mathbb{H}_n}(z)|^{3/2}},$$

and thus

$$\|\widetilde{H}(\varphi_\gamma(\zeta))\| \le \frac{c}{2|u_{\mathbb{H}_n}(\varphi_\gamma(\zeta))|^{3/2}} = \frac{c}{2\,\mathrm{Im}(\zeta)^{3/2}}.$$

Consequently,

$$\lim_{y\to\infty} y|\bar{\gamma}^T \widetilde{H}(\varphi_\gamma(iy))| = 0.$$

On the other hand, we know from Proposition 3.4 that $h_\gamma \in \mathcal{K}(\mathbb{H}, c)$, which implies (see Remark 1.5)

$$\limsup_{y\to\infty} y|h_\gamma(iy)| = \limsup_{y\to\infty} y\left|H_1(\varphi(iy)) - 2i\bar{\gamma}^T \widetilde{H}(\varphi_\gamma(iy))\right| \le c,$$

which gives us

$$\limsup_{y\to\infty} |iy \cdot H_1(\varphi_\gamma(iy))| \le c. \tag{4-13}$$

Now we use the assumption of the smoothness of H_1:

Because of (4-13), all coefficients $b_{k_1,\dots,k_n}$ from (4-12) with $k_1 + \dots + k_n \le 2$ have to be 0. Thus,

$$\lim_{y\to\infty} iy \cdot H_1(\varphi_\gamma(iy)) =: K(\gamma)$$

exists and is a polynomial in $\gamma = (\gamma_2, \dots, \gamma_n)$:

$$K(\gamma) = \sum_{k_1+\dots+k_n=3} b_{k_1,\dots,k_n} \gamma_2^{k_2} \cdot \dots \cdot \gamma_n^{k_n}.$$

As $K(\gamma)$ is bounded, it has to be constant.

If $K(\gamma) \equiv 0$, then all slices of H are zero; hence $H = 0$ by Remark 3.5 and f is the identity, a contradiction.

Hence $K(\gamma)$ is a nonzero constant and $h_\gamma(\zeta)$ is not identically zero, which implies (e.g., by using the representation (1-5)) that $h_\gamma(\zeta)$ has no zeros. □

Question 4.8. Is ∞ the Denjoy–Wolff point for every $f \in \mathfrak{P}_n$?

Appendix: Proof of Lemma 3.2

Lemma 3.2. *Let $a \in \mathbb{C}$, $p, v \in \mathbb{C}^{n-1}$ and $z = (z_1, \tilde{z}) \in \mathbb{H}_n$. Then the following formulas hold*:

$$\left\| \begin{pmatrix} a \\ 0 \end{pmatrix} \right\|_{\mathbb{H}_n, z} = \frac{|a|}{|u_{\mathbb{H}_n}(z)|}, \tag{3-4}$$

$$\left\| \begin{pmatrix} 2i\, \bar{p}^T v \\ v \end{pmatrix} \right\|_{\mathbb{H}_n, z} = 2\frac{\sqrt{\|v\|^2 \, |u_{\mathbb{H}_n}(z)| + |\overline{(p-\tilde{z})}^T v|^2}}{|u_{\mathbb{H}_n}(z)|}, \tag{3-5}$$

(3-6)
$$\left\| \begin{pmatrix} a - 2i\, \bar{\tilde{z}}^T v \\ 0 \end{pmatrix} + \begin{pmatrix} 2i\, \bar{\tilde{z}}^T v \\ v \end{pmatrix} \right\|^2_{\mathbb{H}_n, z} = \left\| \begin{pmatrix} a - 2i\, \bar{\tilde{z}}^T v \\ 0 \end{pmatrix} \right\|^2_{\mathbb{H}_n, z} + \left\| \begin{pmatrix} 2i\, \bar{\tilde{z}}^T v \\ v \end{pmatrix} \right\|^2_{\mathbb{H}_n, z}.$$

Proof. We write $\tilde{z} = (z_2, \dots, z_n)$, $v = (v_2, \dots, v_n)$, $p = (p_2, \dots, p_n)$.

An explicit formula of the Kobayashi metric for the unit ball is given in [Abate 2004, Theorem 3.4].[4] It coincides with the Bergman metric and by using the Cayley map, we get the following formula for the upper half-space:

$$\|w\|^2_{\mathbb{H}_n, z} = w^T \cdot (g_{j,k})_{j,k} \cdot \bar{w},$$

where $w \in \mathbb{C}^n$ and $(g_{j,k})_{j,k}$ is an $n \times n$-matrix with

$$g_{j,k} = -4 \frac{\partial^2}{\partial z_j \, \partial \bar{z}_k} \log\left(\operatorname{Im}(z_1) - \sum_{l=2}^{n} |z_l|^2 \right),$$

and we get for $j, k \geq 2$,

$$g_{1,1} = \frac{1}{u_{\mathbb{H}_n}(z)^2}, \qquad g_{1,k} = \frac{2i\, z_k}{u_{\mathbb{H}_n}(z)^2}, \qquad g_{j,1} = \frac{-2i\, \bar{z}_j}{u_{\mathbb{H}_n}(z)^2},$$

$$g_{j,j} = 4 \frac{\operatorname{Im}(z_1) - \sum_{l=2, l \neq j}^{n} |z_l|^2}{u_{\mathbb{H}_n}(z)^2}, \qquad g_{j,k} = \frac{4 z_k \bar{z}_j}{u_{\mathbb{H}_n}(z)^2}, \quad k \neq j.$$

The formulas (3-4) and (3-5) are now straightforward calculations. We obtain

$$\|(a, 0)\|_{\mathbb{H}_n, z} = \sqrt{(a, 0) \cdot (g_{j,k})_{j,k} \cdot \overline{(a, 0)^T}} = \sqrt{a \cdot g_{1,1} \cdot \bar{a}} = \frac{|a|}{|u_{\mathbb{H}_n}(z)|},$$

[4] Note, however, that the Kobayashi metric in [Abate 2004] differs by a factor of 2 from the one we are using here.

and

$$
\begin{aligned}
&u_{\mathbb{H}_n}(z)^2\cdot\|(2i\,\bar{p}^T v,v)\|^2_{\mathbb{H}_n,z}\\
&\quad=u_{\mathbb{H}_n}(z)^2\cdot(2i\,\bar{p}^T v,v^T)\cdot(g_{j,k})_{j,k}\cdot\overline{(2i\,\bar{p}^T v,v^T)}^T\\
&\quad=u_{\mathbb{H}_n}(z)^2\cdot\biggl(\sum_{j=2}^n g_{j,j}|v_j|^2+g_{1,1}|2i\,\bar{p}^T v|^2\\
&\qquad\qquad+\sum_{j=2}^n g_{j,1}v_j\overline{2i\,\bar{p}^T v}+\sum_{k=2}^n g_{1,k}\bar{v}_j\,2i\,\bar{p}^T v+\sum_{j,k\ge 2,j\ne k} g_{j,k}v_j\bar{v}_k\biggr)\\
&\quad=4\sum_{j=2}^n(\mathrm{Im}(z_1)-\|\tilde{z}\|^2)\cdot|v_j|^2+4\sum_{j=2}^n|z_j|^2\cdot|v_j|^2+4\sum_{j,k\ge 2}p_j\,\bar{p}_k v_j\bar{v}_k\\
&\qquad\qquad-4\sum_{j,k\ge 2}\bar{z}_j\,p_k v_j\bar{v}_k-4\sum_{j,k\ge 2}z_j\,\bar{p}_k\bar{v}_j v_k+4\sum_{j,k\ge 2,j\ne k}\bar{z}_j z_k v_j\bar{v}_k\\
&\quad=4\|v\|^2\cdot|u_{\mathbb{H}_n}(z)|+4\sum_{j=2}^n z_j\bar{z}_j v_j\bar{z}_j\\
&\qquad\qquad+4\sum_{j,k\ge 2}(p_j\,\bar{p}_k v_j\bar{v}_k-\bar{z}_j\,p_k v_j\bar{v}_k-z_j\,\bar{p}_k\bar{v}_j v_k)+4\sum_{j,k\ge 2,j\ne k}\bar{z}_j z_k v_j\bar{v}_k\\
&\quad=4\|v\|^2\cdot|u_{\mathbb{H}_n}(z)|+4\sum_{j,k\ge 2}(p_j\,\bar{p}_k v_j\bar{v}_k-\bar{z}_j\,p_k v_j\bar{v}_k-z_j\,\bar{p}_k\bar{v}_j v_k+\bar{z}_j z_k v_j\bar{v}_k)\\
&\quad=4\|v\|^2\cdot|u_{\mathbb{H}_n}(z)|+4|\overline{(p-\tilde{z})}^T v|^2.
\end{aligned}
$$

For formula (3-6) we just need to show that

$$(2i\bar{\tilde{z}}^T v,v^T)\cdot(g_{j,k})_{j,k}\cdot\overline{(a-2i\bar{\tilde{z}}^T v,0)}^T=0.$$

Indeed, we have

$$
\begin{aligned}
&u_{\mathbb{H}_n}(z)^2\cdot(g_{j,k})_{j,k}\cdot\overline{(a-2i\bar{\tilde{z}}^T v,0)}^T\\
&\qquad\qquad=(\bar{a}+2i\tilde{z}^T\bar{v},-2i\bar{z}_2\bar{a}+4\bar{z}_2\tilde{z}^T\bar{v},\dots,-2i\bar{z}_n\bar{a}+4\bar{z}_n\tilde{z}^T\bar{v})^T
\end{aligned}
$$

and

$$
\begin{aligned}
&(2i\bar{\tilde{z}}^T v,v^T)(\bar{a}+2i\tilde{z}^T\bar{v},-2i\bar{z}_2\bar{a}+4\bar{z}_2\tilde{z}^T\bar{v},\dots,-2i\bar{z}_n\bar{a}+4\bar{z}_n\tilde{z}^T\bar{v})^T\\
&\qquad\qquad=2i\bar{a}\bar{\tilde{z}}^T v-4|\tilde{z}^T\bar{v}|^2-2i\bar{a}\bar{\tilde{z}}^T v+4|\tilde{z}^T\bar{v}|^2=0.\qquad\square
\end{aligned}
$$

References

[Abate 1989] M. Abate, *Iteration theory of holomorphic maps on taut manifolds*, Mediterranean Press, Rende, 1989. MR 92i:32032 Zbl 0747.32002

[Abate 1992] M. Abate, "The infinitesimal generators of semigroups of holomorphic maps", *Ann. Mat. Pura Appl.* (4) **161** (1992), 167–180. MR 93i:32029 Zbl 0758.32013

[Abate 2004] M. Abate, "Angular derivatives in several complex variables", pp. 1–47 in *Real methods in complex and CR geometry* (Martina Franca, 2002), Lecture Notes in Mathematics **1848**, Springer, Berlin, 2004. MR 2005g:32019 Zbl 1068.32006

[Abate et al. 2010] M. Abate, F. Bracci, M. D. Contreras, and S. Díaz-Madrigal, "The evolution of Loewner's differential equations", *Eur. Math. Soc. Newsl.* 78 (2010), 31–38. MR 2768999 Zbl 1230.30003

[Arosio and Bracci 2011] L. Arosio and F. Bracci, "Infinitesimal generators and the Loewner equation on complete hyperbolic manifolds", *Anal. Math. Phys.* **1**:4 (2011), 337–350. MR 2887104 Zbl 1254.32037

[Bauer 2005] R. O. Bauer, "Chordal Loewner families and univalent Cauchy transforms", *J. Math. Anal. Appl.* **302**:2 (2005), 484–501. MR 2005g:30009 Zbl 1067.30035

[Berkson and Porta 1978] E. Berkson and H. Porta, "Semigroups of analytic functions and composition operators", *Michigan Math. J.* **25**:1 (1978), 101–115. MR 58 #1112 Zbl 0382.47017

[Bracci and Patrizio 2005] F. Bracci and G. Patrizio, "Monge–Ampère foliations with singularities at the boundary of strongly convex domains", *Math. Ann.* **332**:3 (2005), 499–522. MR 2006j:32048 Zbl 1086.32028

[Bracci and Shoikhet 2014] F. Bracci and D. Shoikhet, "Boundary behavior of infinitesimal generators in the unit ball", *Trans. Amer. Math. Soc.* **366**:2 (2014), 1119–1140. MR 3130328 Zbl 06265559

[Bracci et al. 2010] F. Bracci, M. D. Contreras, and S. Díaz-Madrigal, "Pluripotential theory, semigroups and boundary behavior of infinitesimal generators in strongly convex domains", *J. Eur. Math. Soc. (JEMS)* **12**:1 (2010), 23–53. MR 2011c:32025 Zbl 1185.32010

[Bracci et al. 2014] F. Bracci, M. Elin, and D. Shoikhet, "Growth estimates for pseudo-dissipative holomorphic maps in Banach spaces", *J. Nonlinear Convex Anal.* **15**:1 (2014), 191–198. MR 3184757 Zbl 1297.46033

[Casavecchia 2010] T. Casavecchia, "A rigidity condition for generators in strongly convex domains", *Complex Var. Elliptic Equ.* **55**:12 (2010), 1131–1142. MR 2012e:32025 Zbl 1213.32010

[Contreras et al. 2010] M. D. Contreras, S. Díaz-Madrigal, and P. Gumenyuk, "Geometry behind chordal Loewner chains", *Complex Anal. Oper. Theory* **4**:3 (2010), 541–587. MR 2011h:30037 Zbl 1209.30010

[Goryaĭnov and Ba 1992] V. V. Goryaĭnov and I. Ba, "Semigroup of conformal mappings of the upper half-plane into itself with hydrodynamic normalization at infinity", *Ukraïn. Mat. Zh.* **44**:10 (1992), 1320–1329. In Russian; translated in *Ukrainian Math. J.* **44**:10 (1992), 1209–1217. MR 94b:30013 Zbl 0873.30006

[Graham and Kohr 2003] I. Graham and G. Kohr, *Geometric function theory in one and higher dimensions*, Monographs and Textbooks in Pure and Applied Mathematics **255**, Marcel Dekker, New York, 2003. MR 2004i:32002 Zbl 1042.30001

[Maassen 1992] H. Maassen, "Addition of freely independent random variables", *J. Funct. Anal.* **106**:2 (1992), 409–438. MR 94g:46069 Zbl 0784.46047

[Pommerenke 1975] C. Pommerenke, *Univalent functions*, Vandenhoeck & Ruprecht, Göttingen, 1975. MR 58 #22526 Zbl 0298.30014

[Poreda 1987a] T. Poreda, “On the univalent holomorphic maps of the unit polydisc in $\mathbf{C}^n$ which have the parametric representation, I: The geometrical properties”, *Ann. Univ. Mariae Curie-Skłodowska Sect. A* **41** (1987), 105–113. MR 91m:32021 Zbl 0698.32004

[Poreda 1987b] T. Poreda, “On the univalent holomorphic maps of the unit polydisc in $\mathbf{C}^n$ which have the parametric representation, II: The necessary conditions and the sufficient conditions”, *Ann. Univ. Mariae Curie-Skłodowska Sect. A* **41** (1987), 115–121. MR 91m:32022 Zbl 0698.32005

[Reich and Shoikhet 2005] S. Reich and D. Shoikhet, *Nonlinear semigroups, fixed points, and geometry of domains in Banach spaces*, Imperial College Press, London, 2005. MR 2006g:47105 Zbl 1089.46002

Received January 16, 2015. Revised May 4, 2015.

SEBASTIAN SCHLEISSINGER
DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI ROMA TOR VERGATA
VIA DELLA RICERCA SCIENTIFICA 1
00133 ROMA
ITALY

sebastian.schleissinger@mathematik.uni-wuerzburg.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.227

ON A QUESTION OF A. BALOG

ILYA D. SHKREDOV

We give a partial answer to a conjecture of A. Balog concerning the size of $AA+A$, where A is a finite subset of real numbers. We also prove several new results on the cardinality of $A:A+A$, $AA+AA$ and $A:A+A:A$.

1. Introduction

Let $A \subset \mathbb{R}$ be a finite set. Define the *sumset*, and respectively the *product set*, by

$$A+A := \{a+b : a, b \in A\}$$

and

$$AA := \{ab : a, b \in A\}.$$

The Erdős–Szemerédi conjecture [1983] states that for all $\varepsilon > 0$,

$$\max\{|A+A|, |AA|\} \gg |A|^{2-\varepsilon}.$$

Loosely speaking, the conjecture says that any set of reals (or integers) cannot be highly structured in both a multiplicative and an additive sense. The best result in this direction is due to Solymosi [2009].

Theorem 1. *Let $A \subset \mathbb{R}$ be a set. Then*

$$\max\{|A+A|, |AA|\} \gg |A|^{4/3} \log^{-1/3}|A|.$$

If one considers the set

$$AA+A = \{ab+c : a, b, c \in A\}$$

then the Erdős–Szemerédi conjecture implies that $AA+A$ has size at least $|A|^{2-\varepsilon}$ (we assume for simplicity that $1 \in A$). Balog [2011] formulated the weaker hypothesis that for all $\varepsilon > 0$ one has

$$|AA+A| \gg |A|^{2-\varepsilon}.$$

In that paper he proved the following result, which implies, in particular, that $|AA+A| \gg |A|^{3/2}$ and $|AA+AA| \gg |A||A:A|^{1/2}$.

This work was supported by Russian Scientific Foundation grant 14-11-00433.
MSC2010: 11B13, 11B75.
Keywords: sum-products, Szemerédi–Trotter theorem.

Theorem 2. *For every finite set of reals A, B, C, $D \subset \mathbb{R}$, we have*

$$|AC + A||BC + B| \gg |A||B||C| \tag{1}$$

and

$$|AC + AD||BC + BD| \gg |B : A||C||D|. \tag{2}$$

More precisely (see [Schoen and Shkredov 2013]),

$$|(A \times B) \cdot \Delta(C) + A \times B| \gg |A||B||C|$$

and

$$|(A \times B) \cdot \Delta(C) + (A \times B) \cdot \Delta(D)| \gg |B : A||C||D|,$$

where

$$\Delta(A) := \{(a, a) : a \in A\}.$$

Murphy et al. [2015] obtained a partial answer to a "dual" question on the size of $A(A+A)$. The main result of this paper is the following new bound for $A : A + A$ and $AA + A$, stated more precisely in Theorem 12.

Theorem 3. *Let A be a finite subset of positive reals. Then there is $\varepsilon_1 > 0$ such that*

$$|A : A + A| \gg |A|^{3/2+\varepsilon_1}. \tag{3}$$

Moreover, if $|A : A| \ll |AA|$ then there exists $\varepsilon_2 > 0$ such that

$$|AA + A| \gg |A|^{3/2+\varepsilon_2}. \tag{4}$$

We also prove several results on the cardinality of $AA + AA$ and $A : A + A : A$; see Theorem 14 and Proposition 15 below.

Roche-Newton and Zhelezov [2015] conjectured there exist absolute constants c and c' such that for any finite $A \subset \mathbb{C}$,

$$\left|\frac{A+A}{A+A}\right| \le c|A|^2 \Longrightarrow |A + A| \le c'|A|.$$

Similar conjectures were made for the sets $(A - A)/(A - A)$, $(A - A)(A - A)$, $A(A + A + A + A)$ and so on. We conclude this paper by giving a partial answer to a variant of Roche-Newton and Zhelezov's conjecture:

$$|(A + A)(A + A) + (A + A)(A + A)| \ll |A|^2 \Longrightarrow |A \pm A| \ll |A| \log|A|;$$

see Corollary 17.

The main idea of the proof of Theorem 3 is the following. We need to estimate from below the sumset of two sets A and $A : A$. As in many problems of this type, the usual applications of the Szemerédi–Trotter theorem [Tao and Vu 2006] or Solymosi's method [Balog 2011] give us a lower bound of the form $|A : A + A| \gg |A|^{3/2}$. In [Schoen and Shkredov 2011] the exponent $3/2$ was improved in the particular case of sumsets of convex sets. After that the method was developed by several authors; see, e.g., [Konyagin and Rudnev 2013; Li 2011; Li and Roche-Newton 2012; Schoen 2014; Schoen and Shkredov 2013; Shkredov 2013a; 2013b; 2015]. In [Shkredov 2015] it was proved that the bound $|A+B| \gg |A|^{3/2+c}$, $c > 0$, holds for a wide class of *different* sets having roughly comparable sizes. For example, such a bound holds if A and B have small multiplicative doubling. It turns out that if (3) cannot be improved then there is some large set C such that $|AC| \ll |A|$. This allows us to apply results from [Shkredov 2015].

2. Notation

Let $\boldsymbol{G}$ be an abelian group and $+$ be the group operation. We use the same letter to denote a set $S \subseteq \boldsymbol{G}$ and its characteristic function $S : \boldsymbol{G} \to \{0, 1\}$. By $|S|$ denote the cardinality of S.

Let $f, g : \boldsymbol{G} \to \mathbb{C}$ be two functions with finite supports. Put

$$(5)\qquad (f * g)(x) := \sum_{y \in \boldsymbol{G}} f(y) g(x-y) \quad \text{and} \quad (f \circ g)(x) := \sum_{y \in \boldsymbol{G}} f(y) g(y+x).$$

Let $A \subseteq \boldsymbol{G}$ be a set. For any real $\alpha > 0$ let

$$(6)\qquad \mathsf{E}^+_\alpha(A) = \sum_{x \in \boldsymbol{G}} (A \circ A)^\alpha(x)$$

be the *higher energy* of A. In the particular case $\alpha = 2$ we write $\mathsf{E}^+(A) = \mathsf{E}^+_2(A)$ and $\mathsf{E}(A, B)$ for $\sum_{x \in \boldsymbol{G}} (A \circ A)(x)(B \circ B)(x)$. The quantity $\mathsf{E}^+(A)$ is called the *additive energy* of a set; see, e.g., [Tao and Vu 2006]. For a sequence $s = (s_1, \dots, s_{k-1})$ put $A^+_s = A \cap (A - s_1) \cap \dots \cap (A - s_{k-1})$. Then

$$\mathsf{E}^+_k(A) = \sum_{s_1, \dots, s_{k-1} \in \boldsymbol{G}} |A^+_s|^2.$$

If we have a group $\boldsymbol{G}$ with multiplication instead of addition, then we use the symbol $\mathsf{E}^\times_\alpha(A)$ for the corresponding energy of a set A and we write $A^\times_s$ for $A \cap (As_1^{-1}) \cap \dots \cap (As_{k-1}^{-1})$. In the case of a unique operation we write just $\mathsf{E}_k(A)$, $\mathsf{E}(A)$ and A_s.

Let $A, B \subseteq \boldsymbol{G}$ be two finite sets. The *magnification ratio* $R_B[A]$ of the pair (A, B) (see, e.g., [Tao and Vu 2006]) is defined by

$$(7)\qquad R_B[A] = \min_{\varnothing \neq Z \subseteq A} \frac{|B+Z|}{|Z|}.$$

A beautiful result on the magnification ratio was proven by Petridis [2012].

Theorem 4. *For any* $A, B, C \subseteq \boldsymbol{G}$, *we have*

$$|B + C + X| \le R_B[A] \cdot |C + X|, \tag{8}$$

where $X \subseteq A$ *and* $|B + X| = R_B[A]|X|$.

We conclude the section with Ruzsa's triangle inequality; see, e.g., [Tao and Vu 2006]. Interestingly, our proof (developing some ideas of [Schoen and Shkredov 2013; Murphy et al. 2015]) describes the situation when the triangle inequality is sharp, namely, when $|B \cap (A - z) - C| \approx |C|$ for many $z \in A - B$.

Lemma 5. *Let* $A, B, C \subseteq \boldsymbol{G}$ *be any sets. Then*

$$|C||A - B| \le |A \times B - \Delta(C)| \le |A - C||B - C|. \tag{9}$$

Proof. We have

$$|A \times B - \Delta(C)| = \sum_{z \in A - B} |B \cap (A - z) - C| \ge |A - B||C|.$$

The inequality above is trivial and the identity follows by the projection of points $(x, y) \in A \times B - \Delta(C)$, $(x, y) = (a - c, b - c)$, $a \in A$, $b \in B$, $c \in C$, onto $z := x - y = a - b \in A - B$. If z is fixed we see that the result of the projection is the intersection of the line $z = x - y$ with our set and moreover the ordinates of the points from the intersection belong to $B \cap (A - z) - C$. It is easy to check that the converse is also true. □

All logarithms are base 2. The signs $\ll$ and $\gg$ are the usual Vinogradov symbols.

3. Preliminaries

As we discussed in the introduction our proof uses some notions from [Shkredov 2015]. So, let us recall the main definition of that paper.

Definition 6. A set $A \subset \boldsymbol{G}$ has *SzT-type* (in other words, A is called a *Szemerédi–Trotter set*) with parameter $\alpha \ge 1$ if for any set $B \subset \boldsymbol{G}$ and an arbitrary $\tau \ge 1$,

$$|\{x \in A + B : (A * B)(x) \ge \tau\}| \ll c(A)|B|^{\alpha} \cdot \tau^{-3}, \tag{10}$$

where $c(A) > 0$ is a constant that depends on the set A only.

Simple calculations (or see [Shkredov 2015, Lemma 7]) give us some connections between various energies of SzT-type sets. Formula (11) below is due to Li [2011].

Lemma 7. *Suppose that* $A, B, C \subseteq \boldsymbol{G}$ *have SzT-type with the same parameter* α. *Then*

$$\mathsf{E}^3(A) \ll \mathsf{E}^2_{3/2}(A) c(A) |A|^{\alpha}, \tag{11}$$

$$\mathsf{E}(A) \ll c^{1/2}(A)|A|^{1+\alpha/2}, \tag{12}$$

and

$$\sum_x (A\circ A)(x)(B\circ B)(x)(C\circ C)(x) \ll \big(c(A)c(B)c(C)\big)^{1/3}\big(|A||B||C|\big)^{\alpha/3} \times \log\big(\min\{|A|,|B|,|C|\}\big). \tag{13}$$

Proof. We prove just (11); estimates (12) and (13) can be established by similar arguments.

Let us arrange the convolutions $(A\circ A)(x)$ in decreasing order: $(A\circ A)(x_1) \ge (A\circ A)(x_2) \ge \cdots$. By assumption A has SzT-type with parameter α, which implies that $(A\circ A)(x_j) \ll (c(A)|A|^\alpha j^{-1})^{1/3}$. Choosing the parameter $\Delta^{3/2} = c(A)|A|^\alpha \mathsf{E}^{-1}_{3/2}(A)$ and applying the obtained bound, we get

$$\mathsf{E}(A) = \sum_{j=1}^{|A-A|} (A\circ A)^2(x_j) \le \Delta^{1/2}\mathsf{E}_{3/2}(A) + \sum_{j:(A\circ A)(x_j)\ge\Delta} (c(A)|A|^\alpha j^{-1})^{2/3}.$$

The condition $(A\circ A)(x_j) \ge \Delta$ implies $j^{1/3} \ll (c(A)|A|^\alpha)^{1/3}\Delta^{-1}$. Thus by our choice of Δ, we have

$$\mathsf{E}(A) \ll \Delta^{1/2}\mathsf{E}_{3/2}(A) + c(A)|A|^\alpha\Delta^{-1} \ll \mathsf{E}^{2/3}_{3/2}(A)(c(A)|A|^\alpha)^{1/3}$$

as required. □

We need Lemma 7 from [Raz et al. 2015] (see also Lemma 27 from [Schoen and Shkredov 2013]).

Lemma 8. *Any set $A\subset\mathbb{R}$, $\mathbb{R}=(\mathbb{R},+)$, has SzT-type with $\alpha=2$ and $c(A)=|A|d(A)$, where*

$$d(A) := \min_{C\neq\varnothing} \frac{|AC|^2}{|A||C|}. \tag{14}$$

So, any set with small multiplicative doubling or, more precisely, with small quantity (14) has SzT-type, relative to addition, in an effective way. The interested reader can check that the minimum in (14) is actually attained. Careful analysis of our proof shows that we do not need this. Other examples of SzT-type sets can be found in [Shkredov 2015].

Now let us prove a simple result on $d(A)$ that follows from Petridis's Theorem 4.

Lemma 9. *Let $A\subseteq\mathbb{R}^+$ be a set. Then $d(A)=d(A^{-1})$, and for any nonempty C we have*

$$d(AA) \le \frac{|AC|^4}{|AA||C|^3}, \quad d(A:A) \le \frac{|AC|^4}{|A:A||C|^3}. \tag{15}$$

In particular,

$$d(AA) \le \frac{|A|^2 d^2(A)}{|AA||C|}, \quad d(A:A) \le \frac{|A|^2 d^2(A)}{|A:A||C|}, \tag{16}$$

where C is a set where the minimum in (14) is attained.

Proof. The identity $d(A) = d(A^{-1})$ is obvious. Let us prove (16). By Theorem 4 there is $X \subseteq C$ such that $|AAX| \le R|AX|$, where $R = R_A[C]$ is defined by formula (7). We have

$$d(AA) \le \frac{|AAX|^2}{|AA||X|} \le R^2 \frac{|AX|^2}{|AA||X|} = \frac{|AX|^4}{|AA||X|^3} \le \frac{|AC|^4}{|AA||C|^3} \tag{17}$$

and the first bound of (15) is obtained. Similarly, let $Y \subseteq C$ be as given by Theorem 4 and put $R = R_A[C^{-1}]$. Then $|(A:A)Y| \le R|A^{-1}Y| \le R|A^{-1}C|$, $R = |AY^{-1}|/|Y| \le |AC^{-1}|/|C|$, and arguments similar to (17) can be applied. $\square$

Finally, we formulate a full version of Theorem 1.

Theorem 10. *Let $A, B \subseteq \mathbb{R}$ be sets, and let $\tau > 0$ be a real number. Then*

$$\big|\{x : |A \cap xB| \ge \tau\}\big| \ll \frac{|A+A||B+B|}{\tau^2}. \tag{18}$$

In particular,

$$\mathsf{E}^{\times}(A, B) \ll |A+A||B+B| \cdot \log\big(\min\{|A|, |B|\}\big). \tag{19}$$

4. Proof of the main results

Our proof relies on a partial case of Theorem 14 from [Shkredov 2015].

Theorem 11. *Suppose $A, A_* \subset \mathbb{R}$ have SzT-type with the same parameter $\alpha = 2$. Then*

$$\begin{aligned} |A \pm A_*| \gg\ & \max\Big\{d(A_*)^{-1/3} d(A)^{-2/9} |A_*|^{8/9} |A|^{2/3},\ d(A)^{-1/3} d(A_*)^{-2/9} |A|^{8/9} |A_*|^{2/3}, \\ & \min\big\{d(A_*)^{-2/27} d(A)^{-13/27} |A_*|^{14/9},\ d(A)^{-2/27} d(A_*)^{-13/27} |A|^{14/9}\big\}\Big\} \\ & \times (\log(|A||A_*|))^{-2/9}. \end{aligned} \tag{20}$$

Now we can prove the main result of the paper.

Theorem 12. *Let A be a finite subset of positive reals. Then*

$$|A : A + A| \gg |A|^{3/2+1/82} \cdot (\log|A|)^{-2/41}, \tag{21}$$

and

$$|AA + A| \gg |AA|^{11/41} |A:A|^{-11/41} |A|^{3/2+1/82} (\log|A|)^{-2/41}. \tag{22}$$

Proof. Put $l = \log|A|$. We will assume that $|A:A+A| \ll M|A|^{3/2}$ and that $|AA+A| \ll M|A|^{3/2}$, where M is a small power of $|A|$, that is, $M = |A|^{\varepsilon}$, and obtain a contradiction. Let us begin with (21) because the proof of the second inequality requires some additional steps.

Recall the arguments from [Balog 2011] or see the proof of Theorem 31 from [Schoen and Shkredov 2013]. Let l_i be the line $y = q_i x$. Thus, $(x, y) \in l_i \cap A^2$ if and only if $x \in A^{\times}_q$. Let $q_1, \dots, q_n \in \Pi \subseteq A:A$ be such that $q_1 < q_2 < \dots < q_n$. Here Π is a set which can vary, in principle, and at the moment we choose Π such that $|A^{\times}_{q_i}| \ge 2^{-1}|A|^2/|A:A|$ for all $q_i \in \Pi$. Thus, $\sum_{q_i \in \Pi}|A^{\times}_{q_i}| \ge \frac{1}{2}|A|^2$. We multiply all points of A^2 lying on the line l_i by $\Delta(A^{-1})$, so we obtain $|A^{\times}_{q_i}:A|$ points still belonging to the line l_i, and then we consider the sumset of the resulting set with $l_{i+1} \cap A^2$. Clearly, we get $|A^{\times}_{q_i}:A||A^{\times}_{q_{i+1}}|$ points from the set $(A:A+A)^2$ lying between the lines l_i and l_{i+1}. Put

$$d(A) \le \tilde{d}(A) := \min_{i=2,\dots,n} \frac{|AA^{\times}_{q_i}|^2}{|A||A^{\times}_{q_i}|}. \tag{23}$$

Therefore, using the definition of $\tilde{d}(A)$, we have

$$\begin{aligned} M^2|A|^3 &\gg |A:A+A|^2 \\ &\ge \sum_{i=1}^{n-1}|A^{\times}_{q_i}||A^{\times}_{q_{i+1}}:A| \\ &\ge |A|^{1/2}\tilde{d}^{1/2}(A)\sum_{i=1}^{n-1}|A^{\times}_{q_i}||A^{\times}_{q_{i+1}}|^{1/2} \\ &\gg |A|^{3/2}\tilde{d}^{1/2}(A)|A:A|^{-1/2}\sum_{i=1}^{n-1}|A^{\times}_{q_i}| \\ &\gg |A|^{7/2}\tilde{d}^{1/2}(A)|A:A|^{-1/2}. \end{aligned} \tag{24}$$

Thus,

$$d(A) \le \tilde{d}(A) = \min_{i=2,\dots,n} \frac{|AA^{\times}_{q_i}|^2}{|A||A^{\times}_{q_i}|} \ll \frac{M^4|A:A|}{|A|}. \tag{25}$$

To estimate $d(A:A)$ and $d(AA)$ we use Lemma 9. In other words, taking our $C = A^{\times}_{q_i}$ to minimize (25), we get

$$d(AA) \le \frac{|AA^{\times}_{q_i}|^4}{|AA||A^{\times}_{q_i}|^3}, \qquad d(A:A) \le \frac{|AA^{\times}_{q_i}|^4}{|A:A||A^{\times}_{q_i}|^3} \ll \frac{M^8|A:A|}{|C|}. \tag{26}$$

Applying the first inequality of Theorem 11 with $A = A$ and $A_* = A:A$, we obtain

$$\begin{aligned} M|A|^{3/2} &\geq |A:A+A| \\ &\gg |A:A|^{8/9}|A|^{2/3}d^{-2/9}(A)\left(\frac{M^8|A:A|}{|C|}\right)^{-1/3} l^{-2/9} \\ &= |A:A|^{5/9}|A|^{2/3}d^{-2/9}(A)|C|^{1/3}M^{-8/3}l^{-2/9} \\ &\gg |A|^{14/9}M^{-32/9}l^{-2/9}, \end{aligned}$$

and hence $M \gg l^{-2/41}|A|^{1/82}$. This implies (21).

It remains to prove (22). In this case we multiply all points of A^2 lying on the line l_i by $\Delta(A)$, so we obtain $|AA^\times_{q_i}|$ points still belonging to the line l_i, and then we consider the sumset of the resulting set with $l_{i+1} \cap A^2$. Clearly, we obtain $|AA^\times_{q_i}||A^\times_{q_{i+1}}|$ points from the set $(AA+A)^2$. Thus,

$$M^2|A|^3 \gg |AA+A|^2 \geq \sum_{i=1}^{n-1} |A^\times_{q_i}||AA^\times_{q_{i+1}}|, \tag{27}$$

and we repeat the arguments above. The proof gives us

$$|AA+A| \gg |AA|^{11/41}|A|^{-4/41}(\mathsf{E}^\times_{3/2}(A))^{22/41}l^{-2/41}. \tag{28}$$

Here we have chosen the set Π as $\sum_{q\in\Pi}|A^\times_q|^{3/2} \gg \mathsf{E}^\times_{3/2}(A)$ or, in other words, $|A^\times_q| \gg (\mathsf{E}^\times_{3/2}(A))^2|A|^{-4}$. Using the Hölder inequality, combined with (28), we get

$$|AA+A| \gg |AA|^{11/41}|A:A|^{-11/41}|A|^{62/41}l^{-2/41}. \qquad \square$$

Remark 13. Using the full power of Theorem 14 from [Shkredov 2015], one can obtain further results connecting $|AA:A|$ and $|A:AA|$ with $|AA+A|$ and $|A:A+A|$ and so on. We do not make such calculations.

The same method allows us to improve the result of Balog concerning the size of $AA+AA$ and $A:A+A:A$.

Theorem 14. *Let $A \subset \mathbb{R}$ be a set. Then*

$$|A:A+A:A| \gg |A:A|^{14/29}|A|^{30/29}(\log|A|)^{-2/29}, \tag{29}$$

and

$$|AA+AA| \gg |AA|^{19/29}|A:A|^{-5/29}|A|^{30/29}(\log|A|)^{-2/29}. \tag{30}$$

Proof. As in the proof of Theorem 12, we define l_i to be the line $y = q_i x$ and let $q_1, \dots, q_n \in \Pi \subseteq A:A$ be such that $q_1 < q_2 < \dots < q_n$ and $|A^\times_{q_i}| \geq 2^{-1}|A|^2/|A:A|$ for any $q_i \in \Pi$. Thus, $\sum_i |A^\times_{q_i}| \geq \frac{1}{2}|A|^2$. We multiply all points of A^2 lying on the line l_i by $\Delta(A^{-1})$, so we obtain $|A^\times_{q_i}:A|$ points still belonging to the line l_i, and then we consider the sumset of the resulting set with itself. Clearly, we get

$|A^{\times}_{q_i} : A||A^{\times}_{q_{i+1}} : A|$ points from the set $(A : A + A : A)^2$ lying between the lines l_i and l_{i+1}. Therefore, we have

$$\begin{aligned}\sigma^2 &:= |A : A + A : A|^2 \\ &\geq \sum_{i=1}^{n-1} |A^{\times}_{q_i} : A||A^{\times}_{q_{i+1}} : A| \\ &\geq \tilde{d}(A)|A| \sum_{i=1}^{n-1} |A^{\times}_{q_i}|^{1/2} |A^{\times}_{q_{i+1}}|^{1/2} \\ &\gg |A|^3 \tilde{d}(A),\end{aligned} \tag{31}$$

where

$$\tilde{d}(A) := \min_{i=1,\dots,n} \frac{|A^{\times}_{q_i} : A|^2}{|A||A^{\times}_{q_i}|}. \tag{32}$$

This gives us $d(A) \leq \tilde{d}(A) \ll \sigma^2 |A|^{-3}$. Using Theorem 11 with $A = A_* = A : A$, we obtain

$$\begin{aligned}\sigma &\gg |A : A|^{14/9} \left(\frac{\sigma^4}{|A|^4 |A : A||C|} \right)^{-5/9} l^{-2/9} \\ &\gg |A : A|^{19/9} |A|^{20/9} |C|^{5/9} \sigma^{-20/9} l^{-2/9} \\ &\gg |A : A|^{14/9} \sigma^{-20/9} |A|^{10/3} l^{-2/9}.\end{aligned} \tag{33}$$

After some calculations, we get $\sigma \gg |A : A|^{14/29} |A|^{30/29} l^{-2/29}$.

To obtain (30) we use the previous arguments. We have

$$\begin{aligned}\sigma^2 &:= |AA + AA|^2 \\ &\geq \sum_{i \in \Pi} |AA^{\times}_{q_i}||AA^{\times}_{q_{i+1}}| \\ &\geq d(A)|A| \sum_{i \in \Pi} |A^{\times}_{q_i}|^{1/2} |A^{\times}_{q_{i+1}}|^{1/2} \\ &\gg d(A)|A||\Pi|\Delta,\end{aligned} \tag{34}$$

choosing $\Pi \subseteq A : A$ such that for any $q \in \Pi$ one has $|A|^2/|A : A| \ll \Delta \leq |A^{\times}_q|$. Clearly, such a set Π exists by simple average arguments. Calculations like those in (33) give us

$$\sigma \gg |AA|^{14/9} \left(\frac{\sigma^4}{|AA||\Pi|^2 \Delta^3} \right)^{-5/9} l^{-2/9} \gg |AA|^{19/9} (|\Pi| \Delta^{3/2})^{10/9} \sigma^{-20/9} l^{-2/9}.$$

After some computations, we obtain

$$\sigma \gg |AA|^{19/29} |A : A|^{-5/29} |A|^{30/29} l^{-2/29}. \qquad \square$$

Finally, let us obtain a result on $AA+A$ and $AA+AA$ of another type.

Proposition 15. *Let $A \subset \mathbb{R}$ be a set. Then*

$$|AA+A|^4,\ |A:A+A|^4 \gg |A|^{-2}(\mathsf{E}^{\times}_{3/2}(A))^2\,\mathsf{E}^{+}_3(A)\log^{-3}|A|, \tag{35}$$

and

$$|AA+AA|^2,\ |A:A+A:A|^2 \gg \mathsf{E}^{+}_3(A)\log^{-3}|A|. \tag{36}$$

Moreover,

$$|AA+A|^4,\ |A:A+A|^4 \gg \frac{|A|^{10}}{|A:A||A-A|^2}, \tag{37}$$

and

$$|AA+AA|^2,\ |A:A+A:A|^2 \gg \frac{|A|^6}{|A-A|^2}. \tag{38}$$

Proof. Put $l=\log|A|$. Using Lemma 7, we obtain that for any A, B and C

$$\sum_x (A\circ A)(x)(B\circ B)(x)(C\circ C)(x) \ll |A||B||C|\big(d(A)d(B)d(C)\big)^{1/3}\log\big(|A||B||C|\big). \tag{39}$$

In the particular case $A=B=C$, the definition of $d(A)$ gives us

$$|AA^{\times}_s|^2,\ |A:A^{\times}_s|^2 \gg |A|^{-2}|A^{\times}_s|\,\mathsf{E}^{+}_3(A)\,l^{-1} \tag{40}$$

for any $s \in A:A$. Using pigeonholing, choose $\Pi \subseteq A:A$ such that $|A^{\times}_q|$ differs at most twice from Π and such that $\sum_{q\in\Pi}|A^{\times}_q|^{3/2} \gg \mathsf{E}^{\times}_{3/2}(A)\,l^{-1}$. Applying (24), (27), (40) and the last bound, we obtain (35). Using (40) one more time and Katz–Koester inclusion [2010], namely,

$$AA^{\times}_s \subseteq AA \cap sAA, \quad A:A^{\times}_s \subseteq (A:A)\cap s^{-1}(A:A), \tag{41}$$

as well as formula (18) of Solymosi's result, we get (36). Another way to prove (36) is just to use formulas (31) and (34), combined with (40).

Inequalities (37) and (38) follow similarly to (35) and (36) from a direct application of Definition 6 and the Hölder inequality. For example, let us show how to get the first estimate of (37). Taking $B = -A$ and the parameter $\tau = |A|^2/(2|A-A|)$ in Definition 6, we obtain

$$d(A) \gg \frac{|A|^3}{|A-A|^2}.$$

Applying (24) and the lower bound for $d(A)$, we get

$$\begin{aligned} |A:A+A|^2 &\ge \sum_{i=1}^{n-1} |A^\times_{q_i}||A^\times_{q_{i+1}}:A| \\ &\ge |A|^{1/2} d^{1/2}(A) \sum_{i=1}^{n-1} |A^\times_{q_i}||A^\times_{q_{i+1}}| \\ &\gg |A|^{1/2}(|A|^3|A-A|^{-2})^{1/2}(|A|^2/|A:A|)^{1/2} \sum_{i=1}^{n-1} |A^\times_{q_i}| \\ &\gg |A|^{5/2}(|A|^3|A-A|^{-2})^{1/2}(|A|^2/|A:A|)^{1/2} \end{aligned}$$

as required. □

Remark 16. Applying arguments in the proof of (36) as well as formula (12) of Lemma 7, we obtain a similar bound, namely,

$$\mathsf{E}^+(A) \ll |A||AA+AA|$$

(actually, using methods from [Shkredov 2013a] one can improve the inequality). It is interesting to compare this estimate with Solymosi's upper bound for the multiplicative energy (19). Using formula (11) of Lemma 7, we also have

$$(\mathsf{E}^+(A))^{3/2}\,\mathsf{E}^\times_{3/2}(A) \ll \mathsf{E}^+_{3/2}(A)|A||AA+A|^2.$$

Combining inequality (36) with some estimates from [Shkredov 2014], we obtain a result in the spirit of [Roche-Newton and Zhelezov 2015].

Corollary 17. *Let $A \subset \mathbb{R}$ be a set. Suppose that*

(42) $$|(A+A)(A+A)+(A+A)(A+A)| \ll |A|^2 \quad \textit{and} \quad \mathsf{E}^+(A)|A-A| \ll |A|^4.$$

Then

(43) $$|A-A| \ll |A| \log^{12/7}|A|.$$

The same holds if one replaces addition with subtraction and multiplication with division in the first condition of (42).

If just the first condition of (42) holds (with plus) then

$$|A \pm A| \ll |A| \log^3 |A|, \tag{44}$$

and if it holds with minus then

$$|A - A| \ll |A| \log^3 |A|. \tag{45}$$

Again, one can replace multiplication with division in the first condition of (42).

Proof. Let us deal with the situation of the sum and the product. Other cases can be considered similarly. By Theorem 30 from [Shkredov 2014] and our second condition, one has

$$\mathsf{E}_3^+(A \pm A) \geq |A|^{45/4} |A - A|^{-1/2} (\mathsf{E}^+(A))^{-9/4} \gg |A|^{9/4} |A - A|^{7/4}.$$

On the other hand, using formula (36) from Proposition 15 and our first condition, we get

$$|A|^4 \log^3 |A| \gg \mathsf{E}_3^+(A \pm A) \gg |A|^{9/4} |A - A|^{7/4}$$

as required.

Finally, using the additive variant of Katz–Koester inclusion (41) (or see Proposition 29 from [Shkredov 2014]), we obtain

$$|A|^3 |A \pm A| \leq \mathsf{E}_3^+(A + A) \ll |A|^4 \log^3 |A|,$$

and

$$|A|^3 |A - A| \leq \mathsf{E}_3^+(A - A) \ll |A|^4 \log^3 |A|. \qquad \square$$

A simpler proof of a stronger result was kindly pointed out to the author by Oliver Roche-Newton. Indeed applying estimate (2) with $A = B = A + A$, $C = A$ and $D = A + A$, we obtain

$$\begin{aligned} |A|^4 &\gg |(A + A)A + (A + A)(A + A)|^2 \\ &\gg |(A + A) : (A + A)| \, |A| \, |A + A| \\ &\gg |A|^3 |A + A|, \end{aligned}$$

and the result follows. Here we have used the estimate $|(A + A) : (A + A)| \geq |A|^2$ from [Balog and Roche-Newton 2015]. Applying the well-known Ungar bound $|(A - A) : (A - A)| \geq |A|^2$ and taking $C = A^{-1}$ and $D = (A + A)^{-1}$, one can replace division with multiplication.

Acknowledgements

The author is grateful to the referee for careful reading, to Tomasz Schoen for useful discussions and to Oliver Roche-Newton, who pointed out a short way to obtain a strong form of Corollary 17.

References

[Balog 2011] A. Balog, "A note on sum-product estimates", *Publ. Math. Debrecen* **79**:3-4 (2011), 283–289. MR 2907965 Zbl 1249.11035

[Balog and Roche-Newton 2015] A. Balog and O. Roche-Newton, "New sum-product estimates for real and complex numbers", *Discrete Comput. Geom.* **53**:4 (2015), 825–846. MR 3341581 Zbl 06450890

[Erdős and Szemerédi 1983] P. Erdős and E. Szemerédi, "On sums and products of integers", pp. 213–218 in *Studies in pure mathematics: to the memory of Paul Turán*, edited by P. Erdős et al., Birkhäuser, Basel, 1983. MR 86m:11011 Zbl 0526.10011

[Katz and Koester 2010] N. H. Katz and P. Koester, "On additive doubling and energy", *SIAM J. Discrete Math.* **24**:4 (2010), 1684–1693. MR 2012d:11020 Zbl 1226.05247

[Konyagin and Rudnev 2013] S. V. Konyagin and M. Rudnev, "On new sum-product-type estimates", *SIAM J. Discrete Math.* **27**:2 (2013), 973–990. MR 3056747 Zbl 1272.68328

[Li 2011] L. Li, "On a theorem of Schoen and Shkredov on sumsets of convex sets", preprint, 2011. arXiv 1108.4382

[Li and Roche-Newton 2012] L. Li and O. Roche-Newton, "Convexity and a sum-product type estimate", *Acta Arith.* **156**:3 (2012), 247–255. MR 2999071 Zbl 1279.11013

[Murphy et al. 2015] B. Murphy, O. Roche-Newton, and I. D. Shkredov, "Variations on the sum-product problem", *SIAM J. Discrete Math.* **29**:1 (2015), 514–540. MR 3323540 Zbl 06437890

[Petridis 2012] G. Petridis, "New proofs of Plünnecke-type estimates for product sets in groups", *Combinatorica* **32**:6 (2012), 721–733. MR 3063158 Zbl 1291.11127

[Raz et al. 2015] O. E. Raz, O. Roche-Newton, and M. Sharir, "Sets with few distinct distances do not have heavy lines", *Discrete Math.* **338**:8 (2015), 1484–1492. MR 3336119 Zbl 1310.05039

[Roche-Newton and Zhelezov 2015] O. Roche-Newton and D. Zhelezov, "A bound on the multiplicative energy of a sum set and extremal sum-products problems", *Mosc. J. Comb. Number Theory* **5**:1–2 (2015). arXiv 1410.1156

[Schoen 2014] T. Schoen, "On convolutions of convex sets and related problems", *Canad. Math. Bull.* **57**:4 (2014), 877–883. MR 3270808 Zbl 06377673

[Schoen and Shkredov 2011] T. Schoen and I. D. Shkredov, "On sumsets of convex sets", *Combin. Probab. Comput.* **20**:5 (2011), 793–798. MR 2012g:11021 Zbl 1306.11013

[Schoen and Shkredov 2013] T. Schoen and I. D. Shkredov, "Higher moments of convolutions", *J. Number Theory* **133**:5 (2013), 1693–1737. MR 3007128 Zbl 1300.11018

[Shkredov 2013a] I. D. Shkredov, "Some new inequalities in additive combinatorics", *Mosc. J. Comb. Number Theory* **3**:3-4 (2013), 189–239. MR 3284125 Zbl 06367620

[Shkredov 2013b] I. D. Shkredov, "Несколько новых результатов о старших энергиях", *Tr. Mosk. Mat. Obs.* **74**:1 (2013), 35–73. Translated as "Some new results on higher energies" in *Trans. Mosc. Math. Soc.* **74** (2013), 31–63. MR 3235789 Zbl 06371555

[Shkredov 2014] I. D. Shkredov, "Energies and structure of additive sets", *Electron. J. Combin.* **21**:3 (2014), Paper #P3.44. MR 3262281 Zbl 1301.11010

[Shkredov 2015] I. D. Shkredov, "О суммах множеств Семереди–Троттера", *Tr. Mat. Inst. Steklova* **289** (2015), 318–327. Translated as "On sums of Szemerédi–Trotter sets" in *Proc. Steklov Inst. Math.* **289**:1 (2015), 300–309.

[Solymosi 2009] J. Solymosi, "Sumas contra productos", *Gac. R. Soc. Mat. Esp.* **12**:4 (2009), 707–719. MR 2589308 Zbl 1284.11019

[Tao and Vu 2006] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics **105**, Cambridge University Press, 2006. MR 2008a:11002 Zbl 1127.11002

Received February 2, 2015. Revised May 23, 2015.

ILYA D. SHKREDOV
STEKLOV MATHEMATICAL INSTITUTE
UL. GUBKINA, 8
MOSCOW
119991
RUSSIA

and

IITP RAS
BOLSHOY KARETNY PER. 19
MOSCOW
127994
RUSSIA

ilya.shkredov@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.241

UNIQUENESS RESULT ON NONNEGATIVE SOLUTIONS OF A LARGE CLASS OF DIFFERENTIAL INEQUALITIES ON RIEMANNIAN MANIFOLDS

YUHUA SUN

We consider a large class of differential inequalities on complete connected Riemannian manifolds and provide a sufficient condition in terms of volume growth for the uniqueness of nonnegative solutions to the differential inequalities.

1. Introduction

The purpose of this paper is to give a sufficient condition for the uniqueness of nonnegative solutions to a large class of the differential inequalities

$$Lu + V(x)u^{\sigma} \leq 0 \tag{1-1}$$

on a connected geodesically complete noncompact N-dimensional Riemannian manifold M^N with $N \geq 2$. Here the operator L is defined by

$$Lu = \operatorname{div}(A(x, u, \nabla u)), \tag{1-2}$$

where $A(x, \eta, \xi) = (A_i(x, \eta, \xi))$ is a vector field on M^N, and for $i = 1, \ldots, N$ the $A_i(x, \eta, \xi)$ are Carathéodorian functions defined on $M^N \times [0, \infty) \times TM^N$, and TM^N is the tangent bundle of M^N. The function V is positive, measurable, and locally integrable on M^N.

Let $m \geq 1$ be an arbitrary given number. We say that the operator L belongs to the class $A(m)$ if there exists a positive constant C such that, for almost all $x \in M^N$, all $\eta \in [0, \infty)$, and all $\xi, \zeta \in T_x M^N$, the following conditions hold:

$$\begin{cases} (A(x, \eta, \xi), \xi) \geq 0, \\ |(A(x, \eta, \xi), \zeta)| \leq C(A(x, \eta, \xi), \xi)^{\frac{m-1}{m}} |\zeta|, \end{cases} \tag{1-3}$$

where $(\cdot\,, \cdot)$ is the inner product given by the Riemannian metric, and $|\zeta|$ is the norm of ζ in $T_x M^N$.

MSC2010: primary 35J99; secondary 58J05.
Keywords: differential inequalities, Riemannian manifolds, volume growth, uniqueness.

The definition of such a class operator $A(m)$ was first introduced by Mīklyukov [1979; 1980]. Actually, the operators of such a class are quite common. Let us mention some examples:

(1) m-Laplacian operator:

$$L_1 u = \operatorname{div}(|\nabla u|^{m-2}\nabla u), \quad m > 1. \tag{1-4}$$

(2) Mean curvature type operators:

$$L_2 u = \operatorname{div}\left(\frac{|\nabla u|^{m-2}\nabla u}{\sqrt{1+|\nabla u|^m}}\right), \quad m > 1. \tag{1-5}$$

and

$$L_3 u = \operatorname{div}\left(\frac{|\nabla u|^{m-2}\nabla u}{\sqrt{1+|\nabla u|^2}}\right), \quad m > 1. \tag{1-6}$$

(3) Nonlinear operator:

$$L_4 u = \operatorname{div}(a(x, u, \nabla u)|\nabla u|^{m-2}\nabla u), \quad m > 1. \tag{1-7}$$

The definition of L in (1-3) is less restrictive than the one defined by

$$|A(x, \eta, \xi)| \le C_1|\xi|^{m-1}, \quad |(A(x, \eta, \xi), \xi)| \ge C_2|\xi|^m, \tag{1-8}$$

for some positive constants C_1, C_2. For example, by choosing $a(x, \eta, \xi)$ of (1-7) appropriately, the operator L_4 belongs to $A(m)$ but does not necessarily satisfy (1-8).

Generally speaking, the operator Lu defined by (1-3) may meanwhile belong to several classes denoted by $A(m_1), \dots, A(m_k)$, where $m_1 \le m_2 \le \cdots \le m_k$. For example, the operators of L_2, L_3 belong to both $A(m-1)$ and $A(m)$. Throughout the paper, when we say that L belongs to the class of $A(m)$, we always mean m is the largest value m_k.

The purpose of this paper is to provide a very simple geometric condition of volume growth on M^N to suffice that the only nonnegative solution u of (1-1) is identically zero. Let us emphasize that there is no curvature assumption on manifolds throughout the paper.

First, let us give our setting on manifolds. Let M^N be a connected geodesically complete noncompact Riemannian manifold. Denote by μ the Riemannian measure, and by $B(x, r)$ the geodesic ball on M^N of radius r centered at $x \in M^N$. Given that $d(\cdot, \cdot)$ is the geodesic distance and that x_0 is a reference point on M, define $B_r := B(x_0, r)$ for simplicity, where $r = d(x, x_0)$. Assume also throughout the paper that $V(x) \in L^\infty_{\mathrm{loc}}(M^N)$.

The problem of investigating the uniqueness of nonnegative solutions has attracted a lot of attention, especially in the Euclidean space. For example, if $M^N = \mathbb{R}^N$ with $N \ge 2$, in the case of $V(x) \equiv 1$, the problem (1-1) was systematically investigated by

Kurta [1999]. By using the nonlinear capacity arguments, he obtained nonexistence results concerning different differential inequalities. For a specific operator L, let us recommend a series of papers of Mitidieri and Pokhozhaev [1998; 1999; 2001] for a more comprehensive description. Related problems have also been studied in massive literatures; see [Caristi et al. 2008; Caristi and Mitidieri 1997; D'Ambrosio 2009; D'Ambrosio and Mitidieri 2010; Ni and Serrin 1985; 1986] and the references therein.

Let us turn to the results in the Riemannian manifolds setting. The celebrated idea of studying the uniqueness of nonnegative solutions in terms of the volume of the geodesic ball was due to Cheng and Yau [1975]. They obtained the following marvelous result: if the volume estimate

$$\mu(B_r) \le Cr^2$$

holds for all large enough r, then any positive solution to $\Delta u \le 0$ is identically constant.

The amazing point of Cheng and Yau's result is that there is no assumption on either curvature or the behavior of the solution near infinity, only in terms of volume growth.

Very recently, this idea was used and developed in [Grigor'yan and Kondratiev 2010; Grigor'yan and Sun 2014; Sun 2014] to investigate the differential inequality of the form

$$\operatorname{div}(A(x)\nabla u) + V(x)u^\sigma \le 0, \tag{1-9}$$

where $\sigma > 1$. Particularly, when $A(x) = \mathrm{Id}$ and $V(x) = 1$, (1-9) becomes

$$\Delta u + u^\sigma \le 0. \tag{1-10}$$

In [Grigor'yan and Sun 2014] it is proved that if

$$\mu(B_r) \le Cr^{\frac{2\sigma}{\sigma-1}} \ln^{\frac{1}{\sigma-1}} r$$

holds for all large enough r, then the only nonnegative solution of (1-10) is identically zero. Moreover, the exponents $2\sigma/(\sigma-1)$ and $1/(\sigma-1)$ are sharp and cannot be relaxed.

Let us define the weak nonnegative solution of (1-1). For convenience, we introduce the notation

$$A_u = (A(x, u, \nabla u), \nabla u) \tag{1-11}$$

and

$$W^{1,m}_{\mathrm{loc}}(M^N) := \{f \mid f \in L^m_{\mathrm{loc}}(M^N),\ \nabla f \in L^m_{\mathrm{loc}}(M^N)\}, \tag{1-12}$$

and denote by $W_c^{1,m}(M^N)$ the subspace of $W_{\mathrm{loc}}^{1,m}(M^N)$ of functions with compact support.

Definition 1.1. A function u on M^N is called a weak nonnegative solution of (1-1) if $u \in W_{\mathrm{loc}}^{1,m}(M^N)$ and $A_u \in L_{\mathrm{loc}}^1(M^N)$ and if, for any nonnegative function $\psi \in W_c^{1,m}(M^N)$, the following inequality holds:

$$-\int_{M^N} (A(x,u,\nabla u),\nabla\psi)\,d\mu + \int_{M^N} V(x)u^\sigma \psi\,d\mu \le 0, \tag{1-13}$$

where $(\cdot\,,\cdot)$ is the inner product in $T_x(M^N)$ given by a Riemannian metric.

Remark 1.2. If u is a weak nonnegative solution of (1-1), and the operator L belongs to the class $A(m)$, we know

$$\begin{aligned}\int_{M^N} (A(x,u,\nabla u),\nabla\psi)\,d\mu &\le C\int_{M^N} |\nabla\psi| A_u^{\frac{m-1}{m}}\,d\mu \\ &\le C\left(\int_{M^N} |\nabla\psi|^m\,d\mu\right)^{\frac{1}{m}}\left(\int_{\mathrm{supp}(\psi)} A_u\,d\mu\right)^{\frac{m-1}{m}} < \infty.\end{aligned}$$

Hence, by the definition of the solution, we know the second integral in (1-13) is bounded.

Define

$$p = \frac{m\sigma}{\sigma - m + 1}, \qquad q = \frac{m-1}{\sigma - m + 1}, \tag{1-14}$$

and introduce a new measure ν defined by

$$d\nu = V^{-\frac{m-1}{\sigma-m+1}}\,d\mu. \tag{1-15}$$

Assume that V satisfies the following condition: for some nonnegative constants δ_1, δ_2, the estimate

$$cr^{-\delta_1} \le V(x) \le Cr^{\delta_2} \tag{V}$$

holds for all large enough r.

Theorem 1.3. *Assume that operator L in (1-1) belongs to the class of $A(m)$ with $1 < m < \sigma + 1$. Assume also that (V) holds with $\delta_1, \delta_2 \ge 0$. If the inequality*

$$\nu(B_r \setminus \overline{B_1}) \le Cr^p \ln^q r \tag{1-16}$$

holds for all large enough r, then the only nonnegative solution of (1-1) is identically zero.

Remark 1.4. It is not clear that the sharpness of exponents p and q in (1-16) holds for all the operators of the class $A(m)$. However, in many specific cases, the exponents p, q are sharp; one can refer to [Grigor'yan and Sun 2014; Sun 2014; 2015].

Notation. The letters C, C', C_0, $C_1, \ldots$ denote positive constants whose values are unimportant and may vary at different occurrences.

In Section 2, we show the proof of Theorem 1.3. In Section 3, we present two examples to show that our result is very inclusive.

2. Proof of Theorem 1.3

Let u be a nonnegative solution of (1-1). Fix some ball B_R, where $R > 0$ is to be chosen later. Take a Lipschitz function φ on M^N with compact support, such that $0 \le \varphi \le 1$ and $\varphi \equiv 1$ in a neighborhood of $\overline{B_R}$. Particularly, $\varphi \in W_c^{1,m}(M^N)$. We use the following test function for (1-13):

$$\psi_\rho(x) = \varphi(x)^s (u+\rho)^{-t}, \tag{2-1}$$

where $\rho > 0$ is a parameter near zero, and s will be chosen to be a large enough fixed constant, and t will take arbitrarily small positive values near zero.

Since $1/(u+\rho)$ is bounded, ψ_ρ has compact support and is bounded. The identity

$$\nabla \psi_\rho = -t\varphi^s (u+\rho)^{-t-1} \nabla u + s\varphi^{s-1}(u+\rho)^{-t} \nabla \varphi$$

implies that $\nabla \psi_\rho \in L^m(M^N)$, hence, $\psi_\rho \in W_c^{1,m}(M^N)$. We obtain from (1-13) that

$$\begin{aligned} t \int_{M^N} \varphi^s (u+\rho)^{-t-1} A_u \, d\mu + \int_{M^N} \varphi^s V u^\sigma (u+\rho)^{-t} \, d\mu \qquad & \\ \le s \int_{M^N} \varphi^{s-1} (u+\rho)^{-t} (A(x,u,\nabla u), \nabla \varphi) \, d\mu. & \end{aligned} \tag{2-2}$$

Estimate the right-hand side of (2-2) by the Young inequality

$$\int_{M^N} fg \, d\mu \le \epsilon \int_{M^N} |f|^{p_0} \, d\mu + C_\epsilon \int_{M^N} |g|^{p_0'} \, d\mu, \tag{2-3}$$

where $1/p_0 + 1/p_0' = 1$. Letting $p_0 = m/(m-1)$, and using (1-3), we obtain

$$\begin{aligned} s \int_{M^N} & \varphi^{s-1} (u+\rho)^{-t} (A(x,u,\nabla u), \nabla \varphi) \, d\mu \\ & \le Cs \int_{M^N} \varphi^{s-1} (u+\rho)^{-t} A_u^{\frac{m-1}{m}} |\nabla \varphi| \, d\mu \\ & = C \int_{M^N} \Big[t^{\frac{1}{p_0}} \varphi^{\frac{s}{p_0}} (u+\rho)^{-\frac{t+1}{p_0}} A_u^{\frac{m-1}{m}} \Big] \Big[\frac{s}{t^{p_0}} \varphi^{\frac{s}{p_0'}-1} (u+\rho)^{1-\frac{t+1}{p_0'}} |\nabla \varphi| \Big] \, d\mu \\ & \le \frac{t}{2} \int_{M^N} \varphi^s (u+\rho)^{-t-1} A_u \, d\mu + C \frac{s^m}{t^{m-1}} \int_{M^N} \varphi^{s-m} (u+\rho)^{m-t-1} |\nabla \varphi|^m \, d\mu. \end{aligned}$$

Substituting the above into (2-2), and canceling out half of the first term in (2-2), we obtain

$$\frac{t}{2}\int_{M^N}\varphi^s(u+\rho)^{-t-1}A_u\,d\mu+\int_{M^N}\varphi^s Vu^\sigma(u+\rho)^{-t}\,d\mu \leq C\frac{s^m}{t^{m-1}}\int_{M^N}\varphi^{s-m}(u+\rho)^{m-t-1}|\nabla\varphi|^m\,d\mu. \tag{2-4}$$

Using the Young inequality again in the right-hand side of (2-4) with

$$p_1=\frac{\sigma-t}{m-t-1},\qquad p_1'=\frac{\sigma-t}{\sigma-m+1},$$

we obtain

$$\begin{aligned}\frac{s^m}{t^{m-1}}\int_{M^N}&\varphi^{s-m}(u+\rho)^{m-t-1}|\nabla\varphi|^m\,d\mu\\ &=\int_{M^N}\Big[\varphi^{\frac{s}{p_1}}V^{\frac{1}{p_1}}(u+\rho)^{\frac{\sigma-t}{p_1}}\Big]\Big[\frac{s^m}{t^{m-1}}\varphi^{\frac{s}{p_1'}-m}V^{-\frac{1}{p_1}}|\nabla\varphi|^m\Big]\,d\mu\\ &\leq\frac12\int_{M^N}\varphi^sV(u+\rho)^{\sigma-t}\,d\mu\\ &\qquad+C\Big(\frac{s^m}{t^{m-1}}\Big)^{\frac{\sigma-t}{\sigma-m+1}}\int_{M^N}\varphi^{s-\frac{m(\sigma-t)}{\sigma-m+1}}V^{-\frac{m-t-1}{\sigma-m+1}}|\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\,d\mu.\end{aligned}\tag{2-5}$$

Using in the right-hand side of (2-5) the simple inequality

$$\Big(\frac{s^m}{t^{m-1}}\Big)^{\frac{\sigma-t}{\sigma-m+1}}\leq\Big(\frac{s^m}{t^{m-1}}\Big)^{\frac{\sigma}{\sigma-m+1}}$$

and combining (2-5) with (2-4), we obtain that

$$\begin{aligned}\frac{t}{2}\int_{M^N}\varphi^s(u+\rho)^{-t-1}A_u\,d\mu&+\int_{M^N}\varphi^s Vu^\sigma(u+\rho)^{-t}\,d\mu\\ &\leq\frac12\int_{M^N}\varphi^sV(u+\rho)^{\sigma-t}\,d\mu\\ &\qquad+Ct^{-\frac{\sigma(m-1)}{\sigma-m+1}}\int_{M^N}\varphi^{s-\frac{m(\sigma-t)}{\sigma-m+1}}V^{-\frac{m-t-1}{\sigma-m+1}}|\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\,d\mu,\end{aligned}\tag{2-6}$$

where the value of s is absorbed into the constant C.

It is easy to obtain from the definition of the solution the boundedness of the term

$$\int_{M^N}\varphi^sVu^\sigma(u+\rho)^{-t}\,d\mu.$$

Then the boundedness of $\int_{M^N}\varphi^sV(u+\rho)^{\sigma-t}\,d\mu$ follows by the boundedness of

$$\int_{M^N}\varphi^sVu^\sigma(u+\rho)^{-t}\,d\mu,$$

and by the fact that $V \in L^1_{\mathrm{loc}}(M^N)$.

By the dominated convergence theorem, we know

$$\lim_{\rho \downarrow 0} \int_{M^N} \varphi^s V (u+\rho)^{\sigma - t}\, d\mu = \int_{M^N} \varphi^s V u^{\sigma - t}\, d\mu.$$

Letting $\rho \downarrow 0$ in (2-6) and applying the monotone convergence theorem, we have

$$\frac{t}{2}\int_{M^N} \varphi^s u^{-t-1} A_u\, d\mu + \int_{M^N} \varphi^s V u^{\sigma-t}\, d\mu \\ \leq \frac{1}{2}\int_{M^N} \varphi^s V u^{\sigma-t}\, d\mu + C t^{-\frac{\sigma(m-1)}{\sigma-m+1}} \int_{M^N} \varphi^{s-\frac{m(\sigma-t)}{\sigma-m+1}} V^{-\frac{m-t-1}{\sigma-m+1}} |\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\, d\mu,$$

which is

$$\frac{t}{2}\int_{M^N} \varphi^s u^{-t-1} A_u\, d\mu + \frac{1}{2}\int_{M^N} \varphi^s V u^{\sigma-t}\, d\mu \\ \leq C t^{-\frac{\sigma(m-1)}{\sigma-m+1}} \int_{M^N} \varphi^{s-\frac{m(\sigma-t)}{\sigma-m+1}} V^{-\frac{m-t-1}{\sigma-m+1}} |\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\, d\mu. \tag{2-7}$$

Applying (1-13) once more, using another test function $\psi = \varphi^s$, we obtain

$$\begin{aligned} \int_{M^N} \varphi^s V u^{\sigma}\, d\mu &\leq s\int_{M^N} \varphi^{s-1} (A(x,u,\nabla u), \nabla\varphi)\, d\mu \\ &\leq Cs \int_{M^N} \varphi^{s-1} A_u^{\frac{m-1}{m}} |\nabla\varphi|\, d\mu \\ &\leq Cs \left(\int_{M^N} \varphi^s u^{-t-1} A_u\, d\mu\right)^{\frac{m-1}{m}} \left(\int_{M^N} \varphi^{s-m} u^{(t+1)(m-1)} |\nabla\varphi|^m\, d\mu\right)^{\frac{1}{m}}. \end{aligned} \tag{2-8}$$

From (2-7), we obtain

$$\int_{M^N} \varphi^s u^{-t-1} A_u\, d\mu \leq C t^{-1-\frac{\sigma(m-1)}{\sigma-m+1}} \int_{M^N} \varphi^{s-\frac{m(\sigma-t)}{\sigma-m+1}} V^{-\frac{m-t-1}{\sigma-m+1}} |\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\, d\mu.$$

Substituting into (2-8) yields

$$\int_{M^N} \varphi^s V u^{\sigma}\, d\mu \leq C\left[t^{-1-\frac{\sigma(m-1)}{\sigma-m+1}} \int_{M^N} \varphi^{s-\frac{m(\sigma-t)}{\sigma-m+1}} V^{-\frac{m-t-1}{\sigma-m+1}} |\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\, d\mu\right]^{\frac{m-1}{m}} \\ \times \left[\int_{M^N} \varphi^{s-m} u^{(t+1)(m-1)} |\nabla\varphi|^m\, d\mu\right]^{\frac{1}{m}}. \tag{2-9}$$

Recalling that $\nabla\varphi = 0$ on B_R and applying the Hölder inequality to the last term of (2-9) with the Hölder couple

$$p_2 = \frac{\sigma}{(t+1)(m-1)}, \quad p_2' = \frac{\sigma}{\sigma-(t+1)(m-1)},$$

we obtain

$$\begin{aligned}(2\text{-}10)\quad &\int_{M^N}\varphi^{s-m}u^{(t+1)(m-1)}|\nabla\varphi|^m\,d\mu\\ &=\int_{M^N\setminus B_R}\bigl(\varphi^{\frac{s}{p_2}}V^{\frac{1}{p_2}}u^{(t+1)(m-1)}\bigr)\bigl(\varphi^{\frac{s}{p_2'}-m}V^{-\frac{1}{p_2}}|\nabla\varphi|^m\bigr)\,d\mu\\ &\le\Bigl(\int_{M^N\setminus B_R}\varphi^s Vu^\sigma\,d\mu\Bigr)^{\frac{(t+1)(m-1)}{\sigma}}\\ &\quad\times\Bigl(\int_{M^N\setminus B_R}\varphi^{s-\frac{m\sigma}{\sigma-(t+1)(m-1)}}V^{-\frac{(t+1)(m-1)}{\sigma-(t+1)(m-1)}}|\nabla\varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}}\,d\mu\Bigr)^{\frac{\sigma-(t+1)(m-1)}{\sigma}}.\end{aligned}$$

Substituting (2-10) into (2-9), choosing s large enough, and noting that $\varphi\le 1$, we obtain

$$\begin{aligned}(2\text{-}11)\quad \int_{M^N}\varphi^s Vu^\sigma\,d\mu &\le Ct^{-\frac{m-1}{m}-\frac{\sigma(m-1)^2}{m(\sigma-m+1)}}\Bigl(\int_{M^N}V^{-\frac{m-1-t}{\sigma-m+1}}|\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\,d\mu\Bigr)^{\frac{m-1}{m}}\\ &\quad\times\Bigl(\int_{M^N}V^{-\frac{(t+1)(m-1)}{\sigma-(t+1)(m-1)}}|\nabla\varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}}\,d\mu\Bigr)^{\frac{\sigma-(t+1)(m-1)}{m\sigma}}\\ &\quad\times\Bigl(\int_{M^N\setminus B_R}\varphi^s Vu^\sigma\,d\mu\Bigr)^{\frac{(t+1)(m-1)}{m\sigma}}.\end{aligned}$$

From the definition of the solution, we know $\int_{M^N}\varphi^s Vu^\sigma\,d\mu$ is finite. It follows from (2-11) that

$$\begin{aligned}(2\text{-}12)\quad &\Bigl(\int_{M^N}\varphi^s Vu^\sigma\,d\mu\Bigr)^{1-\frac{(t+1)(m-1)}{m\sigma}}\\ &\le Ct^{-\frac{m-1}{m}-\frac{\sigma(m-1)^2}{m(\sigma-m+1)}}\Bigl(\int_{M^N}V^{-\frac{m-1-t}{\sigma-m+1}}|\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\,d\mu\Bigr)^{\frac{m-1}{m}}\\ &\quad\times\Bigl(\int_{M^N}V^{-\frac{(t+1)(m-1)}{\sigma-(t+1)(m-1)}}|\nabla\varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}}\,d\mu\Bigr)^{\frac{\sigma-(t+1)(m-1)}{m\sigma}}.\end{aligned}$$

Note that the first integral in the right-hand side of (2-12) has the estimate

$$(2\text{-}13)\quad \int_{M^N}V^{-\frac{m-1-t}{\sigma-m+1}}|\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}\,d\mu\le\int_{M^N}|\nabla\varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}}V^{\frac{t}{\sigma-m+1}}\,d\nu,$$

where we have used that $d\nu=V^{-\frac{m-1}{\sigma-m+1}}\,d\mu$. Similarly, the second integral in the right-hand side of (2-12) can be estimated as follows:

$$\begin{aligned}(2\text{-}14)\quad &\int_{M^N}V^{-\frac{(t+1)(m-1)}{\sigma-(t+1)(m-1)}}|\nabla\varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}}\,d\mu\\ &\le\int_{M^N}|\nabla\varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}}V^{-\frac{t\sigma(m-1)}{[\sigma-(t+1)(m-1)](\sigma-m+1)}}\,d\nu.\end{aligned}$$

Substituting (2-13) and (2-14) into (2-11), we have

$$
\begin{aligned}
\int_{M^N} \varphi^s V u^\sigma \, d\mu
&\le C t^{-\frac{m-1}{m} - \frac{\sigma(m-1)^2}{m(\sigma-m+1)}} \left(\int_{M^N} |\nabla \varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}} V^{\frac{t}{\sigma-m+1}} \, d\nu \right)^{\frac{m-1}{m}} \\
&\quad \times \left(\int_{M^N} |\nabla \varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}} V^{-\frac{t\sigma(m-1)}{[\sigma-(t+1)(m-1)](\sigma-m+1)}} \, d\nu \right)^{\frac{\sigma-(t+1)(m-1)}{m\sigma}} \\
&\quad \times \left(\int_{M^N \setminus B_R} \varphi^s V u^\sigma \, d\mu \right)^{\frac{(t+1)(m-1)}{m\sigma}}.
\end{aligned}
\tag{2-15}
$$

Substituting (2-13) and (2-14) into (2-12), we obtain

$$
\begin{aligned}
\left(\int_{M^N} \varphi^s V u^\sigma \, d\mu \right)^{1-\frac{(t+1)(m-1)}{m\sigma}}
&\le C t^{-\frac{m-1}{m} - \frac{\sigma(m-1)^2}{m(\sigma-m+1)}} \left(\int_{M^N} |\nabla \varphi|^{\frac{m(\sigma-t)}{\sigma-m+1}} V^{\frac{t}{\sigma-m+1}} \, d\nu \right)^{\frac{m-1}{m}} \\
&\quad \times \left(\int_{M^N} |\nabla \varphi|^{\frac{m\sigma}{\sigma-(t+1)(m-1)}} V^{-\frac{t\sigma(m-1)}{[\sigma-(t+1)(m-1)](\sigma-m+1)}} \, d\nu \right)^{\frac{\sigma-(t+1)(m-1)}{m\sigma}}.
\end{aligned}
\tag{2-16}
$$

Let $\{\tilde{\varphi}_k\}_{k\in\mathbb{N}}$ be a sequence for which each $\tilde{\varphi}_k$ is a Lipschitz function such that $\operatorname{supp}(\tilde{\varphi}_k) \subset B_{2^k}$, and $\tilde{\varphi}_k = 1$ in a neighborhood of $B_{2^{k-1}}$, and

$$
|\nabla \tilde{\varphi}_k| \begin{cases} \le \frac{C}{2^{k-1}} & \text{for } x \in B_{2^k} \setminus B_{2^{k-1}}, \\ = 0 & \text{otherwise}, \end{cases}
\tag{2-17}
$$

where C does not depend on k.

Fix some $n \in \mathbb{N}$ and set

$$
t = \frac{1}{n}
\tag{2-18}
$$

and

$$
\varphi_n = \frac{\sum_{k=n+1}^{2n} \tilde{\varphi}_k}{n}.
\tag{2-19}
$$

Note that $\varphi_n = 1$ on B_{2^n}, and $\varphi_n = 0$ outside $B_{2^{2n}}$, and $0 \le \varphi_n \le 1$ on M^N. Note that, for any $a \ge 1$, using that $\operatorname{supp}(\nabla \tilde{\varphi}_k)$ are disjoint, we have

$$
|\nabla \varphi_n|^a = \frac{\sum_{k=n+1}^{2n} |\nabla \tilde{\varphi}_k|^a}{n^a}.
\tag{2-20}
$$

It is easy to see that

$$
\varphi_n \in W^{1,m}_{\mathrm{loc}}(M^N).
$$

Consider the integral

$$J_n(a,b)=\int_{M^N}|\nabla\varphi_n|^a V^b\,d\mu, \tag{2-21}$$

where a, b are taking values from

$$(a,b)=\begin{cases}\left(\frac{m(\sigma-t)}{\sigma-m+1},\ \frac{t}{\sigma-m+1}\right),\\ \left(\frac{m\sigma}{\sigma-(t+1)(m-1)},\ -\frac{t\sigma(m-1)}{[\sigma-(t+1)(m-1)](\sigma-m+1)}\right).\end{cases} \tag{2-22}$$

We write a in the form

$$a=p+lt, \tag{2-23}$$

with the corresponding two values of l,

$$l_1=-\frac{m}{\sigma-m+1},\quad l_2=\frac{m\sigma(m-1)}{[\sigma-(t+1)(m-1)](\sigma-m+1)}, \tag{2-24}$$

where $p=m\sigma/(\sigma-m+1)$.

For $b\geq 0$, we know

$$\begin{aligned} J_n(a,b)&=\int_{M^N}|\nabla\varphi_n|^a V^b\,d\nu\\ &=\int_{M^N}\frac{\sum_{k=n+1}^{2n}|\nabla\tilde{\varphi}_k|^a}{n^a}V^b\,d\nu\\ &\leq\sum_{k=n+1}^{2n}\int_{B_{2^k}\setminus B_{2^{k-1}}}\frac{|\nabla\tilde{\varphi}_k|^a}{n^a}V^b\,d\nu\\ &\leq C\sum_{k=n+1}^{2n}\int_{B_{2^k}\setminus B_{2^{k-1}}}\left(\frac{2^{1-k}}{n}\right)^a r^{\delta_2 b}\,d\nu\\ &\leq C\sum_{k=n+1}^{2n}\left(\frac{2^{1-k}}{n}\right)^a(2^k)^{\delta_2 b}\nu(B_{2^k}\setminus B_1)\end{aligned} \tag{2-25}$$

Note that $a=p+lt$, and $n+1\leq k\leq 2n$, and

$$\begin{aligned}\left(\frac{2^{1-k}}{n}\right)^a(2^k)^{\delta_2 b}&=\left(\frac{2^{-k}}{n}\right)^p\left(\frac{2^{-k}}{n}\right)^{lt}(2^k)^{\delta_2 b}\\ &\leq\left(\frac{2^{-k}}{n}\right)^p(2^k)^{\delta_2 b}\sup_{n+1\leq k\leq 2n}\left(\frac{2^{-k}}{n}\right)^{lt}\\ &\leq C\left(\frac{2^{-k}}{n}\right)^p(2^k)^{\delta_2 b}.\end{aligned} \tag{2-26}$$

Substituting (2-26) into (2-25), and using the volume growth (1-16), we obtain

$$J_n(a,b) \le C \sum_{k=n+1}^{2n} \left(\frac{2^{-k}}{n}\right)^p (2^k)^{\delta_2 b} \nu(B_{2^k} \setminus B_1) \tag{2-27}$$
$$\le C \sum_{k=n+1}^{2n} \left(\frac{2^{-k}}{n}\right)^p (2^k)^{\delta_2 b} (2^k)^p \ln^q(2^k)$$
$$\le C \frac{1}{n^p} \sum_{k=n+1}^{2n} k^q 2^{k\delta_2 b} \le C n^{q+1-p} 2^{2n\delta_2 b} \le C n^{-\frac{\sigma(m-1)}{\sigma-m+1}} 2^{2n\delta_2 b}.$$

Similarly, for the case of $b \le 0$, we obtain

$$J_n(a,b) \le C n^{-\frac{\sigma(m-1)}{\sigma-m+1}} 2^{-2n\delta_1 b}. \tag{2-28}$$

Taking the sequence $\{\varphi_n\}$ in (2-16), we obtain

$$\left(\int_{M^N} \varphi_n^s V u^\sigma \, d\mu\right)^{1-\frac{(t+1)(m-1)}{m\sigma}} \le C t^{-\frac{m-1}{m}-\frac{\sigma(m-1)^2}{m(\sigma-m+1)}} \left(J_n\left(\frac{m(\sigma-t)}{\sigma-m+1}, \frac{t}{\sigma-m+1}\right)\right)^{\frac{m-1}{m}} \times \left(J_n\left(\frac{m\sigma}{\sigma-(t+1)(m-1)}, \frac{-t\sigma(m-1)}{[\sigma-(t+1)(m-1)](\sigma-m+1)}\right)\right)^{\frac{\sigma-(t+1)(m-1)}{m\sigma}}. \tag{2-29}$$

Substituting (2-27) and (2-28) and noting that $t = 1/n$, we obtain

$$\left(\int_{M^N} \varphi_n^s V u^\sigma \, d\mu\right)^{1-\frac{(\frac{1}{n}+1)(m-1)}{m\sigma}} \tag{2-30}$$
$$\le C n^{\frac{m-1}{m}+\frac{\sigma(m-1)^2}{m(\sigma-m+1)}} \left(n^{-\frac{\sigma(m-1)}{\sigma-m+1}} 2^{2n\delta_2 \frac{\frac{1}{n}}{\sigma-m+1}}\right)^{\frac{m-1}{m}} \times \left(n^{-\frac{\sigma(m-1)}{\sigma-m+1}} 2^{2n\delta_1 \frac{\frac{1}{n}\sigma(m-1)}{[\sigma-(\frac{1}{n}+1)(m-1)](\sigma-m+1)}}\right)^{\frac{\sigma-(\frac{1}{n}+1)(m-1)}{m\sigma}}$$
$$\le C n^{\frac{(m-1)^2}{n(\sigma-m+1)}} 2^{\frac{2(\delta_1+\delta_2)(m-1)}{m(\sigma-m+1)}}.$$

Noting that $\varphi_n = 1$ on B_{2^n} and taking the lim sup of both sides in (2-30) as $n \to \infty$, we obtain

$$\int_{M^N} V u^\sigma \, d\mu \le C < \infty. \tag{2-31}$$

Applying similar arguments to (2-15), we obtain that

$$\int_{M^N} \varphi_n^s V u^\sigma \, d\mu \le C\left(\int_{M^N \setminus B_{2^n}} \varphi_n^s V u^\sigma \, d\mu\right)^{\frac{(\frac{1}{n}+1)(m-1)}{m\sigma}}. \tag{2-32}$$

Since $\varphi_n = 1$ on B_{2^n}, we have

$$\int_{B_{2^n}} V u^\sigma \, d\mu \le C\left(\int_{M^N \setminus B_{2^n}} \varphi_n^s V u^\sigma \, d\mu\right)^{\frac{(\frac{1}{n}+1)(m-1)}{m\sigma}}. \tag{2-33}$$

Combining this with (2-31) and letting $n \to \infty$, we obtain that

$$\int_{M^N} V u^\sigma \, d\mu = 0,$$

since $V > 0$ for almost all $x \in M^N$. Thus $u \equiv 0$. □

3. Examples

Our result can cover many known results in the case of $M^N = R^N$. Let us mention two of these examples.

Example 1. Let us investigate the inequality

$$\operatorname{div}(|\nabla u|^{m-2}\nabla u) + V(x)u^\sigma \le 0, \quad \text{in } \mathbb{R}^N, \tag{3-1}$$

where $V(x) = 1/|x|^\gamma$ for $|x| \ge 1$, and $N > m > \max\{1, \gamma\}$, and $\sigma > m-1$.

By [Filippucci 2009, Corollary 1.5], we know if

$$\sigma \le \frac{(N-\gamma)(m-1)}{N-m}, \tag{3-2}$$

then (3-1) has no positive solutions in some natural class. Compared to our result of Theorem 1.3, we know for large r

$$\nu(B_r \setminus \overline{B_1}) = \int_{B_r \setminus \overline{B_1}} V^{-\frac{m-1}{\sigma-m+1}} \, d\mu = \omega_N \int_1^r s^{\frac{\gamma(m-1)}{\sigma-m+1}} s^{N-1} \, ds \approx C r^{N+\frac{\gamma(m-1)}{\sigma-m+1}}, \tag{3-3}$$

where ω_N is the surface area of the unit ball in $\mathbb{R}^N$, and μ is the Lebesgue measure, and the sign $\approx$ means that both the inequalities $\le$ and $\ge$ are satisfied but with different values of different constants c, C.

By (3-3), it follows that the condition (1-16) is equivalent to

$$N + \frac{\gamma(m-1)}{\sigma-m+1} \le p = \frac{m\sigma}{\sigma-m+1}, \tag{3-4}$$

which in turn is equivalent to (3-2).

Example 2. Consider the differential inequality

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right)+u^\sigma \le 0, \quad \text{in } \mathbb{R}^N, \tag{3-5}$$

where $N>2$, $\sigma>1$. This problem was investigated in [Mitidieri and Pokhozhaev 1999]. They obtained that if

$$\sigma \le \frac{N}{N-2}, \tag{3-6}$$

then (3-5) has no positive solutions. Note that the operator in (3-5) belongs to the class of $A(2)$, and that $\nu(B_r\setminus\overline{B_1})=\mu(B_r\setminus\overline{B_1})\approx Cr^N$. By Theorem 1.3, we know if

$$N \le \frac{2\sigma}{\sigma-1}, \tag{3-7}$$

then (3-5) has no positive solution. It is easy to check that (3-6) and (3-7) are equivalent.

Acknowledgements

The author would like to express his gratitude to the International Graduate College of Bielefeld University for the financial support. The author is also grateful to the anonymous referee for careful reading and valuable suggestions which helped improve the quality of the paper.

References

[Caristi and Mitidieri 1997] G. Caristi and E. Mitidieri, "Nonexistence of positive solutions of quasilinear equations", *Adv. Differential Equations* **2**:3 (1997), 319–359. MR 98a:34011 Zbl 1023.34500

[Caristi et al. 2008] G. Caristi, L. D'Ambrosio, and E. Mitidieri, "Liouville theorems for some nonlinear inequalities", *Proc. Steklov Inst. Math.* **260** (2008), 90–111. In Russian; translated in *Tr. Mat. Inst. Steklova* **260** (2008), 97–118. Zbl 1233.35207

[Cheng and Yau 1975] S. Y. Cheng and S. T. Yau, "Differential equations on Riemannian manifolds and their geometric applications", *Comm. Pure Appl. Math.* **28**:3 (1975), 333–354. MR 52 #6608 Zbl 0312.53031

[D'Ambrosio 2009] L. D'Ambrosio, "Liouville theorems for anisotropic quasilinear inequalities", *Nonlinear Anal.* **70**:8 (2009), 2855–2869. MR 2010e:35097 Zbl 1177.35262

[D'Ambrosio and Mitidieri 2010] L. D'Ambrosio and E. Mitidieri, "A priori estimates, positivity results, and nonexistence theorems for quasilinear degenerate elliptic inequalities", *Adv. Math.* **224**:3 (2010), 967–1020. MR 2011e:35109 Zbl 1206.35265

[Filippucci 2009] R. Filippucci, "Nonexistence of positive weak solutions of elliptic inequalities", *Nonlinear Anal.* **70**:8 (2009), 2903–2916. MR 2010f:35450 Zbl 1165.35487

[Grigor'yan and Kondratiev 2010] A. Grigor'yan and V. A. Kondratiev, "On the existence of positive solutions of semilinear elliptic inequalities on Riemannian manifolds", pp. 203–218 in *Around the research of Vladimir Maz'ya, II*, edited by A. Laptev, Int. Math. Ser. (N. Y.) **12**, Springer, New York, 2010. MR 2011m:58030 Zbl 1185.35344

[Grigor'yan and Sun 2014] A. Grigor'yan and Y. Sun, "On nonnegative solutions of the inequality $\Delta u + u^\sigma \le 0$ on Riemannian manifolds", *Comm. Pure Appl. Math.* **67**:8 (2014), 1336–1352. MR 3225632 Zbl 1296.58011

[Kurta 1999] V. V. Kurta, "The nonexistence of positive solutions to some elliptic equations", *Mat. Zametki* **65**:4 (1999), 552–561. In Russian; translated in *Math. Notes* **65**:4 (1999), 462–469. MR 2000f:35045 Zbl 0964.35043

[Mīkljukov 1979] V. M. Mīkljukov, "A new approach to the Bernšteĭn theorem and to related questions of equations of minimal surface type", *Mat. Sb.* (*N.S.*) **108(150)**:2 (1979), 268–289. In Russian, translated in *Math. USSR, Sb.* **36**:2 (1980), 251–271. MR 80e:53005 Zbl 0488.49029

[Mīklyukov 1980] V. M. Mīklyukov, "Capacity and a generalized maximum principle for quasilinear equations of elliptic type", *Dokl. Akad. Nauk SSSR* **250**:6 (1980), 1318–1320. In Russian; translated in *Sov. Math., Dokl.* **21** (1980), 320–322. MR 83d:35019 Zbl 0553.35026

[Mitidieri and Pokhozhaev 1998] È. Mitidieri and S. I. Pokhozhaev, "Absence of global positive solutions of quasilinear elliptic inequalities", *Dokl. Akad. Nauk* **359**:4 (1998), 456–460. In Russian; translated in *Dokl. Math.* **57**:2 (1998), 250–253. MR 2000a:35263 Zbl 0976.35100

[Mitidieri and Pokhozhaev 1999] È. Mitidieri and S. I. Pokhozhaev, "Absence of positive solutions for quasilinear elliptic problems in $\mathbf{R}^N$", *Tr. Mat. Inst. Steklova* **227** (1999), 192–222. In Russian; translated in *Proc. Steklov Inst. Math.* **227**:4 (1999), 186–216. MR 2001g:35082 Zbl 1056.35507

[Mitidieri and Pokhozhaev 2001] È. Mitidieri and S. I. Pokhozhaev, "A priori estimates and the absence of solutions of nonlinear partial differential equations and inequalities", *Tr. Mat. Inst. Steklova* **234** (2001), 1–384. In Russian; translated in *Proc. Steklov Inst. Math.* **234**:3 (2001), 1–362. MR 2005d:35004 Zbl 0987.35002

[Ni and Serrin 1985] W.-M. Ni and J. Serrin, "Nonexistence theorems for quasilinear partial differential equations", *Rend. Circ. Mat. Palermo* (2) *Suppl.* **8** (1985), 171–185. MR 88d:35069 Zbl 0625.35028

[Ni and Serrin 1986] W.-M. Ni and J. Serrin, "Existence and nonexistence theorems for ground states of quasilinear partial differential equations: The anomalous case", *Accad. Naz. Lincei* **77** (1986), 231–257.

[Sun 2014] Y. Sun, "Uniqueness result for non-negative solutions of semi-linear inequalities on Riemannian manifolds", *J. Math. Anal. Appl.* **419**:1 (2014), 643–661. MR 3217172 Zbl 1297.35297

[Sun 2015] Y. Sun, "On nonexistence of positive solutions of quasi-linear inequality on Riemannian manifolds", *Proc. Amer. Math. Soc.* **143**:7 (2015), 2969–2984. MR 3336621 Zbl 06428975

Received July 18, 2014. Revised March 11, 2015.

YUHUA SUN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BIELEFELD
D-33501 BIELEFELD
GERMANY
Current address:
SCHOOL OF MATHEMATICAL SCIENCES AND LPMC
NANKAI UNIVERSITY
94 WEIJIN ROAD
TIANJIN 300071
CHINA
sunyuhua@nankai.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.280.255

CORRECTION TO THE ARTICLE CLOSED ORBITS OF A CHARGE IN A WEAKLY EXACT MAGNETIC FIELD

WILL J. MERRY

Volume **247**:1 (2010), 189–212

Theorem 5.1 of the titular article is incorrect, as pointed out by Gabriele Benedetti. We describe the error and supply an alternative proof for the article's main result (Theorem 5.8).

1. Introduction

In this erratum I use the notation and numbering from [Merry 2010]. The problem, pointed out to me by G. Benedetti, resides in its Theorem 5.1; embarrassingly, the function $f : \mathbb{R}^+ \to \mathbb{R}$ defined by $f(x) := e^{-x}$ already provides a counterexample. One can take $\mathscr{F}_n$ to be the set of singletons $\{x\}$ for $x \in (0, n)$. Theorem 5.1 then erroneously concludes that f has a critical point x_∞ with $f(x_\infty) = 0$, which is, of course, incorrect.

Luckily, the error in Theorem 5.1 does not affect the main result (Theorem 5.8). In fact, whilst attempting to salvage the proof of Theorem 5.8, I realised that the entire argument could be dramatically simplified by the following observation: *Theorem 3.2 still holds in the case* $c(g, \sigma) = \infty$. The proof of this statement is explained below. Once this is established, Contreras' original argument [2006, Proposition 7.1] can be used directly to obtain [Merry 2010, Theorem 5.8].

L. Asselle and G. Benedetti [2015, Lemma 3.5] independently noticed that Theorem 5.8 could be proved by making use of this observation. In their paper, however, they take these ideas considerably further and extend the main result of [Merry 2010] to cover cases in which the magnetic form is *not* weakly exact.

2. The correction

All references in this section are to [Merry 2010]. Let us explain why Theorem 3.2 continues to hold even in the case $c(g, \sigma) = \infty$. We need only verify that the

MSC2010: 37J45, 70H12.
Keywords: magnetic flow, twisted geodesic flow, periodic orbits, Mañé critical value.

additional hypothesis in Proposition 3.7 — which deals specifically with the case $c(g, \sigma) = \infty$ — is superfluous. More precisely, we show that the hypotheses of Theorem 3.2 automatically imply that the hypotheses of Proposition 3.7 are satisfied, which therefore implies that Theorem 3.2 continues to hold in the case $c(g, \sigma) = \infty$.

Thus, we are given a sequence $(x_n, T_n) \subset \mathbb{D}(A, B, k, 0)$, and we must show that there always exists a compact subset $K \subset \widetilde{M}$ such that $x_n \in \Lambda_0^K$ for all $n \in \mathbb{N}$. For this it is enough to show that the energy e_n of (x_n, T_n) (defined on the bottom of page 197) is uniformly bounded. This then implies that the length l_n of x_n is bounded (compare Equation (3-1)), which immediately implies that such a compact set $K \subset \widetilde{M}$ exists. To see that e_n is bounded, we use Equation (2-6), which tells us

$$\frac{1}{n} \geq \left| \frac{\partial}{\partial T} S_k(x_n, T_n) \right| = \left| \frac{1}{T_n} \int_0^{T_n} (k - E(y_n, \dot{y}_n))\, dt \right| = \left| k - \frac{e_n}{T_n} \right|.$$

Since $|T_n| \leq B$ by assumption, e_n is necessarily bounded, as required.

Acknowledgement

I thank Gabriele Benedetti for patiently and repeatedly explaining to me why my Theorem 5.1 was false. I am particularly grateful for the considerable tact he showed while outlining to me how the quintessential function one uses to teach students the necessity of the Palais–Smale condition — namely $x \mapsto e^{-x}$ — provided a counterexample.

References

[Asselle and Benedetti 2015] L. Asselle and G. Benedetti, "The Lusternik–Fet theorem for autonomous Tonelli Hamiltonian systems on twisted cotangent bundles", 2015. arXiv 1412.0531v3

[Contreras 2006] G. Contreras, "The Palais–Smale condition on contact type energy levels for convex Lagrangian systems", *Calc. Var. Partial Differential Eq.* **27**:3 (2006), 321–395. MR 2007i:37116 Zbl 1105.37037

[Merry 2010] W. J. Merry, "Closed orbits of a charge in a weakly exact magnetic field", *Pacific J. Math.* **247**:1 (2010), 189–212. MR 2012b:37168 Zbl 1246.37082

Received January 8, 2015.

WILL J. MERRY
DEPARTEMENT MATHEMATIK
ETH ZÜRICH
RÄMISTRASSE 101
CH-8092 ZÜRICH
SWITZERLAND
merry@math.ethz.ch

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 280 No. 1 January 2016
