Pacific Journal of Mathematics

Volume 298 No. 1 January 2019

PACIFIC JOURNAL OF MATHEMATICS

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

msp.org/pjm

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

The subscription price for 2019 is US $490/year for the electronic version, and $665/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.1

ON SOME REFINEMENTS OF THE EMBEDDING OF CRITICAL SOBOLEV SPACES INTO BMO

ALMAZ BUTAEV

We introduce the nonhomogeneous analogs of function spaces studied by Van Schaftingen. We show that these classes refine the embedding $W^{1,n}(\mathbb{R}^n) \subset \mathrm{bmo}(\mathbb{R}^n)$. The analogous results are established on bounded Lipschitz domains and Riemannian manifolds with bounded geometry.

1. Introduction

Let f be a locally integrable function on $\mathbb{R}^n$. Given a cube $Q \subset \mathbb{R}^n$ (henceforth by a cube we will understand a cube with sides parallel to the axes), we denote the average of f over Q by f_Q, i.e.,

$$f_Q = \frac{1}{|Q|} \int_Q f(x)\,dx,$$

where $|Q|$ is the Lebesgue measure of Q.

In 1961, John and Nirenberg introduced the space of functions of bounded mean oscillation (BMO).

Definition 1.1. We say that $f \in \mathrm{BMO}(\mathbb{R}^n)$ if

$$\|f\|_{\mathrm{BMO}} := \sup_Q \frac{1}{|Q|} \int_Q |f(x) - f_Q|\,dx < \infty.$$

Note that $\|\cdot\|_{\mathrm{BMO}}$ is a norm on the quotient space of functions modulo constants.

Functions of bounded mean oscillations turned out to be the right substitute for L^∞ functions in a number of questions in analysis. In particular, the embedding theorem of Gagliardo, Nirenberg and Sobolev (see, e.g., [Stein 1970, Chapter V]) asserts that for any $p \in [1, n)$ there exists C_p such that

$$\|f\|_{L^{np/(n-p)}} \le C_p \|\nabla f\|_{L^p}, \quad \text{for all } f \in \mathcal{D}.$$

The inequality fails for $p = n$, so we do not have the embedding $W^{1,n}$ into L^∞. However, it follows from the Poincaré inequality that for some constant $C > 0$,

$$\|f\|_{\mathrm{BMO}} \le C \|\nabla f\|_{L^n}, \quad \text{for all } f \in \mathcal{D},$$

MSC2010: 42B35, 46E35.
Keywords: critical Sobolev spaces, BMO.

and therefore $\mathring{W}^{1,n}$ is continuously embedded into $\mathrm{BMO}(\mathbb{R}^n)$.

Based on an inequality established by Bourgain and Brezis [2004], Van Schaftingen [2006] defined a scale of spaces D_k using the k-differential forms

$$\Phi(x) = \sum_{1 \le i_1 < \cdots < i_k \le n} \phi_{i_1,\dots,i_k}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

as follows:

Definition 1.2. For $1 \le k \le n$, D_k is defined as

$$D_k(\mathbb{R}^n) = \{u \in \mathcal{D}'(\mathbb{R}^n) : \|u\|_{D_k} < \infty\},$$

where

$$\|u\|_{D_k} := \sup\{|u(\phi_{i_1,\dots,i_k})| : \Phi \in \mathcal{D}(\mathbb{R}^n; \Lambda^k(\mathbb{R}^n)),\ d\Phi = 0,\ \|\Phi\|_{L^1} \le 1\}.$$

It was shown in [Van Schaftingen 2006] that the D_k classes lie strictly between the critical Sobolev spaces and $\mathrm{BMO}(\mathbb{R}^n)$, refining the classical embedding $\mathring{W}^{1,n} \subset$ BMO. More precisely, the following proper inclusions are continuous:

$$\mathring{W}^{1,n} \subset D_{n-1} \subset \cdots \subset D_1 \subset \mathrm{BMO}.$$

From the point of view of some applications to PDEs, as function spaces D_k ($k < n$) lack certain "useful" properties: multiplications by smooth cut-off functions are not necessarily bounded operators on D_k and D_k are not invariant under all smooth changes of variables.

In this paper, we introduce the nonhomogeneous analogs of Van Schaftingen's classes D_k, which we denote by $d^k(\mathbb{R}^n)$.

Definition 1.3. Let $1 \le k \le n$. We say that $u \in \mathcal{D}'(\mathbb{R}^n)$ belongs to $d^k(\mathbb{R}^n)$ if

$$\sup_{\|\Phi\|_{\Upsilon^1_k(\mathbb{R}^n)} \le 1} \max_I |u(\phi_I)| < \infty, \tag{1-1}$$

where the supremum is taken over all k-differential forms $\Phi = \sum_I \phi_I dx^I$, $\phi_I \in \mathcal{D}(\mathbb{R}^n)$ and $\|\Phi\|_{\Upsilon^1_k} = \|\Phi\|_{L^1} + \|d\Phi\|_{L^1}$. We will denote this supremum by $\|u\|_{d^k}$.

It is useful to compare the defined classes $d^k(\mathbb{R}^n)$ with $D_k(\mathbb{R}^n)$. First of all, $d^k(\mathbb{R}^n) \subset D_k(\mathbb{R}^n)$, $k = 1, 2, \dots, n$ as sets. As Banach spaces $D_k(\mathbb{R}^n)$ are classes of functions modulo constants, while in $d^k(\mathbb{R}^n)$ two functions that differ by a nonzero constant are considered as different elements.

In contrast to D_k spaces, the smooth change of variables and multiplications by cut-off functions are invariant operations on d^k. In particular, this allows us to define d^k on certain Riemannian manifolds. In Section 2, we recall some facts from the theory of local Hardy spaces, which will be used later. In Section 3, we prove the following theorem:

Theorem 1.4. *The space $d^1(\mathbb{R}^n)$ is continuously embedded into the space* $\mathrm{bmo}(\mathbb{R}^n)$ *and there exists $C > 0$ such that for any $u \in d^k(\mathbb{R}^n)$, $1 \le k \le n$,*

$$\|u\|_{\mathrm{bmo}} \le C\|u\|_{d^k}.$$

Combining this theorem with the result of Van Schaftingen [2004a] shows that the d^k classes refine the embedding $W^{1,n}(\mathbb{R}^n) \subset \mathrm{bmo}(\mathbb{R}^n)$, where bmo is the local BMO space of Goldberg [1979] in the sense that

$$W^{1,n} \subset d^{n-1} \subset \cdots \subset d^1 \subset \mathrm{bmo}.$$

We also prove that continuous d^{n-1} functions can be characterized in terms of line integrals, as is the case for the inequality of Bourgain, Brezis and Mironescu [Bourgain et al. 2004]

Theorem 1.5. *Let $u \in \mathcal{D}(\mathbb{R}^n)$. Then $u \in d^{n-1}(\mathbb{R}^n)$ if and only if*

$$\sup_{\partial\gamma=\varnothing} \frac{1}{|\gamma|}\left|\int_\gamma u(t)\tau(t)\,dt\right| + \sup_{|\gamma|\ge 1} \frac{1}{|\gamma|}\left|\int_\gamma u(t)\tau(t)\,dt\right| < \infty,$$

where the suprema are taken over smooth curves γ with finite lengths $|\gamma|$, boundaries $\partial\gamma$ and unit tangent vectors τ.

As an application of d^k classes for PDEs, the following fact is established:

Theorem 1.6. *Let $n \ge 2$, $F \in L^1(\mathbb{R}^n;\mathbb{R}^n)$ and $\operatorname{div} F \in L^1(\mathbb{R}^n)$. Then the system $(I-\Delta)U = F$ admits a unique solution U such that*:

- *If $n = 2$, then*

$$\|U\|_\infty + \|\nabla U\|_2 \le C(\|F\|_1 + \|\operatorname{div} F\|_1).$$

- *If $n \ge 3$, then*

$$\|U\|_{n/(n-2)} + \|\nabla U\|_{n/n-1} \le C(\|F\|_1 + \|\operatorname{div} F\|_1).$$

In Section 4, we introduce the localized versions of d^k spaces on bounded Lipschitz domains Ω. The main result of Section 4 is the proof of the following fact, which was conjectured by Van Schaftingen [2006] for the bmo spaces on domains (see Definition 2.13 below).

Theorem 1.7. *Any $u \in d^1(\Omega)$ is a* $\mathrm{bmo}_r(\Omega)$ *function as there exists $C > 0$ such that*

$$\|u\|_{\mathrm{bmo}_r(\Omega)} \le C\|u\|_{d^1(\Omega)} \quad \text{for all } u \in d^1(\Omega).$$

In Section 5, we define d^k classes on Riemannian manifolds with bounded geometry, and based on the results of Section 3 we prove the refined embeddings between critical Sobolev space and bmo on such manifolds.

Theorem 1.8. *Let M be the Riemannian manifold with bounded geometry. Then the following continuous embeddings hold*:

$$W^{1,n}(M) \subset d^{n-1}(M) \subset \cdots \subset d^1(M) \subset \mathrm{bmo}(M).$$

2. Preliminaries

Let $\Omega \subset \mathbb{R}^n$ be open. We will use the Schwartz notation: $\mathcal{E}(\Omega)$ will denote the class of smooth functions on Ω, $\mathcal{D}(\Omega)$ and $\mathcal{S}(\Omega)$ will stand for compactly supported smooth functions and smooth functions rapidly decaying at infinity with all their derivatives. By $\mathcal{D}^k(\Omega)$ we denote the class of k-differential forms with $\mathcal{D}(\Omega)$ components. All L^p spaces in this paper are considered relative to the Lebesgue measure. For the differential form of order k, $\Phi = \sum_{|I|=k} \phi_I dx^I$, we will use the notation

$$\|\Phi\|_{L^1_k} = \sum_I \|\phi_I\|_{L^1}.$$

However, often when it does not create confusion we will omit the subscript k and simply write $\|\Phi\|_{L^1}$ or $\|\Phi\|_1$.

2A. ***Local Hardy and* BMO *spaces of Goldberg.*** We recall the definition and basic properties of the local Hardy space $h^1(\mathbb{R}^n)$ introduced by Goldberg [1979].

Let us fix $\phi \in \mathcal{S}(\mathbb{R}^n)$ such that $\int \phi \neq 0$. For $f \in L^1(\mathbb{R}^n)$, we define the local maximal function $m_\phi f(x)$ by

$$m_\phi f(x) = \sup_{0<t<1} |\phi_t * f(x)|,$$

where $\phi_t(y) = t^{-n}\phi(y/t)$.

Definition 2.1. We say that f belongs to the local Hardy space $h^1(\mathbb{R}^n)$ if $m_\phi f$ lies in $L^1(\mathbb{R}^n)$; in this case we put

$$\|f\|_{h^1} := \|m_\phi f\|_{L^1}.$$

It is useful to compare h^1 with the classic real Hardy space $H^1(\mathbb{R}^n)$, which can be defined using the global maximal function M_ϕ,

$$M_\phi f(x) := \sup_{t>0} |\phi_t * f(x)|, \quad f \in L^1(\mathbb{R}^n).$$

Definition 2.2. We say that f belongs to the Hardy space $H^1(\mathbb{R}^n)$ if $M_\phi f \in L^1(\mathbb{R}^n)$, and we put

$$\|f\|_{H^1} := \|M_\phi f\|_{L^1}.$$

It follows from the definitions of the maximal functions that $m_\phi f(x) \leq M_\phi f(x)$ for any $f \in L^1$ and $x \in \mathbb{R}^n$. Therefore $H^1 \subset h^1$. One of the reasons it is often more convenient to deal with a larger space h^1 instead of H^1 is that $\mathcal{S}(\mathbb{R}^n) \subset h^1(\mathbb{R}^n)$, while any $f \in H^1(\mathbb{R}^n)$ has to satisfy $\int_{\mathbb{R}^n} f = 0$. Moreover:

Lemma 2.3 [Goldberg 1979]. *The space $\mathcal{D}(\mathbb{R}^n)$ is dense in $h^1(\mathbb{R}^n)$.*

It is important to note that $f \in h^1(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} f = 0$ do not imply that $f \in H^1(\mathbb{R}^n)$ (see Theorem 3 in [Goldberg 1979]). However, the following is true:

Lemma 2.4. *If* $f \in h^1(\mathbb{R}^n)$, $\int_{\mathbb{R}^n} f(x)\,dx = 0$ *and* $\operatorname{supp} f \subset B$, *where* B *is a bounded subset of* $\mathbb{R}^n$, *then there exists* $C_B > 0$ *such that*

$$\|f\|_{H^1} \le C_B \|f\|_{h^1}.$$

Definition 2.5 [Goldberg 1979]. We say that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ belongs to $\mathrm{bmo}(\mathbb{R}^n)$ if

$$\|f\|_{\mathrm{bmo}} := \sup_{l(Q)\le 1} \frac{1}{|Q|}\int_Q |f(x) - f_Q|\,dx + \sup_{l(Q)\ge 1} \frac{1}{|Q|}\int_Q |f(x)|\,dx < \infty,$$

where $f_Q = 1/|Q| \int_Q f(y)\,dy$ and Q are cubes with sides parallel to the axes, of side-length $l(Q)$.

It is clear that $\mathrm{bmo}(\mathbb{R}^n)$ is a subspace of $\mathrm{BMO}(\mathbb{R}^n)$. Moreover, if $\|f\|_{\mathrm{bmo}} = 0$, then $f = 0$ almost everywhere on $\mathbb{R}^n$, unlike in $\mathrm{BMO}(\mathbb{R}^n)$, where constant functions are identified with $f \equiv 0$.

The following theorem of Goldberg shows the relation between h^1 and bmo and the boundedness of pseudodifferential operators of degree zero on h^1.

Theorem 2.6 [Goldberg 1979]. *The space* $\mathrm{bmo}(\mathbb{R}^n)$ *is isomorphic to the space of continuous linear functionals on* $h^1(\mathbb{R}^n)$.

Theorem 2.7 [Goldberg 1979]. *If* $T \in \mathrm{OPS}^0$, *then there exists a constant* $C > 0$ *such that*

$$\|Tf\|_{h^1} \le C\|f\|_{h^1} \text{ for any } f \in \mathcal{S}(\mathbb{R}^n).$$

Therefore, any $T \in \mathrm{OPS}^0$ *can be extended to a continuous linear operator on* $h^1(\mathbb{R}^n)$.

2B. ***Local Hardy and* BMO *spaces on Lipschitz domains.*** The BMO and Hardy spaces on bounded Lipschitz domains were studied in [Chang et al. 1993; 1999; Miyachi 1990] (see also [Jones 1980; Strichartz 1972]).

Definition 2.8 [Chang et al. 1993; Miyachi 1990]. Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. The space $h^1_r(\Omega)$ consists of elements of $L^1(\Omega)$ which are the restrictions to Ω of elements of $h^1(\mathbb{R}^n)$, i.e.,

$$h^1_r(\Omega) = \{f \in L^1(\Omega) : \text{ there exists } F \in h^1(\mathbb{R}^n) : F = f \text{ on } \Omega\}.$$

We can consider this as a quotient space equipped with the quotient norm

$$\|f\|_{h^1_r(\Omega)} := \inf\{\|F\|_{h^1(\mathbb{R}^n)} : F = f \text{ on } \Omega\}.$$

Definition 2.9 [Chang et al. 1999]. The space $h^1_z(\Omega)$ is defined to be the subspace of $h^1(\mathbb{R}^n)$ consisting of those elements which are supported on $\overline{\Omega}$.

Like in the case of $\mathbb{R}^n$, smooth and compactly supported functions are dense in these spaces:

Lemma 2.10 [Caetano 2000]. *Let Ω be a bounded Lipschitz domain. Then the space $\mathcal{D}(\Omega)$ is dense in $h^1_r(\Omega)$.*

Lemma 2.11. *Let Ω be a domain of $\mathbb{R}^n$. Then the set of $\mathcal{D}(\Omega)$ functions is dense in $h^1_z(\Omega)$.*

The BMO analogs on Ω are defined as follows.

Definition 2.12. The space $\mathrm{bmo}_z(\Omega)$ is defined to be a subspace of $\mathrm{bmo}(\mathbb{R}^n)$ consisting of those elements which are supported on $\overline{\Omega}$, i.e.,

$$\mathrm{bmo}_z(\Omega) = \{g \in \mathrm{bmo}(\mathbb{R}^n) : g = 0 \text{ on } \mathbb{R}^n \setminus \overline{\Omega}\}$$

with

$$\|g\|_{\mathrm{bmo}_z(\Omega)} = \|g\|_{\mathrm{bmo}(\mathbb{R}^n)}.$$

Definition 2.13 [Chang et al. 1999]. Let Ω be a bounded Lipschitz domain. A function $g \in L^1_{\mathrm{loc}}(\Omega)$ is said to belong to $\mathrm{bmo}_r(\Omega)$ if

$$\|g\|_{\mathrm{bmo}_r(\Omega)} = \sup_{|Q|\le 1} \frac{1}{|Q|}\int_Q |g(x) - g_Q|\,dx + \sup_{|Q|>1} \frac{1}{|Q|}\int_Q |g(x)|\,dx < \infty,$$

where suprema are taken over all cubes $Q \subset \Omega$. The space of such functions equipped with norm $\|\cdot\|_{\mathrm{bmo}_r(\Omega)}$ is called $\mathrm{bmo}_r(\Omega)$.

Theorem 2.14 [Chang 1994; Miyachi 1990]. *The space* $\mathrm{bmo}_z(\Omega)$ *is isomorphic to the dual of $h^1_r(\Omega)$.*

Theorem 2.15 [Chang 1994; Jonsson et al. 1984]. *The space* $\mathrm{bmo}_r(\Omega)$ *is isomorphic to the dual of $h^1_z(\Omega)$.*

2C. *$H^1_z(\Omega)$ space.* We will also need the following function space:

Definition 2.16 [Chang et al. 1993]. The space $H^1_z(\Omega)$ is defined to be the subspace of $H^1(\mathbb{R}^n)$ consisting of those elements which are supported on $\overline{\Omega}$.

One of the alternative ways to define $H^1_z(\Omega)$ is to invoke the notion of atoms.

Definition 2.17. An $H^1_z(\Omega)$ atom is a Lebesgue measurable function a on $\mathbb{R}^n$, supported on a cube $Q \subset \Omega$, such that

$$\|a\|_{L^2(Q)} \le |Q|^{-1/2}$$

and

$$\int_Q a(x)\,dx = 0.$$

Any $H^1_z(\Omega)$ function can be represented as a series of H^1_z atoms in the following sense:

Theorem 2.18 [Chang et al. 1993, Theorem 3.3]. *Let Ω be a bounded Lipschitz domain and $f \in L^1(\Omega)$. Then $f \in H^1_z(\Omega)$ if and only if there exist a sequence of $H^1_z(\Omega)$ atoms $\{a_k\}$ and real numbers $\{\lambda_k\} \subset \mathbb{R}$ such that $\sum |\lambda_k| < \infty$ and*

$$\sum_k \lambda_k a_k \to f \text{ in } \mathcal{D}'(\Omega).$$

Furthermore,

$$\|f\|_{H^1} \approx \inf\left\{\sum_k |\lambda_k| : f = \sum_k \lambda_k a_k\right\},$$

where the infimum is taken over all atomic decompositions of f.

3. d^k spaces on $\mathbb{R}^n$

Definition 3.1. Let $1 \le k \le n$. We say that $u \in \mathcal{D}'(\mathbb{R}^n)$ belongs to $d^k(\mathbb{R}^n)$ if

$$\sup_{\|\Phi\|_{\Upsilon^1_k(\mathbb{R}^n)} \le 1} \max_{|I|=k} |u(\phi_I)| < \infty, \tag{3-1}$$

where the supremum is taken over all $\Phi = \sum_{|I|=k} \phi_I dx^I \in \mathcal{D}^k(\mathbb{R}^n)$ and $\|\Phi\|_{\Upsilon^1_k} = \|\Phi\|_{L^1_k} + \|d\Phi\|_{L^1_{k+1}}$. We will denote this supremum by $\|u\|_{d^k}$.

Remark 3.2. It is not difficult to show that the class of compactly supported $\Upsilon^1_k(\mathbb{R}^n)$ forms is dense in $\Upsilon^1_k(\mathbb{R}^n)$. This suggests that the domain of $u \in d^k(\mathbb{R}^n)$ can be extended to include components of all $\Upsilon^1_k(\mathbb{R}^n)$ forms. Let $u \in \mathcal{D}'(\Omega)$ and $\tilde{u}$ be a linear map from $\mathcal{D}^k(\Omega)$ to

$$\big(\mathbb{R}^{\binom{n}{k}}, \|\cdot\|_{\max}\big),$$

associated to u by

$$\tilde{u}\bigg(\sum_{|I|=k} \phi_I dx^I\bigg) = (u(\phi_I)).$$

Then $u \in \mathcal{D}'(\mathbb{R}^n)$ belongs to $d^k(\mathbb{R}^n)$, if and only if $\tilde{u}$ can be extended to a bounded linear map from $\Upsilon^1_k(\mathbb{R}^n)$ to $\binom{n}{k}$-dimensional Euclidean space equipped with the max norm.

Note that $\Upsilon^1_n(\mathbb{R}^n) = L^1(\mathbb{R}^n)$, so $d^n(\mathbb{R}^n)$ is isomorphic to $L^\infty(\mathbb{R}^n)$.

Lemma 3.3. *Let $1 \le k < l \le n$ and $u \in d^l(\mathbb{R}^n)$. Then $u \in d^k(\mathbb{R}^n)$ and $\|u\|_{d^k(\mathbb{R}^n)} \le \|u\|_{d^l(\mathbb{R}^n)}$. In other words, the following embeddings are continuous*:

$$d^n(\mathbb{R}^n) \subset d^{n-1}(\mathbb{R}^n) \subset \cdots \subset d^1(\mathbb{R}^n).$$

Proof. It is enough to consider the case $k = l - 1$, because the general case will follow from it by induction. Let $1 \le l \le n,\ u \in d^l(\mathbb{R}^n)$ and

$$\Phi(x) = \sum_{|I|=l-1} \phi_I(x) dx^I \in \mathcal{D}^{l-1}(\mathbb{R}^n).$$

We need to show that for any component ϕ_I,

$$|u(\phi_I)| \le \|u\|_{d^l} \|\Phi\|_{\Upsilon^1_{l-1}}.$$

Fix any such I. Since $|I| = l-1 < n$, there exists $j \in [1, n]$ such that $dx^I \wedge dx^j \neq 0$. Put $\tilde{\Phi}(x) = \Phi(x) \wedge dx^j$. Then $\tilde{\Phi} \in \mathcal{D}^l$ and $\| \tilde{\Phi} \|_{\Upsilon^1_l} \le \| \Phi \|_{\Upsilon^1_{l-1}}$. Moreover, by construction, one of the components of $\tilde{\Phi}$ equals $\pm\phi_I dx^I \wedge dx^j$. Since $u \in d^l(\mathbb{R}^n)$,

$$|u(\phi_I)| \le \|u\|_{d^l} \|\tilde{\Phi}\|_{\Upsilon^1_l} \le \|u\|_{d^l} \|\Phi\|_{\Upsilon^1_{l-1}}. \qquad \square$$

The following theorem follows immediately from the definition of d^k spaces and the result of Van Schaftingen [2004a].

Theorem 3.4. *$W^{1,n}(\mathbb{R}^n)$ is continuously embedded into $d^{n-1}(\mathbb{R}^n)$ as there exists $C > 0$ such that for any $u \in W^{1,n}$,*

$$\|u\|_{d^{n-1}} \le C \|u\|_{W^{1,n}}.$$

One of main results in this section is the following:

Theorem 3.5. *The space $d^1(\mathbb{R}^n)$ is continuously embedded into the space* $\mathrm{bmo}(\mathbb{R}^n)$ *and there exists $C > 0$ such that for any $u \in d^k(\mathbb{R}^n),\ 1 \le k \le n$,*

$$\|u\|_{\mathrm{bmo}} \le C \|u\|_{d^k}.$$

Remark 3.6. This result is a nonhomogeneous analog of the main theorem in [Van Schaftingen 2006]. We adapt the proof of that theorem to the nonhomogeneous setting.

Proof. By Lemma 3.3, it is enough to prove the case $k = 1$. The argument is based on the fact that $\mathrm{bmo}(\mathbb{R}^n)$ is the dual space of $h^1(\mathbb{R}^n)$. We claim that given $f \in \mathcal{D}(\mathbb{R}^n)$, there exist n differential forms $\{\Phi^j\}_{j=1}^n \subset \Upsilon^1_1(\mathbb{R}^n)$ such that for some C independent of f,

$$\|\Phi^j\|_{\Upsilon^1_1} \le C \|f\|_{h^1}, \tag{3-2}$$

$$f = \sum_{i=1}^n \phi^i_i, \tag{3-3}$$

where

$$\Phi^j = \sum_{i=1}^n \phi^j_i dx^i.$$

Assuming the claim, the proof is easy. Let $u \in d^1(\mathbb{R}^n)$. For arbitrary $f \in \mathcal{D}(\mathbb{R}^n)$, let Φ^j be such that (3-2) and (3-3) are true. Then by the Remark 3.2 we can apply u to ϕ_i^i to have

$$|u(f)| \le \sum_{i=1}^{n} |u(\phi_i^i)| \le \sum_{i=1}^{n} \|u\|_{d^1} \|\Phi^i\|_{\Upsilon_1^1} \le Cn\|u\|_{d^1}\|f\|_{h^1}. \tag{3-4}$$

By the density of $\mathcal{D}$ in h^1 and the duality $\mathrm{bmo} = (h^1)'$, we conclude that $u \in \mathrm{bmo}(\mathbb{R}^n)$.

In order to prove the claim, let $f \in \mathcal{D}$ be arbitrary and consider the equation

$$(I - \Delta)v = f \quad \text{in } \mathbb{R}^n.$$

Then $v = \mathcal{J}(f)$, where $\mathcal{J}$ is a convolution operator whose kernel is the Bessel potential of order 2, G_2. For $j \in [1, n]$, let

$$\Phi^j = \sum_{i=1}^{n} \left(\frac{\mathcal{J}}{n} - \partial_i \partial_j \mathcal{J} \right)(f) dx^i.$$

Since $f \in \mathcal{D} \subset \mathcal{S}$, all components of Φ^j are $\mathcal{S}$ functions and

$$d\Phi^j = \sum_{1 \le i < k \le n} \left(\frac{\partial_i \mathcal{J} - \partial_k \mathcal{J}}{n} \right)(f)\, dx^i \wedge dx^k.$$

It is clear that,

$$\frac{\mathcal{J}}{n} - \partial_i \partial_j \mathcal{J} \in \mathrm{OPS}^{-2}(\mathbb{R}^n) + \mathrm{OPS}^0(\mathbb{R}^n) \subset \mathrm{OPS}^0(\mathbb{R}^n)$$

and

$$\left(\frac{\partial_i \mathcal{J} - \partial_k \mathcal{J}}{n} \right) \in \mathrm{OPS}^{-1}(\mathbb{R}^n) \subset \mathrm{OPS}^0(\mathbb{R}^n).$$

Recalling Theorem 2.7, we see that the components of Φ^j and $d\Phi^j$ are h^1 functions and for some C independent of f,

$$\|\Phi^j\|_{L_1^1} + \|d\Phi^j\|_{L_2^1} \le C\|f\|_{h^1},$$

which proves (3-2). Finally, $\{\Phi^j\}$ satisfy (3-3) for

$$\sum_{i=1}^{n} \left(\frac{\mathcal{J}}{n} - \partial_i \partial_i \mathcal{J} \right) f = \mathcal{J}(f) - \Delta \mathcal{J}(f) = (I - \Delta)\mathcal{J}(f) = f. \qquad \square$$

Corollary 3.7. *For $1 \le k \le n$, the space $d^k(\mathbb{R}^n)$ equipped with the norm $\|\cdot\|_{d^k}$ is a Banach space.*

Proof. Let $\{u_m\}_{m=0}^{\infty}$ be a Cauchy sequence in d^k. Theorem 3.5 shows that u_m is a Cauchy sequence in $\mathrm{bmo}(\mathbb{R}^n)$. Since bmo is a complete Banach space, there exists $u \in \mathrm{bmo}(\mathbb{R}^n)$, such that $u_m \to u$ in $\|\cdot\|_{\mathrm{bmo}}$. Moreover, for any $\Phi = \sum_{|I|=k} \phi_I dx^I \in \mathcal{D}^k(\mathbb{R}^n)$ and $j \geq 0$, using the duality of bmo and h^1 and the fact that each ϕ_I is in $\mathcal{D} \subset h^1$,

$$\left| \int (u_j - u)\phi_I \right| = \lim_{m\to\infty} \left| \int (u_j - u_m)\phi_I \right| \leq \lim_{m\to\infty} \|u_j - u_m\|_{d^k} \|\Phi\|_{\Upsilon^1_k},$$

which shows that $u \in d^k(\mathbb{R}^n)$, and $\|u_j - u\|_{d^k} \to 0$, as $j \to \infty$. □

Summing up the results of this section, we can now say that for $1 \leq k \leq n$,

$$W^{1,n}(\mathbb{R}^n) \subset d^{n-1}(\mathbb{R}^n) \subset \cdots \subset d^1(\mathbb{R}^n) \subset \mathrm{bmo}(\mathbb{R}^n).$$

3A. ***v^k classes.***

Definition 3.8. Let $1 \leq k \leq n$. We define the class $v^k(\mathbb{R}^n)$ as the closure of $C_0(\mathbb{R}^n)$ functions in the norm $\|\cdot\|_{d^k}$. Here,

$$C_0(\mathbb{R}^n) = \left\{ u :\in C(\mathbb{R}^n) : \lim_{|x|\to\infty} u(x) = 0 \right\}.$$

First of all we notice that by Lemma 3.3, the $v^k(\mathbb{R}^n)$ form a monotone family of spaces

$$v^n(\mathbb{R}^n) \subset v^{n-1}(\mathbb{R}^n) \subset \cdots \subset v^1(\mathbb{R}^n).$$

The appropriate subspace that will contain all v^k functions was studied by Dafni [2002] and Bourdaud [2002].

Definition 3.9 [Dafni 2002]. The space $\mathrm{vmo}(\mathbb{R}^n)$ is the subspace of $\mathrm{bmo}(\mathbb{R}^n)$ functions satisfying

$$\lim_{\delta\to 0} \sup_{l(Q)\leq\delta} \frac{1}{|Q|} \int_Q |f(x) - f_Q|\, dx = 0 \tag{3-5}$$

and

$$\lim_{R\to\infty} \sup_{\substack{l(Q)>1,\\ Q\cap B(0,R)=\varnothing}} \frac{1}{|Q|} \int_Q |f(x)|\, dx = 0. \tag{3-6}$$

Theorem 3.10 [Dafni 2002]. *The space* $\mathrm{vmo}(\mathbb{R}^n)$ *is the closure of* $C_0(\mathbb{R}^n)$ *in* $\mathrm{bmo}(\mathbb{R}^n)$.

An immediate consequence of this result and Theorem 3.5 is:

Theorem 3.11. *For* $1 \leq k \leq n$, *the space* $v^k(\mathbb{R}^n)$ *is embedded into* $\mathrm{vmo}(\mathbb{R}^n)$.

Corollary 3.12. *The space* $v^1(\mathbb{R}^n)$ *does not contain* $d^n(\mathbb{R}^n)$ *as a subspace. In particular, the* $v^k(\mathbb{R}^n)$ *are proper subspaces of* $d^k(\mathbb{R}^n)$ *for* $k = 1, \ldots, n$.

Proof. Recall that $d^n(\mathbb{R}^n)$ coincides with $L^\infty(\mathbb{R}^n)$. If L^∞ was a subspace of $v^1(\mathbb{R}^n)$, then by Theorem 3.11 we would have $L^\infty \subset \mathrm{vmo}(\mathbb{R}^n)$. However, choosing f as a characteristic function of the quadrant $\{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_i > 0\}$, we have an example of an L^∞ function that does not satisfy (3-5). So $L^\infty \not\subset \mathrm{vmo}(\mathbb{R}^n)$. □

Finally, we recall that $\mathcal{D}(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$ for any $p \in [1, \infty)$. Therefore by Theorem 3.4, we have $W^{1,n} \subset v^{n-1}(\mathbb{R}^n)$.

All in all, we conclude that the following embeddings hold:

$$W^{1,n}(\mathbb{R}^n) \subset v^{n-1}(\mathbb{R}^n) \subset \cdots \subset v^1(\mathbb{R}^n) \subset \mathrm{vmo}(\mathbb{R}^n).$$

3B. ***Intrinsic definition of the space v^{n-1}.***

Definition 3.13. For $u \in d^{n-1}(\mathbb{R}^n) \cap C(\mathbb{R}^n)$, we will use the notation

$$\|u\|_* = \sup_{\partial\gamma=\varnothing} \frac{1}{|\gamma|}\left|\int_\gamma u(t)\tau(t)\,dt\right| + \sup_{|\gamma|\ge 1} \frac{1}{|\gamma|}\left|\int_\gamma u(t)\tau(t)\,dt\right|,$$

where the suprema are taken over smooth curves γ with finite lengths $|\gamma|$ and boundaries $\partial\gamma$, and τ is the unit tangent vector to the curve γ.

Our goal is to prove the following result.

Theorem 3.14. *There are constants c_1 and c_2 great than 0 such that for every $u \in d^{n-1}(\mathbb{R}^n) \cap C(\mathbb{R}^n)$,*

$$c_1\|u\|_* \le \|u\|_{d^{n-1}} \le c_2\|u\|_*.$$

The proof is based on the following three lemmas:

Lemma 3.15. *There exists $C > 0$ such that for any γ with $\partial\gamma = \varnothing$ or $|\gamma| \ge 1$,*

$$\frac{1}{|\gamma|}\left|\int_\gamma u(y)\tau(y)\,dy\right| \le C\|u\|_{d^{n-1}}.$$

Proof. The proof is based on the argument of Bourgain and Brezis [2004].

Let $\eta \ge 0$ be a smooth radial function on $\mathbb{R}^n$, compactly supported in $|x| \le 1$, such that $\|\eta\|_{L^1} = 1$. As usual we put $\eta_\epsilon(x) = \epsilon^{-n}\eta(x/\epsilon)$. Let us define the $(n-1)$-form

$$\Phi^\epsilon(x) = \sum_{j=1}^{n}\left(\int_\gamma \eta_\epsilon(t-x)\tau_j(t)\,dt\right)dx^{I_j}, \quad x \in \mathbb{R}^n,$$

where $I_j = (i_1, \dots, i_{n-1}),\ i_k \ne j$.

The reason to introduce this differential form is the equality

$$\left|\int_\gamma u(t)\tau(t)\,dt\right| = \lim_{\epsilon\to 0}\left|\int_\gamma \tau(t)\int_{\mathbb{R}^n} u(x)\eta_\epsilon(x-t)\,dx\,dt\right| = \lim_{\epsilon\to 0}\left|\int u(x)\Phi^\epsilon(x)\,dx\right|.$$

By Remark 3.2, we need to estimate $\|\Phi^\epsilon\|_{\Upsilon^1_{n-1}}$. It is clear that

$$\|\Phi^\epsilon\|_{L^1_{n-1}} \leq n\|\eta_\epsilon\|_{L^1}|\gamma| = n|\gamma|.$$

Moreover,

$$\begin{aligned} d\Phi^\epsilon(x) &= -\left(\int_\gamma \nabla\eta_\epsilon(y-x)\cdot\tau(y)\,dy\right)dx^1\wedge\cdots\wedge dx^n \\ &= [\eta_\epsilon(a-x)-\eta_\epsilon(b-x)]dx^1\wedge\cdots\wedge dx^n. \end{aligned}$$

Therefore $\|d\Phi^\epsilon\|_{L^1_n}$ is 0 if γ is closed or ≤ 2 if γ is not closed. Finally,

$$\frac{1}{|\gamma|}\left|\int_\gamma u(s)\tau(s)\,ds\right| \leq \frac{1}{|\gamma|}\limsup_{\epsilon\to 0}\left|\int u(x)\Phi^\epsilon\,dx\right| \leq \|u\|_{d^k}(2+n),$$

because $|\gamma|\geq 1$ for nonclosed γ. So we have proved the lemma with $C=n+2$. □

In order to prove the converse estimate, Bourgain and Brezis evoked the decomposition theorem of Smirnov.

Theorem 3.16 [Smirnov 1993]. *For any compactly supported $\Phi\in L^1_{n-1}(\mathbb{R}^n)$, with $d\Phi=0$, there exists a sequence of positive numbers $\{\mu^m_j\}$ and closed smooth curves $\{\gamma^m_j\}$ such that for all $m\geq 1$,*

$$\sum_{j=1}^\infty |\mu^m_j|\,|\gamma^m_j| \leq \|\Phi\|_{L^1_{n-1}},$$

and for every $u\in C(\mathbb{R}^n)$ and $1\leq i\leq n$,

$$\sum_{j=1}^\infty \mu^m_j\int_{\gamma^m_j} u(s)\tau_i(s)\,ds \to \int u(x)\phi_i(x)\,dx,\ \text{as } m\to\infty,$$

where ϕ_i are the components of Φ.

In our case $d\Phi\in L^1_{n-1}(\mathbb{R}^n)$ does not necessarily vanish and we need a more general version of Smirnov's theorem, which we formulate in the following form:

Theorem 3.17 [Smirnov 1993]. *Let $\Phi\in\Upsilon^1_{n-1}(\mathbb{R}^n)$. Then there exist $P\in\Upsilon^1_{n-1}(\mathbb{R}^n)$ and $Q\in\Upsilon^1_{n-1}(\mathbb{R}^n)$ such that*

- $\|\Phi\|_{L^1_{n-1}} = \|P\|_{L^1_{n-1}} + \|Q\|_{L^1_{n-1}}$,
- *$dP=0$ and we can apply the previous theorem to P,*
- *$dQ=d\Phi$.*

Moreover, there exist $\{\lambda^l_j\}$ *and smooth curves* $\tilde{\gamma}^l_j$ *(not necessarily closed) such that for all* $l \geq 1$,

$$\sum_{j=1}^{\infty} |\lambda^l_j||\tilde{\gamma}^l_j| \leq \|Q\|_{L^1_{n-1}},$$

$$\sum_{j=1}^{\infty} |\lambda^l_j| \leq \|dQ\|_{L^1_n}$$

and for $1 \leq i \leq n$,

$$\sum_{j=1}^{\infty} \lambda^l_j \int_{\tilde{\gamma}^l_j} u(s)\tau_i(s)\,ds \to \int u(x)q_i(x)\,dx, \text{ as } l \to \infty,$$

where q_i *are the components of* Q.

Let us introduce an auxiliary norm for $u \in C(\mathbb{R}^n)$:

$$\|u\|_{**} = \sup_{\partial\gamma=\varnothing} \frac{1}{|\gamma|}\left|\int_\gamma u(s)\tau(s)\,ds\right| + \sup_{|\gamma|<1}\left|\int_\gamma u(s)\tau(s)\,ds\right| + \sup_{|\gamma|\geq 1} \frac{1}{|\gamma|}\left|\int_\gamma u(s)\tau(s)\,ds\right|.$$

Lemma 3.18. *For any* $u \in d^{n-1}(\mathbb{R}^n) \cap C(\mathbb{R}^n)$,

$$\|u\|_{d^{n-1}(\mathbb{R}^n)} \leq 2\|u\|_{**}.$$

Proof. By the definition of $d^{n-1}(\mathbb{R}^n)$, there exists

$$\Phi = \sum_{i=1}^{n} \phi_i dx^1 \wedge \cdots \widehat{dx}^i \wedge \cdots dx^n \in \mathcal{D}^{n-1}(\mathbb{R}^n)$$

such that

$$\|\Phi\|_{L^1_{n-1}} + \|d\Phi\|_{L^1} \leq 1$$

and

$$\|u\|_{d^{n-1}} \leq 2 \max_I |u(\phi_I)|. \tag{3-7}$$

Let us apply Theorem 3.17 to Φ. Then Φ can be decomposed into the sum of P and Q such that $d\Phi = dQ$, $\|\Phi\|_{L^1_{n-1}} = \|P\|_{L^1_{n-1}} + \|Q\|_{L^1_{n-1}}$ and Q is a weak limit of the linear combination of the curves $\tilde{\gamma}^l_j$ in the sense that

$$\sum_{j=1}^{\infty} \tilde{\lambda}^l_j \int_{\tilde{\gamma}^l_j} u(s)\tau_i(s)\,ds \to \int u(x)q_i(x)\,dx, \quad \text{as } l \to \infty,$$

where

$$\sum_{j=1}^{\infty} |\tilde{\lambda}^l_j|(1 + |\tilde{\gamma}^l_j|) \leq \|Q\|_{L^1_{n-1}} + \|dQ\|_{L^1} \leq 1, \quad \text{for all } l \geq 1.$$

Moreover, applying Theorem 3.16 to P, we get a sequence of closed curves γ_j^l and numbers λ_j^l such that

$$\sum_{j=1}^{\infty} \lambda_j^l \int_{\gamma_j^l} u(s)\tau_i(s)\,ds \to \int u(x)p_i(x)\,dx, \qquad \text{as } l \to \infty,$$

and

$$\sum_{j=1}^{\infty} |\lambda_j^l||\gamma_j^l| \le \|P\|_{L^1_{n-1}} \le 1, \qquad \text{for all } l \ge 1.$$

All in all,

$$\int u(x)\phi_i(x)\,dx = \lim_{l\to\infty} \sum_{j=1}^{\infty} \lambda_j^l \int_{\gamma_j^l} u(s)\tau_i(s)\,ds + \sum_{j=1}^{\infty} \tilde{\lambda}_j^l \int_{\tilde{\gamma}_j^l} u(s)\tau_i(s)\,ds$$

and

$$\left|\int u(x)\phi_i(x)\,dx\right| \le \sup_{l,j}\left|\frac{1}{|\gamma_j^l|}\int_{\gamma_j^l} u(s)\tau_i(s)\,ds\right| + \sup_{l,|\tilde{\gamma}_j^l|<1}\left|\int_{\tilde{\gamma}_j^l} u(s)\tau_i(s)\,ds\right| + \sup_{l,|\tilde{\gamma}_j^l|\ge 1}\left|\frac{1}{|\tilde{\gamma}_j^l|}\int_{\tilde{\gamma}_j^l} u(s)\tau_i(s)\,ds\right| \le \|u\|_{**}. \tag{3-8}$$

The result follows from (3-7) and (3-8). □

Lemma 3.19. *For any $u \in C(\mathbb{R}^n)$,*

$$\|u\|_* \le \|u\|_{**} \le 4\|u\|_*.$$

Proof. The first inequality follows from the definitions of the norms. In order to see the second one, we need to show

$$\sup_{|\gamma|<1}\left|\int_\gamma u(s)\tau(s)\,ds\right| \le \sup_{\partial\gamma=\varnothing} \frac{3}{|\gamma|}\left|\int_\gamma u(s)\tau(s)\,ds\right| + \sup_{|\gamma|\ge 1} \frac{3}{|\gamma|}\left|\int_\gamma u(s)\tau(s)\,ds\right|.$$

Let us consider any γ with $|\gamma| < 1$ and $\partial\gamma = \{a, b\}$. We can always find γ' such that $1 < |\gamma'| < 2$ and $\gamma'' := \gamma + \gamma'$ is a closed curve.

Then

$$\begin{aligned}\left|\int_\gamma u(s)\tau(s)\,ds\right| &\le \left|\int_{\gamma''} u(s)\tau(s)\,ds\right| + \left|\int_{\gamma'} u(s)\tau(s)\,ds\right| \\ &\le \left|\frac{3}{|\gamma''|}\int_{\gamma''} u(s)\tau(s)\,ds\right| + \left|\frac{3}{|\gamma'|}\int_{\gamma'} u(s)\tau(s)\,ds\right|.\end{aligned}$$ □

3C. *Examples of $d^k(\mathbb{R}^n)$ functions.* In this section, we want to show that there are more functions in $d^k(\mathbb{R}^n)$ besides those in $W^{1,n}(\mathbb{R}^n)$.

3C1. *Triebel–Lizorkin and Besov functions.* We recall that the Sobolev space $W^{s,p}(\mathbb{R}^n),\ 1 < p < \infty$, is a special case of more general classes of functions

$$W^{s,p}(\mathbb{R}^n) = F_p^{s,p}(\mathbb{R}^n) = B_p^{s,p}(\mathbb{R}^n),$$

where $F_q^{s,p},\ s \in \mathbb{R},\ 0 < p, q < \infty$, is the Triebel–Lizorkin space and $B_q^{s,p}(\mathbb{R}^n)$, $s \in \mathbb{R},\ 0 < p, q \le \infty$, is the Besov space (see, e.g., [Grafakos 2009] or [Triebel 1992] for definitions).

It was shown in [Van Schaftingen 2010, Proposition 2.1] that $\mathring{F}_q^{s,p} \subset D_{n-1}$ for all $sp = n,\ 1 < p < \infty,\ 0 < q < \infty$ (here $\mathring{F}_q^{s,p}$ is a homogeneous Triebel–Lizorkin space). Recalling the embedding theorems (see, e.g., [Grafakos 2009, Example 6.5.2])

$$\mathring{B}^{s,p}_{\min(p,q)} \subset \mathring{F}_q^{s,p} \subset \mathring{B}^{s,p}_{\max(p,q)},$$

and

$$\mathring{B}_q^{s,p} \subset \mathring{B}_q^{s',p'},\ \text{ if } sp = s'p' \text{ and } s > s',$$

one can obtain the embedding $\mathring{B}_q^{s,p} \subset D_{n-1}$ for $0 < q < \infty$. The case $q = \infty$ remains open (see [Van Schaftingen 2014, Open problem 1]).

One can notice that the proof of Proposition 2.1 in [Van Schaftingen 2010] is exactly the same as the proof of Theorem 1.5 in [Van Schaftingen 2004b]. In fact it can be extended to the nonhomogeneous setting as follows:

Theorem 3.20. *Let $1 < p < \infty,\ 1 < q < \infty$. Then there exists constants C_1 and C_2 such that*

$$\|u\|_{d^{n-1}} \le C_1 \|u\|_{F_q^{n/p,p}}$$

and

$$\|u\|_{d^{n-1}} \le C_2 \|u\|_{B_q^{n/p,p}}.$$

3C2. *Locally Lipschitz functions.* The following proposition provides a simple sufficient condition to ensure that $u \in d^{n-1}(\mathbb{R}^n)$.

Proposition 3.21. *Let $u \in W^{1,1}_{loc}(\mathbb{R}^n \setminus \{0\})$. If $|x|(u(x) + \nabla u(x)) \in L^\infty(\mathbb{R}^n)$, then $u \in d^{n-1}(\mathbb{R}^n)$ and*

$$\|u\|_{d^{n-1}} \le C \| \, |x|(|u| + |\nabla u|)\|_{L^\infty}.$$

Proof. The proof follows from integration by parts as in the proof of Proposition 4.3 in [Van Schaftingen 2006].

We need to show that for any $\Phi = \sum_{j=1}^n \phi_j(x) dx^1 \wedge \cdots \widehat{dx}^j \wedge \cdots dx^n \in \mathcal{D}^{n-1}(\mathbb{R}^n)$, we have

$$\left| \int u(x)\phi_j(x)\, dx \right| \le C \| \, |x|(u(x) + \nabla u(x))\|_{L^\infty} \|\Phi\|_{\Upsilon^1_{n-1}}.$$

Note that

$$\int x_j\Big(\sum_i \phi_i\partial_i u\Big)\,dx = -\int \phi_j u\,dx - \int x_j u\cdot\Big(\sum_i \partial_i\phi_i\Big)\,dx.$$

So

$$\left|\int u(x)\phi_j(x)\,dx\right| \le n\|\,|x|\nabla u\|_{L^\infty}\|\Phi\|_{L^1_{n-1}} + \|\,|x|u\|_{L^\infty}\|d\Phi\|_{L^1_n}. \qquad \square$$

The proposition allows us to give an example of $u \in d^{n-1}$ which is not covered by the previous classes of functions, the Bessel potential G_n.

Remark 3.22. A typical example of $u \in D^{n-1}\setminus W^{1,n}$ in [Van Schaftingen 2006] is the function $u(x) = \log|x|$. However, this function does not belong to $\mathrm{bmo}(\mathbb{R}^n)$ and therefore is not in any d^k, $1 \le k \le n$, as

$$\sup_{|Q|>1}\frac{1}{|Q|}\int_Q |\log|y||\,dy = \infty.$$

Example 3.23. Let $G_n(x)$ be the Bessel potential of order n, i.e., the function whose Fourier transform is given by $\widehat{G}_n(\xi) = (1+|\xi|^2)^{-n/2}$.

The fact that G_n satisfies the conditions of the last proposition follows from the fact that G_n is a continuously differentiable function on $\mathbb{R}^n\setminus\{0\}$ and the asymptotic formulas for the Bessel potentials (see, e.g., [Aronszajn and Smith 1961], pp. 415–417):

$$\begin{aligned} G_n(x) &\sim C_1\log|x|, && \text{as } x\to 0,\\ G_n(x) &\sim C_2|x|^{-1/2}e^{-|x|}, && \text{as } x\to\infty. \end{aligned}$$

Moreover,

$$\frac{\partial}{\partial x_i}G_n(x) = C_s'\cdot\frac{x_i}{|x|}K_1(|x|),$$

where K_1 is the Bessel–Macdonald function of order 1, with the asymptotics

$$\begin{aligned} K_1(r) &\sim C_3 r^{-1}, && \text{as } r\to 0+,\\ K_1(r) &\sim C_4 r^{-1/2}e^{-r}, && \text{as } r\to\infty. \end{aligned}$$

3D. ***Application to PDEs.*** We will illustrate how nonhomogeneous d^k spaces can be used in the analysis of classic PDEs.

The following result was shown in [Bourgain and Brezis 2007, Theorems 2 and 3]: if $\Delta U = F$ in $\mathbb{R}^n$ and $\operatorname{div} F = 0$, then

$$\|U\|_\infty + \|\nabla U\|_2 \le C\|F\|_1,\ \text{ if } n = 2,$$

and

$$\|U\|_{n/(n-2)} + \|\nabla U\|_{n/(n-1)} \le C\|F\|_1,\ \text{ if } n \ge 3.$$

A more general result of Bourgain and Brezis ([Bourgain and Brezis 2007, Theorem 4′] and [Brezis and Van Schaftingen 2007, Remark 2.1]) implies that one can relax the condition $\operatorname{div} F = 0$ to $\operatorname{div} F \in L^1$ to obtain

$$\|\nabla U\|_{n/(n-1)} \leq C(\|F\|_1 + \|\operatorname{div} F\|_1).$$

Note that for $n \geq 3$ this can be combined with a Sobolev embedding theorem to produce

$$\|U\|_{n/(n-2)} + \|\nabla U\|_{n/(n-1)} \leq C(\|F\|_1 + \|\operatorname{div} F\|_1).$$

However (as noted in [Brezis and Van Schaftingen 2007]), if $n = 2$ then U may no longer be an L^∞ vector field. Let us explain why this may happen using Theorem 3.5.

Let $g(x) = \log|x|$. Then $g * F$ is continuous for any $F \in \Upsilon_1^1$ and if

$$\|U\|_\infty = (2\pi)^{-1}\|g * F\|_\infty \leq C(\|F\|_1 + \|\operatorname{div} F\|_1)$$

were true for any $F \in \mathcal{D}^1(\mathbb{R}^2)$, then we would have

$$|g * F(0)| = \left| \int g(x)F(x)\,dx \right| \leq C\|F\|_{\Upsilon_1^1},$$

and $g(x) = \log|x|$ would be a d^1 function and by Theorem 3.5, $\log|x| \in \operatorname{bmo}(\mathbb{R}^2)$. However, this is false by Remark 3.22.

So the solution of equation $\Delta U = F \in \mathbb{R}^2$ can be essentially unbounded even if $\operatorname{div} F \in L^1$, because the fundamental solution of Δ in $\mathbb{R}^2$ is not an element of $d^1(\mathbb{R}^2)$. Based on the examples of $d^{n-1}(\mathbb{R}^n)$ functions, one can guess that the situation should be better in the case of the Helmholtz equation. Indeed, the following proposition shows that solutions to the Helmholtz equation can be fully controlled even under relaxed conditions.

Theorem 3.24. *Let* $n \geq 2$, $F \in L^1(\mathbb{R}^n; \mathbb{R}^n)$ *and* $\operatorname{div} F \in L^1(\mathbb{R}^n)$. *Then the system* $(I - \Delta)U = F$ *admits a unique solution* U *such that*:

- *If* $n = 2$, *then*
$$\|U\|_\infty + \|\nabla U\|_2 \leq C(\|F\|_1 + \|\operatorname{div} F\|_1).$$
- *If* $n \geq 3$, *then*
$$\|U\|_{n/(n-2)} + \|\nabla U\|_{n/(n-1)} \leq C(\|F\|_1 + \|\operatorname{div} F\|_1).$$

Proof. Without loss of generality we can assume that $F \in \mathcal{S}(\mathbb{R}^2; \mathbb{R}^2)$.

Case 1: If $n \geq 3$, then

$$\Delta U = \frac{\Delta}{I - \Delta} F =: \tilde{F}.$$

As $\Delta/(I-\Delta)$ is an operator of convolution against a finite measure (see, e.g., Chapter 5 in [Stein 1970]), $\tilde{F} \in L^1$ and $\operatorname{div} \tilde{F} = \Delta/(I-\Delta) \operatorname{div} F \in L^1$, with

$$\|\tilde{F}\|_1 + \|\operatorname{div} \tilde{F}\|_1 \le C(\|F\|_1 + \|\operatorname{div} F\|_1).$$

Hence by Theorem 4′ in [Bourgain and Brezis 2007],

$$\|\nabla U\|_{n/(n-1)} \le C(\|\tilde{F}\|_1 + \|\operatorname{div} \tilde{F}\|_1) \le C'(\|F\|_1 + \|\operatorname{div} F\|_1)$$

and the application of Sobolev's embedding theorem completes the proof.

Case 2: If $n = 2$, then solution U has the form $U(x) = G_2 * F(x)$, where $G_2(x)$ is the Bessel potential of order 2. By Example 3.23, $G_2 \in d^1(\mathbb{R}^2)$. Thus for any $x \in \mathbb{R}^2$,

$$|U(x)| = |G_2 * F(x)| \le \|G_2\|_{d^1} \|\tau_x F\|_{\Upsilon_1^1(\mathbb{R}^2)} = \|G_2\|_{d^1} \|F\|_{\Upsilon_1^1(\mathbb{R}^2)},$$

where τ_x is the translation operator defined by $(\tau_x f)(y) = f(y-x)$. In other words

$$\|U\|_\infty \le C(\|F\|_1 + \|\operatorname{div} F\|_1). \tag{3-9}$$

In order to control ∇U, notice that the decay of F and G_2 implies

$$\int |\nabla U_i(x)|^2\,dx = -\int U_i(x)\Delta U_i(x)\,dx = \int U_i(x) F_i(x)\,dx - \int U_i^2(x)\,dx.$$

Hence, recalling that U is a convolution of the L^1 functions G_2 and F,

$$\|\nabla U\|_2 \le C\|U\|_\infty^{1/2}(\|F\|_1 + \|U\|_1)^{1/2} \le C\|U\|_\infty^{1/2}\|F\|_1^{1/2}.$$

Using (3-9) we complete the proof. □

4. d^k spaces on Lipschitz domains

In this section we define d^k classes on domains. Everywhere in this section we assume Ω to be a bounded Lipschitz domain in $\mathbb{R}^n$.

Definition 4.1. Let $1 \le k \le n$. A distribution $u \in \mathcal{D}'(\Omega)$ is said to belong to $d^k(\Omega)$ if there exists $C > 0$ such that $|u(\phi_I)| \le C\|\Phi\|_{\Upsilon_k^1(\Omega)}$ for any

$$\Phi = \sum_{|I|=k} \phi_I dx^I \in \mathcal{D}^k(\Omega).$$

We denote the space of such distributions by $d^k(\Omega)$ and equip it with the norm

$$\|u\|_{d^k(\Omega)} := \sup\{|u(\phi_I)| : \Phi \in \mathcal{D}^k(\Omega);\ \|\Phi\|_{\Upsilon_k^1(\Omega)} \le 1\}.$$

Remark 4.2. Let $1 \le k \le n$. We want to consider distributions $u \in \mathcal{E}'(\Omega)$ such that $|u(\phi_I)| \le C\|\Phi\|_{\Upsilon_k^1(\Omega)}$ for some finite $C > 0$ and any

$$\Phi = \sum_{|I|=k} \phi_I dx^I \in \mathcal{D}^k(\mathbb{R}^n).$$

The class of $\mathcal{E}'(\Omega) \cap d^k(\mathbb{R}^n)$, equipped with the norm $\|\cdot\|_{d^k(\mathbb{R}^n)}$ forms an incomplete normed space. Therefore we define $d_z^k(\Omega)$ as follows.

Remark 4.3. The definitions we use were suggested by Van Schaftingen [2006]. It is also possible to define $d^k(\Omega)$ as we did in Remark 3.2. Any $u \in \mathcal{D}'(\Omega)$ defines a linear map $\tilde{u} : \mathcal{D}^k(\Omega) \to \mathbb{R}^{\binom{n}{k}}$ by

$$\tilde{u}\Big(\sum_{|I|=k} \phi_I dx^I\Big) = (u(\phi_I))_I$$

and $u \in d^k(\Omega)$ if and only if $\tilde{u}$ can be extended to a bounded linear map from $\Upsilon_{k,0}^1(\Omega)$ to $\big(\mathbb{R}^{\binom{n}{k}}, \|\cdot\|_{\max}\big)$, where $\Upsilon_{k,0}^1(\Omega) = \overline{\mathcal{D}^k(\Omega)}$ and the closure is taken with respect to the Υ_k^1 norm.

4A. *$d_z^k(\Omega)$ spaces.* All properties of $d_z^k(\Omega)$ spaces can be deduced from the previous results and the following definition:

Definition 4.4. For $1 \le k \le n$, set $d_z^k(\Omega) = \{u \in d^k(\mathbb{R}^n) : \text{supp } u \in \overline{\Omega}\}$.

Remark 4.5. It is clear that $d_z^k(\Omega)$ is a closed subspace of $d^k(\mathbb{R}^n)$, and therefore complete, and that $\mathcal{E}'(\Omega) \cap d^k(\mathbb{R}^n) \subset d_z^k(\Omega)$. Conversely, any $u \in d_z^k$ is the weak limit of $\mathcal{E}'(\Omega) \cap d^k(\mathbb{R}^n)$. Indeed, consider any $u \in d^k(\mathbb{R}^n)$ supported in $\overline{\Omega}$. By Theorem 3.5 and the definition of $\text{bmo}_z(\overline{\Omega})$, $u \in \text{bmo}_z(\Omega)$. In particular $u \in L^1(\Omega)$. Let η_j be a sequence of $\mathcal{D}(\Omega)$ functions such that $\lim_{j\to\infty} \eta_j = \chi_\Omega$, the characteristic function of Ω. Then by Lebesgue's dominated convergence theorem, for any $\Phi \in \mathcal{D}^k(\overline{\Omega})$ and I,

$$\int_\Omega u(x)\phi_I(x)\,dx = \lim_{j\to\infty} \int_\Omega (\eta_j u)(x)\phi_I(x)\,dx.$$

This shows that $u = \lim\limits_{j\to\infty} (\eta_j u)$ is a weak limit.

Combining this definition with Lemma 3.3 we obtain:

Proposition 4.6. *The spaces $d_z^k(\Omega)$ form a monotone family, i.e., the following embeddings hold*:

$$d_z^n(\Omega) \subset d_z^{n-1}(\Omega) \subset \cdots \subset d_z^1(\Omega).$$

Proposition 4.7. *Let Ω be a bounded Lipschitz domain and $W_0^{1,n}(\Omega)$ be the closure of $\mathcal{D}(\Omega)$ functions in the norm $\|\cdot\|_{W^{1,n}(\Omega)}$. Then $W_0^{1,n}(\Omega)$ is continuously embedded into $d_z^{n-1}(\Omega)$.*

Proof. The space $W_0^{1,n}(\Omega)$ can be characterized (see, e.g., Theorem 5.29 in [Adams and Fournier 2003]) as follows: let $f \in W^{1,n}(\Omega)$, then $f \in W_0^{1,n}(\Omega)$ if and only if the extension of f by zero to $\mathbb{R}^n \setminus \overline{\Omega}$ belongs to $W^{1,n}(\mathbb{R}^n)$. Using this characterization, we can identify any $u \in W_0^{1,n}(\Omega)$ with $\tilde{u} \in W^{1,n}(\mathbb{R}^n)$ supported in $\overline{\Omega}$. By Van Schaftingen's theorem such $\tilde{u}$ is an element of $d^{n-1}(\mathbb{R}^n)$ and is supported in $\overline{\Omega}$. Therefore by Definition 4.4, $\tilde{u} \in d_z^{n-1}(\Omega)$. □

Proposition 4.8. *The space $d_z^1(\Omega)$ is a proper subspace of* $\mathrm{bmo}_z(\Omega)$.

Proof. This follows immediately from Theorem 3.5, Definition 4.4 and the definition of $\mathrm{bmo}_z(\Omega)$. □

All in all, we can see that the spaces $d_z^k(\Omega)$ form a family of intermediate spaces between $W_0^{1,n}(\Omega)$ and $\mathrm{bmo}_z(\Omega)$.

4B. *$d^k(\Omega)$ spaces.* It follows directly from the definitions of $d^k(\mathbb{R}^n)$ and $d^k(\Omega)$, that $u \to u|_\Omega$ maps $d^k(\mathbb{R}^n)$ to $d^k(\Omega)$ and

$$\|u|_\Omega\|_{d^k(\Omega)} \leq \|u\|_{d^k(\mathbb{R}^n)}, \tag{4-1}$$

where $u|_\Omega$ stands for the restriction of u to Ω.

Repeating verbatim the proof of Lemma 3.3, one obtains:

Proposition 4.9. *Let $1 \leq k < l \leq n$ and $u \in d^l(\Omega)$. Then $u \in d^k(\Omega)$ and $\|u\|_{d^k(\Omega)} \leq \|u\|_{d^l(\Omega)}$. In other words*:

$$d^n(\Omega) \subset d^{n-1}(\Omega) \subset \cdots \subset d^1(\Omega).$$

In order to show that $W^{1,n}(\Omega) \subset d^{n-1}(\Omega)$, we recall the extension property of Sobolev spaces. It is well known (see, e.g., Theorem 5.24 in [Adams and Fournier 2003]) that if Ω is a Lipschitz domain, then there exists a bounded linear operator $E : W^{l,p}(\Omega) \to W^{l,p}(\mathbb{R}^n)$ such that $Eu = u$ almost everywhere in Ω for all $u \in W^{l,p}(\Omega)$. If we consider such an extension E on $W^{1,n}(\Omega)$ and recall (4-1) and Theorem 3.4, then

$$\|u\|_{d^{n-1}(\Omega)} = \|Eu|_\Omega\|_{d^{n-1}(\Omega)} \leq \|Eu\|_{d^{n-1}(\mathbb{R}^n)} \leq \|Eu\|_{W^{1,n}(\mathbb{R}^n)} \leq \|E\|\|u\|_{W^{1,n}(\Omega)}.$$

In other words:

Proposition 4.10. *If Ω is a bounded Lipschitz domain, then $W^{1,n}(\Omega)$ is continuously embedded into $d^{n-1}(\Omega)$.*

The following result is the analog of Theorem 3.5 on Lipschitz domains.

Theorem 4.11. *Any $u \in d^1(\Omega)$ is a* $\mathrm{bmo}_r(\Omega)$ *function and*

$$\|u\|_{\mathrm{bmo}_r(\Omega)} \leq C\|u\|_{d^1(\Omega)}.$$

The proof is more technical than that of Theorem 3.5 because of the presence of $\partial\Omega$. Firstly, we state a corollary of the Nečas inequality:

$$\|f\|_{L^2(\Omega)} \le C(\|f\|_{W^{-1,2}(\Omega)} + \|\nabla f\|_{W^{-1,2}(\Omega)}) \quad \text{for all } f \in L^2(\Omega).$$

Lemma 4.12 [Auscher et al. 2005, Lemma 10]. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^n$. If $g \in L^2(\Omega)$ and $\int g = 0$, then there exists a vector-valued function $F \in W_0^{1,2}(\Omega, \mathbb{R}^n)$ such that*

$$\begin{cases} \operatorname{div} F = g \;\; \text{in } \Omega, \\ \|DF\|_{L^2} \le C\|g\|_2. \end{cases}$$

Here DF is a matrix $\partial_j F_i$ and $C > 0$ depends only on the Lipschitz constant of Ω.

Using this lemma we prove the following:

Lemma 4.13. *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^n$. If $g \in H_z^1(\Omega)$, then there exists a vector-valued function $F \in W_0^{1,1}(\Omega, \mathbb{R}^n)$ such that*

$$\begin{cases} \operatorname{div} F = g \;\; \text{in } \Omega, \\ \|DF\|_{L^1} \le C\|g\|_{H^1}. \end{cases}$$

Proof. Let $g \in H_z^1(\Omega)$. Then by Theorem 2.18, it can be decomposed into $H_z^1(\Omega)$ atoms $a_i \in L^2(\mathbb{R}^n)$ as

$$g = \sum_{i=1}^{\infty} \lambda_i a_i$$

and

$$\sum_{i=1}^{\infty} |\lambda_i| \le 2\|g\|_{H^1}.$$

For each $i \ge 1$, using Lemma 4.12, we can find $V^i \in W_0^{1,2}(Q_i, \mathbb{R}^n)$, such that

$$\begin{cases} \operatorname{div} V^i = a_i \;\; \text{in } Q_i, \\ \|DV^i\|_{L^2} \le C\|a_i\|_{L^2}. \end{cases}$$

As $W_0^{1,2}(Q_i)$ fields, V^i can be continuously extended by 0 to $W^{1,2}(\Omega)$. We denote these extensions by the same V^i. We claim that $F = \sum_{i=1}^{\infty} \lambda_i V^i$ is the solution we seek.

Indeed, since the a_i are atoms, we have

$$\|DV^i\|_{L^1} \le |Q_i|^{1/2}\|DV^i\|_{L^2} \le C|Q_i|^{1/2}\|a_i\|_{L^2} \le C_1 \text{ for all } i \ge 1.$$

Therefore, the partial sums $\sum_{i=1}^{N} \lambda_i V^i$, supported in Ω, converge to an element F of $W_0^{1,1}(\Omega, \mathbb{R}^{n\times n})$ and

$$\|DF\|_{L^1} \le C_1 \sum_i |\lambda_i| \le C\|g\|_{H^1}.$$

Finally, by the construction of F,

$$\operatorname{div} F = \sum_i \lambda_i \cdot \operatorname{div} V^i = \sum_i \lambda_i a_i = g. \qquad \square$$

Now we can prove the last theorem of this section:

Proof of Theorem 4.11. We will use the duality between $h^1_z(\Omega)$ and $\mathrm{bmo}_r(\Omega)$ asserted by Theorem 2.15. By Lemma 2.11, it is enough to show that for any $f \in \mathcal{D}(\Omega)$ and $u \in d^1(\Omega)$,

$$|u(f)| \le C\|u\|_{d^1}\|f\|_{h^1}. \tag{4-2}$$

Given $f \in \mathcal{D}(\Omega)$, we write f as the sum $f = g + \theta$, where

$$g = f - \int f(x)\,dx \cdot \psi \quad \text{and} \quad \theta = \int f(x)\,dx \cdot \psi,$$

where $\psi \in \mathcal{D}(\Omega)$ is any function with $\int \psi = 1$.

Note that $\theta \in \mathcal{D}(\Omega)$ with $\|\theta\|_{h^1} \le \|\psi\|_{L^1}\|f\|_{h^1}$ and $\|\theta\|_{W^{1,1}} \le \|f\|_{h^1}\|\psi\|_{W^{1,1}}$. Moreover if we define $\Theta = \sum_{i=1}^n \theta\, dx^i \in \mathcal{D}^1(\Omega)$, then $\|\Theta\|_{\Upsilon^1_1(\Omega)} \le C\|\psi\|_{W^{1,1}}\|f\|_{h^1}$. Therefore

$$|u(\theta)| \le \|u\|_{d^1(\Omega)}\|\Theta\|_{\Upsilon^1_1(\Omega)} \le C_\psi\|u\|_{d^1(\Omega)}\|f\|_{h^1}. \tag{4-3}$$

On the other hand, for $g \in \mathcal{D}(\Omega)$, we recall Lemma 2.4 to see that $g \in H^1_z(\Omega)$ and

$$\|g\|_{H^1} \le C_\Omega\|g\|_{h^1} \le C'_\psi\|f\|_{h^1}. \tag{4-4}$$

Hence, Lemma 4.13 is applicable and there exists $F \in W^{1,1}_0(\Omega;\mathbb{R}^n)$ such that

$$\begin{cases} \operatorname{div} F = g \ \text{ in } \Omega, \\ \|DF\|_{L^1(\Omega;\mathbb{R}^{n\times n})} \le C\|g\|_{H^1}. \end{cases}$$

Using this F, we introduce n differential forms

$$\Phi^j = \sum_{i=1}^n \partial_i F_j dx^i.$$

We claim that all the Φ^j are in $\Upsilon^1_{1,0}(\Omega)$ and $\|\Phi^j\|_{\Upsilon^1_1(\Omega)} \le C'_\psi\|f\|_{h^1}$ (recall that $\Upsilon^1_{k,0}(\Omega) = \overline{\mathcal{D}^k(\Omega)}$ where the closure is taken with respect to the Υ^1_k norm). Assuming the claim and recalling that u is well defined on components of $\Upsilon^1_{1,0}(\Omega)$ forms (see Remark 4.3), one has

$$\begin{aligned} |u(g)| = \left|u\left(\sum_{i=1}^n \partial_i F_i\right)\right| &\le \sum_{i,j=1}^n |u(\partial_i F_j)| \le n\|u\|_{d^1(\Omega)} \max_{1\le j\le n}\|\Phi^j\|_{\Upsilon^1_1(\Omega)} \\ &\le C\|u\|_{d^1(\Omega)}\|f\|_{h^1}. \end{aligned} \tag{4-5}$$

We complete the proof by deducing (4-2) from (4-3), (4-5) and the triangle inequality.

In order to prove the claim, we note that $d\Phi^j = 0$ by construction and all components of Φ^j are $L^1(\Omega)$ functions, bounded in the L^1-norm by a multiple of $\|g\|_{H^1}$. Recalling (4-4), we may conclude that

$$\|\Phi^j\|_{\Upsilon^1_1(\Omega)} = \|\Phi^j\|_{L^1_1(\Omega)} \le C\|f\|_{h^1}.$$

Furthermore, $F_j \in W^{1,1}_0(\Omega)$ for $j = 1, \dots, n$, which means that there exist sequences $\{F^m_j\}^\infty_{m=1} \subset \mathcal{D}(\Omega)$ such that $\|\partial_i F^m_j - \partial_i F_j\|_{L^1(\Omega)} \to 0$, as $m \to \infty$. Hence, by forming closed $\mathcal{D}^1(\Omega)$-forms,

$$\Phi^{j,m} = \sum_{i=1}^{n} \partial_i F^m_j \, dx^i,$$

we can construct $\mathcal{D}^1(\Omega)$ approximations of Φ^j, such that, as $m \to \infty$,

$$\|\Phi^{j,m} - \Phi^j\|_{\Upsilon^1_1(\Omega)} = \|\Phi^{j,m} - \Phi^j\|_{L^1_1(\Omega)} \to 0,$$

which shows that $\Phi^j \in \Upsilon^1_{1,0}(\Omega)$ for $j = 1, \dots, n$. □

5. d^k spaces on Riemannian manifolds

Let (M, g) be a complete Riemannian manifold. Then $\exp_p$ is defined on T_pM and, as mentioned earlier, for sufficiently small $r_p > 0$, maps $B_{r_p}(0) \in T_pM$ diffeomorphically onto an open subset of M. Let us denote by $\mathrm{inj}_M(p)$, the supremum of all such $r_p > 0$ and define the injectivity radius of M as

$$\mathrm{inj}_M := \inf\{\mathrm{inj}_M(p) : p \in M\}.$$

Definition 5.1. A Riemannian manifold (M, g) is called a manifold with bounded geometry if

(1) M is complete and connected;

(2) $\mathrm{inj}_M > 0$;

(3) for every multi-index α, there exists $C_\alpha > 0$ such that $|D^\alpha g_{i,j}| \le C_\alpha$ in the normal geodesic coordinates $(\Omega_p(r_p), \exp_p^{-1})$.

Examples of manifolds with bounded geometry include compact Riemannian manifolds, $\mathbb{R}^n$ and $\mathbb{H}^n$ (see, e.g., [Eldering 2013]).

5A. *Tame partition of unity.* Let (M, g) be a Riemannian manifold with bounded geometry. For $\delta \in (0, \mathrm{inj}_M)$, we denote by $\Omega_\delta(p)$ the image $B_\delta(0)$ by the map $\exp_p$ which is called a geodesic ball with radius δ centered at p.

Proposition 5.2 [Triebel 1992, p. 284]. *For sufficiently small $\delta > 0$ there exists a uniformly locally finite covering of M by a sequence of geodesic balls*

$\{\Omega_\delta(p_j)\}_{j\in\mathbb{Z}_+}$ *and a corresponding smooth partition of unity* $\{\psi_j\}_{j\in\mathbb{Z}_+}$ *subordinate to* $\{\Omega_\delta(p_j)\}_{j\in\mathbb{Z}_+}$.

Following Taylor [2009], we will call such a covering and partition of unity a tame covering and a tame partition of unity.

5B. $W^{s,p}(M)$, $h^1(M)$ *and* bmo(M).

Definition 5.3 [Triebel 1992, Chapter 7]. Let (M, g) be a Riemannian manifold with bounded geometry and let $\{\psi_j\}$ be a tame partition of unity subordinate to a tame cover by geodesic balls $\{\Omega_\delta(p_j)\}$. The Sobolev space $W^{s,p}(M)$, $1 < p < \infty$, $s > 0$, is defined as

$$W^{s,p}(M) = \left\{ f \in \mathcal{D}'(M) : \sum_{j\in\mathbb{Z}_+} \|\psi_j f \circ \exp_{p_j}\|^p_{W^{s,p}(\mathbb{R}^n)} < \infty \right\}$$

Taylor [2009], introduced versions of Hardy spaces and bmo on manifolds with bounded geometry. One way to define $h^1(M)$ is as follows:

Definition 5.4 [Taylor 2009, Corollary 2.4]. Let $f \in \mathcal{D}'(M)$ and let $\{\psi_j\}$ be a tame partition of unity subordinate to a tame cover by geodesic balls $\{\Omega_\delta(p_j)\}$. We say that $f \in h^1(M)$ if $\sum_j \|(\psi_j f)\circ\exp_{p_j}\|_{h^1(\mathbb{R}^n)} < \infty$. We equip the space $h^1(M)$ with the norm

$$\|f\|_{h^1(M)} = \sum_j \|(\psi_j f)\circ\exp_{p_j}\|_{h^1(\mathbb{R}^n)}.$$

The space bmo(M) is defined similarly.

Definition 5.5 [Taylor 2009, Corollary 3.4]. Let $f \in L^1_{\text{loc}}(M)$ and let $\{\psi_j\}$ be a tame partition of unity subordinate to a tame cover by geodesic balls $\{\Omega_\delta(p_j)\}$. We say that $f \in \text{bmo}(M)$ if $\sum_j \|(\psi_j f)\circ\exp_{p_j}\|_{\text{bmo}(\mathbb{R}^n)} < \infty$. We equip the space bmo$(M)$ with the norm

$$\|f\|_{\text{bmo}(M)} = \sum_j \|(\psi_j f)\circ\exp_{p_j}\|_{\text{bmo}(\mathbb{R}^n)}.$$

Remark 5.6. All these classes of functions have equivalent global definitions. However, for our purposes it is more convenient to use the introduced versions. We refer to [Taylor 2009; Aubin 1982; Triebel 1992] for alternative definitions and the proofs of their equivalence.

5C. $d^k(M)$ *spaces and the embedding into* bmo(M).

Definition 5.7. Let $\{\psi_j\}$ be a tame partition of unity subordinate to a tame cover by geodesic balls $\{\Omega_\delta(p_j)\}$. We say that $u \in \mathcal{D}'(M) \in d^k(M)$ if, for each j, $(\psi_j u)\circ\exp_{p_j} \in d^k(\mathbb{R}^n)$ and

$$\|u\|_{d^k(M)} := \sum_j \|(\psi_j u)\circ\exp_{p_j}\|_{d^k(\mathbb{R}^n)} < \infty.$$

We complete this part with the result which immediately follows from the definitions of the spaces $W^{1,n}(M)$, $d^k(M)$, $\mathrm{bmo}(M)$, and the results of Section 3B: Lemma 3.3 and Theorems 3.4 and 3.5.

Theorem 5.8. *Let M be the Riemannian manifold with bounded geometry. Then the following continuous embeddings hold*:

$$W^{1,n}(M) \subset d^{n-1}(M) \subset \cdots \subset d^1(M) \subset \mathrm{bmo}(M).$$

Acknowledgements

The author is very grateful to Galia Dafni for careful reading of the manuscript and many useful discussions concerning it. He would also like to show his gratitude to the referee for valuable suggestions and comments that considerably helped to improve the quality of this paper.

References

[Adams and Fournier 2003] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, 2nd ed., Pure and Applied Math. **140**, Elsevier, Amsterdam, 2003. MR Zbl

[Aronszajn and Smith 1961] N. Aronszajn and K. T. Smith, "Theory of Bessel potentials, I", *Ann. Inst. Fourier* (*Grenoble*) **11** (1961), 385–475. MR Zbl

[Aubin 1982] T. Aubin, *Nonlinear analysis on manifolds: Monge–Ampère equations*, Grundlehren der Math. Wissenschaften **252**, Springer, 1982. MR Zbl

[Auscher et al. 2005] P. Auscher, E. Russ, and P. Tchamitchian, "Hardy Sobolev spaces on strongly Lipschitz domains of $\mathbb{R}^n$", *J. Funct. Anal.* **218**:1 (2005), 54–109. MR Zbl

[Bourdaud 2002] G. Bourdaud, "Remarques sur certains sous-espaces de $\mathrm{BMO}(\mathbb{R}^n)$ et de $\mathrm{bmo}(\mathbb{R}^n)$", *Ann. Inst. Fourier* (*Grenoble*) **52**:4 (2002), 1187–1218. MR

[Bourgain and Brezis 2004] J. Bourgain and H. Brezis, "New estimates for the Laplacian, the div-curl, and related Hodge systems", *C. R. Math. Acad. Sci. Paris* **338**:7 (2004), 539–543. MR Zbl

[Bourgain and Brezis 2007] J. Bourgain and H. Brezis, "New estimates for elliptic equations and Hodge type systems", *J. Eur. Math. Soc.* **9**:2 (2007), 277–315. MR Zbl

[Bourgain et al. 2004] J. Bourgain, H. Brezis, and P. Mironescu, "$H^{1/2}$ maps with values into the circle: minimal connections, lifting, and the Ginzburg–Landau equation", *Publ. Math. Inst. Hautes Études Sci.* **99** (2004), 1–115. MR Zbl

[Brezis and Van Schaftingen 2007] H. Brezis and J. Van Schaftingen, "Boundary estimates for elliptic systems with L^1-data", *Calc. Var. Partial Differential Equations* **30**:3 (2007), 369–388. MR Zbl

[Caetano 2000] A. M. Caetano, "Approximation by functions of compact support in Besov–Triebel–Lizorkin spaces on irregular domains", *Studia Math.* **142**:1 (2000), 47–63. MR Zbl

[Chang 1994] D.-C. Chang, "The dual of Hardy spaces on a bounded domain in $\mathbb{R}^n$", *Forum Math.* **6**:1 (1994), 65–81. MR Zbl

[Chang et al. 1993] D.-C. Chang, S. G. Krantz, and E. M. Stein, "H^p theory on a smooth domain in $\mathbb{R}^N$ and elliptic boundary value problems", *J. Funct. Anal.* **114**:2 (1993), 286–347. MR Zbl

[Chang et al. 1999] D.-C. Chang, G. Dafni, and E. M. Stein, "Hardy spaces, BMO, and boundary value problems for the Laplacian on a smooth domain in $\mathbb{R}^n$", *Trans. Amer. Math. Soc.* **351**:4 (1999), 1605–1661. MR Zbl

[Dafni 2002] G. Dafni, "Local VMO and weak convergence in h^1", *Canad. Math. Bull.* **45**:1 (2002), 46–59. MR Zbl

[Eldering 2013] J. Eldering, *Normally hyperbolic invariant manifolds: the noncompact case*, Atlantis Studies in Dynamical Systems **2**, Atlantis, Paris, 2013. MR Zbl

[Goldberg 1979] D. Goldberg, "A local version of real Hardy spaces", *Duke Math. J.* **46**:1 (1979), 27–42. MR Zbl

[Grafakos 2009] L. Grafakos, *Modern Fourier analysis*, 2nd ed., Graduate Texts in Math. **250**, Springer, 2009. MR Zbl

[Jones 1980] P. W. Jones, "Extension theorems for BMO", *Indiana Univ. Math. J.* **29**:1 (1980), 41–66. MR Zbl

[Jonsson et al. 1984] A. Jonsson, P. Sjögren, and H. Wallin, "Hardy and Lipschitz spaces on subsets of $\mathbb{R}^n$", *Studia Math.* **80**:2 (1984), 141–166. MR Zbl

[Miyachi 1990] A. Miyachi, "H^p spaces over open subsets of $\mathbb{R}^n$", *Studia Math.* **95**:3 (1990), 205–228. MR Zbl

[Smirnov 1993] S. K. Smirnov, "Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows", *Algebra i Analiz* **5**:4 (1993), 206–238. In Russian; translated in *St. Peters. Math.* **5**:4 (1994), 841–867. MR Zbl

[Stein 1970] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Math. Series **30**, Princeton Univ. Press, 1970. MR Zbl

[Strichartz 1972] R. S. Strichartz, "The Hardy space H^1 on manifolds and submanifolds", *Canad. J. Math.* **24** (1972), 915–925. MR Zbl

[Taylor 2009] M. Taylor, "Hardy spaces and BMO on manifolds with bounded geometry", *J. Geom. Anal.* **19**:1 (2009), 137–190. MR Zbl

[Triebel 1992] H. Triebel, *Theory of function spaces, II*, Monographs in Mathematics **84**, Birkhäuser, Basel, 1992. MR Zbl

[Van Schaftingen 2004a] J. Van Schaftingen, "Estimates for L^1-vector fields", *C. R. Math. Acad. Sci. Paris* **339**:3 (2004), 181–186. MR Zbl

[Van Schaftingen 2004b] J. Van Schaftingen, "A simple proof of an inequality of Bourgain, Brezis and Mironescu", *C. R. Math. Acad. Sci. Paris* **338**:1 (2004), 23–26. MR Zbl

[Van Schaftingen 2006] J. Van Schaftingen, "Function spaces between BMO and critical Sobolev spaces", *J. Funct. Anal.* **236**:2 (2006), 490–516. MR Zbl

[Van Schaftingen 2010] J. Van Schaftingen, "Limiting fractional and Lorentz space estimates of differential forms", *Proc. Amer. Math. Soc.* **138**:1 (2010), 235–240. MR Zbl

[Van Schaftingen 2014] J. Van Schaftingen, "Limiting Bourgain–Brezis estimates for systems of linear differential equations: theme and variations", *J. Fixed Point Theory Appl.* **15**:2 (2014), 273–297. MR Zbl

Received March 28, 2018. Revised July 12, 2018.

ALMAZ BUTAEV
DEPARTMENT OF MATHEMATICS AND STATISTICS
CONCORDIA UNIVERSITY
MONTREAL QC
CANADA
almaz.butaev@concordia.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.27

A COUNTEREXAMPLE TO THE EASY DIRECTION OF THE GEOMETRIC GERSTEN CONJECTURE.

DAVID BRUCE COHEN

For finitely generated groups H and G, equipped with word metrics, a translation-like action of H on G is a free action such that each element of H acts by a map which has finite distance from the identity map in the uniform metric. For example, if H is a subgroup of G, then right translation by elements of H yields a translation-like action of H on G. Whyte asked whether a group having no translation-like action by a Baumslag–Solitar group must be hyperbolic, where the free abelian group of rank 2 is understood to be a Baumslag–Solitar group. We show that the converse question has a negative answer, and in particular the fundamental group of a closed hyperbolic 3-manifold admits a translation-like action by the free abelian group of rank 2.

1. Introduction

A metric space X is said to be uniformly discrete if it has a minimum distance, meaning

$$\inf\{d(x, y) : x, y \in X; x \neq y\} > 0,$$

and said to have bounded geometry if for all $r > 0$, there is some $N_r > 0$ such that every r-ball has cardinality at most N_r. If X satisfies both of these conditions, it is said to be a UDBG space [Whyte 1999, §2]. For example, a finitely generated group equipped with a word metric is a UDBG space. More generally, if X is the vertex set of a connected graph of bounded degree, equipped with the metric that assigns length 1 to each edge, then X is a UDBG space.

Definition [Whyte 1999, Definition 6.1]. Let X be a UDBG space. A translation-like action of a group H on X is a free action by maps at a finite distance from the identity. That is, the action satisfies the following rules.

- For $x \in X$ and $h \in H$, if $h \cdot x = x$, then $h = 1_H$.
- For all $h \in H$, the set $\{d(x, h \cdot x) : x \in X\}$ is bounded.

MSC2010: 20F65, 20F67.
Keywords: geometric group theory, translation-like actions, hyperbolic groups.

We will mostly be interested in the case where H is finitely generated and the UDBG space X is a finitely generated group G equipped with a word metric. In this case, a translation-like action of H on G is just a vertex-surjective embedding of a disjoint union of copies of a Cayley graph of H into a Cayley graph of G (since an orbit of a translation-like action of H on G embeds the Cayley graph of H into some Cayley graph of G).

Translation-like actions generalize subgroups. If H is a finitely generated subgroup of G, then H acts translation-like on G via

$$h \cdot g = gh^{-1}$$

for $h \in H$ and $g \in G$. Many properties which pass to subgroups of G also pass to groups which act translation-like on G. For instance, Jeandel [2015, Theorem 3] has shown that if G has no weakly aperiodic subshift of finite type, then the same is true for finitely presented groups acting translation-like on G. By considering that many properties which pass to subgroups of G also pass to groups which act translation-like on G, Whyte was able to give "geometric" versions of several famous conjectures about how the properties of G constrain its subgroups [Whyte 1999, §6].

The geometric von Neumann–Day conjecture. The von Neumann–Day conjecture (disproven by Olshanskii [1980]) asserts that a group G should be nonamenable if and only if G contains a free subgroup. Whyte [1999, Theorem 6.1] used translation-like actions to formulate and prove a geometric version of this conjecture — namely, that G is nonamenable if and only if $\mathbb{Z} * \mathbb{Z}$ acts translation-like on G.

The geometric Burnside problem. The Burnside problem (answered in the negative by Golod and Shafarevich [Golod 1964]) asks whether every infinite finitely generated group contains a $\mathbb{Z}$-subgroup. The geometric Burnside problem asks whether every infinite, finitely generated group admits a translation-like action of $\mathbb{Z}$. Seward [2014, Theorem 1.4] proved that the answer to this question is yes.

The geometric Gersten conjecture. Recall that for $m, n \in \mathbb{Z}_{\neq 0}$, the Baumslag–Solitar group $\mathrm{BS}(m, n)$ is the group presented by $\langle a, b | ab^m a^{-1} = b^n \rangle$, and in particular $\mathrm{BS}(1, 1) \cong \mathbb{Z}^2$. It is known that these groups do not embed in hyperbolic groups. The Gersten conjecture [Bestvina 2000, Q 1.1] — usually attributed to Gromov — roughly states that for a group satisfying some finiteness properties, hyperbolicity should be equivalent to having no Baumslag–Solitar subgroup. We do not know whether Gersten actually asked this question, although [Gersten 1996] asks whether every finitely presented subgroup of a hyperbolic group must be hyperbolic. [Brady 1999] showed that this was false, and hence that the Gersten conjecture is false for finitely presented groups (weaker versions remain open).

The geometric Gersten conjecture states being hyperbolic is equivalent to having no translation-like action by any $\mathrm{BS}(m, n)$. In point of fact, Whyte only asked about the “hard” direction — whether a group which is not hyperbolic must admit a translation-like action of a Baumslag–Solitar group — and only for 2-dimensional groups. By an observation of Jeandel [2015, §5], knowing the hard direction for all amenable groups would imply that every group (except for virtually cyclic groups) has a weakly aperiodic subshift of finite type, as conjectured by Carroll and Penland [2015]. In a recent preprint Jiang [2017] has shown that the lamplighter group admits no translation-like actions by Baumslag–Solitar groups. Since the lamplighter is not hyperbolic, this disproves the hard direction of the geometric Gersten conjecture, although finitely presented counterexamples remain unknown.

Seward [2014, §1.(3’)] asked about the other direction — whether Baumslag–Solitar groups may act translation-like on hyperbolic groups. Our main theorem gives a negative answer to this question.

Theorem 1.1. Let G be the fundamental group of a closed hyperbolic 3-manifold. Then $\mathbb{Z}^2$ acts translation-like on G.

2. Proof of Theorem 1.1

Let G be the fundamental group of a closed hyperbolic 3-manifold. We will prove Theorem 1.1 by showing that G is bilipschitz to a UDBG space which admits a translation-like action of $\mathbb{Z}^2$ — the following lemma says that this is sufficient.

Lemma 2.1. If H acts translation-like on X_1, and X_1 is bilipschitz to X_2 then H acts translation-like on X_2.

Proof. Let $\psi : X_1 \to X_2$ be a bilipschitz map. We define a translation-like action of H on X_2 by conjugating the action as follows. For $x \in X_2$, take

$$h \cdot x = \psi(h \cdot \psi^{-1}(x)).$$

It is clear that this is a free action, and it is translation-like because

$$d(x, h \cdot x) \leq \mathrm{Lip}(\psi) d(\psi^{-1}(x), h \cdot \psi^{-1}(x)). \qquad \square$$

Lemma 2.2. There exists a UDBG space X such that $\mathbb{Z}^2$ acts translation-like on X and X is bilipschitz to G.

Proof. Consider the set of points

$$X = \{(2^c a, 2^c b, 2^c) : a, b, c \in \mathbb{Z}\}$$

in the upper half space model of $\mathbb{H}^3$. (See [Ratcliffe 1994, §4.6] for details on the upper half space model). The reader may verify that this is indeed a UDBG space (the shortest distance is $\log(2)$ and it is not hard to see that the size of r-balls in X is roughly exponential in r).

To define a translation-like action of $\mathbb{Z}^2$ on X, let the generators e_1, e_2 of $\mathbb{Z}^2$ act by

$$e_1 \cdot (2^c a, 2^c b, 2^c) = (2^c(a+1), 2^c b, 2^c)$$

and

$$e_2 \cdot (2^c a, 2^c b, 2^c) = (2^c a, 2^c(b+1), 2^c).$$

These maps commute, each moves points by a distance of 1, and the $\mathbb{Z}^2$-action they induce is clearly free, so it is translation-like.

Observe that X is quasi-isometric to $\mathbb{H}^3$ because every point of $\mathbb{H}^3$ lies within a bounded distance of $X \subset \mathbb{H}^3$. Thus, by the Svarc–Milnor theorem [Bridson and Haefliger 1999, Proposition I.8.19], X is quasi-isometric to G. By combining [Whyte 1999, Theorem 1] and [Whyte 1999, Theorem 5.1], one sees that any quasi-isometry between nonamenable UDBG spaces is at a bounded distance from a bilipschitz map, so X is bilipschitz to G. □

Combining Lemmas 2.1 and 2.2, we have proved Theorem 1.1

3. Questions

We close with three questions.

Other Baumslag–Solitar groups. Do any hyperbolic groups admit translation-like actions of Baumslag–Solitar groups $\mathrm{BS}(m, n)$ with $m \geq 2$?

Other hyperbolic groups. Which hyperbolic groups admit translation-like actions of $\mathbb{Z}^2$? Jiang [2017] recently observed that one may use results of [Benjamini et al. 2012] to show that $\mathbb{Z}^2$ cannot act translation-like on free groups, and it appears that this technique may be used to rule out translation-like actions of $\mathbb{Z}^2$ on hyperbolic surface groups [Benjamini et al. 2012, Proposition 4.1], but we have no idea whether such actions exist on hyperbolic one-relator groups or on random groups.

Gromov–Furstenberg for returns of the horospherical flow in a hyperbolic 3-manifold. (See [Burago and Kleiner 2002] for context). Let Γ be a cocompact lattice in $\mathrm{PSL}(2; \mathbb{C})$, let H be an ϵ-neighborhood of some horosphere H_0 in $\mathbb{H}^3$, let $* \in \mathbb{H}^3$, and consider the intersection $\mathcal{O} = (\Gamma \cdot *) \cap H$. If we equip $\mathcal{O}$ with the metric inherited from H, then $\mathcal{O}$ is quasi-isometric to $H_0 \cap (\Gamma \cdot B_\epsilon(*))$, where $B_\epsilon(*)$ denotes the ϵ ball around $*$ in $\mathbb{H}^3$. From Ratner's theorem [1991], it then follows (with some thought) that $\mathcal{O}$ is quasi-isometric to $\mathbb{Z}^2$. Must $\mathcal{O}$ be bilipschitz to $\mathbb{Z}^2$? This was our original attempt at finding a translation-like action of $\mathbb{Z}^2$ on Γ.

Acknowledgments

We wish to thank Kevin Whyte, Benson Farb, and Andy Putman for conversations. We wish to thank Yongle Jiang for sharing an early draft of his preprint with us. This work has been supported by NSF award 1502608.

References

[Benjamini et al. 2012] I. Benjamini, O. Schramm, and Á. Timár, "On the separation profile of infinite graphs", *Groups Geom. Dyn.* **6**:4 (2012), 639–658. MR Zbl

[Bestvina 2000] M. Bestvina, "Questions in geometric group theory", problem list, 2000, available at https://tinyurl.com/mlbestv.

[Brady 1999] N. Brady, "Branched coverings of cubical complexes and subgroups of hyperbolic groups", *J. London Math. Soc.* (2) **60**:2 (1999), 461–480. MR Zbl

[Bridson and Haefliger 1999] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Math. Wissenschaften **319**, Springer, 1999. MR Zbl

[Burago and Kleiner 2002] D. Burago and B. Kleiner, "Rectifying separated nets", *Geom. Funct. Anal.* **12**:1 (2002), 80–92. MR Zbl

[Carroll and Penland 2015] D. Carroll and A. Penland, "Periodic points on shifts of finite type and commensurability invariants of groups", *New York J. Math.* **21** (2015), 811–822. MR Zbl

[Gersten 1996] S. M. Gersten, "Subgroups of word hyperbolic groups in dimension 2", *J. London Math. Soc.* (2) **54**:2 (1996), 261–283. MR Zbl

[Golod 1964] E. S. Golod, "On nil-algebras and finitely approximable p-groups", *Izv. Akad. Nauk SSSR Ser. Mat.* **28**:2 (1964), 273–276. In Russian; translated in *Amer. Math. Soc. Trans.* (2) **48** (1965), 103–106. MR Zbl

[Jeandel 2015] E. Jeandel, "Translation-like actions and aperiodic subshifts on groups", preprint, 2015. arXiv

[Jiang 2017] Y. Jiang, "Translation-like actions yield regular maps", preprint, 2017. arXiv

[Olshanskii 1980] A. Y. Olshanskii, "On the question of the existence of an invariant mean on a group", *Uspekhi Mat. Nauk* **35**:4 (1980), 199–200. In Russian; translated in *Russ. Math. Surv.* **35**:4 (1980), 180–181. MR

[Ratcliffe 1994] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*, Graduate Texts in Math. **149**, Springer, 1994. MR Zbl

[Ratner 1991] M. Ratner, "On Raghunathan's measure conjecture", *Ann. of Math.* (2) **134**:3 (1991), 545–607. MR Zbl

[Seward 2014] B. Seward, "Burnside's problem, spanning trees and tilings", *Geom. Topol.* **18**:1 (2014), 179–210. MR Zbl

[Whyte 1999] K. Whyte, "Amenability, bi-Lipschitz equivalence, and the von Neumann conjecture", *Duke Math. J.* **99**:1 (1999), 93–112. MR Zbl

Received March 6, 2018. Revised June 18, 2018.

DAVID BRUCE COHEN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CHICAGO
CHICAGO, IL
UNITED STATES
davidbrucecohen@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.33

THE GENERAL LINEAR 2-GROUPOID

MATÍAS DEL HOYO AND DAVIDE STEFANI

We deal with the symmetries of a (2-term) graded vector space or bundle. Our first theorem shows that they define a (strict) Lie 2-groupoid in a natural way. Our second theorem explores the construction of nerves for Lie 2-categories, showing that it yields simplicial manifolds if the 2-cells are invertible. Finally, our third and main theorem shows that smooth pseudofunctors into our general linear 2-groupoid classify 2-term representations up to homotopy of Lie groupoids.

1. Introduction

A Lie group G can be thought of as a smooth collection of symmetries of an abstract object. A linear representation $G \curvearrowright V$ is therefore a way to realize these symmetries on a concrete vector space V, which we will assume to be finite-dimensional and real. Such a representation can be defined either as a smooth map $\rho : G \times V \to V$ satisfying $\rho^h \rho^g = \rho^{hg}$ and $\rho^1 = \mathrm{id}$, or as a Lie group morphism $G \to \mathrm{GL}(V)$ into the general linear group. We can then study the group G by looking at its representations $G \curvearrowright V$, and this approach turns out to be very profitable.

Following the previous philosophy, a Lie groupoid $G \rightrightarrows M$ should be thought of as a smooth collection of symmetries of an abstract family parametrized by M. Lie groupoids have received much attention lately, as they provide a unifying framework for classic geometries, and also serve as models for spaces with singularities such as orbifolds and, more generally, differentiable stacks. The infinitesimal versions

MSC2010: 18G30, 22A22, 57R22.
Keywords: Lie 2-groupoids, nerve, simplicial manifolds, representation up to homotopy.

of Lie groupoids are Lie algebroids, geometric objects intertwining Lie algebra bundles and (singular) foliations. Differentiation and integration set up a fruitful interaction between the two theories.

A linear representation $(G \rightrightarrows M) \curvearrowright (V \to M)$ of a Lie groupoid over a vector bundle associates to each arrow $y \xleftarrow{g} x$ a linear isomorphism $\rho^g : V^x \to V^y$ between the corresponding fibers, in a way compatible with identities and compositions. It can be presented either as a partially defined map $G \times V \to V$ or as a Lie groupoid map $G \to \mathrm{GL}(V)$ into the general linear groupoid [del Hoyo 2013]. The problem with Lie groupoid representations is that they are rather scarce, they impose strong conditions on V, and they do not provide us with enough information on $G \rightrightarrows M$. This reflects in the lack of an adjoint representation, or in the limitations when establishing a Tannaka duality result for Lie groupoids (see [Trentinaglia 2010]).

A solution for these problems, proposed by C. Arias Abad and M. Crainic [2013], involves representations up to homotopy $G \curvearrowright V$ of a Lie groupoid over a graded vector bundle. Such objects can be easily defined as differentials on certain bigraded algebras of sections; alternatively, they can be regarded as a sequence of tensors: a differential ∂ on V, followed by chain maps $\rho^g : V^x \to V^y$ between the fibers, then chain homotopies $\gamma^{h,g}$ relating ρ^{hg} and $\rho^h\rho^g$, etc. Representation up to homotopy has proved to be a useful concept in dealing, for instance, with cohomology theory [loc. cit.], deformations [Crainic et al. 2018] and Morita equivalences [del Hoyo and Ortiz 2018].

When $V = V_1 \oplus V_0$ is a 2-term graded vector bundle, a representation up to homotopy $G \curvearrowright V$ leads to a VB-groupoid, a double structure mixing Lie groupoids and vector bundles, via a semidirect product construction $G \ltimes V \to G$. It turns out that any VB-groupoid can be split as a semidirect product, by choosing a horizontal lift of arrows, as proven first in [Gracia-Saz and Mehta 2017]. This yields a one-to-one correspondence between VB-groupoids and 2-term representations up to homotopy, which can be extended to maps, and respect equivalence classes [del Hoyo and Ortiz 2018]. Prominent examples of VB-groupoids are the tangent and cotangent constructions. They encode the adjoint and coadjoint representations, respectively.

A VB-groupoid is an instance of a fibration of groupoids, and according to classic Grothendieck correspondence, after choosing a horizontal lift of arrows, a groupoid fibration $E \to G$ is the same as a pseudofunctor $G \dashrightarrow \{\text{Groupoids}\}$ (see [SGA 1 1971]). It follows that 2-term representations up to homotopy should, in some sense, be the same as pseudofunctors. The main purpose of the present paper is to shed light on this. To take care of the smooth and the linear structure, we are led to fix a 2-term graded vector bundle V and restrict our attention to pseudofunctors involving the several fibers of V. The resulting $G \dashrightarrow \mathrm{GL}(V)$ is a suitable generalization of the classification map $G \to \mathrm{GL}(V)$ for actual representations.

Given a graded vector bundle $V = V_1 \oplus V_0 \to M$, we construct a general linear

2-groupoid $\mathrm{GL}(V)$, consisting of differentials on the fibers, quasi-isomorphisms between them, and chain homotopies. There are several nonequivalent notions of Lie 2-groupoids in the literature, some of them too strict and some others too lax for our purposes. After discussing some variants, we introduce a notion of Lie 2-groupoid, and prove our Theorem 5.5, asserting that $\mathrm{GL}(V)$ is indeed a Lie 2-groupoid. It is remarkable that even for a 2-term graded vector space V, its general linear 2-groupoid $\mathrm{GL}(V)$ is not a 2-group; it has more than one object, so groupoids arise naturally.

In the set-theoretic context there is a nerve for 2-categories that relates lax functors with simplicial maps [Bullejos et al. 2005; Lack 2010]. We develop the smooth version of it, and our Theorem 6.3 shows that, even though NC is not always a simplicial manifold, it is so when the Lie 2-category C has invertible 2-arrows, in particular for a Lie 2-groupoid. This nerve construction relates our notion of Lie 2-groupoids with the simplicial approach to Lie 2-groupoids, based on the horn-filling condition, which has received much attention lately. This can be seen as a piece of evidence supporting our definitions for Lie 2-groupoids and smooth pseudofunctors. We also compare our construction with that of [Mehta and Tang 2011].

Building on the previous results, which we believe are of interest in their own right, we finally establish our Theorem 7.7, setting an equivalence of categories between 2-term representations up to homotopy $G \curvearrowright V$ and pseudofunctors $G \dashrightarrow \mathrm{GL}(V)$ commuting with basic projections. Combining this with the main theorem of [Gracia-Saz and Mehta 2017], and its extension in [del Hoyo and Ortiz 2018], we get what we might call a smooth linear variant of Grothendieck correspondence (see Remark 2.5):

$$\left\{\begin{matrix}\text{VB-groupoids}\\ \Gamma \to G\end{matrix}\right\} \rightleftharpoons \left\{\begin{matrix}\text{2-term representations up to homotopy}\\ G \curvearrowright V_1 \oplus V_0\end{matrix}\right\} \rightleftharpoons \left\{\begin{matrix}\text{pseudofunctors}\\ G \dashrightarrow \mathrm{GL}(V)\end{matrix}\right\}.$$

It seems natural to extend this result for higher degrees, relating positively graded representations up to homotopy and maps into a general linear ∞-groupoid. Also, as potential applications of our theorem, we believe it is possible to relate our correspondence with the infinitesimal version announced in [Mehta 2014], and to frame the main theorem from [del Hoyo and Ortiz 2018] as a result about maps between differentiable 2-stacks. These problems will be explored elsewhere.

Organization. In Sections 2 and 3 we quickly review 2-categories and their nerves, to fix notation and provide a reference for the tools needed later. Section 4 introduces our notion of a Lie 2-groupoid and compares it with other important ones found in the literature. In Section 5 we prove our first theorem, which constructs the fundamental example: the general linear 2-groupoid. Section 6 explores the combinatorics behind the nerve of 2-categories, and exploits it to establish our second theorem: the nerve of a Lie 2-category whose 2-cells are invertible is a simplicial manifold. In Section 7 we prove our main theorem, realizing representations up to homotopies as maps, and we discuss further questions and applications.

2. Basics on 2-categories

We review here definitions and basic facts on set-theoretic 2-categories that are fundamental for the rest of the paper. We give a definition of 2-groupoid, compare it with others in the literature, and discuss the notion of lax functors. We refer to [Borceux 1994; Lack 2010; Mac Lane 1998] for further details. The material here is preparatory, to set notation and conventions and to serve as a quick reference.

A *2-category* C is a category enriched over the category of small categories. It has three levels of structure: objects, arrows between objects, and arrows between arrows or *2-cells*, whose collections we denote by C_0, C_1, C_2, respectively. We use letters $x, y, \dots$ for objects, $f, g, \dots$ for arrows, and $\alpha, \beta, \dots$ for 2-cells:

$$y \underset{g}{\overset{f}{\leftleftarrows}} x \quad (\Downarrow \alpha)$$

The arrows and 2-cells between two fixed objects x, y form a category $C(y, x)$, whose composition we denote by $\bullet$. For each triple x, y, z there is a composition functor $C(z, y) \times C(y, x) \xrightarrow{\circ} C(z, x)$ and a unit $\mathrm{id}_x \in C(x, x)$ satisfying the axioms encoded in the following commutative diagrams:

$$\begin{array}{ccccc} & & C(w,z) \times C(z,y) \times C(y,x) & & \\ & \swarrow^{\circ \times \mathrm{id}} & & \searrow^{\mathrm{id} \times \circ} & \\ C(w,y) \times C(y,x) & & & & C(w,z) \times C(z,x) \\ & \searrow_{\circ} & & \swarrow_{\circ} & \\ & & C(w,x) & & \end{array}$$

$$\begin{array}{ccccc} & & C(y,x) & & \\ & \swarrow^{\mathrm{id}_y \times \mathrm{id}} & & \searrow^{\mathrm{id} \times \mathrm{id}_x} & \\ C(y,y) \times C(y,x) & & \downarrow \mathrm{id} & & C(y,x) \times C(x,x) \\ & \searrow_{\circ} & & \swarrow_{\circ} & \\ & & C(y,x) & & \end{array}$$

Example 2.1. The paradigmatic example of a 2-category is that of small categories, functors and natural transformations. Another basic example is that of spaces, continuous maps and (homotopy classes of) homotopies.

We are interested in 2-groupoids. For us, a 2-*groupoid* G is a 2-category such that (i) it is small, in the sense that G_0 is a set, (ii) every 2-cell is invertible, and (iii) every arrow $y \xleftarrow{f} x$ is invertible up to homotopy, namely, there exist $x \xleftarrow{g} y$ and 2-cells $fg \cong \mathrm{id}_y$ and $gf \cong \mathrm{id}_x$. Some references demand the arrows be invertible on the nose. We call such 2-groupoids *strict*. Let us remark that our fundamental example, that of the general linear 2-groupoid, is not strict.

Example 2.2. A topological space X yields a 2-groupoid $\pi_2(X)$ whose objects are the points of X, whose arrows are the continuous paths $I \to X$, and whose 2-cells are (homotopy classes of) path homotopies. Composition is given by juxtaposition, moving through each path at double speed. A nonconstant path is only invertible up to homotopy, hence $\pi_2(X)$ is not strict.

A simple characterization of (small) 2-categories and strict 2-groupoids is by using *double structures*, namely diagrams of compatible structures as below, where compatible means that the horizontal structural maps are functorial with respect to the vertical structures:

$$\begin{array}{ccc} G_2 & \rightrightarrows & G_0 \\ \downdownarrows & & \downdownarrows \\ G_1 & \rightrightarrows & G_0 \end{array}$$

However, our notion of 2-groupoid does not benefit much from this perspective. The following lemma, which is automatic for strict groupoids but works in general, will be useful later.

Lemma 2.3. *If G is a 2-groupoid and $y \xleftarrow{f} x$ is an arrow in G, then the right multiplication functor $R_f : G(z, y) \to G(z, x)$ is an equivalence of categories for any z. The same holds for left multiplication.*

Proof. A 2-cell $\alpha : f \Rightarrow g$ defines a natural isomorphism $R_f \Rightarrow R_g$, for the 2-cells are invertible. Then, given an arbitrary f, and picking g a quasi-inverse, we have $\mathrm{id}_{G(x,x)} = R_{\mathrm{id}_x} \cong R_g R_f$ and analogously $\mathrm{id}_{G(y,y)} = R_{\mathrm{id}_y} \cong R_f R_g$. □

A *functor* $\phi : C \to D$ between 2-categories consists of functions $\phi_i : C_i \to D_i$ preserving all the structure *on the nose*. This notion is sometimes too rigid, for it involves many identities between functors. A useful variant is that of a (normal) *lax functor* $\phi : C \dashrightarrow D$, which consists of three maps $\phi_i : C_i \to D_i$ preserving source, target, units and the composition $\bullet$, but only preserving $\circ$ up to a given natural transformation. More precisely, also given is a map

$$\phi_{1,1} : C_1 \times_{C_0} C_1 \to D_2, \qquad \phi_{1,1}(g, f) : \phi_1(gf) \Rightarrow \phi_1(g) \circ \phi_1(f),$$

ruling the failure of associativity of $\circ$ and satisfying these coherence axioms:

(i) $\phi_{1,1}(\mathrm{id}_y, f) = \mathrm{id}_f = \phi_{1,1}(f, \mathrm{id}_x)$, where $y \xleftarrow{f} x$ (normality).

(ii) $(\phi_2(\beta) \circ \phi_2(\alpha)) \bullet \phi_{1,1}(g, f) = \phi_{1,1}(g', f') \bullet \phi_2(\beta \circ \alpha)$, where

$$z \underset{g'}{\overset{g}{\Leftarrow\!\!\Downarrow \beta}} y \underset{f'}{\overset{f}{\Leftarrow\!\!\Downarrow \alpha}} x .$$

(iii) $(\phi_{1,1}(h, g) \circ \phi_1(f)) \bullet \phi_{1,1}(hg, f) = (\phi_1(h) \circ \phi_{1,1}(g, f)) \bullet \phi_{1,1}(h, gf)$, where

$$w \xleftarrow{h} z \xleftarrow{g} y \xleftarrow{f} x.$$

When the structure 2-cells $\phi_{1,1}(g, f)$ are invertibles the lax functor is called a *pseudofunctor*. These notions are very interesting even when C is a usual category, viewed as a 2-category with only identity 2-cells. To ease the notation we will often write ϕ instead of ϕ_i, etc.

Example 2.4. Given an epimorphism of groups $\pi : G \to H$, a set-theoretic section $\sigma : H \to G$, $\sigma(1_G) = 1_H$, leads to a pseudofunctor $\phi : H \dashrightarrow \{\text{Groups}\}$, where G is viewed as a 2-groupoid with one object and only identity 2-cells, and {Groups} is the 2-category of groups, morphisms, and inner automorphisms as 2-cells. Here $\phi(*) = K$ is the kernel of π, $\phi(h)$ is given by conjugation by $\sigma(h)$, and $\phi(h', h)$ is the conjugation by $\sigma(h')\sigma(h)\sigma(h'h)^{-1}$. The lax functor is an actual functor if and only if σ is a morphism.

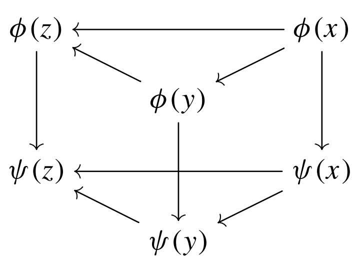
We also need to deal with morphisms between lax functors (see [Borceux 1994]). Given lax functors $\phi, \psi : C \dashrightarrow D$ between 2-categories, a *lax transformation* $H : \phi \Rightarrow \psi$ associates to each $x \in C_0$ an arrow $H_x : \phi(x) \to \psi(x)$ and to each arrow $f : x \to y$ a 2-cell $H_f : H_y\phi(f) \Rightarrow \psi(f)H_x$ satisfying these conditions:

(i) $H_{\mathrm{id}_x} = \mathrm{id}_{H_x}$ (normality),

(ii) $(\psi(\alpha) \circ \mathrm{id}_{H_x}) \bullet H_f = H_g \bullet (\mathrm{id}_{H_y} \circ \phi(\alpha))$. where

$$y \underset{g}{\overset{f}{\Leftarrow\!\!\Downarrow \alpha}} x .$$

(iii) For each pair of composable arrows $z \xleftarrow{g} y \xleftarrow{f} x$ there is a commutative prism with vertical faces H_g, H_f, H_{gf} and horizontal faces given by the structural 2-cells of ϕ, ψ:



Such an H is a *lax equivalence* if the H_x are invertible up to a 2-cell and the H_f are invertible.

Remark 2.5. Example 2.4 can be easily extended to suitable epimorphisms between categories, known as *fibered categories* [Borceux 1994; SGA 1 1971]. The outcome is the *Grothendieck correspondence* between equivalence classes of fibered categories $E \to C$ and pseudofunctors $C \dashrightarrow$ {Categories}. This is the first and most important example of lax functors. The main goal of the present paper can be considered to be a smooth linear variant of this correspondence.

3. The nerve of a 2-category

After reviewing the classic nerve construction, we discuss here the nerve for 2-categories and 2-groupoids. We explain its behavior with respect to lax functors, and we use it to relate 2-groupoids with the weak approach to higher categories based on the horn-filling condition. Some references for this are [Bullejos and Cegarra 2003; Bullejos et al. 2005; Henriques 2008; Lack 2010].

As usual, let $[n] = \{n, n-1, \dots, 1, 0\}$ denote the ordinal of cardinality $n+1$, and let Δ be the category of finite ordinals and order-preserving maps, spanned by the elementary maps

$$d^i : [n-1] \to [n], \qquad d^i(k) = \begin{cases} k & \text{if } k < i, \\ k+1 & \text{if } k \geq i, \end{cases}$$

$$s^j : [n+1] \to [n], \qquad s^j(k) = \begin{cases} k & \text{if } k \leq j, \\ k-1 & \text{if } k > j, \end{cases}$$

which satisfy the so-called simplicial identities. Then a *simplicial set* is a contravariant functor $X : \Delta^\circ \to$ {Sets}. It can be described as a sequence of sets $X_n = X([n])$ and a collection of *face* $d_i = X(d^i)$ and *degeneracy* $s_j = X(s^j)$ operators satisfying the (dual) simplicial identities. Maps of simplicial sets are natural transformations, or equivalently, sequences of maps compatible with the faces and degeneracies. Simplicial objects on a category $\mathcal{C}$ are defined analogously.

Example 3.1. A simple but fundamental example is the *n-simplex* Δ^n. From the functorial viewpoint, it is the one represented by the ordinal $[n]$. Thinking of Δ^n as a graded set with further structure, it is freely generated by an element of type $[n]$, namely $\mathrm{id}_{[n]}$. By Yoneda's lemma, a map $\Delta^n \to X$ is the same as an element in X_n. The *border* $\partial\Delta^n \subset \Delta^n$ is spanned by all the faces of the generator, and the *horn* $\Lambda^n_k \subset \Delta^n$ by all the faces but the k-th.

Given a category $\mathcal{C}$ and a covariant functor $\phi : \Delta \to \mathcal{C}$, which should be thought of as a model for simplices in $\mathcal{C}$, we can define a *singular functor*

$$\phi^* : \mathcal{C} \to \{\text{Simplicial sets}\}$$

that associates to each object $X \in \mathcal{C}$ a simplicial set by the formula $(\phi^* X)_n = \hom_{\mathcal{C}}(\phi([n]), X)$. In other words, $\phi^* X$ is the restriction of the contravariant functor represented by X to Δ via ϕ.

Example 3.2. When $\mathcal{C}$ is the category of topological spaces and $\phi([n])$ is the topological n-simplex, then $\phi^* X = SX$ is the *singular simplicial set* associated to X, used to define its homology. When $\mathcal{C}$ is the category of (small) categories and $\phi([n]) = [n]$, where we see an ordinal as a category by setting an arrow $i \to j$ if $i \leq j$, then $\phi^* C = \mathrm{N}C$ is the *nerve* of the category, whose n-simplices are chains of n composable arrows and whose faces and degeneracies are given by dropping an extremal arrow, composing two consecutive ones, or inserting an identity.

We are concerned with the nerve construction for 2-categories, namely the singular functor defined when $\mathcal{C}$ is the category of 2-categories and lax functors, and $\phi([n]) = [n]$ is viewed as a 2-category with only identity 2-cells. Thus, if C is a 2-category, then its *nerve* $\mathrm{N}C$ has as n-simplices the lax functors $u : [n] \dashrightarrow C$, and its simplicial operators are given by precomposition. Note that $\mathrm{N}C_0 = C_0$ and $\mathrm{N}C_1 = C_1$ consist of the objects and arrows of C, respectively, and $\mathrm{N}C_2$ consists of triangles that are commutative up to a given 2-cell:

$$\begin{array}{ccccc} & & y & & \\ & {\scriptstyle g}\swarrow & \Uparrow \alpha & \nwarrow{\scriptstyle f} & \\ z & & \xleftarrow{\quad h \quad} & & x \end{array}$$

To describe the higher simplices, note that a lax functor $u : [n] \dashrightarrow C$ can be thought of as a labeling in an abstract n-simplex, where u_i are objects at the vertices, $u_{j,i}$ are arrows at its edges, and $u_{k,j,i}$ are 2-cells corresponding to each triangle. For each tetrahedron on the simplex the following equation among 2-cells must hold:

$$\begin{array}{ccccc} & & u_{l,i} & & \\ & {\scriptstyle u_{l,k,i}}\swarrow\!\!\!\!\swarrow & & \searrow\!\!\!\!\searrow{\scriptstyle u_{l,j,i}} & \\ u_{l,k}u_{k,i} & & & & u_{l,j}u_{j,i} \\ & {\scriptstyle u_{l,k}u_{k,j,i}}\searrow\!\!\!\!\searrow & & \swarrow\!\!\!\!\swarrow{\scriptstyle u_{l,k,j}u_{j,i}} & \\ & & u_{l,k}u_{k,j}u_{j,i} & & \end{array}$$

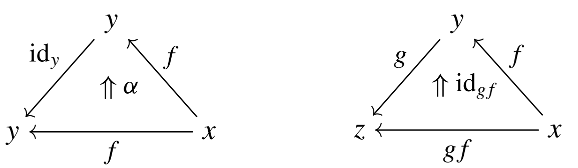
The above data completely determines the nerve $\mathrm{N}C$ in the sense that it is 3-*coskeletal*, that is, for $k > 3$ we have

$$\mathrm{N}C_k = \{\partial \Delta^k \to \mathrm{N}C\}.$$

A fundamental feature of the classic nerve for 1-categories is that it defines a fully faithful functor; it embeds the category of (small) categories into that of simplicial sets. Extending this, there is the following proposition for the nerve of 2-categories, which also provides information about the 2-cells. Here, by a *simplicial homotopy* we mean a simplicial map $X \times \Delta^1 \to Y$.

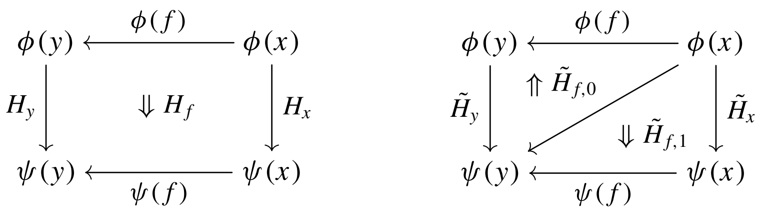
Proposition 3.3 [Bullejos et al. 2005]. *The nerve* $C \mapsto \mathrm{N}C$ *defines a fully faithful functor from the category of* (*small*) *2-categories and* (*normal*) *lax functors to the category of simplicial sets. Moreover*, *if* $\phi, \psi : C \dashrightarrow D$ *are lax functors and every 2-cell in* D *is invertible*, *then there is a lax transformation* $H : \phi \Rightarrow \psi$ *if and only if there is a simplicial homotopy* $\tilde{H} : N\phi \cong N\psi$.

Sketch of proof. Given a simplicial map $\tilde{\phi} : \mathrm{N}C \to \mathrm{N}D$, we can define a lax functor $\phi : C \dashrightarrow D$ such that $N\phi = \tilde{\phi}$ by setting $\phi_0 = \tilde{\phi}_0$, $\phi_1 = \tilde{\phi}_1$, and defining ϕ_2 and $\phi_{1,1}$ as restrictions of $\tilde{\phi}_2$ to the following types of triangles:



The simplicial identities on $\tilde{\phi}$ imply the axioms of a lax functor on ϕ, and that $N\phi = \tilde{\phi}$, proving the first assertion.

Regarding the second triangle, given lax functors $\phi, \psi : C \dashrightarrow D$, while a lax transformation $H : \phi \cong \psi$ associates to an arrow $y \xleftarrow{f} x$ a 2-cell filling a commutative square, a simplicial homotopy $\tilde{H} : N\phi \cong N\psi$ should provide a triangulation of that square:



where $\tilde{H}_{f,0}$ and $\tilde{H}_{f,1}$ are short for $\tilde{H}(s_1(f), s_0(\mathrm{id}_{[1]}))$ and $\tilde{H}(s_0(f), s_1(\mathrm{id}_{[1]}))$. The lax transformation H induces a simplicial homotopy $\tilde{H}$ by setting $\tilde{H}_{f,0} = \mathrm{id}$ and $\tilde{H}_{f,1} = H_f$. Conversely, if every 2-cell on D is invertible, we can define an H out of $\tilde{H}$ by setting

$$H_f = \tilde{H}_{f,1} \bullet (\tilde{H}_{f,0})^{-1}. \qquad \square$$

Another fundamental feature of the classic nerve is the following characterization of its image: a simplicial set is the nerve of a category if and only if every inner horn $(0 < k < n)$ admits a filling, and this filling is unique for $n > 1$. Similarly, it is the nerve of a groupoid if and only if the same holds for every horn, inner or not:

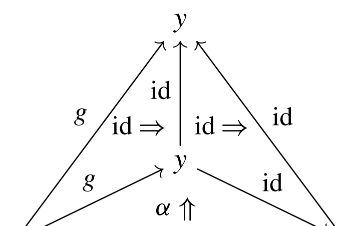
$$\begin{array}{ccc} \Lambda^n_k & \xrightarrow{\ \forall\ } & X \\ \downarrow & \nearrow_{\exists(!)} & \\ \Delta^n & & \end{array}$$

This motivates an approach to higher category theory that has received much attention lately. A simplicial set X is then a *weak m-category* if every inner horn in X admits a filling, and the filling is unique for $n > m$, and X is a *weak m-groupoid* if the same holds for every horn, inner or not. The missing face of the horn, provided by the filling, should be thought of as a *composition*, defined up to homotopy, of the remaining faces. The next proposition relates 2-groupoids with weak 2-groupoids via the nerve functor. Similar results are discussed in [Duskin 2002].

Proposition 3.4. *Let C be a 2-category.* NC *is a weak 2-category if and only if every 2-cell of C is invertible*, *and* NC *is a weak 2-groupoid if and only if C is a 2-groupoid.*

Proof. Since NC is 3-coskeletal, every (n, k)-horn has a unique filling for $n \geq 5$. For $n = 2$ the horizontal composition of arrows provides inner horn-fillings, and the fillings of the outer horns correspond to the existence of quasi-inverses. So let us study the cases $n = 3, 4$.

For $n = 3$, given a 2-cell $\alpha : f \Rightarrow g : x \to y$, we can build a (3, 1)-horn with faces thus:



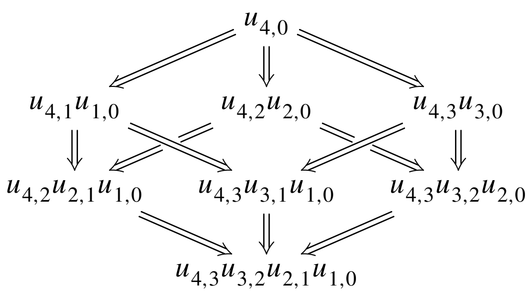
and the remaining face of a filling will give a right inverse $\beta : g \Rightarrow f$ to α, showing that inner-horn filling implies that every 2-cell is invertible. Conversely, a horn gives three 2-cells, which correspond to three sides of this square:

$$\begin{array}{ccc} u_{3,0} & \Longrightarrow & u_{3,1}u_{1,0} \\ \Downarrow & & \Downarrow \\ u_{3,2}u_{2,0} & \Longrightarrow & u_{3,2}u_{2,1}u_{1,0} \end{array}$$

In an inner horn, either the 2-cell on the top or in the left is missing, but since every 2-cell is invertible, we can fill the square by taking the obvious composition.

In an outer horn, either the 2-cell on the bottom or on the right is missing, and assuming C is a 2-groupoid, we can get the missing face by factoring the triple composition by either $u_{3,2}$ or $u_{1,0}$ as it follows from Lemma 2.3.

For $n = 4$, the 2-skeleton of a 4-simplex u gives the edges of a cube as below:

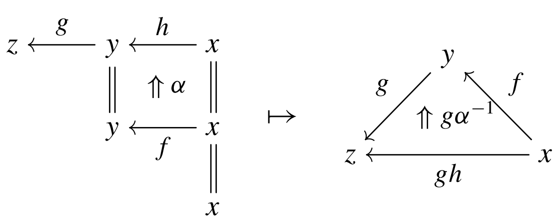


Each face of the 4-simplex corresponds to the commutativity of the corresponding face of the cube. The bottom face commutes because of the compatibility between horizontal and vertical composition. Since every 2-cell is invertible, five commuting faces on the cube imply that the other is commutative as well, thus every horn admits a unique filling, concluding the proof. □

Remark 3.5. Other ways to associate a simplicial set to a 2-category C are by regarding it as a double category with a trivial side, applying twice the classic nerve, or reducing the resulting bisimplicial set by using the *diagonal* d and the *total functor* T, also known as the bar or codiagonal:

$$\text{2-categories} \xrightarrow{N^2} \text{bisimplicial sets} \overset{d,T}{\rightrightarrows} \text{simplicial sets}.$$

It is shown in [Bullejos and Cegarra 2003] that TN^2C and dN^2C are equivalent to NC from a homotopy viewpoint. We remark here that, when C is a strict 2-groupoid there is actually an isomorphism $TN^2C \cong \text{NC}$, which is completely determined by the following formula for 2-cells:



4. Defining Lie 2-groupoids

We discuss here the smooth versions of 2-categories and 2-groupoids we will work with, provide some examples, and discuss other uses for those terms in the literature.

A *Lie 2-category* C is, roughly speaking, a 2-category internal to the category of smooth manifolds. It consists of a (small) 2-category as defined before, on which (i) the sets of objects C_0, arrows C_1 and 2-cells C_2 are equipped with manifold structures; (ii) the source and target maps $s,t : C_i \to C_{i-1}$ of 2-cells and arrows are surjective submersions, and (iii) the units $u : C_{i-1} \to C_i$ and the multiplications $\circ$ and $\bullet$ are smooth. Functors $\phi : C \to D$ between Lie 2-categories are easy to define, as 2-functors for which the three maps $\phi_i : C_i \to D_i$ are smooth.

Example 4.1. Let $(\mathbb{R},\cdot)$ be the multiplicative monoid of real numbers, viewed as a Lie 2-category with a single object, space of arrows $\mathbb{R}$, and both horizontal and vertical composition equal to the multiplication. This is a Lie 2-category on which not every 2-cell is invertible.

Let G be a Lie 2-category that, from the set-theoretic viewpoint, is also a 2-groupoid, as defined in the previous sections. In order to define when G is a Lie 2-groupoid we have to make sense of smooth inversions. For 2-cells this is clear, for there is an inversion map $i : G_2 \to G_2$, and we can require it to be smooth. For arrows this is less clear, for inversion is only defined up to homotopy: there is no inversion map in general. Note that, since source and target $G_2 \to G_1$ are surjective submersions, the sets of 2-horns $N_{2,i}G = \hom(\Lambda^2_i, \mathrm{NG})$ define manifolds:

$$N_{2,0}G = \left\{ \begin{array}{ccc} & y & \\ & \nwarrow^{f} & \\ z & \xleftarrow[h]{} & x \end{array} \right\}, \quad N_{2,1}G = \left\{ \begin{array}{ccc} & y & \\ {}^{g}\swarrow & & \nwarrow^{f} \\ z & & x \end{array} \right\}, \quad N_{2,2}G = \left\{ \begin{array}{ccc} & y & \\ {}^{g}\swarrow & & \\ z & \xleftarrow[h]{} & x \end{array} \right\}.$$

We will discuss a smooth structure on the whole nerve NG in the following sections. For now, we just endow N_2G with a manifold structure using the fibered product

$$\begin{array}{ccc} N_2G & \longrightarrow & N_{2,1}G \\ \downarrow & & \downarrow m \\ G_2 & \xrightarrow[t]{} & G_1 \end{array}$$

We define G to be a *Lie 2-groupoid* if, besides being a Lie 2-category and a 2-groupoid, (i) the inversion of 2-cells $i : G_2 \to G_2$ is smooth, and (ii) the following restriction maps are surjective submersions:

$$d_{2,0} : N_2G \to N_{2,0}G, \qquad d_{2,2} : N_2G \to N_{2,2}G.$$

We say that the Lie 2-groupoid is *strict* if it is set-theoretic strict and the inversion of arrows $i : G_1 \to G_1$ is smooth. The smooth structure on N_2G also allow us to make sense of lax functors in the smooth setting. We define a *smooth lax functor* between Lie 2-categories $\phi : C \dashrightarrow D$ as a lax functor such that ϕ_0, ϕ_1 and the map $(\phi_2, \phi_{1,1}) : N_2C \to N_2D$ are smooth. A *smooth lax transformations* $H : \phi \Rightarrow \psi$ is one on which the maps $C_0 \to D_1,\ C_1 \to D_2$ are smooth.

Example 4.2. Given an abelian Lie group K, we can see it as the 2-cells of a Lie 2-category with one object and one arrow, and where both multiplications $\bullet$ and $\circ$ agree with that of K. The resulting 2-category $K \rightrightarrows * \rightrightarrows *$ is in fact a Lie 2-groupoid. A similar thing can be done with a bundle of abelian Lie groups $G \rightrightarrows M$, such as a torus bundle. This *delooping* construction stays within the finite-dimensional setting and plays a key role for instance in the theory of *gerbes*.

We would like to quickly review the *Dold–Kan construction*. When $\mathcal{C}$ is an abelian category, e.g., that of vector spaces, then a simplicial object $X : \Delta^\circ \to \mathcal{C}$ gives rise to a chain complex (X'_n, ∂) by defining $X'_n = \bigcap_{i>0} \ker(d_i : X_n \to X_{n-1})$ and $\partial = d_0$. It turns out that this construction yields an equivalence of categories between simplicial objects and positively graded chain complexes. The horn-filling condition translates into the abelian setting, in such a way that categories and groupoids both correspond to 2-term complexes, and linear natural transformations correspond to chain homotopies.

Example 4.3. By a *linear 2-category* we mean a Lie 2-category V on which the V_i are (real finite-dimensional) vector spaces and the structure maps are linear. They are examples of Lie 2-groupoids. Viewing them as double linear categories, and applying Dold–Kan correspondence both horizontally and vertically, we encode such a V into a 3-term complex thus:

$$\begin{array}{ccc} V'_2 & \longrightarrow & 0 \\ \downarrow & & \downarrow \\ V'_1 & \longrightarrow & V_0 \end{array}$$

Remark 4.4. The term "Lie 2-groupoid" is used in the literature in senses other than the one we have introduced, which is suitable for our fundamental example. In [Mehta and Tang 2011] and other references, it refers to what we called a strict Lie 2-groupoid and presupposes the existence of inverse arrows, whereas our notion is more general. In [Zhu 2009] and other references, a Lie 2-groupoid is defined as a smooth version of a weak 2-groupoid; the existence of a well-defined composition is not required. We will see later that a smooth version of the nerve functor for Lie 2-categories allows us to regard our Lie 2-groupoids as examples of them.

5. The general linear 2-groupoid

Here we show our first main theorem, asserting that the symmetries of a (2-term) graded vector space or bundle can be endowed with the structure of a Lie 2-groupoid, which we call the general linear 2-groupoid. This construction extends the general linear groupoid of a vector bundle without a grading (see, e.g., [del Hoyo 2013]).

Throughout this section, let $V = V_1 \oplus V_0 \to M$ be a graded vector bundle over a smooth manifold. We will first describe the set-theoretic structure of its general linear 2-groupoid $\mathrm{GL}(V)$ and then take care of the smoothness. From the set-theoretic viewpoint we have:

(i) An object $\partial^x \in \mathrm{GL}(V)_0$ is a differential $\partial^x : V_1^x \to V_0^x$ on the fiber $V^x = V_0^x \oplus V_1^x$.

(ii) An arrow $\alpha : \partial^x \to \partial^y \in \mathrm{GL}(V)_1$ is a couple of linear maps $\alpha_1 : V_1^x \to V_1^y$, $\alpha_0 : V_0^x \to V_0^y$, defining a quasi-isomorphism between V^x and V^y:

$$\begin{array}{ccc} V_1^x & \xrightarrow{\alpha_1} & V_1^y \\ {\scriptstyle \partial^x}\big\downarrow & & \big\downarrow{\scriptstyle \partial^y} \\ V_0^x & \xrightarrow[\alpha_0]{} & V_0^y \end{array}$$

(iii) A 2-cell $R : \alpha \to \alpha' : \partial^x \to \partial^y$ on $\mathrm{GL}(V)_2$ is a chain homotopy, given by a linear map $R : V_0^x \to V_1^y$ such that $R\partial^x = \alpha_1 - \alpha_1'$ and $\partial^y R = \alpha_0 - \alpha_0'$.

$$\begin{array}{ccc} V_1^x & \xrightarrow{\alpha_1} & V_1^y \\ {\scriptstyle \partial^x}\big\downarrow & \overset{R}{\nearrow} & \big\downarrow{\scriptstyle \partial^y} \\ V_0^x & \xrightarrow[\alpha_0]{} & V_0^y \end{array}$$

The multiplication $\circ$ in $\mathrm{GL}(V)$ is the composition of maps, and the multiplication $\bullet$ is the composition of chain homotopies, which is just the sum of the corresponding maps R. Every 2-cell is invertible, and every arrow is invertible up to a 2-cell. Thus we have a well-defined 2-groupoid $\mathrm{GL}(V)$. Via Dold–Kan we can embed it into the 2-category of linear categories.

Remark 5.1. Even when $M = *$ our construction $\mathrm{GL}(V)$ yields a 2-groupoid and not what one might call a 2-group, for there are many objects and not just one. Fixing an object ∂ on $\mathrm{GL}(V)$, its isotropy 2-groupoid $\mathrm{GL}(V)_\partial$ can be compared with the construction studied in [Sheng and Zhu 2012].

Next we show that $\mathrm{GL}(V)$ inherits a smooth structure from certain vector bundles. To ease the notation, given vector bundles $A, B \to M$, we write $[A, B] \to M$ for the inner-hom vector bundle. Then:

(i) $\mathrm{GL}(V)_0$ identifies with the total space of $[V_1, V_0] \to M$.

(ii) $\mathrm{GL}(V)_1$ is a subspace of

$$E = [\pi_1^* V_1, \pi_1^* V_0] \oplus [\pi_2^* V_1, \pi_2^* V_0] \oplus [\pi_1^* V_1, \pi_2^* V_1] \oplus [\pi_1^* V_0, \pi_2^* V_0],$$

a vector bundle over $M \times M$, where $\pi_i : M \times M \to M$ are the obvious projections.

(iii) $\mathrm{GL}(V)_2$ is the set-theoretic fiber product $\mathrm{GL}(V)_1 \times_{M\times M} [\pi_1^* V_0, \pi_2^* V_1]$.

The issue here is to show that $\mathrm{GL}(V)_1 \subset E$ is a submanifold. Then $\mathrm{GL}(V)_2$ will identify with a fibered product along a submersion, in fact with a pullback vector bundle. This issue is rather subtle and will require a careful analysis. The first step in our argument is to provide a simple system of equations describing $\mathrm{GL}(V)_1 \subset E$.

Lemma 5.2. *We can write* $\mathrm{GL}(V)_1 = F \cap U_1 \cap U_0$, *where*

$$F = \{(\partial^x, \partial^y, \alpha_0, \alpha_1) \in E : \alpha_0 \partial^x = \partial^y \alpha_1\},$$
$$U_1 = \{(\partial^x, \partial^y, \alpha_0, \alpha_1) \in E : \ker(\partial^x) \cap \ker(\alpha_1) = 0\},$$
$$U_0 = \{(\partial^x, \partial^y, \alpha_0, \alpha_1) \in E : \mathrm{im}(\partial^y) + \mathrm{im}(\alpha_0) = V_0^y\}.$$

Proof. An element $(\partial^x, \partial^y, \alpha_0, \alpha_1)$ belongs to F if and only if the corresponding square of vector space maps commutes, it belongs to U_1 if and only if the morphism between the fibers is injective in degree 1 homology, and belongs to U_0 if and only if it is surjective in degree 0 homology. Since both fibers V^x, V^y, as 2-term complexes, have the same Euler characteristic $\dim V_0 - \dim V_1$, then so do their homologies, and therefore the two inequalities $\dim H_1(V^x) \leq \dim H_1(V^y)$ and $\dim H_0(V^x) \geq \dim H_0(V^y)$ imply that α is in fact a quasi-isomorphism. □

The subset F can be seen as the preimage of the zero section of the following map between the total space of vector bundles over $M \times M$, where $E' = [\pi_1^* V_1, \pi_2^* V_0]$:

$$\phi : E \to E', \qquad \phi(\partial^x, \partial^y, \rho_1, \alpha_0) = \alpha_0 \partial^x - \partial^y \alpha_1.$$

This map is quadratic and its rank is not constant in general, as the next example shows.

Example 5.3. Let $M = *$ and $V_0 = V_1 = \mathbb{R}$. Then $\mathrm{GL}(V)_0 \cong \mathbb{R}$, $E \cong \mathbb{R}^4$ and F identifies with $\{(x, y, z, w) \in \mathbb{R}^4 : xy - zw = 0\}$, which is not a submanifold of $\mathbb{R}^4$. This example shows that if we define the general linear 2-category $\mathrm{gl}(V)$ as we have defined $\mathrm{GL}(V)$, but without imposing the quasi-isomorphism axiom, then $\mathrm{gl}(V)$ cannot be made a Lie 2-category in a reasonable way.

Next we show that the map ϕ above has maximal rank over the opens U_i, and since the zero section $0_{M\times M} \subset E'$ is closed embedded, the same holds for $\mathrm{GL}(V)_1$.

Proposition 5.4. *The map* $\phi : E \to E'$ *has maximal rank over the opens* U_i.

Proof. Let $p = (\partial^x, \partial^y, \alpha_1, \alpha_0) \in U_1$ and let $q = \phi(p) = \alpha_0 \partial^x - \partial^y \alpha_1$. To show that $d\phi_p : T_p E \to T_q E'$ is surjective we argue by realizing vectors as 1-jets of curves. Given $\gamma(t) \in E'$, $\gamma(0) = q$, we want to lift the curve γ to a curve on E through p. By using local trivializations of V we can assume that $x(t) = x$ and

$y(t) = y$ are fixed. Let us suppose that $p \in U_1$; the other case is analogous. Since $\ker \partial^x \cap \ker \alpha_1 = 0$, the linear map $(\partial^x, \alpha_1) : V_1^x \to V_0^x \oplus V_1^y$ is a monomorphism, and therefore it admits a linear retraction $(\tilde{\partial}^x, \tilde{\alpha}_1) : V_0^x \oplus V_1^y \to V_1^x$. Then the curve $\tilde{\gamma}(t) = (\partial^x, \gamma(t)\tilde{\alpha}_1, \alpha_1, \gamma(t)\tilde{\partial}^x) \in E$ is a lift as required. □

Theorem 5.5. *Given a graded vector bundle $V = V_1 \oplus V_0$, the general linear 2-groupoid $\mathrm{GL}(V)$ inherits a natural structure of a Lie 2-groupoid.*

Proof. As we have already discussed, $\mathrm{GL}(V)_0$ identifies $[V_1, V_0]$, $\mathrm{GL}(V)_1 \subset E$ with the preimage of a closed embedded submanifold along a maximal rank map, and $\mathrm{GL}(V)_2$ is a fiber product along a submersion. It is straightforward to check that with these definitions the structure maps of $\mathrm{GL}(V)$ are smooth, including the inversion of 2-cells. It only remains to show that the following restriction maps are surjective submersions:

$$d_{2,0} : N_2G \to N_{2,0}G, \qquad d_{2,2} : N_2G \to N_{2,2}G.$$

Let us show it for $d_{2,0}$, the other case is analogous. We argue again by lifting curves. We start with $\alpha(t) : \partial^{x(t)} \to \partial^{y(t)}$ and $\gamma(t) : \partial^{x(t)} \to \partial^{z(t)} \in \mathrm{GL}(V)_1$, defining a curve on $N_{(2,0)}G$, and in order to lift it to N_2G, we want to define $\beta(t) : \partial^{y(t)} \to \partial^{z(t)}$ and $R(t) : \gamma(t) \Rightarrow \beta(t)\alpha(t)$. Working locally we can again assume $x = x(t)$, $y = y(t)$, $z = z(t)$ are fixed. The monomorphism $(\alpha_1(t), \partial^x(t)) : V_1^x \to V_1^y \oplus V_0^y$ admits a retraction $\tilde{\alpha}_1(t)$, $\tilde{\partial}^x(t)$, and by basic arguments on linear algebra, we can take it smooth on t. Then the short exact sequence

$$0 \to V_1^x \xrightarrow{(\alpha_1(t), \partial^x(t))} V_1^y \oplus V_0^y \xrightarrow{(\partial^y(t), \alpha_0(t))} V_0^y \to 0$$

splits smoothly and we gain a section $(\tilde{\partial}^y(t), \tilde{\alpha}_0(t))$. We can then define

$$\beta_i(t) = \gamma_i(t)\tilde{\alpha}_i(t), \quad R(t) = \gamma_1(t)\tilde{\partial}^x.$$ □

Remark 5.6. Let us denote by $\mathrm{GL}'(V) \subset \mathrm{GL}(V)$ the open Lie 2-groupoid with the same objects, arrows the invertible chain maps, and 2-cells the chain homotopies. This is a strict Lie 2-groupoid, somehow simpler than our version, and both agree around the units, thus both should behave in the same way with respect to *differentiation*, even though this process is not yet clear. See [Sheng and Zhu 2012] for a related discussion. But regarding our purposes, this simpler construction $\mathrm{GL}(V)'$ is not satisfactory; there are representations up to homotopy of Lie groupoids that cannot be invertible. An example is the adjoint representation of the pair groupoid of the sphere $\mathrm{Pair}(S^2)$, or of any other nonparallelizable manifold. We will come back to this later.

6. The nerve of a Lie 2-category

We deal here with the problem of endowing the nerve NC of a Lie 2-category C with a reasonable smooth structure. We show with a simple example that for general C this may not be possible. Our second main theorem shows that if every 2-cell is invertible then NC is indeed a simplicial manifold, and this happens for instance if C is a Lie 2-groupoid.

Given a Lie 2-category C, we define its *ambient* simplicial manifold AC for the nerve NC, roughly speaking, by considering arbitrary collections $\{u_{k,j,i}\}$ of 2-cells and disregarding any compatibility. More precisely, we define AC by

$$A_nC = \prod_{[2]\xrightarrow{a}[n]} C_2, \qquad u \in A_nC,\ b:[m]\to[n] \ \Rightarrow\ b^*(u)_a = u_{b\circ a} \in A_mC$$

This way AC is a well-defined simplicial manifold, and every face map is a surjective submersion, for it is just the projection onto some of the coordinates. There is a *canonical inclusion* $\phi : \mathrm{NC} \to \mathrm{AC}$ defined by the formula $\phi(u)_a = (u \circ a)_{2,1,0}$, where $u \in N_nC,\ u : [n] \dashrightarrow C$, and $a : [2] \to [n]$. In other words, $\phi(u)$ keeps track of the 2-cells corresponding to each triangle, and by means of the identities, the arrows on the edges and the objects on the vertices. Since every simplex in NC is determined by its 2-skeleton, the map ϕ is injective. We are concerned with the question of whether $\phi(N_nC) \subset A_nC$ is a submanifold, which is not the case in general.

Example 6.1. Let $(\mathbb{R}, \cdot)$ be the multiplicative monoid viewed as a Lie 2-category as described in Example 4.1. Then $N_0C = \{*\},\ N_1C = \{\mathrm{id}_*\}$, and $N_2C = \mathbb{R}$, but $N_3C \subset A_3C$ is not a submanifold. Disregarding the degenerate coordinates, we can identify N_3C with tuples $(x, y, z, w) \in \mathbb{R}^4$ such that $xy = zw$, the equation corresponding to the commutativity of the tetrahedron.

For C a 1-category, a simplex $u \in N_nC$ is the same as a chain of n composable arrows, so we can write N_nC as an iterated fiber product, and use this to define a smooth structure on it. Next we develop a similar combinatorial description for simplices $u \in N_nC$, where C is a 2-category whose 2-cells are invertible.

We see Δ^{n-1} inside Δ^n by using the face d_n, and define a decreasing filtration

$$\Delta^n = F_0\Delta^n \supset F_1\Delta^n \supset \cdots \supset F_{n-1}\Delta^n \supset \Delta^{n-1}$$

by setting $F_k\Delta^n = \{a : [m] \to [n]/a(m) < n \text{ or } a(0) \geq k\}$, namely $F_k\Delta^n$ is the union of Δ^{n-1} with the last face of dimension k. As an example, Figure 1 depicts the filtration for $n = 3$.

Define $N_n^kC = \{F_k\Delta^n \to \mathrm{NC}\}$. Note that $N_n^0C = N_nC$, that we have projections $N_n^kC \to N_n^{k+1}C$, and that $N_n^{n-1}C = N_{n-1} \times_{C_0} C_1$ is the set-theoretic fiber product over $u \mapsto u_n$ and s.

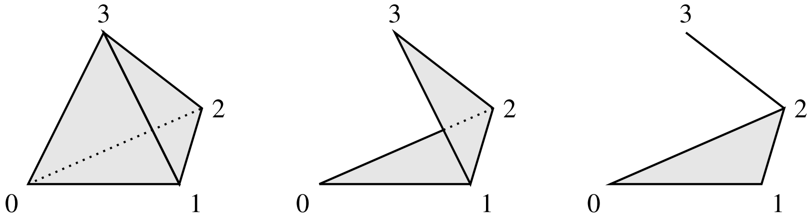


Figure 1. From left to right: $F_0\Delta^3,\ F_1\Delta^3,\ F_2\Delta^3$.

Proposition 6.2. *If every* 2*-cell of* C *is invertible then there are set-theoretic fiber products*:

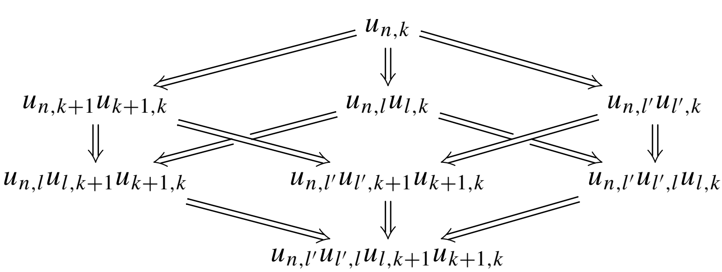
$$\begin{array}{ccc} N_n^{k-1}C & \longrightarrow & C_2 \\ \big\downarrow & & \big\downarrow t \\ N_n^k C & \xrightarrow[\phi_n^k]{} & C_1 \end{array}, \qquad \phi_n^k(u) = u_{n,k} \circ u_{k,k-1}.$$

Proof. The inclusion $F_{k+1}\Delta^n \to F_k\Delta^n$ has all the vertices on its image, all the edges except for (n,k), and all the triangles except for (n,l,k), with $k<l<n$. Thus, given $u : F_k\Delta^n \to \mathrm{NC}$, if we know its restriction u' to $F_{k+1}\Delta^n$ and the 2-cell α corresponding to the triangle $(n,k+1,k)$, then we have all the vertices, we recover the edge (n,k) as the source of α, and we recover the 2-cells corresponding to the triangles (n,l,k) inductively on $l-k$ by means of the equation

$$u_{n,l,k} = (u_{n,l} \circ u_{l,k+1,k})^{-1} \bullet (u_{n,l,k+1} \circ u_{k+1,k}) \bullet u_{n,k+1,k}.$$

This shows that the map $N_n^k C \to N_n^{k+1}C \times_{C_1} C_2$ is injective.

To see that it is also surjective, we need to check that, given $u' : F_{k+1}\Delta^n \to \mathrm{NC}$ and given $\alpha : u'_{n,k} \Rightarrow u'_{n,k+1}u'_{k+1,k}$, the above equations can be used to define a simplicial map $u : F_k\Delta^n \to \mathrm{NC}$. This translates into showing that for every tetrahedron (l,k,j,i) the above equation holds. The only tetrahedrons that deserve an explanation are those of the type (n,l',l,k) with $k<l<l'<n$. Moreover, if $l=k+1$ then the equation holds by the construction of u. So let us assume that $k+1<l$. The 4-simplex corresponding to $(n,l',l,k+1,k)$ yields a cube:



We want to see that the back right face commutes. But we know that: the back left face commutes by definition of $u_{n,l,k}$; the upper face commutes by definition of $u_{n,l',k}$; the left front face commutes for it factors through $u_{k+1,k}$; the right front face commutes for it factors through $u_{n,l'}$; and the bottom face commutes for $\circ$ and $\bullet$ are mutually distributive. Hence the result. □

We can now prove our second main theorem.

Theorem 6.3. *Given a Lie* 2*-category* C *whose* 2*-arrows are* (*smoothly*) *invertible*, *the nerve* NC *is naturally a simplicial manifold.*

Proof. We endow each N_nC with a smooth structure inductively. For $n=0$, 1 we do this by means of the obvious identifications $N_0C = C_0$ and $N_1C = C_1$. For larger n we use the filtration and fiber products of the previous proposition, noting that one of the maps is always a surjective submersion, and using the standard transversality criterion. Hence N_nC is a closed embedded submanifold of the product

$$N_nC \subset N_{n-1}C \times \prod_{(i+1,i)} C_1 \times \prod_{(n,i+1,i)} C_2.$$

We will prove that, for these smooth structures, the canonical inclusion $\phi : N_nC \to A_nC$ into the ambient is a closed embedding. This implies that (i) the smooth structures that we have defined on N_nC do not depend on the particular filtration we have used, and that (ii) the simplicial maps on NC are smooth and NC is a simplicial manifold.

For each triple (k, j, i), we have to show that the composition

$$\phi_{k,j,i} = \pi_{k,j,i}\phi : N_nC \to A_nC \to C_2$$

is smooth. By projecting on the first coordinate of the above product, and using an inductive argument, we solve the case $n > k$. By projecting on the other coordinates we solve the cases $(n, i+1, i)$. It remains to study the other projections $\phi_{n,j,i}$. But such a projection can be written as an expression involving the other coordinates and the multiplications $\circ$ and $\bullet$, which are smooth. A similar argument applies also to the degenerate coordinates. □

It follows from our theorem that the nerve of a Lie 2-groupoid is a simplicial manifold, and that a smooth pseudofunctor $\phi : G \dashrightarrow G'$ is the same as a simplicial smooth map $\phi : \mathrm{NG} \to \mathrm{NG}'$. Next we present a less immediate corollary.

Corollary 6.4. *With the above hypothesis*, *the face maps* $d_i : N_nC \to N_{n-1}C$ *are surjective submersions.*

Proof. This is more a corollary of the proof rather than of the statement. When $i=n$ it follows by factoring d_n through the filtration, for each projection $N_n^kC \to N_n^{k+1}C$ is the base-change of a surjective submersion, as well as $N_n^{n-1}C \to N_{n-1}C$. When

$i \neq n$ we can argue similarly, but now using a different filtration of Δ^n, by complexes containing the face $d_i(\Delta^{n-1})$. □

We finish this section by developing a smooth version of Proposition 3.4, setting a bridge between our theory and that of weak Lie 2-categories and weak Lie 2-groupoids, as defined in [Henriques 2008; Zhu 2009]. A simplicial manifold X is a *weak Lie m-category* or a *weak Lie m-groupoid* if the corresponding restrictions maps $X_n \to X_{n,k}$ are surjective submersions, for some reasonable smooth structure on the space of (n,k)-horns. The space of horns $X_{n,k}$ can be expressed as an equalizer

$$\prod_{i\neq k} X_{n-1} \rightrightarrows \prod_{i,j\neq k} X_{n-1},$$

which may not exist in the category of manifolds. In general this is proved by an inductive argument. In our case, when $X = \mathrm{N}C$ is the nerve of a Lie 2-category with invertible 2-arrows, it follows from our construction that $X_n \to \prod_{i\neq k} X_{n-1}$ is a closed embedded submanifold for $n > 3$ and for $n = 3,\ k = 2$. The case $n = 3$, $k = 1$, follows by using a symmetric filtration on the simplex. Therefore, since X_n is also a set-theoretic equalizer, we conclude that $X_n \cong X_{n,k}$ is a diffeomorphism in these cases. The case $n = 2$ is easy, and therefore we can conclude:

Proposition 6.5. *Let C be a Lie 2-category on which every 2-arrow is invertible. Then* $\mathrm{N}C$ *is a weak Lie 2-category. Moreover,* $\mathrm{N}C$ *is a weak Lie 2-groupoid if and only if C is a Lie 2-groupoid.*

Remark 6.6. The main theorem on [Mehta and Tang 2011] shows that if G is a strict Lie 2-groupoid then TN^2G is a weak Lie 2-groupoid. Thus, in light of the isomorphism described in Remark 3.5, our theorem can be regarded as an extension of that to a nonstrict Lie 2-groupoid. This is crucial for us, for our fundamental example $\mathrm{GL}(V)$ is not strict.

7. Representations as pseudofunctors

In this section we review the notion of representation up to homotopy $G \curvearrowright V$ of a Lie groupoid G, the particular case of 2-term vector bundles $V = V_1 \oplus V_0$, and present our main theorem, stating a one-to-one correspondence between representations $G \curvearrowright V$ and pseudofunctors $G \dashrightarrow \mathrm{GL}(V)$.

Given a Lie groupoid $G \rightrightarrows M$ and a vector bundle $E \to M$, a *representation* $G \curvearrowright E$ can be defined as a map $\rho : G \times_M E \to E$, $\rho(y \xleftarrow{g} x, e) = \rho_g(e)$, such that (i) $\rho_g : E_x \to E_y$ is linear, (ii) $\rho_{\mathrm{id}} = \mathrm{id}$, and (iii) $\rho_h\rho_g = \rho_{hg}$. A *pseudorepresentation* is a sort of nonassociative action; it is defined analogously but just requiring (i) and (ii).

Example 7.1. If $G \rightrightarrows *$ is a Lie group, viewed as a Lie groupoid with a single object, then its representations are the usual ones. If $M \rightrightarrows M$ is a manifold, viewed as a Lie groupoid with only identities arrows, then its representations are the vector bundles over M. More generally, if $G \times M \rightrightarrows M$ is the groupoid arising from a Lie group action $G \curvearrowright M$, then a representation $(G \times M) \curvearrowright E$ is the same as an equivariant vector bundle.

Example 7.2. Given a manifold M, a representation $\mathrm{Pair}(M) \curvearrowright E$ of its pair groupoid is the same as a trivialization of E. Given a surjective submersion $q : M \to N$, a representation $M \times_N M \curvearrowright E$ of the submersion groupoid (see [del Hoyo 2013]) is the same as an isomorphism $E \cong q^*E'$ with a pullback vector bundle. This can be further generalized to a foliation $F \subset \mathrm{TM}$, which yields a holonomy groupoid $\mathrm{Hol}(F) \rightrightarrows M$, whose representations are the same as foliated bundles.

Example 7.3. Let P^2 denote the real projective plane, and let $E \to P^2$ be its tautological line bundle. Since it is not trivial there cannot be a representation of the pair groupoid $\mathrm{Pair}(P^2) \curvearrowright E$. Still, we can define a pseudorepresentation $\mathrm{Pair}(P^2) \curvearrowright E$, by defining for instance $\rho_{(\ell',\ell)}(v)$ as the orthogonal projection of $v \in \ell$ over ℓ'.

By means of the exponential law, a Lie groupoid representation can be described as a Lie groupoid morphism into the *general linear groupoid* (see, e.g., [del Hoyo 2013])

$$\rho^\# : (G \rightrightarrows M) \to (\mathrm{GL}(E) \rightrightarrows M), \qquad \rho^\#(g) = \rho_g,$$

whose objects are the fibers of $E \to M$ and whose arrows are isomorphisms between fibers. In the case of a pseudorepresentation we still have a smooth map $G \to \mathrm{GL}(E)$ between the arrow spaces, compatible with source and target but that may fail to preserve the multiplication. This viewpoint allows one to treat representations as maps, and it is especially useful when dealing with differentiation and integration.

Lie groupoid representations turn out to be very restrictive. A convenient generalization, is that of a *representation up to homotopy* of a Lie groupoid G over a graded vector bundle $V = \bigoplus V_i$. It is defined as a degree 1 differential D on a space of sections $\Gamma(\mathrm{NG}, V)$ of V over the nerve of G inducing a graded module structure. By decomposing $D = \bigoplus D_i$ into bihomogeneous components, we can reinterpret D as a pseudorepresentation over a complex (V, ∂) with homotopies controlling its associativity. See [Arias Abad and Crainic 2013; del Hoyo and Ortiz 2018; Mehta and Tang 2011] for further details. We recall here the 2-term case, the simplest new case, using an homological convention.

Proposition 7.4 [del Hoyo and Ortiz 2018; Gracia-Saz and Mehta 2017]. *If* $V = V_1 \oplus V_0$, *then a representation up to homotopy* $G \curvearrowright V$ *is the same as a tuple* $(\partial, \rho_1, \rho_0, \gamma)$, *where* $\partial : V_1 \to V_0$ *is a linear map*, $\rho_i : G \curvearrowright V_i$ *are pseudorepresentations commuting with* ∂, *and*

$$\gamma : (z \xrightarrow{h} y \xrightarrow{g} x) \mapsto (\gamma^{h,g} : \rho^{hg} \Rightarrow \rho^h \rho^g)$$

is a curvature tensor satisfying

$$\rho_1^{g_3}\circ\gamma^{g_2,g_1}-\gamma^{g_3g_2,g_1}+\gamma^{g_3,g_2g_1}-\gamma^{g_3,g_2}\circ\rho_0^{g_1}=0.$$

A morphism $\theta : V \to V'$ is the same as a triple $(\theta_1, \theta_0, \mu)$ where $\theta = (\theta_1, \theta_0) : V \to V'$ is a vector bundle chain map and $\mu : (y \xrightarrow{g} x) \mapsto (\mu^g : V_0^x \to V_1'^y)$ is a tensor satisfying $\rho'\theta - \theta\rho = \partial'\mu + \mu\partial$, and

$$\theta_1^z\gamma^{h,g}+\mu^h\rho_0^g+\rho'^h_1\mu_g-\mu^{hg}-\gamma^{h,g}\theta_0^x=0.$$

The *point-wise homology* of a 2-term representation $G \curvearrowright V$ consists of $H_1^x(V) = \ker\partial^x$ and $H_0^x(V) = \operatorname{coker}\partial^x$. If the rank of ∂ is constant then $H_1(V)$ and $H_0(V)$ are vector bundles and there is an induced representation over them. A representation up to homotopy V whose point-wise homology vanishes is called *acyclic*. A morphism $\theta : V \to W$ of 2-term representations up to homotopy inducing isomorphisms on the point-wise homology is called a *quasi-isomorphism*.

Example 7.5. for $\rho : \mathrm{Pair}(P^2) \curvearrowright E$ the pseudorepresentation discussed before, we can define an acyclic representation up to homotopy $\mathrm{Pair}(P^2) \curvearrowright E \oplus E$ by setting $\partial = \mathrm{id}$, $\rho_1 = \rho_0 = \rho$ and $\gamma = \rho - \rho\rho$. The same can be done for any pseudorepresentation.

Example 7.6. Given a Lie groupoid $G \rightrightarrows M$ endowed with a connection σ, namely a section of $s : TG \to s^*\mathrm{TM}$, the *adjoint representation* $G \curvearrowright (A \oplus \mathrm{TM})$ has ∂ equal to the anchor map and ρ_0 given by $t\sigma$. The equivalence class does not depend on σ. This generalizes the classical adjoint representation of Lie groups and plays a role in the deformation theory of groupoids. The *coadjoint representation* $G \curvearrowright T^*M \oplus A^*$ is defined by duality.

We are now ready to present our main theorem. Given a Lie groupoid $G \rightrightarrows M$ we have a canonical projection $\pi_G : G \to \mathrm{Pair}(M)$ into the pair groupoid that just remembers the source and target of an arrow. Given a 2-term vector bundle $V \to M$, we have a canonical projection $\pi_V : \mathrm{GL}(V) \to \mathrm{Pair}(M)$ that only remembers the base-points on the vector bundle.

Theorem 7.7. *Given a Lie groupoid $G \rightrightarrows M$ and a graded vector bundle $V = V_1 \oplus V_0 \to M$, there is an equivalence between the category of representations up to homotopy $\rho : G \curvearrowright V$ and quasi-isomorphisms and the category of pseudofunctors $\phi : G \dashrightarrow \mathrm{GL}(V)$ satisfying $\pi_V\phi = \pi_G$ and smooth lax equivalences.*

This result is truly a generalization of the situation for ordinary representations. That is, when V is only in degree 0, then $\mathrm{GL}(V)$ is the usual general linear groupoid, and the pseudofunctors $G \dashrightarrow \mathrm{GL}(V)$ are just morphisms of Lie groupoids.

Proof. This is a direct consequence of the constructions and results collected during our work. In light of the set-theoretical simplicial interpretation in Proposition 3.3, our construction of the general linear 2-groupoid in Theorem 5.5, and our characterization for smooth nerve in Theorem 6.3, a smooth pseudofunctor $\phi : G \dashrightarrow \mathrm{GL}(V)$ is the same as a simplicial map $\phi : \mathrm{N}G \to \mathrm{NGL}(V)$. The degree 0 component ϕ_0 is the same as a differential ∂ on V, the degree 1 component ϕ_1 gives a pseudo-representation ρ on V compatible with ∂, and the degree 2 component ϕ_2 yields a curvature tensor

$$\gamma : (z \xleftarrow{h} y \xleftarrow{g} x) \mapsto (\gamma^{h,g} : \rho^{hg} \Rightarrow \rho^h \rho^g),$$

defining a 2-term representation up to homotopy, as characterized in Proposition 7.4. Similarly, a smooth lax equivalence $H : \phi \Rightarrow \psi : G \dashrightarrow \mathrm{GL}(V)$ consists of smooth maps $M \to \mathrm{GL}(V)_1,\ G \to \mathrm{GL}(V)_2$, corresponding to the components θ and μ of a quasi-isomorphism (see Proposition 7.4). It is straightforward to check that these correspondences between objects and arrows are functorial. □

There are some remarks to be made regarding functoriality. Firstly, even though a quasi-isomorphism $\theta : V \to V$ of representations up to homotopy gives a simplicial homotopy $\mathrm{N}G \times I \to \mathrm{NGL}(V)$, not every such homotopy arises in this way, as can be seen in the proof of Proposition 3.3. Secondly, if we want to consider morphisms $V \to V$ that are not quasi-isomorphisms, then the corresponding lax transformations would involve chain maps that are not within $\mathrm{GL}(V)$. Lastly, since the construction $V \mapsto \mathrm{GL}(V)$ is not functorial, it makes little sense to frame the noninvertible morphism $V \to V'$ between different vector bundles within our theory.

We close this paper by outlining three different problems related to our results, the first related to the infinitesimal picture, the second to the theory of 2-stacks, and the third to higher versions of our results.

Remark 7.8. In [Mehta 2014], an infinitesimal analog to our main theorem was announced. It is commonly accepted that weak higher Lie groupoids and higher Lie algebroids are related by a theory of differentiation and integration, though the details of such a theory are yet to be understood. Within this context, we expect that the differentiation of our general linear 2-groupoid is the object $\mathrm{gl}(V)$ introduced there, and that the differentiation and integration of maps will provide an alternative approach to the integration of 2-term representations up to homotopy, other than that of [Bursztyn et al. 2016].

Remark 7.9. In [del Hoyo and Ortiz 2018], the Morita equivalences of VB-groupoids are discussed. It is proved there that the derived category of VB-groupoids $\mathrm{VB}[G]$ over a fixed base is a Morita invariant, and consequently, the same holds for 2-term representations up to homotopy. This result, from our framework, admits the following interpretation. Our general linear 2-groupoid $\mathrm{GL}(V)$ represents a *differentiable 2-stack*, and the maps into it classify certain VB-groupoids, with prescribed side and core bundle. This should be thought of as an incarnation of the 2-stack Perf_2 appearing in algebraic geometry. Further details demand a better understanding of differentiable 2-stacks, and are postponed to be studied elsewhere.

Remark 7.10. It is natural to expect our results to remain valid on higher degrees. The construction of the general linear groupoid seems suitable to be generalized for more general graded vector bundles. The understanding of pseudofunctors within this context seems to be less clear, though a complete immersion into the simplicial approach would solve this issue. Related to this, a realization of more general representations up to homotopy as higher VB-groupoids is currently being studied [del Hoyo and Trentinaglia ≥ 2019]. Expectations here should be curbed, for even disregarding the smooth and linear structures, such a higher analog for Grothendieck correspondence is still unknown.

Acknowledgements

The authors thank H. Bursztyn, G. Ginot, A. Cabrera, E. Dubuc and M. Anel for many useful conversations that shaped their view on the topic. This collaboration took place during a series of visits of Stefani to IMPA, Rio de Janeiro, where del Hoyo worked as a visiting professor. The authors thank IMPA for this opportunity, and Capes Cofecub and Projeto P.V.E. 88881.030367/2013-01 (CAPES/Brazil) for partial financial support. Del Hoyo was partially supported by grant PQ 303034/2017-3 from National Council for Scientific and Technological Development – CNPq.

References

[Arias Abad and Crainic 2013] C. Arias Abad and M. Crainic, "Representations up to homotopy and Bott's spectral sequence for Lie groupoids", *Adv. Math.* **248** (2013), 416–452. MR Zbl

[Borceux 1994] F. Borceux, *Handbook of categorical algebra, I: Basic category theory*, Encyclopedia of Math. and Its Applications **50**, Cambridge Univ. Press, 1994. MR Zbl

[Bullejos and Cegarra 2003] M. Bullejos and A. M. Cegarra, "On the geometry of 2-categories and their classifying spaces", *K-Theory* **29**:3 (2003), 211–229. MR Zbl

[Bullejos et al. 2005] M. Bullejos, E. Faro, and V. Blanco, "A full and faithful nerve for 2-categories", *Appl. Categ. Structures* **13**:3 (2005), 223–233. MR Zbl

[Bursztyn et al. 2016] H. Bursztyn, A. Cabrera, and M. del Hoyo, "Vector bundles over Lie groupoids and algebroids", *Adv. Math.* **290** (2016), 163–207. MR Zbl

[Crainic et al. 2018] M. Crainic, J. a. N. Mestre, and I. Struchiner, "Deformations of Lie groupoids", *Int. Math. Res. Not.* (published online September 2018). arXiv

[Duskin 2002] J. W. Duskin, "Simplicial matrices and the nerves of weak n-categories, I: Nerves of bicategories", *Theory Appl. Categ.* **9**:10 (2002), 198–308. MR Zbl

[Gracia-Saz and Mehta 2017] A. Gracia-Saz and R. A. Mehta, "$\mathcal{VB}$-groupoids and representation theory of Lie groupoids", *J. Symplectic Geom.* **15**:3 (2017), 741–783. MR Zbl

[Henriques 2008] A. Henriques, "Integrating L_∞-algebras", *Compos. Math.* **144**:4 (2008), 1017–1045. MR Zbl [del Hoyo 2013] M. L. del Hoyo, "Lie groupoids and their orbispaces", *Port. Math.* **70**:2 (2013), 161–209. MR Zbl

[del Hoyo and Ortiz 2018] M. del Hoyo and C. Ortiz, "Morita equivalences of vector bundles", *Int. Math. Res. Not.* (published online June 2018).

[del Hoyo and Trentinaglia ≥ 2019] M. del Hoyo and G. Trentinaglia, "Simplicial vector bundles over Lie groupoids", work in progress.

[Lack 2010] S. Lack, "A 2-categories companion", pp. 105–191 in *Towards higher categories*, edited by J. C. Baez and J. P. May, IMA Vol. Math. Appl. **152**, Springer, 2010. MR Zbl

[Mac Lane 1998] S. Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Math. **5**, Springer, 1998. MR Zbl

[Mehta 2014] R. Mehta, "Representing representations up to homotopy", slides from lecture, Univ. California, Berkeley, 2014, Available at https://tinyurl.com/mehtaucb.

[Mehta and Tang 2011] R. A. Mehta and X. Tang, "From double Lie groupoids to local Lie 2-groupoids", *Bull. Braz. Math. Soc.* (*N.S.*) **42**:4 (2011), 651–681. MR Zbl

[SGA 1 1971] A. Grothendieck, *Revêtements étales et groupe fondamental* (Séminaire de Géométrie Algébrique du Bois Marie 1960–1961), Lecture Notes in Math. **224**, Springer, 1971. MR Zbl

[Sheng and Zhu 2012] Y. Sheng and C. Zhu, "Integration of semidirect product Lie 2-algebras", *Int. J. Geom. Methods Mod. Phys.* **9**:5 (2012), art. id. 1250043. MR Zbl

[Trentinaglia 2010] G. Trentinaglia, "Tannaka duality for proper Lie groupoids", *J. Pure Appl. Algebra* **214**:6 (2010), 750–768. MR Zbl

[Zhu 2009] C. Zhu, "n-groupoids and stacky groupoids", *Int. Math. Res. Not.* **2009**:21 (2009), 4087–4141. MR Zbl

Received July 13, 2017. Revised June 3, 2018.

MATÍAS DEL HOYO
UNIVERSIDADE FEDERAL FLUMINENSE
NITERÓI, RJ
BRAZIL
mldelhoyo@id.uff.br

DAVIDE STEFANI
UNIVERSITÉ PIERRE ET MARIE CURIE
PLACE JUSSIEU
PARIS
FRANCE
davide.stefani@imj-prg.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.59

EQUIVARIANT FORMALITY OF HAMILTONIAN TRANSVERSELY SYMPLECTIC FOLIATIONS

YI LIN AND XIANGDONG YANG

Consider the Hamiltonian action of a compact connected Lie group on a transversely symplectic foliation which satisfies the transverse hard Lefschetz property. We establish an equivariant formality theorem and an equivariant symplectic $d\delta$-lemma in this setting. As an application, we show that if the foliation is also Riemannian, then there exists a natural formal Frobenius manifold structure on the equivariant basic cohomology of the foliation.

1. Introduction

Reinhart [1959b] introduced the basic cohomology of foliations in the late 1950s as a cohomology theory for the space of leaves. It has become one of fundamental topological invariants for foliations, especially for Riemannian foliations. An important subclass of Riemannian foliations are Killing foliations, as any Riemannian foliation on a simply connected manifold is Killing. According to Molino's structure theory [1988], for Killing foliations, the leaf closures are the orbits of leaves under the action of an abelian Lie algebra of transverse Killing fields, called the structural Killing algebra. Goertsches and Töben [2018] introduced the notion of *equivariant basic cohomology*, and used it to study the transverse actions of structural Killing algebras on Killing foliations. Among other things, they proved a Borel type localization theorem, and established the equivariant formality in the presence of a basic Morse–Bott function whose critical set is the union of closed leaves. As a result, they were able to compute the basic Betti number in many concrete examples, and relate the basic cohomology to the dynamical aspects of a foliation.

Let (M, η, g) be a compact K-contact manifold with a Reeb vector field ξ, and let T be the closure of the Reeb flow in the isometry group $\mathrm{Isom}(M, g)$. Then T is a compact connected torus. Moreover, the characteristic Reeb foliation is Killing, with a structural Killing algebra isomorphic to $\mathrm{Lie}(T)/\operatorname{span}\{\xi\}$. It is well known that in this situation a generic component of the contact moment map $\Phi : M \to \mathfrak{t}^*$

Yang is supported by the National Natural Science Foundation of China (grants No. 11571242 and No. 11701051) and the China Scholarship Council.
MSC2010: primary 57S25; secondary 57R91.
Keywords: transversely symplectic foliations, Hamiltonian actions, equivariant formality.

is a Morse–Bott function, whose critical set is the union of closed Reeb orbits. In particular, the results established in [Goertsches and Töben 2018] apply to the transverse actions of the structural Killing algebras on K-contact manifolds, and yield the equivariant formality theorem in this case (see [Goertsches et al. 2012]).

It is noteworthy that the characteristic foliation of the Reeb vector field of a K-contact manifold (M, η, g) is *transversely symplectic*; in addition, the transverse action of the structural Killing algebra is Hamiltonian in the sense of Souriau [1997]. In view of Goertsches and Töben's equivariant formality result on K-contact manifolds, one naturally wonders *if the equivariant formality theorem would continue to hold for a more general class of Hamiltonian actions on transversely symplectic foliations.*

On symplectic manifolds, there are two approaches to proving the Kirwan–Ginzburg equivariant formality theorem. The first approach is Morse theoretic, which works for arbitrary compact Hamiltonian symplectic manifolds (see [Ginzburg 1987; Kirwan 1984]). The second approach is symplectic Hodge theoretic, which needs to assume that the underlying symplectic manifold has the hard Lefschetz property (see [Lin and Sjamaar 2004]). On the upside, it provides an improved version of the equivariant formality theorem, which asserts that any de Rham cohomology class has a canonical equivariant extension.

In an accompanying paper [Lin 2018], the first author extended symplectic Hodge theory to any transversely symplectic manifold with the *transverse s-Lefschetz property*, and established the symplectic $d\delta$-lemma in this framework. In the present article, for Hamiltonian actions of compact connected Lie groups on transversely symplectic foliations, we apply the symplectic Hodge theory to prove the following result.

Theorem 1.1 (Theorem 3.11). *Consider the Hamiltonian action of a compact connected Lie group G on a compact transversely symplectic foliation $(M, \mathcal{F}, \omega)$. Suppose that $(M, \mathcal{F}, \omega)$ satisfies the transverse hard Lefschetz property. Then there is a canonical $S(\mathfrak{g}^*)^G$-module isomorphism from the equivariant basic cohomology $H_G(M, \mathcal{F})$ to $S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F})$.*

It is important to note that on a transversely symplectic foliation, components of a moment map are in general not Morse–Bott functions, unless the action satisfies the so-called *clean condition* discovered by Lin and Sjamaar [2017]. However, a striking feature of our Hodge theoretic approach is that it would continue to work, even when the action is not clean, as long as the transverse hard Lefschetz property is satisfied.

On a compact symplectic manifold with the hard Lefschetz property, Merkulov [1998] established the symplectic $d\delta$-lemma, and used it to produce a formal Frobenius manifold structure on the de Rham cohomology of the symplectic manifold.

Independently, Cao and Zhou [1999; 2000] proved similar results on the ordinary and equivariant de Rham cohomology of Kähler manifolds. For Hamiltonian Lie group actions on transversely symplectic foliations with the transverse hard Lefschetz property, our method yields an equivariant version of the symplectic $d\delta$-lemma on basic forms. As an application of this result, we show that there is a formal Frobenius manifold structure on the equivariant basic cohomology of the foliation (Theorem 4.7). This simultaneously generalizes the constructions of Merkulov and Cao and Zhou.

Transversely symplectic foliations are naturally related to different areas in differential geometry. Reeb characteristic foliations in both contact and cosymplectic geometries are clearly transversely symplectic. Moreover, leaf spaces of transversely symplectic foliations include symplectic orbifolds (in the sense of Satake [1957]) and symplectic quasifolds [Prato 2001] as special examples. In many known cases, transversely symplectic foliations arise as taut Kähler foliations, which are known to have the transverse hard Lefschetz property (see [El Kacimi-Alaoui 1990]). The results proved in this paper apply to these situations, and yield new examples of dGBV-algebras whose cohomologies carry the structure of a formal Frobenius manifold.

This paper is organized as follows. In Section 2 we review symplectic Hodge theory on transversely symplectic foliations. In Section 3, we establish an equivariant formality theorem for the Hamiltonian action of a compact connected Lie group on a transversely symplectic foliation. We also obtain an equivariant version of the symplectic $d\delta$-lemma on transversely symplectic foliations. In Section 4, we show that there exists a formal Frobenius manifold structure on the equivariant basic cohomology of a Hamiltonian transversely symplectic foliation that satisfies the transverse hard Lefschetz property. In Section 5, we present some concrete examples of transversely symplectic foliations, which are also Riemannian, and which satisfy the transverse hard Lefschetz property.

2. Hodge theory on transversely symplectic foliations

In this section, we review the elements of transversely symplectic Hodge theory to set up the stage. We refer to [Brylinski 1988] and [Yan 1996] for general background on symplectic Hodge theory, and to [Lin 2018] for a detailed exposition on symplectic Hodge theory on foliations.

Assume that $\mathcal{F}$ is a foliation on a smooth manifold M of codimension m. Let $\Xi(M)$ be the Lie algebra of smooth vector fields on M, and let $\Xi(\mathcal{F}) \subset \Xi(M)$ be the Lie subalgebra of vector fields which are tangent to the leaves of $\mathcal{F}$. We say that an element $X \in \Xi(M)$ is *foliate*, if $[X, Y] \in \Xi(\mathcal{F})$ for any $Y \in \Xi(\mathcal{F})$. In particular, the set of foliate fields, denoted by $L(M, \mathcal{F})$, is a Lie subalgebra of $\Xi(M)$, since it is the normalizer of $\Xi(\mathcal{F})$ in $\Xi(M)$. A *transverse vector field* is a smooth section of $TM/T\mathcal{F}$ that is induced by a foliate vector field. It is easy to see that the set of transverse fields $l(M,\mathcal{F}) = L(M,\mathcal{F})/\Xi(\mathcal{F})$ also admits a Lie algebra structure with an induced Lie bracket from $L(M,\mathcal{F})$.

The space of *basic forms* on M is defined as follows.

$$\Omega(M,\mathcal{F}) = \big\{\alpha \in \Omega(M) \mid \iota(X)\alpha = \mathcal{L}(X)\alpha = 0, \text{ for all } X \in \Xi(\mathcal{F})\big\}.$$

Since the exterior differential operator d preserves basic forms, we obtain a subcomplex of the de Rham complex $\{\Omega^*(M), d\}$, called the *basic de Rham complex*:

$$\cdots \to \Omega^{k-1}(M,\mathcal{F}) \xrightarrow{d} \Omega^k(M,\mathcal{F}) \xrightarrow{d} \Omega^{k+1}(M,\mathcal{F}) \xrightarrow{d} \cdots.$$

The cohomology of the basic de Rham complex $\{\Omega^*(M,\mathcal{F}), d\}$, denoted by $H(M,\mathcal{F})$, is called the *basic cohomology* of M with respect to the foliation $\mathcal{F}$. If M is connected, then $H^0(M,\mathcal{F}) \cong \mathbb{R}^1$. In general, the group $H^k(M,\mathcal{F})$ may be infinite-dimensional for $k \geq 2$. However, if M is a closed oriented manifold and if $\mathcal{F}$ is a Riemannian foliation, then the basic cohomology is finite-dimensional; moreover, we have either $H^m(M,\mathcal{F}) = 0$ or $H^m(M,\mathcal{F}) = \mathbb{R}$ (see [El Kacimi-Alaoui et al. 1985, Théorème 0.]). In particular, a Riemannian foliation $\mathcal{F}$ on a closed manifold M is said to be *taut* if $H^m(M,\mathcal{F}) = \mathbb{R}$.

Definition 2.1 [Haefliger 1971]. Let $\mathcal{F}$ be a foliation on a smooth manifold M, and let P be the integrable subbundle of TM associated to $\mathcal{F}$. We say that $\mathcal{F}$ is a *transversely symplectic foliation*, if there exists a closed 2-form ω, called the *transversely symplectic form*, such that for each $x \in M$, the kernel of ω_x coincides with the fiber of P at x.

Let $(M,\mathcal{F},\omega)$ be a transversely symplectic foliation of codimension $2n$. The transversely symplectic form ω induces a nondegenerate bilinear paring $B(\cdot,\cdot)$ on $\Omega^p(M,\mathcal{F})$, which in turn gives rise to the *symplectic Hodge star operator* $\star$ on $\Omega^p(M,\mathcal{F})$ as

$$\beta \wedge \star\alpha = B(\alpha,\beta)\frac{\omega^n}{n!},$$

for any $\alpha, \beta \in \Omega^p(M,\mathcal{F})$. The bilinear pairing $B(\cdot,\cdot)$ is symmetric when p is even, and skew-symmetric when p is odd. It follows easily from the definition that

$$\beta \wedge \star\alpha = \star\beta \wedge \alpha, \qquad \star^2 = \mathrm{id}. \tag{1}$$

The transpose operator δ of d is defined by

$$\delta : \Omega^p(M,\mathcal{F}) \to \Omega^{p-1}(M,\mathcal{F}), \qquad \alpha \mapsto (-1)^{p+1} \star d \star \alpha.$$

By definition, it is easy to see that the operator δ satisfies the equations $\delta^2 = 0$ and $d\delta + \delta d = 0$. In this context, a basic form α is called (symplectic) *harmonic* if it satisfies $d\alpha = \delta\alpha = 0$. Set

$$\Omega_{\text{har}}(M,\mathcal{F}) = \{\alpha \in \Omega(M,\mathcal{F}) \mid d\alpha = \delta\alpha = 0\}.$$

There are three important operators acting on the space of basic forms:

$$\begin{aligned} L &: \Omega^*(M,\mathcal{F}) \to \Omega^{*+2}(M,\mathcal{F}), && \alpha \mapsto \alpha \wedge \omega, \\ \Lambda &: \Omega^*(M,\mathcal{F}) \to \Omega^{*-2}(M,\mathcal{F}), && \alpha \mapsto \star L \star \alpha, \\ H &: \Omega^k(M,\mathcal{F}) \to \Omega^k(M,\mathcal{F}), && \alpha \mapsto (n-k)\alpha. \end{aligned}$$

In particular, we have the following result.

Lemma 2.2. *Let f be a basic function, and X a foliate vector field such that $\iota(X)\omega = df$. Then for any basic form α we have*:

(i) $[\Lambda, \iota(X)]\alpha = 0$.

(ii) $\delta(f\alpha) = f\delta\alpha - \iota(X)\alpha$.

(iii) $\delta(df \wedge \alpha) = -df \wedge \delta\alpha + \mathcal{L}(X)\alpha$.

Proof. The assertion (i) is a direct consequence of [Lin 2018, Lemma 3.2], and (ii) can be proved by the same argument as the one used in [Lin and Sjamaar 2004, Proposition 2.5]. It remains to check the assertion (iii). Using (ii) and the identity $d\delta + \delta d = 0$, we have

$$\begin{aligned} \delta(df \wedge \alpha) &= \delta(d(f\alpha) - f d\alpha) \\ &= -d\delta(f\alpha) - \delta(f d\alpha) \\ &= -d(f\delta\alpha - \iota(X)\alpha) - f\delta d\alpha + \iota(X) d\alpha \\ &= -d(f\delta\alpha) - f\delta d\alpha + (d\iota(X) + \iota(X)d)\alpha \\ &= -df \wedge \delta\alpha - f(d\delta + \delta d)\alpha + \mathcal{L}(X)\alpha \\ &= -df \wedge \delta\alpha + \mathcal{L}(X)\alpha. \end{aligned}$$

This proves the assertion (iii). □

A straightforward calculation yields the following commutator relations.

Proposition 2.3 (see [Lin 2018, Lemma 3.2]).

$$[L,d] = 0, \quad [\Lambda,d] = \delta, \quad [\Lambda,\delta] = 0, \quad [L,\delta] = -d,$$
$$[L,\Lambda] = H, \quad [H,L] = -2L, \quad [H,\Lambda] = 2\Lambda.$$

Definition 2.4. Let $(M,\mathcal{F},\omega)$ be a transversely symplectic foliation of codimension $2n$. We say that M satisfies the *transverse hard Lefschetz property*, if for any $0 \le k \le n$, the map

$$L^k : H^{n-k}(M,\mathcal{F}) \to H^{n+k}(M,\mathcal{F})$$

is an isomorphism.

On compact symplectic manifolds, Brylinski [1988] conjectured that every de Rham cohomology class has a symplectic harmonic representative. However, Mathieu [1995] proved that this conjecture is true if and only if the manifold satisfies the hard Lefschetz property. Mathieu's theorem was sharpened by Merkulov [1998] and Guillemin [2001], who independently established the symplectic $d\delta$-lemma. The symplectic $d\delta$-lemma was first extended to transversely symplectic flows by Zhenqi He [2010], and more recently, by the first author [Lin 2018] to arbitrary transversely symplectic foliations. The following results are reformulations of [Lin 2018, Theorems 4.1 and 4.8].

Theorem 2.5. *Let $(M, \mathcal{F}, \omega)$ be a transversely symplectic foliation with the transverse hard Lefschetz property. Then every basic cohomology class has a symplectic harmonic representative.*

Theorem 2.6. *Assume that $(M, \mathcal{F}, \omega)$ is a transversely symplectic foliation that satisfies the transverse hard Lefschetz property. Then on the space of basic forms,*

$$\operatorname{im} d \cap \ker \delta = \ker d \cap \operatorname{im} d = \operatorname{im} d\delta.$$

Let $\Omega_\delta(M, \mathcal{F}) = \ker \delta \cap \Omega(M, \mathcal{F})$. Since d anticommutes with δ, the subspace $\Omega_\delta(M, \mathcal{F})$ forms a subcomplex of the basic de Rham complex $\{\Omega(M, \mathcal{F}), d\}$, the cohomology of which we denote by $H_\delta(M, \mathcal{F})$. The following result is a direct consequence of Theorem 2.6. Here $H(\Omega(M, \mathcal{F}), \delta)$ denotes the homology of $\Omega(M, \mathcal{F})$ with respect to δ.

Theorem 2.7. *Assume that $(M, \mathcal{F}, \omega)$ is a transversely symplectic foliation that satisfies the transverse hard Lefschetz property. Then the d-chain maps in the diagram*

$$\Omega(M, \mathcal{F}) \longleftarrow \Omega_\delta(M, \mathcal{F}) \longrightarrow H(\Omega(M, \mathcal{F}), \delta)$$

are quasi-isomorphisms that induce isomorphisms in cohomology.

3. Equivariant formality and basic $d_G\delta$-lemma

In this section we study the equivariant basic cohomology of Hamiltonian actions on transversely symplectic foliations using the Hodge theoretic approach. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Recall that a *transverse action* of $\mathfrak{g}$ on a foliated manifold $(M, \mathcal{F})$ is defined to be a Lie algebra homomorphism $\mathfrak{g} \to l(M, \mathcal{F})$ (see [Goertsches and Töben 2018, Definition 2.1]). We propose the following definition of transverse actions of a Lie group G.

Definition 3.1. Consider the action of a Lie group G with the Lie algebra $\mathfrak{g}$ on a foliated manifold $(M, \mathcal{F})$. We say that the action of G is *transverse* if the image of the associated infinitesimal action map $\mathfrak{g} \to \Xi(M)$ lies in $L(M, \mathcal{F})$.

Remark 3.2. Suppose that there is a transverse action of a Lie group G with Lie algebra $\mathfrak{g}$ on a foliated manifold $(M, \mathcal{F})$. Then by definition we have the following commutative diagram of Lie algebra homomorphisms:

$$\begin{array}{ccc} \mathfrak{g} & \longrightarrow & L(M,\mathcal{F}) \\ & \searrow & \Big\downarrow \text{pr} \\ & & l(M,\mathcal{F}) \end{array}$$

Here the vertical map is the natural projection. Therefore we also have a transverse $\mathfrak{g}$-action on $(M, \mathcal{F})$ in the sense of [Goertsches and Töben 2018, Definition 2.1].

Lemma 3.3. *Consider the transverse action of a compact connected Lie group G on a foliated manifold $(M, \mathcal{F})$. If α is a basic form, and if X_M is a fundamental vector field induced by an element $X \in \mathfrak{g}$, then $\iota(X_M)\alpha$ and $\mathcal{L}(X_M)\alpha$ are also basic forms.*

Proof. Let $Y \in \Xi(\mathcal{F})$. Since the action of G is transverse, we get $[Y, X_M] \in \Xi(\mathcal{F})$. It follows that

$$\iota(Y)\iota(X_M)\alpha = -\iota(X_M)\iota(Y)\alpha = 0,$$

and that

$$\mathcal{L}(Y)\iota(X_M)\alpha = \iota([Y, X_M])\alpha + \iota(X_M)\mathcal{L}(Y)\alpha = 0.$$

This proves that $\iota(X_M)\alpha$ is a basic form. A similar calculation shows that $\mathcal{L}(X_M)\alpha$ is also basic. □

Suppose that there is a transverse action of a compact connected Lie group G on a foliated manifold $(M, \mathcal{F})$. As an immediate consequence of Lemma 3.3, we see that $\Omega(M, \mathcal{F})$ is a $G^\star$-module in the sense of [Guillemin and Sternberg 1999, Definition 2.3.1]. Therefore, there is a well defined Cartan model of $\Omega(M, \mathcal{F})$ given by

$$\Omega_G(M, \mathcal{F}) := [S(\mathfrak{g}^*) \otimes \Omega(M, \mathcal{F})]^G,$$

which we call the *equivariant basic Cartan complex*.

To simplify the notation, let us write $\Omega_{\text{bas}} = \Omega(M, \mathcal{F})$, and $\Omega_{G,\text{bas}} = \Omega_G(M, \mathcal{F})$. Elements of $\Omega_{G,\text{bas}}$ can be regarded as equivariant polynomial maps from $\mathfrak{g}$ to Ω_{bas}, and are called *equivariant basic differential forms* on M. The equivariant basic Cartan model $\Omega_{G,\text{bas}}$ has a bigrading given by

$$\Omega^{i,j}_{G,\text{bas}} = [S^i(\mathfrak{g}^*) \otimes \Omega^{j-i}_{\text{bas}}]^G;$$

moreover, it is equipped with the vertical differential $1 \otimes d$, which we abbreviate to d, and the horizontal differential ∂, which is defined by

$$\partial(\alpha(\xi)) = -\iota(\xi)\alpha(\xi), \quad \text{for all } \xi \in \mathfrak{g}.$$

Here $\iota(\xi)$ denotes the inner product with the fundamental vector field on M induced by $\xi \in \mathfrak{g}$. As a single complex, $\Omega_{G,\text{bas}}$ has a grading given by

$$\Omega^k_{G,\text{bas}} = \bigoplus_{i+j=k} \Omega^{i,j}_{G,\text{bas}},$$

and a total differential $d_G = d + \partial$, which is called the equivariant exterior differential. We say that an equivariant differential basic form α is equivariantly closed, resp., equivariantly exact, if $d_G \alpha = 0$, resp. $\alpha = d_G \beta$ for some equivariant basic form β.

Definition 3.4. The *equivariant basic cohomology* of the transverse G-action on $(M, \mathcal{F})$ is defined to be the total cohomology of the equivariant basic Cartan complex $\{\Omega_G(M, \mathcal{F}), d_G\}$, which is denoted by $H_G(M, \mathcal{F})$.

We would like to point out that the above definition of equivariant basic cohomology was first introduced by Goertsches and Töben [2018] using the language of equivariant cohomology of $\mathfrak{g}^\star$-algebras. Following Goresky, Kottwitz and MacPherson [Goresky et al. 1998], we propose the following definition of equivariant formality for transverse G-actions.

Definition 3.5. A transverse G-action on $(M, \mathcal{F})$ is *equivariantly formal* if

$$H_G(M, \mathcal{F}) \cong S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F})$$

as graded $S(\mathfrak{g}^*)^G$-modules.

Next, we review the notion of Hamiltonian G-actions on transversely symplectic foliations.

Definition 3.6 [Lin and Sjamaar 2017]. Consider the action of a compact connected Lie group G with the Lie algebra $\mathfrak{g}$ on a transversely symplectic foliation $(M, \mathcal{F}, \omega)$. We say that the G-action on $(M, \mathcal{F}, \omega)$ is *Hamiltonian*, if the G-action preserves the transversely symplectic form ω, and if there exists an equivariant map,

$$\Phi : M \to \mathfrak{g}^*,$$

called a moment map, such that $d\langle \Phi, \xi \rangle = \iota(\xi)\omega$, for each $\xi \in \mathfrak{g}$. Here $\langle \cdot, \cdot \rangle$ denotes the dual pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$.

Remark 3.7. By definition, the Hamiltonian action of a Lie group G on a transversely symplectic manifold $(M, \mathcal{F}, \omega)$ is always transverse. Indeed, since the action preserves the transversely symplectic form ω, it also preserves its null foliation $\mathcal{F}$. It then follows from [Molino 1988, Proposition 2.2] that the G-action must be transverse.

From now on, we assume that $(M, \mathcal{F}, \omega)$ is a compact transversely symplectic foliation that satisfies the transverse hard Lefschetz property, and that there is a compact connected Lie group G acting on $(M, \mathcal{F}, \omega)$ in a Hamiltonian fashion with a moment map $\Phi : M \to \mathfrak{g}^*$, where $\mathfrak{g} = \mathrm{Lie}(G)$. The symplectic Hodge theory gives rise to a third differential $1 \otimes \delta$ on $\Omega_{G,\mathrm{bas}}$, which we will abbreviate to δ.

Lemma 3.8. *On the space of equivariant basic differential forms $\Omega_{G,\mathrm{bas}}$, the following identities hold*:

$$\partial\delta = -\delta\partial, \quad d_G\delta = -\delta d_G.$$

Proof. It was shown in [Lin and Sjamaar 2004, Lemma 3.1] that $\partial\delta = -\delta\partial$ and $d_G\delta = -\delta d_G$ hold on the space of equivariant differential forms. Since d_G, δ and ∂ map basic forms to basic forms, these two identities also hold on the space of equivariant basic differential forms. □

This implies that $\Omega^{\delta}_{G,\mathrm{bas}} := \ker\delta \cap \Omega_{G,\mathrm{bas}}$ is a *double subcomplex* of $\Omega_{G,\mathrm{bas}}$, and that the homology $H(\Omega_{G,\mathrm{bas}}, \delta)$ with respect to δ is a double complex with the differentials induced by d and ∂. Thus we have a diagram of morphisms of double complexes

$$\Omega_{G,\mathrm{bas}} \longleftarrow \Omega^{\delta}_{G,\mathrm{bas}} \longrightarrow H(\Omega_{G,\mathrm{bas}}, \delta). \tag{2}$$

Since δ acts trivially on the polynomial part, these morphisms in (2) are actually morphisms of $S(\mathfrak{g}^*)^G$-modules.

We first establish a preliminary result about the action of $\iota(\xi)$ on invariant basic forms. Let Ω^G_{bas} be the space of G-invariant basic forms on M. The Cartan's identity

$$\mathcal{L}(\xi) = \iota(\xi)d + d\iota(\xi)$$

implies that the morphism $\iota(\xi) : \Omega^G_{\mathrm{bas}} \to \Omega^G_{\mathrm{bas}}$ is a chain map with respect to d. Here $\mathcal{L}(\xi)$ denotes the Lie derivative of the fundamental vector field on M induced by $\xi \in \mathfrak{g}$. Similarly, an application of the identity $\delta\partial + \partial\delta = 0$ to the zeroth column of $\Omega_{G,\mathrm{bas}}$ implies that $\iota(\xi)$ is a chain map with respect to δ.

Lemma 3.9. *Let $\xi \in \mathfrak{g}$ and $\alpha \in \Omega^G_{\mathrm{bas}}$. If α is d-closed, then $\iota(\xi)\alpha$ is d-exact. If α is δ-closed, then $\iota(\xi)\alpha$ is δ-exact.*

Proof. Since the action of G is Hamiltonian, it follows from [Lin and Sjamaar 2004, Proposition 2.5] that

$$\iota(\xi)\alpha = \Phi^{\xi}(\delta\alpha) - \delta(\Phi^{\xi}\alpha), \tag{3}$$

where Φ^{ξ} is the ξ-component of the moment map $\Phi : M \to \mathfrak{g}^*$. If α is δ-closed, then we have that $\iota(\xi)\alpha = -\delta(\Phi^{\xi}\alpha)$. Since Φ^{ξ} is a basic function, we get that $\iota(\xi)\alpha$ is δ-exact in Ω^G_{bas}.

It remains to show that if $\alpha \in \Omega^G_{\mathrm{bas}}$ is a d-closed basic k-form, then $\iota(\xi)\alpha$ is d-exact. Since M satisfies the transverse hard Lefschetz property, by [Lin 2018, Theorem 4.3], for each class $[\alpha] \in H^k(M, \mathcal{F})$ there exists a unique primitive decomposition

$$[\alpha] = \sum_r L^r[\alpha_r].$$

Here $[\alpha_r] \in H^{k-2r}(M, \mathcal{F})$ is a primitive basic cohomology class, i.e., $L^{n-k+2r+1}[\alpha]$ is equal to 0. However, since the action is Hamiltonian, we have

$$\iota(\xi)(\omega \wedge \alpha) = d\Phi^{\xi} \wedge \alpha + \omega \wedge \iota(\xi)\alpha.$$

Thus to finish the proof, it suffices to show that $\iota(\xi)\alpha$ is exact when $[\alpha]$ is a primitive basic cohomology class. We note that the argument given in [Lin and Sjamaar 2004, Lemma 3.2] continues to hold in the present situation to show the exactness of $\iota(\xi)\alpha$. □

Note that the symplectic $d\delta$-lemma, Theorem 2.6, holds for equivariant basic differential forms as well as for ordinary basic differential forms. In particular, the inclusion $\Omega_{\mathrm{bas}}^{G} \hookrightarrow \Omega_{\mathrm{bas}}$ is a deformation retraction for δ as well as for d. The same argument as given in the proof of [Lin and Sjamaar 2004, Lemma 3.3.] provides us the following result.

Lemma 3.10. *The differentials induced by d and ∂ on $H(\Omega_{G,\mathrm{bas}}, \delta)$ are* 0. *Moreover*, *we have the isomorphism*

$$H(\Omega_{G,\mathrm{bas}}, \delta) \cong S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F}). \tag{4}$$

We are now in a position to prove the equivariant formality property of Hamiltonian actions on transversely symplectic foliations.

Theorem 3.11. *Let $(M, \mathcal{F}, \omega)$ be a compact transversely symplectic manifold that satisfies the transverse hard Lefschetz property*, *and let a compact connected Lie group G act on M in a Hamiltonian fashion. Then the morphisms in* (2) *induce isomorphisms of $S(\mathfrak{g}^*)^G$-modules*

$$H_G(M, \mathcal{F}) \xleftarrow{\cong} H(\Omega^{\delta}_{G,\mathrm{bas}}, d_G) \xrightarrow{\cong} H(\Omega_{G,\mathrm{bas}}, \delta).$$

Proof. We first note that since G is connected, the identity $\mathcal{L}(\xi) = d\iota(\xi) + \iota(\xi)d$ together with the identity (3) imply that G acts trivially on both $H(M, \mathcal{F})$ and $H(\Omega(M, \mathcal{F}), \delta)$. Let E be the spectral sequence of $\Omega_{G,\mathrm{bas}}$ relative to the filtration associated to the horizontal grading and E_δ that of $\Omega^{\delta}_{G,\mathrm{bas}}$. The first terms are

$$E_1 = \ker d / \operatorname{im} d = [S(\mathfrak{g}^*) \otimes H(M, \mathcal{F})]^G = S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F}) \tag{5}$$

$$\begin{aligned} (E_\delta)_1 &= (\ker d \cap \ker \delta)/(\operatorname{im} d \cap \ker \delta) \\ &= [S(\mathfrak{g}^*) \otimes H(\Omega(M, \mathcal{F}), \delta)]^G = S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F}). \end{aligned} \tag{6}$$

Here we used the observation we made in the paragraph right before Lemma 3.10, as well as the isomorphism $H(\Omega(M, \mathcal{F}), \delta) \cong H(M, \mathcal{F})$ of Theorem 2.7. By Lemma 3.10, $H(\Omega_{G,\mathrm{bas}}, \delta)$ is a trivial double complex, its spectral sequence is therefore constant with trivial differentials at each stage. The two morphisms in (2) induce morphisms of spectral sequences,

$$E \longleftarrow E_\delta \longrightarrow H(\Omega_{G,\mathrm{bas}}, \delta).$$

It follows from (4), (5) and (6) that these morphisms induce isomorphisms at the first stage. Thus they must induce isomorphisms at every stage. In particular, these three spectral sequences converge to the same limit, and so the morphisms in (2) induce isomorphisms on total cohomology. This completes the proof. □

An argument similar to the one used in [Lin and Sjamaar 2004, Theorem 3.9] gives us the following equivariant version of the symplectic $d\delta$-lemma on transversely symplectic manifolds.

Theorem 3.12. *Let $\alpha \in \Omega_{G,\mathrm{bas}}$ be an equivariant basic form satisfying $d_G\alpha = 0$ and $\delta\alpha = 0$. If α is either d_G-exact or δ-exact, then there exists $\beta \in \Omega_{G,\mathrm{bas}}$ such that $\alpha = d_G\delta\beta$.*

We now discuss the implications of Theorem 3.11. Observe that $\Omega^{0,k}_{G,\mathrm{bas}} = (\Omega^k_{\mathrm{bas}})^G$, the space of G-invariant basic k-forms on M. Thus the zeroth column of the basic Cartan model is the G-invariant basic de Rham complex Ω^G_{bas}, which is a deformation retraction of the basic de Rham complex because G is connected. Therefore, we have an isomorphism $H(\Omega^G_{\mathrm{bas}}) \cong H(M, \mathcal{F})$. The natural projection map $\bar{p} : \Omega_{G,\mathrm{bas}} \to \Omega^G_{\mathrm{bas}}$, defined by $\bar{p}(\alpha) = \alpha(0)$, is a chain map with respect to the equivariant exterior derivative d_G on $\Omega_{G,\mathrm{bas}}$ and the ordinary exterior derivative d on Ω_{bas}. It induces a morphism of cohomology groups $p : H_G(M, \mathcal{F}) \to H(M, \mathcal{F})$. Theorem 3.11 implies that the spectral sequence E degenerates at the first stage, and that the map p is surjective. In other words, every basic cohomology class can be extended to an equivariant basic cohomology class. However, Theorem 3.11 would also imply that there is a canonical choice of such an extension. Let

$$s : H(M, \mathcal{F}) \to H_G(M, \mathcal{F}) \tag{7}$$

be the composition of the map

$$H(M, \mathcal{F}) \to S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F})$$

which sends a cohomology class a to $1 \otimes a$, and the isomorphism

$$S(\mathfrak{g}^*)^G \otimes H(M, \mathcal{F}) \to H_G(M, \mathcal{F})$$

as given by Theorem 3.11. The following result is a direct consequence of Theorems 2.7 and 3.11.

Corollary 3.13. *The map s is a section of p. Thus every basic cohomology class can be extended to an equivariant basic cohomology class in a canonical way.*

Proof. For details of the proof see [Lin and Sjamaar 2004, Corollary 3.5]. □

4. Formal Frobenius manifolds modeled on equivariant basic cohomology

Consider the Hamiltonian action of a compact connected Lie group on a transversely symplectic foliation. In this section, following the approach initiated by Barannikov and Kontsevich [1998], we show that if the foliation satisfies the transverse hard Lefschetz property, and if it is also a Riemannian foliation, then there exists a formal Frobenius manifold structure on its equivariant basic cohomology.

dGBV *algebra in transversely symplectic geometry.* We first give a quick review of *differential Gerstenhaber–Batalin–Vilkovisky* (dGBV) *algebra*. Suppose $(\mathscr{A}, \wedge)$ is a supercommutative graded algebra with identity over a field k, and that there is a k-linear operator $\delta : \mathscr{A}^* \to \mathscr{A}^{*-1}$. Define the bracket $[\bullet]$ by setting

$$[a \bullet b] = (-1)^{|a|}\big(\delta(a \wedge b) - (\delta a) \wedge b - (-1)^{|a|} a \wedge (\delta b)\big),$$

where a and b are homogeneous elements and $|a|$ is the degree of $a \in \mathscr{A}$. We say that $(\mathscr{A}, \wedge, \delta)$ forms a *Gerstenhaber–Batalin–Vilkovisky* (GBV) *algebra* with odd bracket $[\bullet]$ if it satisfies:

(i) δ is a differential, i.e., $\delta^2 = 0$.

(ii) For any homogeneous elements a, b and c we have

$$[a \bullet (b \wedge c)] = [a \bullet b] \wedge c + (-1)^{(|a|+1)|b|} b \wedge [a \bullet c]. \tag{8}$$

Definition 4.1. A GBV-algebra $(\mathscr{A}, \wedge, \delta)$ is called a dGBV-*algebra*, if there exists a differential operator $d : \mathscr{A}^* \to \mathscr{A}^{*+1}$ such that

(i) d is a derivation with respect to the product $\wedge$, i.e.,

$$d(a \wedge b) = da \wedge b + (-1)^{|a|} a \wedge db$$

for any homogeneous elements a and b;

(ii) $d\delta + \delta d = 0$.

An *integral* on a dGBV algebra $\mathscr{A}$ is a k-linear functional

$$\int : \mathscr{A} \to k \tag{9}$$

such that for all $a, b \in \mathscr{A}$, the following equations hold:

$$\int (da) \wedge b = (-1)^{|a|+1} \int a \wedge db,$$
$$\int (\delta a) \wedge b = (-1)^{|a|} \int a \wedge \delta b.$$

Moreover, an integral $\int$ induces a bilinear pairing on $H(\mathscr{A}, d)$ as follows:

$$(\cdot\,,\cdot) : H(\mathscr{A}, d) \times H(\mathscr{A}, d) \to k, \quad ([a], [b]) = \int a \wedge b.$$

In particular, if the above bilinear pairing is nondegenerate, then we say that the integral is *nice*.

The following theorem enables us to use a dGBV algebra as an input to produce a formal Frobenius manifold (see [Barannikov and Kontsevich 1998; Manin 1999]).

Theorem 4.2. *Let* $(\mathscr{A}, \wedge, \delta, d, [\bullet])$ *be a* dGBV *algebra satisfying the following conditions*:

(1) *The dimension of* $H(\mathscr{A}, d)$ *is finite.*

(2) *There exists a nice integral on* $\mathscr{A}$.

(3) *The inclusions* $(\ker \delta, d) \hookrightarrow (\mathscr{A}, d)$ *and* $(\ker d, \delta) \hookrightarrow (\mathscr{A}, \delta)$ *are quasi-isomorphisms.*

Then there is a canonical construction of a formal Frobenius manifold structure on $H(\mathscr{A}, d)$.

As an initial step, we first prove that the equivariant basic Cartan complex of a transversely symplectic manifold carries the structure of a dGBV algebra.

Proposition 4.3. *Suppose that there is a transverse action of a compact connected Lie group* G *on a transversely symplectic manifold* $(M, \mathcal{F}, \omega)$. *Let* δ *be the differential on equivariant basic differential forms as introduced in Section 3, and let* $\wedge$ *denote the wedge product. Then the quadruple* $(\Omega_{G,\mathrm{bas}}, \wedge, \delta, d_G)$ *is a* dGBV *algebra.*

Proof. The only thing that requires a proof is that (8) holds on equivariant basic differential forms. To this end, it suffices to show that (8) holds for ordinary basic differential forms a, b, c on a foliated coordinate neighborhood. So without loss of generality, we may assume that $b = f_0 df_1 \wedge \cdots \wedge df_k$, and that for each $0 \le i \le k$, f_i is a basic functions such that $df_i = \iota(X_i)\omega$ for some foliate vector field X_i. However, it is easy to see that if $b_1, \ldots, b_s$ are basic forms such that for each $1 \le i \le s$, (8) holds for $b = b_i$ and arbitrarily given basic forms a and c, then (8) holds for $b = b_1 \wedge \cdots \wedge b_s$ and arbitrarily given basic forms a and c. Therefore it is enough to show that (8) is true in the following two cases.

Case 1: Assume that $b = f$ is a basic function such that $df = \iota(X)\omega$ for some foliate vector X. Applying Lemma 2.2(ii), we have

$$\begin{aligned}[a \bullet fc] &= (-1)^{|a|}(\delta(a \wedge fc) - \delta(a) \wedge fc - (-1)^{|a|} a \wedge \delta(fc)) \\ &= (-1)^{|a|}\big(f\delta(a \wedge c) - (\iota(X)a) \wedge c - \delta(a) \wedge fc - (-1)^{|a|} a \wedge f\delta c\big) \\ &= f[a \bullet c] - (-1)^{|a|}(\iota(X)a) \wedge c \\ &= f[a \bullet c] + (-1)^{|a|}(\delta(fa) - f\delta a) \wedge c \\ &= f[a \bullet c] + [a \bullet f] \wedge c.\end{aligned}$$

Case 2: Assume that $b = df$ for a basic function f such that $df = \iota(X)\omega$ for some foliate vector X. On the one hand, due to the identity Lemma 2.2(iii), we get

$$
\begin{aligned}
(10)\quad & [a\bullet(df\wedge c)] \\
&= (-1)^{|a|}\big(\delta(a\wedge df\wedge c)-\delta a\wedge df\wedge c-(-1)^{|a|}a\wedge\delta(df\wedge c)\big) \\
&= \mathcal{L}(X)(a\wedge c)-df\wedge\delta(a\wedge c)-(-1)^{|a|}\delta a\wedge df\wedge c+a\wedge df\wedge\delta c-a\wedge\mathcal{L}(X)c \\
&= (\mathcal{L}(X)a)\wedge c-df\wedge\delta(a\wedge c)+df\wedge\delta a\wedge c+a\wedge df\wedge\delta c \\
&= (\mathcal{L}(X)a)\wedge c-df\wedge\big(\delta(a\wedge c)-\delta a\wedge c-(-1)^{|a|}a\wedge\delta c\big) \\
&= (\mathcal{L}(X)a)\wedge c+(-1)^{|a|+1}df\wedge[a\bullet c].
\end{aligned}
$$

On the other hand, applying Lemma 2.2(iii) again, we have

$$
\begin{aligned}
(11)\quad [a\bullet df] &= (-1)^{|a|}\big(\delta(a\wedge df)-\delta a\wedge df-(-1)^{|a|}a\wedge\delta df\big) \\
&= \delta(df\wedge a)-(-1)^{|a|}\delta a\wedge df+a\wedge d\delta f \\
&= -df\wedge\delta a+\mathcal{L}(X)a+df\wedge\delta a \\
&= \mathcal{L}(X)a.
\end{aligned}
$$

It follows immediately from (10) and (11) that (8) holds in this case. □

Formal Frobenius manifolds from* dGBV*-algebras. To show that there is a nice integral on the dGBV-algebra $(\Omega_{G,\mathrm{bas}}, \wedge, \delta, d_G)$, we need the transverse integration theory developed on the space of basic forms on a taut Riemannian foliation (see [Tondeur 1997, Chapter 7; Sergiescu 1985]). Here we follow the method used in [Tondeur 1997], as it may be easier to describe for a general audience.

Recall that a foliation $\mathcal{F}$ on a smooth manifold M is said to be *Riemannian*, if there exists a Riemannian metric g on M, called a *bundle-like* metric for the foliation $\mathcal{F}$, such that for any two foliate vector fields Y and Z on an open subset $U \subset M$ which are perpendicular to the leaves, the function $g(Y, Z)$ is basic on U (see [Reinhart 1959a]). From now on, we assume that M is a closed oriented connected smooth manifold, that $(M, \mathcal{F}, \omega)$ is a transversely symplectic foliation of dimension l and codimension $2n$ which satisfies the transverse hard Lefschetz property, and that there is a Hamiltonian action

$$G\times M\to M,\quad (h,x)\mapsto L_h(x)$$

of a compact connected Lie group G on M. In addition, we also assume that $\mathcal{F}$ is a Riemannian foliation with a bundle-like metric g.

Let P be the integrable subbundle of TM associated to the foliation $\mathcal{F}$ on M. Observe that under our assumption $\mathcal{F}$ is transversely oriented. It follows that $\mathcal{F}$ is also tangentially oriented. That is to say that P is an oriented vector bundle. Fix an orientation on P, and define the *characteristic form* $\chi_{\mathcal{F}}$ for the triple $(M, g, \mathcal{F})$ as follows (see [Tondeur 1997, Chapter 4]):

$$\chi_{\mathcal{F}}(Y_1, \dots, Y_l) = \det(g(Y_i, E_j)), \tag{12}$$

where $Y_1, \dots, Y_l \in T_x M$, and $(E_1, \dots, E_l)$ is an oriented orthonormal frame of P_x. Clearly, when the orientation on P is fixed, the definition of $\chi_{\mathcal{F}}$ depends only on the choice of a bundle-like metric. However, by the transverse hard Lefschetz property, $H^{2n}(M, \mathcal{F}) \cong H^0(M, \mathcal{F}) \cong \mathbb{R}$, which implies that the Riemannian foliation $(M, \mathcal{F})$ is taut (see [Royo Prieto et al. 2009, Theorem 1.4.6]). Thus as explained in [Tondeur 1997, Chapter 7 and Formula 4.26], we can choose a bundle-like metric g such that the corresponding characteristic form $\chi_{\mathcal{F}}$ satisfies

$$\iota(X_1) \cdots \iota(X_l) d\chi_{\mathcal{F}} = 0 \quad \text{for all } X_1, \dots, X_l \in C^\infty(P). \tag{13}$$

Since the action of G preserves the foliation $\mathcal{F}$, it is easy to check that for all $h \in G$, the characteristic form with respect to the pullback metric $L_h^* g$ is $L_h^* \chi_{\mathcal{F}}$. A straightforward check shows that $L_h^* \chi_{\mathcal{F}}$ also satisfies (13). So averaging the bundle-like metric g over the compact Lie group G if necessary, we may assume that the characteristic form $\chi_{\mathcal{F}}$ with respect to the bundle-like metric g is not only G-invariant, but also satisfies (13). In particular, $\chi_{\mathcal{F}}$ can be regarded as an equivariant differential form. Using the usual equivariant integration (see [Guillemin and Sternberg 1999]), we define a $S(\mathfrak{g}^*)^G$-linear operator as

$$\int : \Omega_{G,\mathrm{bas}} \to S(\mathfrak{g}^*)^G, \quad \alpha \mapsto \int_M \alpha \wedge \chi_{\mathcal{F}}. \tag{14}$$

Lemma 4.4. *For all* $\alpha \in \Omega^s_{G,\mathrm{bas}}$, *for all* $\beta \in \Omega^t_{G,\mathrm{bas}}$,

$$\int (d_G \alpha) \wedge \beta = (-1)^{s+1} \int \alpha \wedge d_G \beta, \tag{15}$$

$$\int (\delta\alpha) \wedge \beta = (-1)^s \int \alpha \wedge \delta\beta. \tag{16}$$

Proof. We first prove a preliminary result that for any two ordinary basic differential forms $\alpha \in \Omega^s(M, \mathcal{F})$ and $\beta \in \Omega^t(M, \mathcal{F})$, the following identity holds.

$$\int_M (d\alpha) \wedge \beta \wedge \chi_{\mathcal{F}} = (-1)^{s+1} \int_M \alpha \wedge d\beta \wedge \chi_{\mathcal{F}}. \tag{17}$$

By the Leibniz rule,

$$d(\alpha \wedge \beta \wedge \chi_{\mathcal{F}}) = d\alpha \wedge \beta \wedge \chi_{\mathcal{F}} + (-1)^s \alpha \wedge (d\beta) \wedge \chi_{\mathcal{F}} + (-1)^{s+t} \alpha \wedge \beta \wedge d\chi_{\mathcal{F}}.$$

Since

$$\int_M d(\alpha \wedge \beta \wedge \chi_{\mathcal{F}}) = 0,$$

to prove (17) it suffices to show that

$$\int_M \alpha\wedge\beta\wedge d\chi_{\mathcal{F}}=0. \tag{18}$$

Observe that $\chi_{\mathcal{F}}$ is of degree l; we may assume that $s+t=2n-1$, for otherwise (18) holds for degree reasons. Next recall that by our choice of the bundle-like metric, the characteristic form $\chi_{\mathcal{F}}$ has the property that for any vector fields $X_1,\cdots,X_l$ tangent to the leaves of $\mathcal{F}$, $\iota(X_1)\cdots\iota(X_l)d\chi_{\mathcal{F}}=0$. Since α and β are basic, this would imply that $\alpha\wedge\beta\wedge d\chi_{\mathcal{F}}=0$, from which (17) follows as an immediate consequence.

Since d does not act on the polynomial part of an equivariant basic form, (17) also holds for equivariant basic forms. On the other hand, for each $\alpha\in\Omega^s_G(M,\mathcal{F})$ and $\beta\in\Omega^t_G(M,\mathcal{F})$, a simple degree counting shows that

$$\int_M \partial\alpha\wedge\beta\wedge d\chi_{\mathcal{F}}=\int_M \alpha\wedge\partial\beta\wedge d\chi_{\mathcal{F}}=0. \tag{19}$$

Combing (17) and (19) we get that (15) holds.

To prove that (16) holds, it suffices to show that for any ordinary basic forms $\alpha\in\Omega^s(M,\mathcal{F})$ and $\beta\in\Omega^t(M,\mathcal{F})$,

$$\int_M (\delta\alpha)\wedge\beta\wedge\chi_{\mathcal{F}}=(-1)^s\int \alpha\wedge(\delta\beta)\wedge\chi_{\mathcal{F}}.$$

Without loss of generality, we may assume that $s+t=2n+1$. Using (1) and (17), we have

$$\begin{aligned}\int_M (\delta\alpha)\wedge\beta\wedge\chi_{\mathcal{F}}&=(-1)^{s+1}\int_M (\star d\star\alpha)\wedge\beta\wedge\chi_{\mathcal{F}}\\&=(-1)^{s+1}\int_M (d\star\alpha)\wedge\star\beta\wedge\chi_{\mathcal{F}}\\&=\int_M(\star\alpha)\wedge d\star\beta\wedge\chi_{\mathcal{F}}\\&=(-1)^s\int_M \alpha\wedge\delta\beta\wedge\chi_{\mathcal{F}}.\end{aligned}$$

This completes the proof. □

Note that $S(\mathfrak{g}^*)^G$ is an integral domain. Let $\mathbb{F}=\left\{\frac{f}{g}\mid f,g\in S(\mathfrak{g}^*)^G\right\}$ be the fractional field of $S(\mathfrak{g}^*)^G$. Define

$$\widetilde{\Omega}_{G,\mathrm{bas}}=\Omega_{G,\mathrm{bas}}\otimes_{S(\mathfrak{g}^*)^G}\mathbb{F}.$$

Extend d_G, $\wedge$ and δ to $\widetilde{\Omega}_{G,\mathrm{bas}}$, and define

$$\widetilde{H}_G(M,\mathcal{F})=H(\widetilde{\Omega}_{G,\mathrm{bas}},d_G). \tag{20}$$

As a direct consequence of Theorem 3.11, we have

$$\widetilde{H}_G(M,\mathcal{F}) = H_G(M,\mathcal{F}) \otimes_{S(\mathfrak{g}^*)^G} \mathbb{F}.$$

Applying Proposition 4.3, we see that $(\widetilde{\Omega}_{G,\mathrm{bas}}, \delta, \wedge, d_G)$ is a dGBV-algebra over $\mathbb{F}$. Moreover, the operator defined in (14) naturally extends to a $\mathbb{F}$-linear operator

$$\int : \widetilde{\Omega}_{G,\mathrm{bas}} \to \mathbb{F}. \tag{21}$$

Clearly, Lemma 4.4 implies that the operator (21) defines an integral on the dGBV algebra $(\widetilde{\Omega}_{G,\mathrm{bas}}, \wedge, \delta, d_G)$. To show that this integral is also nice, we need the following result on the basic Poincaré duality.

Theorem 4.5 [Tondeur 1997, Corollary 7.58]. *Let $\mathcal{F}$ be a taut and transversally oriented Riemannian foliation on a closed oriented manifold M. The pairing*

$$\alpha \otimes \beta \mapsto \int_M \alpha \wedge \beta \wedge \chi_{\mathcal{F}}$$

induces a nondegenerate pairing

$$H^r(M,\mathcal{F}) \times H^{q-r}(M,\mathcal{F}) \to \mathbb{R}$$

on finite-dimensional vector spaces, where $q = \operatorname{codim} \mathcal{F}$.

Lemma 4.6. *The integral operator defined in (21) is nice, i.e., it induces a $\mathbb{F}$-bilinear nondegenerate pairing*

$$\widetilde{H}^*_G(M,\mathcal{F}) \times \widetilde{H}^*_G(M,\mathcal{F}) \to \mathbb{F}.$$

Proof. Let $[\alpha]$ be an arbitrary class in $H_G(M,\mathcal{F})$ such that

$$\int_M \alpha \wedge \beta \wedge \chi_{\mathcal{F}} = 0, \quad \text{for each } [\beta] \in H_G(M,\mathcal{F}).$$

To prove Lemma 4.6, it suffices to show that $[\alpha]$ has to vanish.

Let $\{f_1, \ldots, f_k\}$ be a basis of the real vector space $(S\mathfrak{g}^*)^G$. By Theorem 3.11, there exist finitely many cohomology classes $[\gamma_i]$ in $H(M,\mathcal{F})$ such that

$$[\alpha] = \sum_i f_i \otimes s([\gamma_i]).$$

Here $s : H(M,\mathcal{F}) \to H_G(M,\mathcal{F})$ is the canonical section introduced in (7). Let k_i be the degree of the basic form γ_i. After a reshuffling of the index, we may assume that $k_1 \geq k_2 \geq \cdots$. Then for any $[\zeta] \in H^{2n-k_1}(M,\mathcal{F})$,

$$\sum_i f_i \otimes \left(\int_M s([\gamma_i]) \wedge s([\zeta]) \wedge \chi_{\mathcal{F}} \right) = 0,$$

which implies

$$\int_M s([\gamma_1]) \wedge s([\zeta]) \wedge \chi_{\mathcal{F}} = 0.$$

It then follows from a simple counting of degrees that $\int_M \gamma_1 \wedge \zeta \wedge \chi_{\mathcal{F}} = 0$. Since $[\zeta] \in H^{2n-k_1}(M, \mathcal{F})$ is arbitrarily chosen, by Theorem 4.5 we have that $[\gamma_1] = 0$. Thus $s([\gamma_1]) = 0$. Repeating this argument, we see that $[\gamma_i] = 0$ for all i. It follows that $[\alpha]$ must be zero. □

We are ready to state the main result of this section.

Theorem 4.7. *Assume that $(\mathcal{F}, \omega)$ is a transversely symplectic foliation on a closed oriented smooth manifold M that satisfies the transverse hard Lefschetz property, and that a compact connected Lie group G acts on $(M, \mathcal{F}, \omega)$ in a Hamiltonian fashion. If $\mathcal{F}$ is also a Riemannian foliation, then there is a canonical formal Frobenius manifold structure on the equivariant basic cohomology $\widetilde{H}_G(M, \mathcal{F})$ as defined in (20).*

Proof. It remains to show that the following maps induced by the inclusions

$$\rho : H(\ker \delta, d_G) \to H(\Omega_{G,\mathrm{bas}}, d_G) \tag{22}$$

$$\mu : H(\ker d_G, \delta) \to H(\Omega_{G,\mathrm{bas}}, \delta) \tag{23}$$

are isomorphisms. The fact that the map (22) is an isomorphism is a direct consequence of Theorem 3.11. Let $\alpha \in \ker d_G$ be a δ-closed form which represents a class $[\alpha]$ in $H(\ker d_G, \delta)$. Suppose that $[\alpha]$ is trivial in $H(\Omega_{G,\mathrm{bas}}, \delta)$, then there exists a $\beta \in \Omega_{G,\mathrm{bas}}$ such that $\alpha = \delta\beta$. By Theorem 3.12, we have $\alpha = d_G\delta\gamma$ for some $\gamma \in \Omega_{G,\mathrm{bas}}$. This shows that α represents a trivial class in $H(\ker d_G, \delta)$, and that the map (23) is injective.

To see that (23) is surjective, suppose that $\alpha \in \Omega_{G,\mathrm{bas}}$ such that $\delta\alpha = 0$, i.e., $[\alpha]$ is a class in $H(\Omega_{G,\mathrm{bas}}, \delta)$. Let $\gamma = d_G\alpha$. Then γ is both d_G-exact and δ-closed. By Theorem 3.12, there exists a $\beta \in \Omega_{G,\mathrm{bas}}$ such that $\gamma = d_G\delta\beta$. Set $\tilde{\alpha} = \alpha - \delta\beta$. Then $\tilde{\alpha} \in \ker d_G$ and $[\tilde{\alpha}] = [\alpha]$ in $H(\Omega_{G,\mathrm{bas}}, \delta)$. This proves that (23) is surjective. By Theorem 4.2 there exists a formal Frobenius manifold structure on $\widetilde{H}_G(M, \mathcal{F})$. □

When G is a trivial group consisting of one single element, we have the following:

Corollary 4.8. *Assume that $(M, \mathcal{F}, \omega)$ is a transversely symplectic manifold that satisfies the transverse hard Lefschetz property. If $\mathcal{F}$ is also a Riemannian foliation, then there is a canonical formal Frobenius manifold structure on the basic cohomology $H(M, \mathcal{F})$.*

Remark 4.9. When the foliation $\mathcal{F}$ is zero-dimensional, from Corollary 4.8 we recover the Merkulov's construction [1998] of a Frobenius manifold structure on the de Rham cohomology of a symplectic manifold with the hard Lefschetz property. When the foliation $\mathcal{F}$ is zero-dimensional, and when M is a closed Kähler manifold, we recover from Theorem 4.7 the construction by Cao and Zhou [1999], which produces a Frobenius manifold structure on the equivariant cohomology of a Hamiltonian action of a compact connected Lie group on a Kähler manifold. Moreover, we are able to remove the assumption in [Cao and Zhou 1999] that the action is holomorphic.

5. Examples of Frobenius manifolds from transversely symplectic foliations

In this section we present some examples of transversely symplectic foliations which give rise to new examples of dGBV-algebra whose cohomology admits a formal Frobenius manifold structure. We begin with a useful observation on when an action of a compact Lie group gives rise to a G-invariant Riemannian foliation.

Lemma 5.1. *Consider the action of a compact Lie group G on a manifold M. Suppose that $\mathfrak{h}$ is an ideal of the Lie algebra $\mathfrak{g}$ of G, and that the induced infinitesimal action of $\mathfrak{h}$ on M is free. Then it generates a G-invariant Riemannian foliation $\mathcal{F}$ on M.*

Proof. It is clear from our assumption that the foliation $\mathcal{F}$ is G-invariant. Now suppose that g is an G-invariant Riemannian metric. We will show that g must be a bundle-like metric. Let Y and Z be two foliate vector fields which are perpendicular to the leaves, and let ξ_M be the fundamental vector field generated by the infinitesimal action of $\xi \in \mathfrak{h}$. Then,

$$\mathcal{L}(\xi_M)(g(Y, Z)) = (\mathcal{L}(\xi_M)g)(Y, Z) + g([\xi_M, X], Y) + g(X, [\xi_M, Y]).$$

Note that $\mathcal{L}(\xi_M)g = 0$ because g is G-invariant. Moreover, since X is a foliate vector field, $[\xi_M, X]$ must be tangent to the leaves. Thus $g([\xi_M, X], Y) = 0$ as Y is perpendicular to the leaves. For the same reason, $g(X, [\xi_M, Y]) = 0$. It follows that $\mathcal{L}(\xi_M)(g(Y, Z)) = 0$. Since $\xi \in \mathfrak{h}$ is arbitrarily chosen, $g(Y, Z)$ must be a basic function. This completes the proof. □

Now, we will discuss examples of transversely symplectic foliations to which Theorem 4.7 and Corollary 4.8 apply.

Example 5.2 (cooriented contact manifolds). Let M be a $(2n+1)$-dimensional cooriented compact contact manifold with a contact one form η and a Reeb vector ξ. Then the Reeb characteristic foliation $\mathcal{F}_\xi$ induced by ξ is transversely symplectic, with a transversely symplectic form $d\eta$. If there exists a contact metric g such that ξ is a Killing vector field, then (M, η, g) is called a K-contact manifold. It is well known that the Reeb characteristic foliation of a K-contact manifold (M, η, g) is Riemannian. By Corollary 4.8, when M satisfies the transverse hard Lefschetz property, its basic cohomology will carry the structure of a formal Frobenius manifold. In particular, this is the case when (M, η, g) is a Sasakian manifold (see [Boyer and Galicki 2008]). It is also noteworthy that there exist examples of compact K-contact manifolds which do not admit any Sasakian structures, and which satisfy the hard Lefschetz property as introduced in [Cappelletti-Montano et al. 2015; 2016]. By [Lin 2013, Theorem 4.4], these non-Sasakian K-contact manifolds also satisfy the transverse hard Lefschetz property.

Example 5.3 (Hamiltonian actions on contact manifolds). Let M be a $(2n+1)$-dimensional compact contact manifold with a contact one form η and a Reeb vector field ξ, and let G be a compact connected Lie group with the Lie algebra $\mathfrak{g}$. Suppose that G acts on M preserving the contact one form η. Then the η-contact moment map $\Phi : M \to \mathfrak{g}^*$, given by

$$\langle \Phi, X \rangle = \eta(X_M), \quad \text{for all } X \in \mathfrak{g},$$

also defines a moment map for the transverse G-action on the transversely symplectic foliation $(M, \mathcal{F}_\xi, d\eta)$. Here $\langle \cdot\,, \cdot \rangle$ is the dual pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$, and X_M is the fundamental vector field generated by X.

Recall that the action of G is said to be of *Reeb type*, if the Reeb vector ξ is generated by the infinitesimal action of an element in $\mathfrak{g}$ (see [Boyer and Galicki 2008, Definition 8.4.28]). It is clear from Lemma 5.1 that when the action of G is of Reeb type, the Reeb characteristic foliation $\mathcal{F}_\xi$ is Riemannian. If in addition, (M, η, g) is a Sasakian manifold, then $\mathcal{F}_\xi$ satisfies the transverse hard Lefschetz property. In particular, these observations apply to the case when (M, η, g) is a compact toric contact manifold of Reeb type. Therefore, by Theorem 4.7, there is a formal Frobenius manifold structure on the equivariant basic cohomology of toric contact manifolds of Reeb type.

Example 5.4 (cosymplectic manifolds [Li 2008]). Let (M, η, ω) be a $(2n+1)$-dimensional compact cosymplectic manifold. By definition, η is a closed one form, and ω a closed two form ω, such that $\eta \wedge \omega^n$ is a volume form. Then the Reeb characteristic foliation $\mathcal{F}_\xi$ induced by the Reeb vector field ξ (defined by the equations $\iota(\xi)\eta = 1$ and $\iota(\xi)\omega = 0$) is transversely symplectic with the transversely symplectic form ω.

We claim that for any $1 \le k \le n$, the basic form ω^k represents a nontrivial basic cohomology class in $H^{2k}(M, \mathcal{F})$. Assume to the contrary that $[\omega^k] = 0 \in H^{2k}(M, \mathcal{F})$ for some $1 \le k \le n$. Then there exists a basic $(2n-1)$-form β such that $\omega^n = d\beta$. Since $d\eta = 0$, we have

$$\int_M \eta \wedge \omega^n = \int_M \eta \wedge d\beta = \int_M -d(\eta \wedge \beta) = 0,$$

which contradicts the fact that $\eta \wedge \omega^n$ is a volume form. This proves our claim.

The cosymplectic manifold M is called a co-Kähler manifold, if one can associate to (M, η, ω) an almost contact structure (ϕ, ξ, η, g), where ϕ is an $(1, 1)$-tensor, and g a Riemannian metric, such that ϕ is parallel with respect to the Levi-Civita connection of g. It is straightforward to check that if M is co-Kähler, then the Reeb characteristic foliation $\mathcal{F}_\xi$ is transversely Kähler. Due to the claim established in the previous paragraph, it is indeed a taut transversely Kähler foliation, and therefore satisfies the transverse hard Lefschetz property. By Corollary 4.8, the basic cohomology of M has a structure of a formal Frobenius manifold.

Example 5.5 (symplectic orbifolds). Let (X, σ) be an effective symplectic orbifold of dimension $2n$. Then the total space of the orthogonal frame orbibundle $\pi : Fr(X) \to X$ is a smooth manifold on which the structure group $O(2n)$ acts locally free. The form $\omega := \pi^*\sigma$ is a closed 2-form on $Fr(X)$ whose kernel gives rise to a transversely symplectic foliation $\mathcal{F}$. It follows easily from Lemma 5.1 that $\mathcal{F}$ is also Riemannian. When X is a Kähler orbifold, it was shown in [Wang and Zaffran 2009] that $Fr(X)$ satisfies the transverse hard Lefschetz property. Since in this case, the basic differential complex of $(Fr(X), \mathcal{F})$ is isomorphic to the de Rham differential complex on X, Corollary 4.8 implies that there is a formal Frobenius manifold structure on the de Rham cohomology of X.

Now suppose that a compact connected Lie group G acts on (X, σ) in a Hamiltonian fashion with a moment map $\Phi : X \to \mathfrak{g}^*$, where $\mathfrak{g} = \mathrm{Lie}(G)$. By averaging, we may assume that there is a G-invariant Riemannian metric g that is compatible with σ. Then the G-action maps an orthogonal frame to another orthogonal frame; and therefore, lifts to a Hamiltonian G-action on $(Fr(X), \mathcal{F}, \omega)$. Analogous to the discussion in the previous paragraph, when X is Kähler orbifold, Theorem 4.7 implies that there is a formal Frobenius manifold structure on the equivariant de Rham cohomology of X.

Example 5.6 (symplectic quasifolds [Prato 2001]). Assume that (X, σ) is a symplectic manifold on which the torus T acts in a Hamiltonian fashion. We denote the moment map by $\phi : X \to \mathfrak{t}^*$. Let $N \subset T$ be a nonclosed subgroup with Lie algebra $\mathfrak{n}$ and let a be a regular value of the corresponding moment map $\varphi : X \to \mathfrak{n}^*$. Consider the submanifold

$$M = \varphi^{-1}(a) \subset X.$$

The N-action on M yields a transversely symplectic foliation $\mathcal{F}$ with $\omega := i^*\sigma$ being the transversely symplectic form, where i is the inclusion map of M in X. In this case, the leaf space $M/\mathcal{F}$ is a symplectic quasifold in the sense of Prato [2001], at least when N is a connected subgroup of T. It is straightforward to check that the induced T-action on $(M, \mathcal{F}, \omega)$ is Hamiltonian.

It follows from Lemma 5.1 that $\mathcal{F}$ is also a Riemannian foliation. Moreover, using an argument similar to the one given in Example 5.4, it can be shown that $\mathcal{F}$ is a taut Riemannian foliation. The leaf space of $\mathcal{F}$ is called a toric quasifold when $\dim(T/N)$ is half of the dimension of the leaf space. It is shown by [Ishida

2017, Theorem 5.7] that when this is the case, $\mathcal{F}$ is a transversely Kähler foliation. Therefore there exist formal Frobenius manifold structures on the basic cohomology and equivariant basic cohomology of toric quasifolds.

Acknowledgements

Much of this joint work was completed while Lin was visiting Sichuan University in the spring of 2016. He would like to thank the School of Mathematics and the geometry and topology group there for providing him with an excellent working environment. Yang would like to thank Prof. Guosong Zhao for his constant encouragement and moral support over the years. Both authors are grateful to Prof. Xiaojun Chen for his interest in this work, and for many useful discussions. Finally, the authors would like to thank the anonymous referee for helpful comments and suggestions.

References

[Barannikov and Kontsevich 1998] S. Barannikov and M. Kontsevich, "Frobenius manifolds and formality of Lie algebras of polyvector fields", *Int. Math. Res. Not.* **1998**:4 (1998), 201–215. MR Zbl

[Boyer and Galicki 2008] C. P. Boyer and K. Galicki, *Sasakian geometry*, Oxford Univ. Press, 2008. MR Zbl

[Brylinski 1988] J.-L. Brylinski, "A differential complex for Poisson manifolds", *J. Differential Geom.* **28**:1 (1988), 93–114. MR Zbl

[Cao and Zhou 1999] H.-D. Cao and J. Zhou, "Formal Frobenius manifold structure on equivariant cohomology", *Commun. Contemp. Math.* **1**:4 (1999), 535–552. MR Zbl

[Cao and Zhou 2000] H.-D. Cao and J. Zhou, "Frobenius manifold structure on Dolbeault cohomology and mirror symmetry", *Comm. Anal. Geom.* **8**:4 (2000), 795–808. MR Zbl

[Cappelletti-Montano et al. 2015] B. Cappelletti-Montano, A. De Nicola, and I. Yudin, "Hard Lefschetz theorem for Sasakian manifolds", *J. Differential Geom.* **101**:1 (2015), 47–66. MR Zbl

[Cappelletti-Montano et al. 2016] B. Cappelletti-Montano, A. De Nicola, J. C. Marrero, and I. Yudin, "A non-Sasakian Lefschetz K-contact manifold of Tievsky type", *Proc. Amer. Math. Soc.* **144**:12 (2016), 5341–5350. MR Zbl

[El Kacimi-Alaoui 1990] A. El Kacimi-Alaoui, "Opérateurs transversalement elliptiques sur un feuilletage riemannien et applications", *Compositio Math.* **73**:1 (1990), 57–106. MR Zbl

[El Kacimi-Alaoui et al. 1985] A. El Kacimi-Alaoui, V. Sergiescu, and G. Hector, "La cohomologie basique d'un feuilletage Riemannien est de dimension finie", *Math. Z.* **188**:4 (1985), 593–599. MR Zbl

[Ginzburg 1987] V. A. Ginzburg, "Equivariant cohomologies and Kähler's geometry", *Funktsional. Anal. i Prilozhen.* **21**:4 (1987), 19–34. In Russian; translated in *Funct. Anal. Appl.* **21**:4 (1987, 271–283. MR Zbl

[Goertsches and Töben 2018] O. Goertsches and D. Töben, "Equivariant basic cohomology of Riemannian foliations", *J. Reine. Angew. Math.* **745** (2018), 1–40. MR Zbl

[Goertsches et al. 2012] O. Goertsches, H. Nozawa, and D. Töben, "Equivariant cohomology of K-contact manifolds", *Math. Ann.* **354**:4 (2012), 1555–1582. MR Zbl

[Goresky et al. 1998] M. Goresky, R. Kottwitz, and R. MacPherson, "Equivariant cohomology, Koszul duality, and the localization theorem", *Invent. Math.* **131**:1 (1998), 25–83. MR Zbl

[Guillemin 2001] V. Guillemin, "Symplectic Hodge theory and the $d\delta$-lemma", lecture notes, Massachusetts Institute of Technology, 2001.

[Guillemin and Sternberg 1999] V. W. Guillemin and S. Sternberg, *Supersymmetry and equivariant de Rham theory*, Springer, 1999. MR Zbl

[Haefliger 1971] A. Haefliger, "Homotopy and integrability", pp. 133–163 in *Manifolds* (Amsterdam, 1970), edited by N. H. Kuiper, Lecture Notes in Math. **197**, Springer, 1971. MR Zbl

[He 2010] Z. He, *Odd dimensional symplectic manifolds*, Ph.D. thesis, Massachusetts Institute of Technology, 2010, Available at https://search.proquest.com/docview/847033382.

[Ishida 2017] H. Ishida, "Torus invariant transverse Kähler foliations", *Trans. Amer. Math. Soc.* **369**:7 (2017), 5137–5155. MR Zbl

[Kirwan 1984] F. C. Kirwan, *Cohomology of quotients in symplectic and algebraic geometry*, Mathematical Notes **31**, Princeton Univ. Press, 1984. MR Zbl

[Li 2008] H. Li, "Topology of co-symplectic/co-Kähler manifolds", *Asian J. Math.* **12**:4 (2008), 527–543. MR Zbl

[Lin 2013] Y. Lin, "Lefschetz contact manifolds and odd dimensional symplectic geometry", preprint, 2013. arXiv

[Lin 2018] Y. Lin, "Hodge theory on transversely symplectic foliations", *Q. J. Math.* **69**:2 (2018), 585–609. MR

[Lin and Sjamaar 2004] Y. Lin and R. Sjamaar, "Equivariant symplectic Hodge theory and the $d_G\delta$-lemma", *J. Symplectic Geom.* **2**:2 (2004), 267–278. MR Zbl

[Lin and Sjamaar 2017] Y. Lin and R. Sjamaar, "Convexity properties of presymplectic moment maps", 2017. To appear in *J. Symplectic Geom.* arXiv

[Manin 1999] Y. I. Manin, *Frobenius manifolds, quantum cohomology, and moduli spaces*, Amer. Math. Soc. Colloquium Publ. **47**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Mathieu 1995] O. Mathieu, "Harmonic cohomology classes of symplectic manifolds", *Comment. Math. Helv.* **70**:1 (1995), 1–9. MR Zbl

[Merkulov 1998] S. A. Merkulov, "Formality of canonical symplectic complexes and Frobenius manifolds", *Int. Math. Res. Not.* **1998**:14 (1998), 727–733. MR Zbl

[Molino 1988] P. Molino, *Riemannian foliations*, Progress in Math. **73**, Birkhäuser, Boston, 1988. MR Zbl

[Prato 2001] E. Prato, "Simple non-rational convex polytopes via symplectic geometry", *Topology* **40**:5 (2001), 961–975. MR Zbl

[Reinhart 1959a] B. L. Reinhart, "Foliated manifolds with bundle-like metrics", *Ann. of Math.* (2) **69** (1959), 119–132. MR Zbl

[Reinhart 1959b] B. L. Reinhart, "Harmonic integrals on foliated manifolds", *Amer. J. Math.* **81** (1959), 529–536. MR Zbl

[Royo Prieto et al. 2009] J. I. Royo Prieto, M. Saralegi-Aranguren, and R. Wolak, "Cohomological tautness for Riemannian foliations", *Russ. J. Math. Phys.* **16**:3 (2009), 450–466. MR Zbl

[Satake 1957] I. Satake, "The Gauss–Bonnet theorem for V-manifolds", *J. Math. Soc. Japan* **9** (1957), 464–492. MR Zbl

[Sergiescu 1985] V. Sergiescu, "Cohomologie basique et dualité des feuilletages riemanniens", *Ann. Inst. Fourier* (*Grenoble*) **35**:3 (1985), 137–158. MR Zbl

[Souriau 1997] J.-M. Souriau, *Structure of dynamical systems: a symplectic view of physics*, Progress in Math. **149**, Birkhäuser, Boston, 1997. MR Zbl

[Tondeur 1997] P. Tondeur, *Geometry of foliations*, Monographs in Math. **90**, Birkhäuser, Basel, 1997. MR Zbl

[Wang and Zaffran 2009] Z. Z. Wang and D. Zaffran, "A remark on the hard Lefschetz theorem for Kähler orbifolds", *Proc. Amer. Math. Soc.* **137**:8 (2009), 2497–2501. MR Zbl

[Yan 1996] D. Yan, "Hodge structure on symplectic manifolds", *Adv. Math.* **120**:1 (1996), 143–154. MR Zbl

Received July 18, 2017. Revised August 3, 2018.

YI LIN
DEPARTMENT OF MATHEMATICAL SCIENCES
GEORGIA SOUTHERN UNIVERSITY
STATESBORO, GA
UNITED STATES
yilin@georgiasouthern.edu

XIANGDONG YANG
DEPARTMENT OF MATHEMATICS
CHONGQING UNIVERSITY
CHONGQING
CHINA
math.yang@cqu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.83

HEEGAARD FLOER HOMOLOGY OF L-SPACE LINKS WITH TWO COMPONENTS

BEIBEI LIU

We compute different versions of link Floer homology HFL^- and $\widehat{\mathrm{HFL}}$ for any L-space link with two components. The main approach is to compute the h-function of the filtered chain complex which is determined by Alexander polynomials of all sublinks of the L-space link. As an application, the Thurston norm of its complement is explicitly determined by Alexander polynomials of the link and its components.

1. Introduction

Heegaard Floer homology is an invariant for closed, oriented 3-manifolds, defined using Heegaard diagrams [Ozsváth and Szabó 2004b]. This construction can be extended to give an invariant, *Heegaard Floer link homology* (also called *link Floer homology*), for oriented links in S^3 [Ozsváth and Szabó 2008a]. In general, it is very hard to compute the Heegaard Floer link homology HFL^- and $\widehat{\mathrm{HFL}}$. For any L-space link with two components (see Definition 2.2), Yajing Liu [2017] computed the link Floer homology HFL^-. Based on his work, we come up with a method to compute the link Floer homology $\widehat{\mathrm{HFL}}$ of any L-space link with two components. By the work of Ozsváth and Szabó [2008b], we compute the Thurston polytope and the Thurston norm of its complement. For an r-component L-space link with a given generic admissible multipointed Heegaard diagram, one can associate it with *generalized Floer complexes* $A^-(\boldsymbol{s})$ filtered by Alexander gradings [Manolescu and Ozsváth 2010]. In this article, we work over $\mathbb{F} = \mathbb{F}_2$ and $\boldsymbol{s} \in \mathbb{H}$, where $\mathbb{H}$ is some r-dimensional lattice; see Definition 2.3 and [Manolescu and Ozsváth 2010]. If the link L is an L-space link, we have the following result for $A^-(\boldsymbol{s})$:

Proposition 1.1 [Liu 2017, Proposition 1.11]. *For any L-space link*,

$$H_*(A^-(\boldsymbol{s})) = \mathbb{F}[\![U]\!] \quad \textit{with } \boldsymbol{s} \in \mathbb{H}.$$

Here U has homological grading -2. Define $-2h(\boldsymbol{s})$ as the homological grading of the generator in $H_*(A^-(\boldsymbol{s}))$. By the work of Gorsky, Némethi and Yajing Liu,

MSC2010: 57M99.
Keywords: L-space links, link Floer homology, Thurston polytopes.

$h(\boldsymbol{s})$ is determined by Alexander polynomials $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$ and $\Delta_{L_2}(t)$ for any 2-component L-space link and $\boldsymbol{s} \in \mathbb{H}$. There is a spectral sequence which converges to $\mathrm{HFL}^-(L, \boldsymbol{s})$ [Gorsky and Némethi 2015]. It collapses at the E^2-page, and $h(\boldsymbol{s})$ determines its E^1-page; see [Gorsky and Némethi 2015, Theorem 2.2.10; Liu 2017].

The computation of $\widehat{\mathrm{HFL}}(L, \boldsymbol{s})$ is more complicated. We introduce a bigraded "iterated cone" complex $(\mathfrak{C}(s_1, s_2), d + d_1)$ in Section 3. There exists a spectral sequence associated with this bigraded complex where the E^1-page is defined by HFL^- and $E^3 = \widehat{\mathrm{HFL}}(L, s_1, s_2)$. Theorem 3.2 shows that the E^1-page of this spectral sequence is

$$\mathrm{HFL}^-(s_1+1, s_2+1) \oplus \mathrm{HFL}^-(s_1, s_2+1) \oplus \mathrm{HFL}^-(s_1+1, s_2) \oplus \mathrm{HFL}^-(s_1, s_2),$$

and the differential d_1 is induced by actions of U_1 and U_2. Lemma 3.3 indicates how U_i acts on the Heegaard Floer link homology $\mathrm{HFL}^-(L, \boldsymbol{s})$ for any $\boldsymbol{s} \in \mathbb{H}$ and $i = 1, 2$. So we can compute the E^2-page of the spectral sequence. If $d_2 = 0$, the spectral sequence collapses at the E^2-page. If d_2 is nonzero, we need to use another strategy to compute $\widehat{\mathrm{HFL}}(L, \boldsymbol{s})$. We first find all possible cases where d_2 may be nontrivial. In order to compute $\widehat{\mathrm{HFL}}(L, \boldsymbol{s})$, we use the symmetric property of Heegaard Floer link homology: $\widehat{\mathrm{HFL}}(L, \boldsymbol{s}) \cong \widehat{\mathrm{HFL}}(L, -\boldsymbol{s})$, up to some grading shift [Ozsváth and Szabó 2006, Equation 5]. In Section 3, we find that in all cases where d_2 may be nontrivial, the spectral sequence corresponding to $\widehat{\mathrm{HFL}}(L, -\boldsymbol{s})$ collapses at its E^2-page. Then we can compute $\widehat{\mathrm{HFL}}(L, -\boldsymbol{s})$, and hence $\widehat{\mathrm{HFL}}(L, \boldsymbol{s})$. Therefore, we compute $\widehat{\mathrm{HFL}}$ for all L-space links with two components and obtain the main theorem of this paper.

Theorem 1.2. *For any L-space link $L = L_1 \cup L_2$ with two components, $\widehat{\mathrm{HFL}}(s_1, s_2)$ is determined by the h-function and hence determined by symmetrized Alexander polynomials $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$, $\Delta_{L_2}(t)$, and the linking number* lk *of L_1 and L_2.*

Remark 1.3. The Heegaard Floer link homology depends on the orientation of the link. For any L-space link $L = L_1 \cup L_2$, we need to give it an orientation, which determines the linking number of L_1 and L_2.

Yajing Liu [2017] showed that $\mathrm{rank}_{\mathbb{F}}(\mathrm{HFL}^-(L, \boldsymbol{s})) \leq 2$. We show that 4 is a bound for the rank of link Floer homology $\widehat{\mathrm{HFL}}$ for any L-space link with two components. Then we give examples for all possible ranks from 0 to 4 in Section 3.

Corollary 1.4. *For 2-component L-space links $L = L_1 \cup L_2$ and $\boldsymbol{s} \in \mathbb{H}$,*

$$\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(L, \boldsymbol{s})) \leq 4.$$

In particular, $|\chi(\widehat{\mathrm{HFL}}(L, \boldsymbol{s}))| \leq 4$.

In Section 4, we present an application of Theorem 1.2. It is known from [Ozsváth and Szabó 2008b] that $\widehat{\mathrm{HFL}}(L)$ detects the Thurston norm of the link complement. For any compact, oriented surface with boundary $F = \bigcup_{i=1}^{n} F_i$ (maybe disconnected), define its *complexity* as

$$\chi_-(F) = \sum_{\{F_i \mid \chi(F_i) \le 0\}} -\chi(F_i).$$

For any link $L \subseteq S^3$, and any homology class $h \in H_2(S^3, L)$, there exists a compact oriented surface F with boundary embedded in $S^3 \setminus \mathrm{nd}(L)$ which represents this homology class (i.e., $[F] = h$). So for any homology class $h \in H_2(S^3, L; \mathbb{Z})$, we can assign a function

$$x(h) = \min_{\{F \hookrightarrow S^3 \setminus \mathrm{nd}(L),\, [F]=h\}} \chi_-(F).$$

This function can be naturally extended to a seminorm, the *Thurston seminorm*, denoted by $x : H_2(S^3, L; \mathbb{R}) \to \mathbb{R}$ [Ozsváth and Szabó 2008b]. The unit ball for the norm x is called the *Thurston polytope*. Consider the convex hull of lattice points $\boldsymbol{s} \in \mathbb{H}$, where $\widehat{\mathrm{HFL}}(L, \boldsymbol{s}) \neq 0$, which is also called the *link Floer homology polytope*. We can compute the dual Thurston polytope, and thus the Thurston norm by [Ozsváth and Szabó 2008b]. So for any 2-component L-space link $L = L_1 \cup L_2$, the Thurston polytope and the Thurston norm are determined by Alexander polynomials of all sublinks, but in a very nontrivial way.

Theorem 1.5. *If $L = L_1 \cup L_2$ is an L-space link with two components in S^3, then the Thurston norm of its complement is determined by Alexander polynomials $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$, $\Delta_{L_2}(t)$ and the linking number of L_1 and L_2.*

Ozsváth and Szabó pointed out that for any alternating link, up to a scalar, the Thurston polytope is dual to the Newton polytope of its multivariable Alexander polynomial [Ozsváth and Szabó 2008b], which is contained in the dual Thurston polytope by [McMullen 2002]. We compute dual Thurston polytopes of two nonalternating L-space links with two components in Examples 4.4 and 4.5. They both agree with Newton polytopes of their Alexander polynomials. A natural question arises:

Question 1.6. For any 2-component L-space link which is not a split union of two L-space knots, is the Thurston polytope dual to the Newton polytope of its multivariable Alexander polynomial?

Remark 1.7. In Example 4.4, we present a 2-component L-space link where the set $\mathrm{supp}(\widehat{\mathrm{HFL}}) = \{(s_1, s_2) \in \mathbb{H} \mid \widehat{\mathrm{HFL}}(s_1, s_2) \neq 0\}$ is larger than $\mathrm{supp}(\chi(\widehat{\mathrm{HFL}})) = \{(s_1, s_2) \in \mathbb{H} \mid \chi(\widehat{\mathrm{HFL}}(s_1, s_2)) \neq 0\}$. But the convex hull of $\mathrm{supp}(\widehat{\mathrm{HFL}})$ is the same as that of $\mathrm{supp}(\chi(\widehat{\mathrm{HFL}}))$, since lattice points (s_1, s_2) for which $\chi(\widehat{\mathrm{HFL}}(s_1, s_2)) = 0$ and $\widehat{\mathrm{HFL}}(s_1, s_2) \neq 0$ are inside the convex hull of $\mathrm{supp}(\chi(\widehat{\mathrm{HFL}}))$.

For any split L-space link, the answer to Question 1.6 is negative since its Alexander polynomial vanishes, but the dual Thurston polytope is nonempty. Example 5.5 gives the link Floer homology polytope of the split union of two right-handed trefoils. The split union of two L-space knots is an L-space link [Liu 2017], and the h-function of the link satisfies $h(s_1, s_2) = h_1(s_1) + h_2(s_2)$, where h_1 and h_2 are h-functions of L_1 and L_2, respectively. We compute $\widehat{\mathrm{HFL}}$ for any split union of two L-space knots. In general, we compute $\widehat{\mathrm{HFL}}$ for all 2-component L-space links with Alexander polynomials $\Delta(t_1, t_2) = 0$.

Theorem 1.8. *For any* 2*-component L-space link $L = L_1 \cup L_2$ and $(s_1, s_2) \in \mathbb{H}$, if $\Delta_L(t_1, t_2) = 0$, then*

$$\widehat{\mathrm{HFL}}(L, s_1, s_2) \cong \widehat{\mathrm{HFL}}(L_1 \sqcup L_2, s_1, s_2) \cong \widehat{\mathrm{HFL}}(L_1, s_1) \otimes \widehat{\mathrm{HFL}}(L_2, s_2) \otimes (\mathbb{F} \oplus \mathbb{F}_{-1}),$$

where $L_1 \sqcup L_2$ denotes the split union of L_1 and L_2 .

In this paper, we use $L = L_1 \cup L_2$ to denote L-space links with two components L_1, L_2, unless otherwise stated.

2. Heegaard Floer link homology

2A. ***L-space links.*** The concept of L-spaces was introduced in [Ozsváth and Szabó 2005].

Definition 2.1. A 3-manifold Y is an L-space if it is a rational homology sphere and its Heegaard Floer homology has minimal possible rank: for any Spin^c-structure s, $\widehat{\mathrm{HF}}(Y, s) = \mathbb{F}$ has rank 1, and $\mathrm{HF}^-(Y, s)$ is a free $\mathbb{F}[U]$-module of rank 1.

Gorsky and Némethi [2016] defined L-space links in terms of large surgeries.

Definition 2.2. An l-component link $L \subseteq S^3$ is an L-space link if there exist integers $p_1, p_2, \ldots, p_l$ such that for all integers $n_i \geq p_i$, $1 \leq i \leq l$, the $(n_1, n_2, \ldots, n_l)$-surgery $S^3_{n_1, n_2, \ldots, n_l}$ is an L-space.

The computation of Heegaard Floer link homology is not easy. However, L-space links have some nice properties which make the computation of Heegaard Floer link homology easier. In particular, we only consider L-space links $L = L_1 \cup L_2$ with two components in this article.

For a 2-component L-space link $L = L_1 \cup L_2$ in S^3, consider a generic admissible multipointed Heegaard diagram with each component L_i having only two basepoints w_i, z_i. One can associate a generalized Floer complex $A^-(s_1, s_2)$ with $(s_1, s_2) \in \mathbb{H}$, which is introduced in [Manolescu and Ozsváth 2010, Section 4]. It is a free $\mathbb{F}[U_1, U_2]$-module. The operations U_1 and U_2 are homotopic to each other on each $A^-(s_1, s_2)$ (see [Ozsváth and Szabó 2008a]), and both have homological degree -2.

Definition 2.3. For an oriented link $L = L_1 \cup L_2$ with two components, define $\mathbb{H}$ to be an affine lattice over $\mathbb{Z}^2$,

$$\mathbb{H} = \mathbb{H}_1 \oplus \mathbb{H}_2, \quad \mathbb{H}_i = \mathbb{Z} + \frac{\mathrm{lk}(L_1, L_2)}{2} \quad (i = 1, 2),$$

where $\mathrm{lk}(L_1, L_2)$ denotes the linking number of L_1 and L_2.

By Proposition 1.1, for any L-space link L with two components, we have $H_*(A^-(s_1, s_2)) = \mathbb{F}[[U]]$, where $(s_1, s_2) \in \mathbb{H}$. Let $-2h(s_1, s_2)$ denote the homological grading of the generator in $H_*(A^-(s_1, s_2))$. The function $h(s_1, s_2)$ is the HFL-weight function of an L-space link defined in [Gorsky and Némethi 2015]. In this article, we call it the h-function. On each $A^-(s_1, s_2)$, the operations U_1 and U_2 are homotopic, and we denote them by U.

Lemma 2.4 [Gorsky and Némethi 2015, Lemma 2.2.3]. *Let $\boldsymbol{e}_1 = (1, 0)$ and $\boldsymbol{e}_2 = (0, 1)$. For any $\boldsymbol{s} = (s_1, s_2) \in \mathbb{H}$, there are inclusions $j : A^-(s_1, s_2) \hookrightarrow A^-(\boldsymbol{s} + \boldsymbol{e_i})$ for $i = 1, 2$ which induce injections on homology as follows*:

$$j_* : H_*(A^-(s_1, s_2)) \to H_*(A^-(\boldsymbol{s} + \boldsymbol{e_i})),$$

where $j_* = U_i^{\delta(i)}$ and $\delta(i) = 0$ or 1.

Remark 2.5. The actions U_i induce maps $U_i : A^-(\boldsymbol{s} + \boldsymbol{e_i}) \to A^-(\boldsymbol{s})$ for $i = 1, 2$, and induce maps on homology. By Proposition 1.1, $H_*(A^-(\boldsymbol{s})) \cong \mathbb{F}[[U]]$ for any $\boldsymbol{s} \in \mathbb{H}$. Assume that a, b are the generators of $H_*(A^-(\boldsymbol{s}))$ and $H_*(A^-(\boldsymbol{s} + \boldsymbol{e_i}))$. Then $j_*(a) = U^{\delta(i)} b$ and $U_i(b) = U^{1-\delta(i)} a$.

Corollary 2.6. *For any L-space link with two components and $\boldsymbol{s} \in \mathbb{H}$, either $h(\boldsymbol{s}) = h(\boldsymbol{s} + \boldsymbol{e_i})$ or $h(\boldsymbol{s}) = h(\boldsymbol{s} + \boldsymbol{e_i}) + 1$, where $i = 1, 2$, $\boldsymbol{e}_1 = (1, 0)$, and $\boldsymbol{e}_2 = (0, 1)$.*

Proof. By Lemma 2.4, we have $-2h(\boldsymbol{s}) = -2h(\boldsymbol{s} + \boldsymbol{e_i}) - 2\delta(i)$, where $\delta(i) = 0$ or 1. So $h(\boldsymbol{s}) = h(\boldsymbol{s} + \boldsymbol{e_i})$ or $h(\boldsymbol{s}) = h(\boldsymbol{s} + \boldsymbol{e_i}) + 1$. □

Next, we revisit Yajing Liu's work [2017] about how to use the h-function to compute $\mathrm{HFL}^-(L)$ for any 2-component L-space link $L = L_1 \cup L_2$.

Lemma 2.7 [Gorsky and Némethi 2015, Lemma 2.2.9]. *For any $(s_1, s_2) \in \mathbb{H}$, the chain complex* $\mathrm{CFL}^-(s_1, s_2)$ *of the L-space link $L = L_1 \cup L_2$ is quasi-isomorphic to the "iterated cone" complex*

$$\left[\begin{array}{ccc} A^-(s_1 - 1, s_2) & \xrightarrow{i_1} & A^-(s_1, s_2) \\ \uparrow{\scriptstyle i_2} & & \uparrow{\scriptstyle i_2} \\ A^-(s_1 - 1, s_2 - 1) & \xrightarrow{i_1} & A^-(s_1, s_2 - 1) \end{array}\right]$$

where i_1 and i_2 are inclusion maps in Lemma 2.4.

Let d denote the differential in the generalized Floer complex $A^-(s_1, s_2)$ and $i = i_2 - i_1$. The above "iterated cone" complex has two differentials d and i. The differential d acts in Floer complexes on vertices of the cube, and i acts between Floer complexes. Let the cube grading $|K|$ of the upper-right corner of the cube be 0. The differential d decreases the homological grading by 1, and preserves the cube grading. The differential i preserves the homological grading, and decreases $|K|$ by 1. The total grading is defined as the sum of the homological grading and the cube grading. Let $D = d + i$ and $\mathfrak{K}(s_1, s_2)$ denote the "iterated cone" complex. There exists a spectral sequence whose E^∞-page is the homology of $\mathfrak{K}(s_1, s_2)$ under D.

Theorem 2.8 [Gorsky and Némethi 2015, Theorem 2.2.10]. *Let $L = L_1 \cup L_2$ be an L-space link with two components. For any $(s_1, s_2) \in \mathbb{H}$, there exists a spectral sequence which converges to* $\mathrm{HFL}^-(s_1, s_2)$ *and collapses at its E^2-page. Its E^2-page is isomorphic to $H_*(H_*(A^-(s_1, s_2), d), i)$.*

So $\mathrm{HFL}^-(s_1, s_2)$ is isomorphic to $H_*(H_*(A^-(s_1, s_2), d), i)$. By Proposition 1.1, for any $(s_1, s_2) \in \mathbb{H}$, $H_*(A^-(s_1, s_2), d) \cong \mathbb{F}[[U]][-2h(s_1, s_2)]$, where $-2h(s_1, s_2)$ is the homological grading of the generator in $H_*(A^-(s_1, s_2), d)$, and U_1, U_2 act as U, homotopic to each other on $A^-(s_1, s_2)$ [Ozsváth and Szabó 2008a]. To compute $\mathrm{HFL}^-(s_1, s_2)$, we just need to compute the homology of the mapping cone of i:

$$\begin{array}{ccc}
\mathbb{F}[[U]][-2h(s_1-1, s_2)][b] & \xrightarrow{\ i_1\ } & \mathbb{F}[[U]][-2h(s_1, s_2)][a] \\
\uparrow{\scriptstyle i_2} & & \uparrow{\scriptstyle i_2} \\
\mathbb{F}[[U]][-2h(s_1-1, s_2-1)][c] & \xrightarrow{\ i_1\ } & \mathbb{F}[[U]][-2h(s_1, s_2-1)][d]
\end{array}$$

where a, b, c, d denote the generators in $\mathbb{F}[[U]][-2h(s_1, s_2)]$, $\mathbb{F}[[U]][-2h(s_1-1, s_2)]$, $\mathbb{F}[[U]][-2h(s_1-1, s_2-1)]$, and $\mathbb{F}[[U]][-2h(s_1, s_2-1)]$, respectively. Let $h = h(s_1, s_2)$. By Corollary 2.6, there are 6 cases for the h-function corresponding to the mapping cone.

h	h
h	h

Case (1)

h	h
$h+1$	$h+1$

Case (2)

$h+1$	h
$h+1$	h

Case (3)

h	h
$h+1$	h

Case (4)

$h+1$	h
$h+1$	$h+1$

Case (5)

$h+1$	h
$h+2$	$h+1$

Case (6)

Figure 1. Possible local behaviors of the h-function.

According to the h-function in Figure 1, we can compute the corresponding $\mathrm{HFL}^-(s_1, s_2)$ in each case.

Case (1): $i(b) = a, i(c) = b - d, i(d) = a$ and $i(a) = 0$, so $\mathrm{HFL}^-(s_1, s_2) = 0$.

Case (2): $i(b) = a, i(c) = Ub - d, i(d) = Ua$ and $i(a) = 0$, so $\mathrm{HFL}^-(s_1, s_2) = 0$.

Case (3): $i(b) = Ua, i(c) = b - Ud, i(d) = a$ and $i(a) = 0$, so $\mathrm{HFL}^-(s_1, s_2) = 0$.

Case (4): $i(b)=a, i(c)=Ub-Ud, i(d)=a$ and $i(a)=0$, so $\mathrm{HFL}^-(s_1, s_2)=\langle b-d\rangle$. Both b and d have homological grading $-2h$ and cube grading 1. The total grading of $b-d$ is $-2h+1$. Thus $\mathrm{HFL}^-(s_1, s_2) = \mathbb{F}[-2h+1]$.

Case (5): $i(b) = Ua, i(c) = b - d, i(d) = Ua$ and $i(a) = 0$, so $\mathrm{HFL}^- = \langle a\rangle$ with total grading $-2h$. Thus $\mathrm{HFL}^-(s_1, s_2) = \mathbb{F}[-2h]$.

Case (6): $i(b) = Ua, i(c) = Ub - Ud, i(d) = Ua$, and $i(a) = 0$, so in this case $\mathrm{HFL}^-(s_1, s_2) = \langle a, b-d\rangle$. Here a has total grading $-2h$ and $b-d$ has total grading $-2(h+1)+1 = -2h-1$. Thus $\mathrm{HFL}^-(s_1, s_2) = \mathbb{F}[-2h] \oplus \mathbb{F}[-2h-1]$.

Moreover, we also determine the Euler characteristics $\chi(\mathrm{HFL}^-(s_1, s_2))$ in these six cases. In Case (1), Case (2), Case (3) and Case (6), $\chi(\mathrm{HFL}^-(s_1, s_2)) = 0$. In Case (4), $\chi(\mathrm{HFL}^-(s_1, s_2)) = -1$, and in Case (5), $\chi(\mathrm{HFL}^-(s_1, s_2)) = 1$. Thus for any L-space link with two components, once the h-function is determined, we can compute $\mathrm{HFL}^-(s_1, s_2)$ for any $(s_1, s_2) \in \mathbb{H}$.

Corollary 2.9. *For any* 2*-component L-space link and* $(s_1, s_2) \in \mathbb{H}$, $\mathrm{HFL}^-(s_1, s_2)$ *is spanned by a or $b-d$ or both, where a has even grading and $b-d$ has odd grading.*

2B. ***Alexander polynomials of L-space links.*** In this section, we mainly introduce Yajing Liu's work [2017] about how to determine the h-function of any 2-component L-space link $L = L_1 \cup L_2$ by Alexander polynomials $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$, and $\Delta_{L_2}(t)$. Recall that for any L-space link $L = L_1 \cup L_2$, we have

$$\Delta_L(t_1, t_2) \doteq \sum_{(s_1,s_2)\in\mathbb{H}} \chi(\mathrm{HFL}^-(s_1, s_2)) t_1^{s_1} t_2^{s_2},$$

$$\Delta_L(t, 1) \doteq \frac{1-t^{\mathrm{lk}}}{1-t} \Delta_{L_1}(t), \tag{2-1}$$

where $f \doteq g$ means that f and g differ by multiplication by units. Yajing Liu [2017] defined normalization of Alexander polynomials.

Definition 2.10 [Liu 2017, Definition 5.12]. Let the symmetrized Alexander polynomial of L be $\Delta_L(x_1, x_2)$ in the form of

$$\Delta_L(t_1, t_2) = \sum_{i,j} a^L_{i,j} \cdot t_1^i \cdot t_2^j,$$

where t_i corresponds to the link component L_i for $i = 1, 2$. Let the symmetrized Alexander polynomials of L_1 and L_2 be $\Delta_{L_1}(t)$, $\Delta_{L_2}(t)$ in the form of

$$\frac{t}{t-1}\Delta_{L_1}(t) = \sum_{k\in\mathbb{Z}} a_k^{L_1}\cdot t^k, \qquad \frac{t}{t-1}\Delta_{L_2}(t) = \sum_{k\in\mathbb{Z}} a_k^{L_2}\cdot t^k.$$

Let (i_0, j_0) be such that

$$j_0 = \max\Big\{ j \in \mathbb{Z} + \frac{\mathrm{lk}-1}{2} \mid a^L_{i,j} \neq 0\Big\}, \qquad i_0 = \max\Big\{ i \in \mathbb{Z} + \frac{\mathrm{lk}-1}{2} \mid a^L_{i,j_0} \neq 0\Big\}.$$

Then these Alexander polynomials are called *normalized* if

(1) the leading coefficient of $\Delta_{L_i}(t)$ is 1 for both $i = 1, 2$,

(2) if $a^{L_2}_{j_0-\mathrm{lk}/2+1/2} = 1$, then $a^L_{i_0,j_0} = 1$, while if $a^{L_2}_{j_0-\mathrm{lk}/2+1/2} = 0$, then $a^L_{i_0,j_0} = -1$, where lk is the linking number of L_1 and L_2.

For the normalized Alexander polynomials of the 2-component L-space link $L = L_1 \cup L_2$, $\chi(\mathrm{HFL}^-)(s_1, s_2) = a^L_{s_1-1/2,s_2-1/2}$ and $\chi(\mathrm{HFK}^-(L_i, s)) = a^{L_i}_s$ for $i = 1, 2$ [Liu 2017]. Moreover, Yajing Liu gave the following formulas to determine the h-function in [Liu 2017, (5.8)]:

$$h(s_1, s_2-1) - h(s_1, s_2) = a^{L_2}_{s_2-\mathrm{lk}/2} - \sum_{i=1}^{\infty} a^L_{s_1+i-1/2,s_2-1/2} = 0 \quad\text{or}\quad 1. \tag{2-2}$$

Similarly,

$$h(s_1-1, s_2) - h(s_1, s_2) = a^{L_1}_{s_1-\mathrm{lk}/2} - \sum_{i=1}^{\infty} a^L_{s_1-1/2,s_2+i-1/2} = 0 \quad\text{or}\quad 1. \tag{2-3}$$

When $s_1 \to +\infty$ or $s_2 \to +\infty$,

$$h(+\infty, s_2) = h_2(s_2 - \mathrm{lk}/2), \quad h(s_1, +\infty) = h_1(s_1 - \mathrm{lk}/2), \tag{2-4}$$

$$h_1(s-1) - h_1(s) = a^{L_1}_s, \quad h_2(s-1) - h_2(s) = a^{L_2}_s, \tag{2-5}$$

where $h_1(s_1 - \mathrm{lk}/2)$ and $h_2(s_2 - \mathrm{lk}/2)$ are h-functions for link components L_1 and L_2, and $s \in \mathbb{Z}$. For sufficiently large s, $h_1(s) = h_2(s) = 0$. By using the formulas above, we can compute the h-function, and hence $\mathrm{HFL}^-(s_1, s_2)$ for any 2-component L-space link $L = L_1 \cup L_2$.

Remark 2.11. The link components L_1 and L_2 of 2-component L-space links are both L-space knots [Liu 2017, Lemma 1.10].

Corollary 2.12 [Dawra 2015; Gorsky and Némethi 2015; Liu 2017]. *For any L-space link $L = L_1 \cup L_2$ with two components*, $\mathrm{HFL}^-(L)$ *is determined by Alexander polynomials $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$ and $\Delta_{L_2}(t)$.*

3. Computation of $\widehat{\mathrm{HFL}}$ for 2-component L-space links

3A. ***The spectral sequence corresponding to*** $\widehat{\mathbf{HFL}}$**.** In Section 2, we proved that for any L-space link $L = L_1 \cup L_2$ with $(s_1, s_2) \in \mathbb{H}$, $\mathrm{HFL}^-(s_1, s_2)$ is determined by the h-function. Now we are going to prove Theorem 1.2 that the Heegaard Floer link homology $\widehat{\mathrm{HFL}}(s_1, s_2)$ is also determined by the h-function.

Let $\mathfrak{C}(s_1, s_2)$ denote

$$\mathrm{CFL}^-(s_1+1, s_2+1) \oplus \mathrm{CFL}^-(s_1+1, s_2) \oplus \mathrm{CFL}^-(s_1, s_2+1) \oplus \mathrm{CFL}^-(s_1, s_2).$$

For any $(s_1, s_2) \in \mathbb{H}$, we have maps $U_1 : \mathrm{CFL}^-(s_1, s_2) \to \mathrm{CFL}^-(s_1 - 1, s_2)$ and $U_2 : \mathrm{CFL}^-(s_1, s_2) \to \mathrm{CFL}^-(s_1, s_2 - 1)$. The action of U_1 (or U_2) is defined by the h-function (see Lemma 3.3). Let $D = d + d_1$, where d is the differential in chain complex $\mathrm{CFL}^-(s_1, s_2)$ and $d_1 = U_1 - U_2$. Then we have the "iterated cone" complex $(\mathfrak{C}(s_1, s_2), d + d_1)$ in the following form:

$$\begin{array}{ccc}
\mathrm{CFL}^-(s_1, s_2+1) & \xleftarrow[U_1]{} & \mathrm{CFL}^-(s_1+1, s_2+1) \\
{\scriptstyle U_2}\big\downarrow & & {\scriptstyle U_2}\big\downarrow \\
\mathrm{CFL}^-(s_1, s_2) & \xleftarrow[U_1]{} & \mathrm{CFL}^{-1}(s_1+1, s_2)
\end{array}$$

Lemma 3.1. *Suppose that $L = L_1 \cup L_2$ is an L-space link. Let $\widehat{\mathrm{CFL}}(s_1, s_2)$ denote the chain complex of the hat-version of Heegaard Floer link homology of L with $(s_1, s_2) \in \mathbb{H}$. Then $\widehat{\mathrm{CFL}}(s_1, s_2)$ is quasi-isomorphic to the "iterated cone" complex $(\mathfrak{C}(s_1, s_2), d + d_1)$.*

Proof. We can write $\widehat{\mathrm{CFL}}(s_1, s_2)$ as

$$\frac{\mathrm{CFL}^-(s_1, s_2)/U_1(\mathrm{CFL}^-(s_1+1, s_2))}{U_2(\mathrm{CFL}^-(s_1, s_2+1)/U_1(\mathrm{CFL}^-(s_1+1, s_2+1)))}.$$

The quotient $\mathrm{CFL}^-(s_1, s_2)/U_1(\mathrm{CFL}^-(s_1+1, s_2))$ can be realized as the mapping cone of $U_1 : \mathrm{CFL}^-(s_1+1, s_2) \to \mathrm{CFL}^-(s_1, s_2)$, and similarly the quotient $\mathrm{CFL}^-(s_1, s_2+1)/U_1(\mathrm{CFL}^-(s_1+1, s_2+1))$ can be realized as the mapping cone of $U_1 : \mathrm{CFL}^-(s_1+1, s_2+1) \to \mathrm{CFL}^-(s_1, s_2+1)$. Thus $\widehat{\mathrm{CFL}}(s_1, s_2)$ can be realized as the cone of the natural map induced by U_2 between these two cones. $\square$

Theorem 3.2. *Let $L = L_1 \cup L_2$ be an L-space link with two components. For any $(s_1, s_2) \in \mathbb{H}$, there exists a spectral sequence with the following properties*:

(a) *Its E^2-page is isomorphic* (*as a graded $\mathbb{F}$-module*) *to $H_*(H_*(\mathfrak{C}(s_1, s_2), d), d_1)$.*

(b) *Its E^∞-page is isomorphic* (*as a graded $\mathbb{F}$-module*) *to $\widehat{\mathrm{HFL}}(s_1, s_2)$.*

(c) *The spectral sequence collapses at the E^3-page.*

Proof. For the "iterated cone" complex $\mathfrak{C}(s_1, s_2)$, it is doubly graded. One is the homological grading ν in the chain complex $\mathrm{CFL}^-(s_1, s_2)$ with $(s_1, s_2) \in \mathbb{H}$. We define *cube grading* $|C|$ in the cube of the "iterated cone" complex $\mathfrak{C}(s_1, s_2)$. Fix $(s_1, s_2) \in \mathbb{H}$. The cube grading is defined as $(s_1+s_2)-(v_1+v_2)$, where $(v_1, v_2) \in \mathbb{H}$. It is equivalent to saying that the cube grading of the lower left corner is 0, and U_1 (or U_2) increases the cube grading by 1.

The spectral sequence corresponding to the doubly-graded complex $\mathfrak{C}(s_1, s_2)$ with two (anti)commuting differentials d and d_1 converges to $H_*(\mathfrak{C}(s_1, s_2), d+d_1)$. By Lemma 3.1, its E^∞-page is isomorphic to $\widehat{\mathrm{HFL}}(s_1, s_2)$. Its E^1-page is written as $\mathrm{HFL}^-(s_1+1, s_2+1) \oplus \mathrm{HFL}^-(s_1+1, s_2) \oplus \mathrm{HFL}^-(s_1, s_2+1) \oplus \mathrm{HFL}^-(s_1, s_2)$. Its E^2-page is $H_*(H_*(\mathfrak{C}(s_1, s_2), d), d_1)$. The differential $d_0 = d$ preserves the cube grading $|C|$ and decreases the homological degree ν by 1. The differential d_1 in the E^1-page increases $|C|$ by 1 and decreases ν by 2. For any nonnegative integer k, the differential d_k increases $|C|$ by k and decreases ν by $k+1$. The total homological grading is $\nu + |C|$. By grading reasons, the cube grading is less than or equal to 2. Thus, for the integer $k > 2$, $d_k = 0$ and this spectral sequence collapses at the E^3-page. □

By Theorem 3.2, $\widehat{\mathrm{HFL}}(s_1, s_2) \cong E^3$. Then we can compute $\widehat{\mathrm{HFL}}(s_1, s_2)$ by computing the E^3-page of the spectral sequence. The following lemma describes the action of U_1 (or U_2) on the E^1-page.

Lemma 3.3. *Consider the map* $U_1 : \mathrm{HFL}^-(s_1+1, s_2+1) \to \mathrm{HFL}^-(s_1, s_2+1)$. *Let* α *be a generator of* $\mathrm{HFL}^-(s_1+1, s_2+1)$ *with total homological grading* x. *If there exists a generator* β *in* $\mathrm{HFL}^-(s_1, s_2+1)$ *with total homological grading* $x-2$, *then* $U_1(\alpha) = \beta$.

Proof. As shown in Figure 2, let a_1, b_1, c_1 and d_1 denote the generators of $H_*(A^-(s_1, s_2+1))$, $H_*(A^-(s_1-1, s_2+1))$, $H_*(A^-(s_1-1, s_2))$ and $H_*(A^-(s_1, s_2))$, respectively, and likewise a, b, c and d the generators of $H_*(A^-(s_1+1, s_2+1))$, $H_*(A^-(s_1, s_2+1))$, $H_*(A^-(s_1, s_2))$ and $H_*(A^-(s_1+1, s_2))$. Here a_1 and b have different cube gradings as generators of $H_*(A^-(s_1, s_2+1))$ and d_1 and c have different cube gradings as generators of $H_*(A^-(s_1, s_2))$. By the computation of HFL^- in Section 2A, $h(s_1, s_2+1) = h(s_1+1, s_2)$ if $\mathrm{HFL}^-(s_1+1, s_2+1)$ is nonempty. Similarly, $h(s_1-1, s_2+1) = h(s_1, s_2)$ since $\mathrm{HFL}^-(s_1, s_2)$ is also nonempty. Assume that $\alpha = b - d$. Then it has total homological grading $-2h(s_1, s_2+1)+1$. The generator a_1 has total homological grading $-2h(s_1, s_2+1)$, and $b_1 - d_1$ has total homological grading $-2h(s_1-1, s_2+1)+1$. By the assumption of this lemma, the total homological grading of β is $-2h(s_1, s_2+1)-1$. So β can only be $b_1 - d_1$, and $h(s_1-1, s_2+1) = h(s_1, s_2+1)+1$.

Now consider the map $U_1 : H_*(A^-(s_1, s_2+1)) \to H_*(A^-(s_1-1, s_2+1))$, where $H_*(A^-(s_1, s_2+1)) = \langle b \rangle$ and $H_*(A^-(s_1-1, s_2+1)) = \langle b_1 \rangle$. Since U_1 has

$A^-(s_1-1, s_2+1)[b_1]$	$A^-(s_1, s_2+1)[a_1]$
$A^-(s_1-1, s_2)[c_1]$	$A^-(s_1, s_2)[d_1]$

$\mathrm{HFL}^-(s_1, s_2+1)$

$A^-(s_1, s_2+1)[b]$	$A^-(s_1+1, s_2+1)[a]$
$A^-(s_1, s_2)[c]$	$A^-(s_1+1, s_2)[d]$

$\mathrm{HFL}^-(s_1+1, s_2+1)$

Figure 2. Generators for Lemma 3.3.

homological degree -2, $U_1(d) = d_1$ by Lemma 2.4 and Remark 2.5. Similarly, $U_1(c) = c_1$. Then $U_1(\alpha) = U_1(b-d) = b_1 - d_1 = \beta$. If $\alpha = a$, then $\beta = a_1$, and we can use a similar argument to prove $U_1(\alpha) = \beta$ in this case. □

Remark 3.4. The map $U_2 : \mathrm{HFL}^-(s_1+1, s_2+1) \to \mathrm{HFL}^-(s_1+1, s_2)$ can be described similarly to Lemma 3.3.

3B. ***Proof of the main theorem.*** In this subsection, we prove Theorem 1.2, and show that 4 is an upper bound for the rank of link Floer homology $\widehat{\mathrm{HFL}}(s_1, s_2)$ for any 2-component L-space link and $(s_1, s_2) \in \mathbb{H}$. Example 3.8 gives a 2-component L-space link where the rank of $\widehat{\mathrm{HFL}}(s_1, s_2)$ ranges from 0 to 4.

In order to prove Theorem 1.2, we need the symmetric property of Heegaard Floer link homology.

Lemma 3.5 [Ozsváth and Szabó 2006, Equation 5]. *For an oriented L-space link $L = L_1 \cup L_2$ with two components and $\boldsymbol{s} = (s_1, s_2) \in \mathbb{H}$, there exists a relatively graded isomorphism*

$$\widehat{\mathrm{HFL}}(L, \boldsymbol{s}) \cong \widehat{\mathrm{HFL}}(L, -\boldsymbol{s}).$$

Remark 3.6. In particular, the h-functions satisfy $h(-\boldsymbol{s}) = h(\boldsymbol{s}) + |\boldsymbol{s}|$, [Liu 2017, Lemma 5.5], where $|\boldsymbol{s}| = s_1 + s_2$.

Proof of Theorem 1.2. Let $h = h(s_1+1, s_2+1)$. If $d_2 = 0$, then the spectral sequence in Theorem 3.2 collapses at its E^2-page. We can use the computation of HFL^- in Section 2A and Lemma 3.3 to compute $\widehat{\mathrm{HFL}}(s_1, s_2)$. For example, suppose that the h-function corresponding to $\widehat{\mathrm{HFL}}(s_1, s_2)$ is the following:

$h+1$	h	h
$h+1$	$h+1$	h
$h+2$	$h+1$	$h+1$

Then the E^2-page of the spectral sequence is:

$$\begin{array}{ccc} \mathbb{F}[-2h] & \xleftarrow[U_1]{\quad} & \mathbb{F}[-2h+1] \\ {\scriptstyle U_2}\big\downarrow & & {\scriptstyle U_2}\big\downarrow \\ \mathbb{F}[-2h-1] & \xleftarrow[U_1]{\quad} & \mathbb{F}[-2h] \end{array}$$

Since U_1 and U_2 both have homological grading -2, $U_1 = U_2 = 0$. By Theorem 3.2, $d_2 = 0$ since it increases the cube grading by 2, and decreases the homological grading ν by 3. Thus $\widehat{\mathrm{HFL}}(s_1, s_2) \cong \mathbb{F}[-2h-1] \oplus \mathbb{F}[-2h-1] \oplus \mathbb{F}[-2h-1] \oplus \mathbb{F}[-2h-1]$. Here the cube grading for the generator in $\mathbb{F}[-2h-1]$ is 0. We can use this method to compute $\widehat{\mathrm{HFL}}$ in all cases where $d_2 = 0$. Now it suffices to consider cases where d_2 may be nontrivial.

If $d_2 \neq 0$, then $\mathrm{HFL}^-(s_1+1, s_2+1)$ and $\mathrm{HFL}^-(s_1, s_2)$ are both nonzero and contain generators such that their homological grading difference is 3. For nonzero $\mathrm{HFL}^-(s_1+1, s_2+1)$, we have the following three possibilities for the corresponding h-function:

h	h
$h+1$	h

Case (1)

$h+1$	h
$h+1$	$h+1$

Case (2)

$h+1$	h
$h+2$	$h+1$

Case (3)

In Case (1), $\mathrm{HFL}^-(s_1+1, s_2+1) = \mathbb{F}[-2h+1]$. In order to have nontrivial d_2, $\mathrm{HFL}^-(s_1, s_2)$ must contain a generator with homological grading $-2h-2$. So the h-function corresponding to $\mathrm{HFL}^-(s_1, s_2)$ can only have the pattern as in Case (2) or Case (3). Once the h-function in $\mathrm{HFL}^-(s_1, s_2)$ is determined, its values in $\mathrm{HFL}^-(s_1, s_2+1)$ and $\mathrm{HFL}^-(s_1+1, s_2)$ are also determined by Corollary 2.6. Thus there are two possibilities for the h-function corresponding to $\widehat{\mathrm{HFL}}(s_1, s_2)$:

$h+1$	h	h
$h+2$	$h+1$	h
$h+2$	$h+2$	$h+1$

Case (1a)

$h+1$	h	h
$h+2$	$h+1$	h
$h+3$	$h+2$	$h+1$

Case (1b)

In both cases, we have $\mathrm{HFL}^-(s_1+1, s_2+1) = \mathbb{F}[-2h+1]$, $\mathrm{HFL}^-(s_1, s_2+1) = \mathbb{F}[-2h] \oplus \mathbb{F}[-2h-1]$ and $\mathrm{HFL}^-(s_1+1, s_2) = \mathbb{F}[-2h] \oplus \mathbb{F}[-2h-1]$. By Lemma 3.3, $U_1 a = b$ and $U_2 a = c$, where a is the generator in $\mathrm{HFL}^-(s_1+1, s_2+1)$, and b and c are generators with homological grading $-2h-1$ in $\mathrm{HFL}^-(s_1, s_2+1)$ and $\mathrm{HFL}^-(s_1+1, s_2)$, respectively. So the image of a under the differential d_1 is nonzero, and a does not survive in the E^2-page. Thus d_2 is trivial in both Case (1a) and Case (1b).

In Case (2), $HFL^-(s_1+1, s_2+1) = \mathbb{F}[-2h]$. In order to have nontrivial d_2, $HFL^-(s_1, s_2)$ must contain a generator with homological grading $[-2h-3]$. So the h-function in $HFL^-(s_1, s_2)$ must have the pattern in Case (3). Then $HFL^-(s_1, s_2) \cong \mathbb{F}[-2h-2] \oplus \mathbb{F}[-2h-3]$. Corresponding to this case, there are four possibilities for the h-function in $\widehat{HFL}(s_1, s_2)$:

$h+1$	$h+1$	h
$h+2$	$h+1$	$h+1$
$h+3$	$h+2$	$h+1$

Case (2a)

$h+2$	$h+1$	h
$h+2$	$h+1$	$h+1$
$h+3$	$h+2$	$h+2$

Case (2b)

$h+2$	$h+1$	h
$h+2$	$h+1$	$h+1$
$h+3$	$h+2$	$h+1$

Case (2c)

$h+1$	$h+1$	h
$h+2$	$h+1$	$h+1$
$h+3$	$h+2$	$h+2$

Case (2d)

We use the symmetric property of Heegaard Floer link homology to compute $\widehat{HFL}(s_1, s_2)$. Let $h^* = h(-s_1, -s_2)$. By Remark 3.6,

$$h(-s_1, -s_2-1) - h(-s_1, -s_2) = 1 - (h(s_1, s_2) - h(s_1, s_2+1))$$

and

$$h(-s_1-1, -s_2) - h(-s_1, -s_2) = 1 - (h(s_1, s_2) - h(s_1+1, s_2)).$$

So the h-function in $\widehat{HFL}(-s_1, -s_2)$ corresponding to these four subcases are

h^*	h^*	h^*
h^*+1	h^*	h^*
h^*+1	h^*+1	h^*

dual-h (2a)

h^*+1	h^*	h^*
h^*+1	h^*	h^*
h^*+1	h^*+1	h^*+1

dual-h (2b)

h^*	h^*	h^*
h^*+1	h^*	h^*
h^*+1	h^*+1	h^*+1

dual-h (2c)

h^*+1	h^*	h^*
h^*+1	h^*	h^*
h^*+1	h^*+1	h^*

dual-h (2d)

Note that in all these four cases for $\widehat{HFL}(-s_1, -s_2)$, $HFL^-(-s_1+1, -s_2+1) = 0$. So $d_2 = 0$ in the spectral sequence corresponding to $\widehat{HFL}(-s_1, -s_2)$. Now the computation of $\widehat{HFL}(-s_1, -s_2)$ is quite straightforward.

In dual-h (2a),

$$\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\left[\begin{array}{ccc} \mathbb{F}[-2h^*+1] & \xleftarrow[U_1]{\hspace{3em}} & 0\\ {\scriptstyle U_2}\big\downarrow & & {\scriptstyle U_2}\big\downarrow\\ \mathbb{F}[-2h^*] & \xleftarrow[U_1]{} & \mathbb{F}[-2h^*+1]\end{array}\right]$$

By grading reasons, $d_2=U_1=U_2=0$. Then it is easy to obtain $\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\mathbb{F}[-2h^*]\oplus\mathbb{F}[-2h^*]\oplus\mathbb{F}[-2h^*]$, and the Euler characteristic $\chi=3$. By Lemma 3.5, $\widehat{\mathrm{HFL}}(s_1,s_2)$ contains 3 generators with the same total grading. Observe that $\mathrm{HFL}^-(s_1,s_2)=\mathbb{F}[-2h-2]\oplus\mathbb{F}[-2h-3]$. Then the generator with total grading $-2h-2$ survives in $\widehat{\mathrm{HFL}}(s_1,s_2)$. Thus

$$\widehat{\mathrm{HFL}}(s_1,s_2)\cong\mathbb{F}[-2h-2]\oplus\mathbb{F}[-2h-2]\oplus\mathbb{F}[-2h-2],$$

and the Euler characteristic χ is 3.

In dual-h (2b),

$$\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\left[\begin{array}{ccc} 0 & \xleftarrow[U_1]{\hspace{3em}} & 0\\ {\scriptstyle U_2}\big\downarrow & & {\scriptstyle U_2}\big\downarrow\\ \mathbb{F}[-2h^*] & \xleftarrow[U_1]{} & 0\end{array}\right]$$

In this case, $\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\mathbb{F}[-2h^*]$. By an argument similar to the one in dual-h (2a), we obtain that $\widehat{\mathrm{HFL}}(L)(s_1,s_2)\cong\mathbb{F}[-2h-2]$, and the Euler characteristic χ is 1.

In dual-h (2c),

$$\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\left[\begin{array}{ccc} \mathbb{F}[-2h^*+1] & \xleftarrow[U_1]{} & 0\\ {\scriptstyle U_2}\big\downarrow & & {\scriptstyle U_2}\big\downarrow\\ \mathbb{F}[-2h^*] & \xleftarrow[U_1]{} & 0\end{array}\right]$$

By grading reasons, $d_2=U_1=U_2=0$. Then $\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\mathbb{F}[-2h^*]\oplus\mathbb{F}[-2h^*]$. So $\widehat{\mathrm{HFL}}(s_1,s_2)\cong\mathbb{F}[-2h-2]\oplus\mathbb{F}[-2h-2]$, and the Euler characteristic is $\chi=2$.

In dual-h (2d),

$$\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\left[\begin{array}{ccc} 0 & \xleftarrow[U_1]{\hspace{3em}} & 0\\ {\scriptstyle U_2}\big\downarrow & & {\scriptstyle U_2}\big\downarrow\\ \mathbb{F}[-2h^*] & \xleftarrow[U_1]{} & \mathbb{F}[-2h^*+1]\end{array}\right]$$

Hence, $\widehat{\mathrm{HFL}}(L)(s_1,s_2)\cong\mathbb{F}[-2h-2]\oplus\mathbb{F}[-2h-2]$, and the Euler characteristic is $\chi=2$.

Now we consider Case (3). In this case, we have $\mathrm{HFL}^-(s_1+1, s_2+1) \cong \mathbb{F}[-2h] \oplus \mathbb{F}[-2h-1]$. Then there are three possibilities for $\mathrm{HFL}^-(s_1, s_2)$ if d_2 is nontrivial: $\mathrm{HFL}^-(s_1, s_2)$ is either $\mathbb{F}[-2h-4]$ or $\mathbb{F}[-2h-4] \oplus \mathbb{F}[-2h-5]$ or $\mathbb{F}[-2h-3]$. If $\mathrm{HFL}^-(s_1, s_2) = \mathbb{F}[-2h-4]$, its h-function is shown in Case (3a), and if $\mathrm{HFL}^-(s_1, s_2) \cong \mathbb{F}[-2h-4] \oplus \mathbb{F}[-2h-5]$, its h-function is shown in Case (3b):

$h+2$	$h+1$	h
$h+3$	$h+2$	$h+1$
$h+3$	$h+3$	$h+2$

Case (3a)

$h+2$	$h+1$	h
$h+3$	$h+2$	$h+1$
$h+4$	$h+3$	$h+2$

Case (3b)

In Case (3a) and Case (3b), we observe that both generators in $\mathrm{HFL}^-(s_1+1, s_2+1)$ have nontrivial images in $\mathrm{HFL}^-(s_1, s_2+1)$ and $\mathrm{HFL}^-(s_1+1, s_2)$ by Lemma 3.3. So these two generators have nontrivial images under the differential d_1, and cannot survive in the E^2-page. Thus d_2 is trivial in both cases.

If $\mathrm{HFL}^-(s_1, s_2) \cong \mathbb{F}[-2h-3]$, there are four possibilities for the h-function corresponding to $\widehat{\mathrm{HFL}}(s_1, s_2)$:

$h+1$	$h+1$	h
$h+2$	$h+2$	$h+1$
$h+3$	$h+2$	$h+1$

Case (3c)

$h+2$	$h+1$	h
$h+2$	$h+2$	$h+1$
$h+3$	$h+2$	$h+2$

Case (3d)

$h+1$	$h+1$	h
$h+2$	$h+2$	$h+1$
$h+3$	$h+2$	$h+2$

Case (3e)

$h+2$	$h+1$	h
$h+2$	$h+2$	$h+1$
$h+3$	$h+2$	$h+1$

Case (3f)

Let $h^* = h(-s_1, -s_2) = h(s_1, s_2) + s_1 + s_2$. By Remark 3.6, we find the h-function in $\widehat{\mathrm{HFL}}(-s_1, -s_2)$ corresponding to each case:

h^*-1	h^*-1	h^*-1
h^*	h^*	h^*-1
h^*	h^*	h^*-1

dual-h (3c)

h^*	h^*-1	h^*-1
h^*	h^*	h^*-1
h^*	h^*	h^*

dual-h (3d)

h^*	h^*-1	h^*-1
h^*	h^*	h^*-1
h^*	h^*	h^*-2

dual-h (3e)

h^*-1	h^*-1	h^*-1
h^*	h^*	h^*-1
h^*	h^*	h^*

dual-h (3f)

Observe that in these four cases, $\mathrm{HFL}^-(-s_1,-s_2)=0$. So d_2 is trivial in the spectral sequence corresponding to $\widehat{\mathrm{HFL}}(-s_1,-s_2)$. We compute $\widehat{\mathrm{HFL}}(-s_1,-s_2)$, and hence $\widehat{\mathrm{HFL}}(s_1,s_2)$.

In dual-h (3c), $\widehat{\mathrm{HFL}}(-s_1,-s_2)\cong\mathbb{F}[-2h^*+1]$. By Lemma 3.5, $\widehat{\mathrm{HFL}}(s_1,s_2)\cong\mathbb{F}[-2h-3]$ with the Euler characteristic $\chi=-1$.

In Case (3d), $\widehat{\mathrm{HFL}}(L)(s_1,s_2)\cong\mathbb{F}[-2h-3]\oplus\mathbb{F}[-2h-3]\oplus\mathbb{F}[-2h-3]$, and the Euler characteristic is $\chi=-3$ by a similar computation.

In Case (3e), $\widehat{\mathrm{HFL}}(L)(s_1,s_2)\cong\mathbb{F}[-2h-3]\oplus\mathbb{F}[-2h-3]$, and the Euler characteristic is $\chi=-2$.

In Case (3f), $\widehat{\mathrm{HFL}}(L)(s_1,s_2)\cong\mathbb{F}[-2h-3]\oplus\mathbb{F}[-2h-3]$, and the Euler characteristic is $\chi=-2$.

Thus we conclude that for any L-space link $L=L_1\cup L_2$ with two components, if the h-function is determined, we can compute $\widehat{\mathrm{HFL}}(s_1,s_2)$ with any $(s_1,s_2)\in\mathbb{H}$. By equations in Section 2B, the h-function is determined by Alexander polynomials $\Delta_L(x_1,x_2)$, $\Delta_{L_1}(t)$, $\Delta_{L_2}(t)$ and the linking number $\mathrm{lk}(L_1,L_2)$. □

Furthermore, we also get a bound for $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(s_1,s_2))$ and the Euler characteristic $\chi(\widehat{\mathrm{HFL}}(s_1,s_2))$ with any $(s_1,s_2)\in\mathbb{H}$.

Proof of Corollary 1.4. Consider the short exact sequence

$$0\to\mathrm{CFL}^-(s_1+1,s_2+1)\xrightarrow{U_1}\mathrm{CFL}^-(s_1,s_2+1)\to C_1(s_1,s_2+1)\to 0, \tag{3-1}$$

where $C_1(s_1,s_2+1)$ is the quotient complex with $(s_1,s_2+1)\in\mathbb{H}$. By Lemma 3.1,

$$\widehat{\mathrm{CFL}}(s_1,s_2)\cong C_1(s_1,s_2)/U_2(C_1(s_1,s_2+1)). \tag{3-2}$$

Now we claim that $\mathrm{rank}_{\mathbb{F}}(H_*(C_1(s_1,s_2+1)))\le 2$ for any $(s_1,s_2)\in\mathbb{H}$. From the short exact sequence (3-1), we have

$$\begin{aligned}&\mathrm{rank}_{\mathbb{F}}(H_*(C_1(s_1,s_2+1)))\\ &\qquad\le\mathrm{rank}_{\mathbb{F}}(\mathrm{HFL}^-(s_1+1,s_2+1))+\mathrm{rank}_{\mathbb{F}}(\mathrm{HFL}^-(s_1,s_2+1)).\end{aligned} \tag{3-3}$$

If $\mathrm{rank}_{\mathbb{F}}(H_*(C_1(s_1,s_2+1)))\ge 3$, then at least one of $\mathrm{HFL}^-(s_1+1,s_2+1)$ and $\mathrm{HFL}^-(s_1,s_2+1)$ should have rank at least 2, and the other one should have rank at least 1. By the computation in Section 2A, the h-functions corresponding to $\mathrm{HFL}^-(s_1+1,s_2+1)$ and $\mathrm{HFL}^-(s_1,s_2+1)$ have the following possibilities:

$h+1$	h	$h-1$
$h+1$	$h+1$	h

Case (1)

$h+1$	h	h
$h+2$	$h+1$	h

Case (2)

$h+1$	h	$h-1$
$h+2$	$h+1$	h

Case (3)

Here we assume that the generator of $H_*(A^-(s_1, s_2+1))$ has homological grading $-2h$. In Case (1), we have $U_1 : \mathbb{F}[-2h+2] \oplus \mathbb{F}[-2h+1] \to \mathbb{F}[-2h]$. Let α denote the generator of $\mathbb{F}[-2h+2] \subseteq \mathrm{HFL}^-(s_1+1, s_2+1)$, and β the generator of $\mathbb{F}[-2h] \cong \mathrm{HFL}^-(s_1, s_2+1)$. By Lemma 3.3, $U(\alpha) = \beta$. Then $H_*(C_1(s_1)) \cong \mathbb{F}[-2h+1]$, and the rank in this case is 1.

In Case (2), we have $U_1 : \mathbb{F}[-2h+1] \to \mathbb{F}[-2h] \oplus \mathbb{F}[-2h-1]$. Similarly $H_*(C_1(s_1, s_2+1)) \cong \mathbb{F}[-2h]$, and it has rank 1.

In Case (3), we have $U_1 : \mathbb{F}[-2h+2] \oplus \mathbb{F}[-2h+1] \to \mathbb{F}[-2h] \oplus \mathbb{F}[-2h-1]$. By Lemma 3.3, $H_*(C_1(s_1, s_2+1)) = 0$.

Thus for any $(s_1, s_2) \in \mathbb{H}$, $\mathrm{rank}_{\mathbb{F}}(H_*(C_1(s_1, s_2+1))) \leq 2$. By (3-2),

$$\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(s_1, s_2)) \leq \mathrm{rank}_{\mathbb{F}}(H_*(C_1(s_1, s_2+1))) + \mathrm{rank}_{\mathbb{F}}(H_*(C_1(s_1, s_2))) \leq 2+2=4$$

for any $(s_1, s_2) \in \mathbb{H}$. Therefore, $-4 \leq \chi(\widehat{\mathrm{HFL}}(L, s_1, s_2)) \leq 4$. □

In fact, we construct an example with $\chi(\widehat{\mathrm{HFL}}(L, s_1, s_2)) = -4$, given in the proof of Theorem 1.2, where $d_2 = 0$. Similarly, we construct an example with $\chi(\widehat{\mathrm{HFL}}(L, s_1, s_2)) = 4$.

Example 3.7. Assume that the h-function corresponding to $\widehat{\mathrm{HFL}}(s_1, s_2)$ is the following:

$h+1$	$h+1$	h
$h+2$	$h+1$	$h+1$
$h+2$	$h+2$	$h+1$

In this case, $\widehat{\mathrm{HFL}}(s_1, s_2) \cong \mathbb{F}[-2h-2] \oplus \mathbb{F}[-2h-2] \oplus \mathbb{F}[-2h-2] \oplus \mathbb{F}[-2h-2]$, and hence $\chi(\widehat{\mathrm{HFL}}(s_1, s_2)) = 4$.

Example 3.8. Figure 3 depicts the two-bridge link $b(20, -3)$.

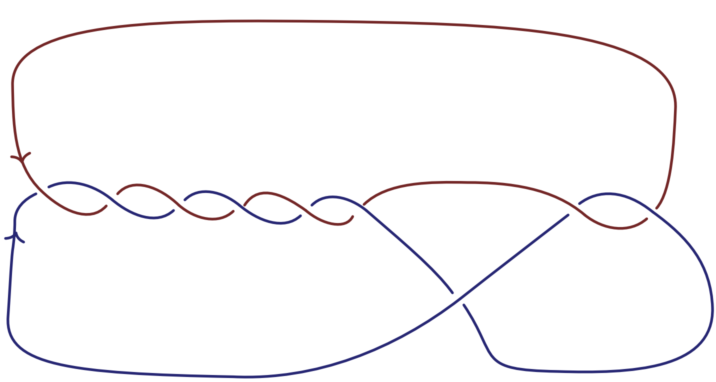


Figure 3. $b(20, -3)$.

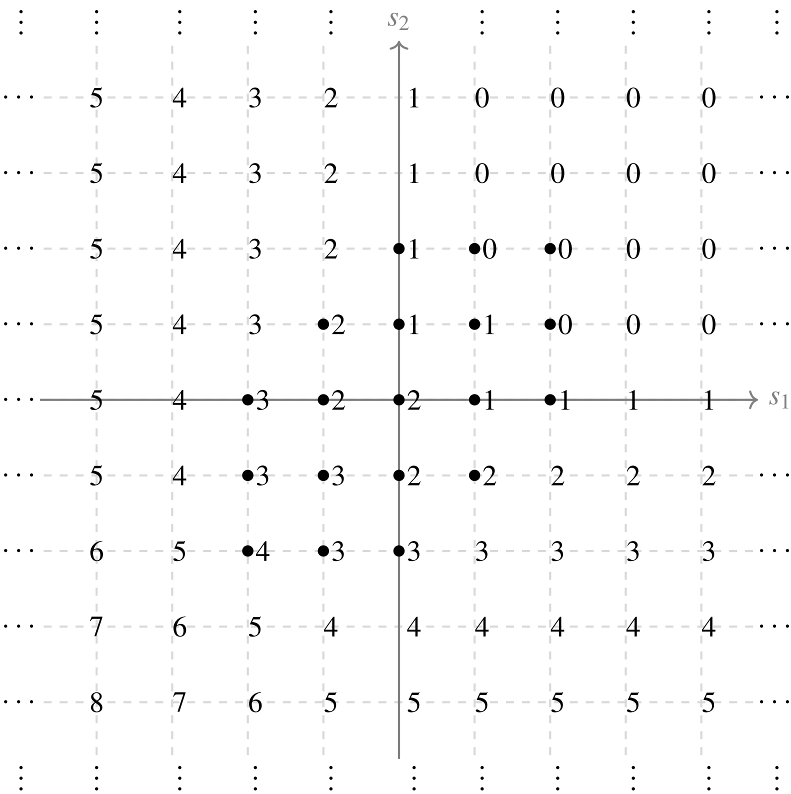


Figure 4. The h-function for $b(20, -3)$.

Yajing Liu proved that $b(20, -3)$ is an L-space link [2017, Theorem 3.8]. Its two components are both unknots with linking number 2. By [Dawra 2015], its normalized multivariable Alexander polynomial is

$$\begin{aligned}(3\text{-}4)\quad \Delta_L(t_1, t_2) = {} & t_1^{1/2}t_2^{3/2} + t_1^{3/2}t_2^{1/2} + t_1^{1/2}t_2^{-1/2} + t_1^{-1/2}t_2^{1/2} + t_1^{-3/2}t_2^{-1/2} \\ & + t_1^{-1/2}t_2^{-3/2} - t_1^{3/2}t_2^{3/2} - t_1^{1/2}t_2^{1/2} - t_1^{-1/2}t_2^{-1/2} - t_1^{-3/2}t_2^{-3/2}.\end{aligned}$$

Let L_1 and L_2 denote the unknot components. We obtain normalized Alexander polynomials of L_1 and L_2:

$$\frac{t}{t-1}\Delta_{L_1}(t) = \frac{t}{t-1}\Delta_{L_2}(t) = 1 + t^{-1} + t^{-2} + t^{-3} + t^{-4} + \cdots .$$

Using results of Section 2B, we compute the h-function for $\widehat{\mathrm{HFL}}(s_1, s_2)$ with any $(s_1, s_2) \in \mathbb{H}$ by Alexander polynomials. The h-function is shown in Figure 4, where numbers denote $h(s_1, s_2)$ for any $(s_1, s_2) \in \mathbb{H}$. For example, $h(0, 0) = h(-1, 0) = 2$. The black dots $\bullet$ denote the lattice points $(s_1, s_2) \in \mathbb{H}$ where $\widehat{\mathrm{HFL}}(s_1, s_2)$ is nonzero.

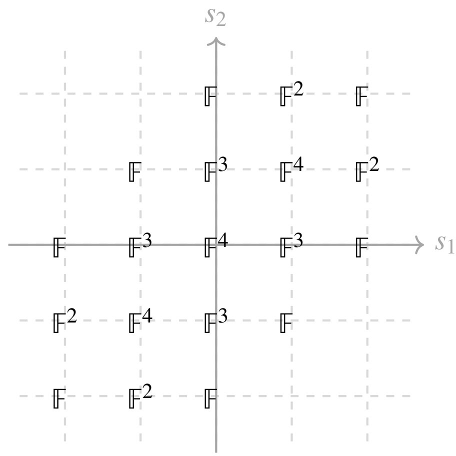


Figure 5. $\widehat{\mathrm{HFL}}(b(20, -3))$.

By an explicit computation, the link Floer homology $\widehat{\mathrm{HFL}}(s_1, s_2)$ is shown in Figure 5. We observe that $|\chi(s_1, s_2)| = \mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(s_1, s_2))$, and the rank of $\widehat{\mathrm{HFL}}(s_1, s_2)$ ranges from 0 to 4. This indicates that the bound for the rank in Corollary 1.4 can be realized by some L-space link with some $(s_1, s_2) \in \mathbb{H}$. More precisely, $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(2, 2)) = 1$, $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(2, 1)) = 2$, $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(1, 0)) = 3$, $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(0, 0)) = 4$ and $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(3, 0)) = 0$.

4. An application of $\widehat{\mathrm{HFL}}$ to the Thurston norm

The Thurston norm was studied by many people, and some lower bounds were obtained in [McMullen 2002; Friedl and Kim 2008; Friedl and Vidussi 2015; Agol and Dunfield 2015]. Ozsváth and Szabó [2008b] showed that the link Floer homology detects the Thurston norm of the link complement. In Section 3, for any 2-component L-space link $L = L_1 \cup L_2$ and $\boldsymbol{s} \in \mathbb{H}$, we computed $\widehat{\mathrm{HFL}}(L, \boldsymbol{s})$ by using Alexander polynomials $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$, $\Delta_{L_2}(t)$ and the linking number $\mathrm{lk}(L_1, L_2)$. Thus we can compute the link Floer homology polytope for L, and also compute the dual Thurston polytope and the Thurston (semi)norm [Ozsváth and Szabó 2008b, Theorem 1.1].

In Section 1, we introduced complexity $\chi_-(F)$ for any compact oriented surface F with boundary. To any link $L \subseteq S^3$, and any homology class $h \in H_2(S^3, L)$, we can assign a function

$$x(h) = \min_{\{F \hookrightarrow S^3 \backslash \mathrm{nd}(L),\, [F]=h\}} \chi_-(F).$$

This function can be naturally extended to a seminorm, the *Thurston seminorm*, denoted by $x : H_2(S^3, L; \mathbb{R}) \to \mathbb{R}$.

Theorem 4.1 [Thurston 1986, Theorem 1]. *The function* $x : H_2(S^3, L; \mathbb{R}) \to \mathbb{R}$ *is a seminorm that vanishes exactly on the subspace spanned by embedded surfaces of nonnegative Euler characteristic.*

Assume that $L \subseteq S^3$ is a link with l components in S^3. Let u_i denote the meridian of the i-th component L_i of L. Recall that every lattice point $\boldsymbol{s} \in \mathbb{H}$ can be written as

$$\sum_{i=1}^{l} s_i \cdot [u_i],$$

where $s_i \in \mathbb{Q}$ satisfies the property that

$$2s_i + \mathrm{lk}(L_i, L - L_i)$$

is an even integer for $i = 1, \ldots, l$.

In [Ozsváth and Szabó 2008b], the Heegaard Floer link homology provides a function $y : H^1(S^3 - L; \mathbb{R}) \to \mathbb{R}$ defined by the formula

$$y(h) = \max_{\{\boldsymbol{s} \in \mathbb{H} \subseteq H_1(S^3 - L; \mathbb{R}) \,|\, \widehat{\mathrm{HFL}}(L, \boldsymbol{s}) \neq 0\}} |\langle s, h\rangle|.$$

Ozsváth and Szabó proved the following formula for the link Floer homology and the Thurston norm.

Theorem 4.2 [Ozsváth and Szabó 2008b, Theorem 1.1]. *For an oriented link* $L \subseteq S^3$ *with no trivial components, the Heegaard Floer link homology detects the Thurston (semi)norm of its complement. For each* $h \in H^1(S^3 - L; \mathbb{R})$, *we have*

$$x(\mathrm{PD}[h]) + \sum_{i=1}^{l} |\langle h, u_i\rangle| = 2y(h),$$

where u_i *is the meridian of the* i*-th component of* L *and* $|\langle h, u_i\rangle|$ *denotes the absolute value of the Kronecker pairing of* $h \in H^1(S^3 - L; \mathbb{R})$ *and* $u_i \in H_1(S^3 - L; \mathbb{R})$.

Remark 4.3. A trivial component of a link L is an unknot component which is also unlinked from the rest of the link.

The unit ball for the norm x is called the *Thurston polytope*, and the unit ball for the norm y is called the *link Floer homology polytope*, which is also the convex hull of those $\boldsymbol{s} \in \mathbb{H}$ for which $\widehat{\mathrm{HFL}}(L, \boldsymbol{s}) \neq 0$. The unit ball for the dual norm x^* of x in $H_1(S^3 - L; \mathbb{R})$ is called the *dual Thurston polytope*. By Theorem 4.2, twice the link Floer homology polytope can be written as the sum of the dual Thurston polytope and an element of the symmetric hypercube in $H^1(S^3 - L)$ with edge-length two [Ozsváth and Szabó 2008b]. We give some examples of L-space links with two components, and compute their link Floer homology polytopes by using Alexander polynomials and linking numbers in detail. Moreover, we compute the

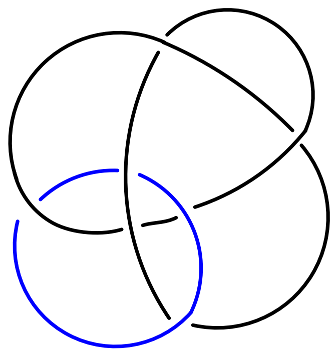


Figure 6. $L7n1$.

dual Thurston polytopes and Thurston norms of their complements by Theorem 4.2. We also compare the link Floer homology polytope and the convex hull of those $\boldsymbol{s} \in \mathbb{H}$ for which $\chi(\widehat{\mathrm{HFL}}(L, \boldsymbol{s})) \neq 0$.

Example 4.4 (the dual Thurston polytope for the L-space link $L7n1$). The link $L7n1$ in Figure 6 is an L-space link [Liu 2017, Example 3.17]. The link component L_1 is an unknot and the other link component L_2 is a right-handed trefoil. The linking number is 2 and its multivariable Alexander polynomial is

$$\Delta_L(t_1, t_2) = t_1^{1/2} t_2^{3/2} + t_1^{-1/2} t_2^{-3/2}.$$

Normalized Alexander polynomials of L_1 and L_2 are

$$\frac{t}{t-1}\Delta_{L_1}(t) = 1 + t^{-1} + t^{-2} + t^{-3} + t^{-4} + \cdots,$$
$$\frac{t}{t-1}\Delta_{L_2}(t) = t + t^{-1} + t^{-2} + t^{-3} + t^{-4} + \cdots.$$

The h-function in $\widehat{\mathrm{HFL}}(s_1, s_2)$ is shown in Figure 7. In this figure, the numbers denote the h-function, and $\bullet$ denotes the lattice points $(s_1, s_2) \in \mathbb{H}$ where $\widehat{\mathrm{HFL}}(s_1, s_2)$ is nonzero. By an explicit computation, the link Floer homology $\widehat{\mathrm{HFL}}(s_1, s_2)$ is shown in Figure 8. Moreover, $\widehat{\mathrm{HFL}}(0, 0) \cong \mathbb{F}[-2] \oplus \mathbb{F}[-3]$, so $\chi(\widehat{\mathrm{HFL}}(0, 0))$ is zero. For any other lattice point (s_1, s_2) labeled by $\bullet$ except $(0, 0)$, $\widehat{\mathrm{HFL}}(s_1, s_2)$ has rank one and $\chi(\widehat{\mathrm{HFL}}(s_1, s_2))$ is also nonzero. Thus in this example, the link Floer homology polytope is the same as the convex hull of those $(s_1, s_2) \in \mathbb{H}$ for which $\chi(\widehat{\mathrm{HFL}}(s_1, s_2))$ are nonzero. By Theorem 4.2, the dual Thurston polytope in $H_1(S^3 - L; \mathbb{R})$ is shown in Figure 9.

In Figure 9, the thick red line is the dual Thurston polytope for $L7n1$. It is the same as the Newton polytope of the Alexander polynomial $\Delta_L(t_1, t_2)$. The

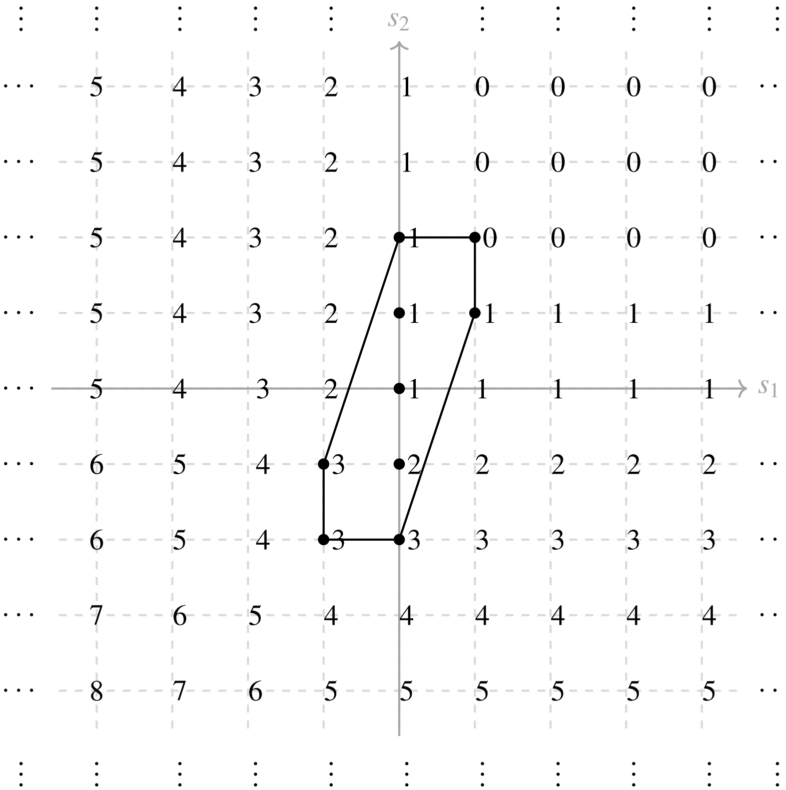


Figure 7. The h-function for $L7n1$.

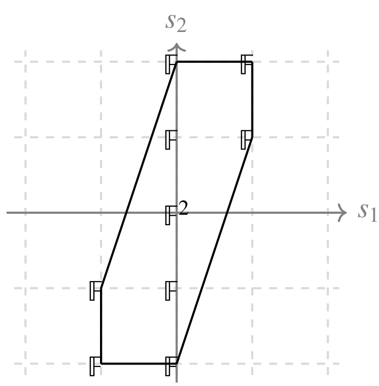


Figure 8. The link Floer homology polytope for $L7n1$.

unknot component of $L7n1$ bounds a surface F_{L_1} with Euler characteristic -1, and the right-handed trefoil link component L_2 bounds a surface F_{L_2} with Euler

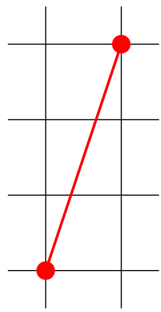


Figure 9. The dual Thurston polytope for $L7n1$.

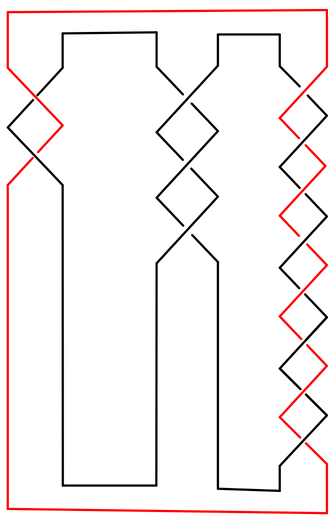


Figure 10. $b(-2, 3, 8)$.

characteristic -3. The surfaces F_{L_1} and F_{L_2} have maximal Euler characteristic in their respective homology classes.

Example 4.5 (the dual Thurston polytope for the pretzel link $L = b(-2, 3, 8)$). We claim that the pretzel link $b(-2, 3, 8)$ is an L-space link with two components. The link component L_1 is an unknot and the other link component L_2 is a right-handed trefoil as shown in Figure 10. The linking number of L_1 and L_2 is 5. Let P_1 be the knot obtained from $b(-2, 3, 8)$ by 1-Dehn surgery on L_1. It is the twisted torus knot $K(5, 6; 2, 1)$ [Remigio-Juárez and Rieck 2012, Proposition 3.1], and it is an L-space knot as proved by F. Vafaee [2015, Theorem 1]. Then for sufficiently large d, $S^3_{1,d}(L) = S^3_{d-25}(P_1)$ is an L-space. The link components L_1 and L_2 are

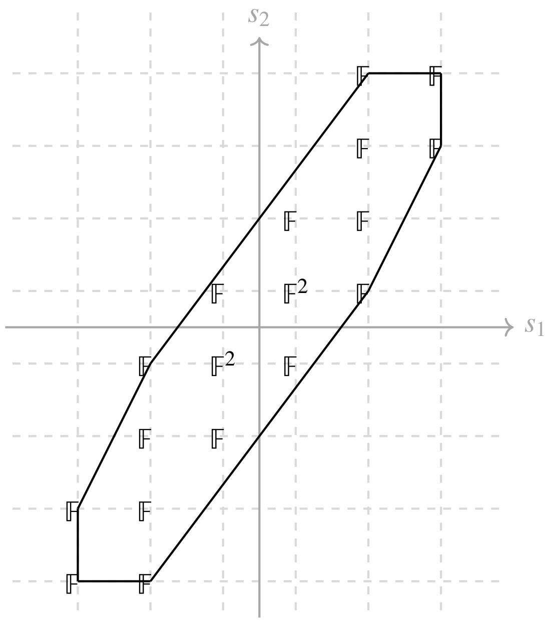


Figure 11. The link Floer homology polytope for $b(-2, 3, 8)$.

L-space knots, so $S^3_1(L_1)$ and $S^3_d(L_2)$ are both L-spaces. Observe that $d - 25 > 0$, so the pretzel link $b(-2, 3, 8)$ is an L-space link by L-space surgery criterion [Liu 2017, Lemma 2.6]. The symmetrized Alexander polynomial of $b(-2, 3, 8)$ is

$$\Delta_L(t_1, t_2) = t_1^{-2}t_2^{-3} + t_1^{-1}t_2^{-2} + 1 + t_1t_2 + t_1^2t_2^3.$$

The h-function corresponding to $\widehat{\mathrm{HFL}}(s_1, s_2)$ with $(s_1, s_2) \in \mathbb{H}$ is shown in Figure 13. By an explicit computation, the link Floer homology $\widehat{\mathrm{HFL}}(s_1, s_2)$ is as shown in Figure 11. We have $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(1/2, 1/2)) = \chi(\widehat{\mathrm{HFL}}(1/2, 1/2)) = 2$, and $\mathrm{rank}_{\mathbb{F}}(\widehat{\mathrm{HFL}}(-1/2, -1/2)) = \chi(\widehat{\mathrm{HFL}}(-1/2, -1/2)) = 2$. Observe that the link Floer homology polytope is the same as the convex hull of those $(s_1, s_2) \in \mathbb{H}$ for which $\chi(\widehat{\mathrm{HFL}}(s_1, s_2))$ are nonzero. By Theorem 4.2, the dual Thurston polytope is the shaded area in Figure 12.

Remark 4.6. For L-space links $L7n1$ and $b(-2, 3, 8)$, the Thurston polytopes are both dual to Newton polytopes of their symmetrized Alexander polynomials $\Delta_L(t_1, t_2)$. Ozsváth and Szabó [2008b] pointed out that the Thurston polytope of an alternating link is dual to the Newton polytope of its multivariable Alexander polynomial. This is also true for L-space knots. A natural question is whether the Thurston polytope of an L-space link with two components (which is not a split union of two L-space knots) is dual to the Newton polytope of its symmetrized Alexander polynomial.

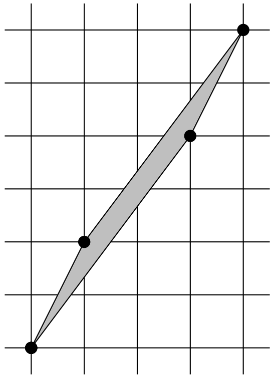


Figure 12. The dual Thurston polytope for $b(-2, 3, 8)$.

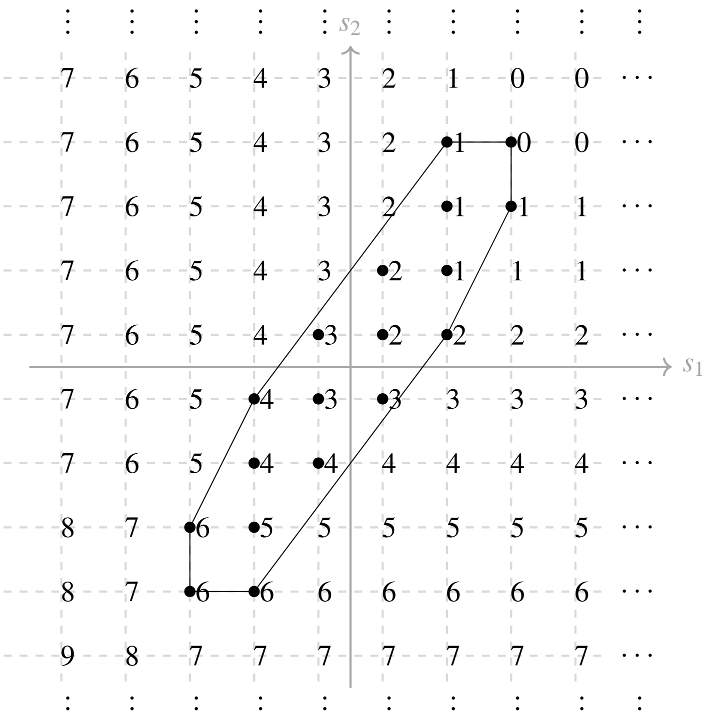


Figure 13. The h-function for $b(-2, 3, 8)$.

5. Two-component L-space links with vanishing Alexander polynomials

In Section 4, we have given examples of L-space links where Thurston polytopes are dual to Newton polytopes of their symmetrized Alexander polynomials. In this section, we mainly discuss 2-component L-space links with vanishing Alexander polynomials, especially split L-space links. Recall that multivariable Alexander polynomials for split links are 0. So Newton polytopes for split L-space links are empty, but link Floer homology polytopes may be nontrivial. To see this in detail, we need some lemmas first.

Lemma 5.1 [Liu 2017, Example 1.13(A)]. *Split disjoint unions of L-space knots are L-space links.*

Lemma 5.2 [Borodzik and Gorsky 2016, Proposition 3.11]. *For a split L-space link $L = L_1 \sqcup L_2$ with two components which are both L-space knots and $(s_1, s_2) \in \mathbb{H}$, the h-function $h(s_1, s_2)$ satisfies*

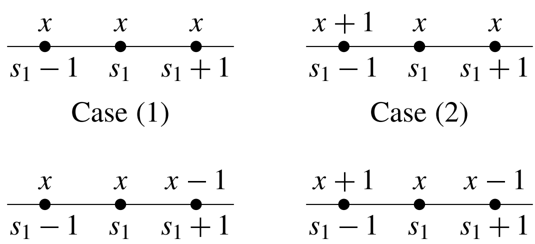
$$h(s_1, s_2) = h_1(s_1) + h_2(s_2),$$

where $h_1(s_1)$ and $h_2(s_2)$ are h-functions of L_1 and L_2, respectively.

Remark 5.3. L-space knots can be regarded as special L-space links with just one component. For any L-space knot $K \subseteq S^3$, we can associate it with a chain complex $A^-(s_1)$ filtered by the Alexander grading, and $H_*(A^-(s_1))$ has a unique generator for any s_1. Let $-2h(s_1)$ be the homological grading of the generator.

Proposition 5.4. *Let $L = L_1 \sqcup L_2$ be a split union of two L-space knots L_1 and L_2. Then $\widehat{\mathrm{HFL}}(L, s_1, s_2) \cong \widehat{\mathrm{HFL}}(L_1, s_1) \otimes \widehat{\mathrm{HFL}}(L_2, s_2) \otimes (\mathbb{F} \oplus \mathbb{F}_{(-1)})$ for any $(s_1, s_2) \in \mathbb{H}$.*

Proof. The proof is quite straightforward using our computation of $\widehat{\mathrm{HFL}}(s_1, s_2)$ in Section 3. For any $(s_1, s_2) \in \mathbb{H}$, the h-function corresponding to $\widehat{\mathrm{HFK}}(L_1, s_1)$ has the following possibilities:



Case (1) Case (2)
Case (3) Case (4)

Here $h_1(s_1) = x$ and x is any positive integer. Observe that

$$H_*(A^-(s_1)/A^-(s_1 - 1)) \cong \mathrm{HFK}^-(L_1, s_1),$$

$$\cdots \to \mathrm{HFK}^-_{i+2}(s_1 + 1) \xrightarrow{U} \mathrm{HFK}^-_i(s_1) \to \widehat{\mathrm{HFK}}_i(s_1) \to \mathrm{HFK}^-_{i+1}(s_1 + 1) \xrightarrow{U} \mathrm{HFK}^-_{i-1}(s_1) \cdots .$$

The long exact sequence is induced by the short exact sequence

$$0 \to \mathrm{CFK}^-(s_1+1) \xrightarrow{U} \mathrm{CFK}^-(s_1) \to \widehat{\mathrm{CFK}}(s_1) \to 0.$$

By the long exact sequence, we compute $\widehat{\mathrm{HFL}}(L_1, s_1)$ as follows:

Case (1) $\widehat{\mathrm{HFK}}(L_1, s_1) \cong 0$.

Case (2) $\widehat{\mathrm{HFK}}(L_1, s_1) \cong \mathbb{F}[-2x]$.

Case (3) $\widehat{\mathrm{HFK}}(L_1, s_1) \cong \mathbb{F}[-2x+1]$.

Case (4) $\widehat{\mathrm{HFK}}(L_1, s_1) \cong 0$.

Similarly, for the link component L_2, we assume that $h_2(s_2) = y$. There are also four possibilities for the h-function corresponding to $\widehat{\mathrm{HFK}}(L_2, s_2)$. By Lemma 5.2, $h(s_1, s_2) = h_1(s_1) + h_2(s_2)$. We find that there are only four possibilities for the h-function such that $\widehat{\mathrm{HFL}}(L, s_1, s_2) \neq 0$:

$h+1$	h	h
$h+1$	h	h
$h+2$	$h+1$	$h+1$

Case (a)

$h+1$	h	h
$h+2$	$h+1$	$h+1$
$h+2$	$h+1$	$h+1$

Case (b)

$h+1$	$h+1$	h
$h+1$	$h+1$	h
$h+2$	$h+2$	$h+1$

Case (c)

$h+1$	$h+1$	h
$h+2$	$h+2$	$h+1$
$h+2$	$h+2$	$h+1$

Case (d)

In Case (a), h-functions for L_1 and L_2 are both like Case (2): $(x+1) \quad x \quad x$ and $(y+1) \quad y \quad y$. Then $\widehat{\mathrm{HFL}}(s_1, s_2) \cong \mathbb{F}[-2(x+y)] \oplus \mathbb{F}[-2(x+y)-1]$, $\widehat{\mathrm{HFK}}(L_1, s_1) \cong \mathbb{F}[-2x]$ and $\widehat{\mathrm{HFK}}(L_2, s_2) \cong \mathbb{F}[-2y]$. So

$$\widehat{\mathrm{HFL}}(s_1, s_2) \cong \widehat{\mathrm{HFK}}(L_1, s_1) \otimes \widehat{\mathrm{HFK}}(L_2, s_2) \otimes (\mathbb{F} \oplus \mathbb{F}_{(-1)}). \tag{5-1}$$

In Case (b), the h-function for L_1 is like Case (2): $(x+1) \quad x \quad x$, and the h-function for L_2 is like Case (3): $y \quad y \quad y-1$. In Case (c), the h-function for L_1 is like Case (3), and for L_2, the h-function is like Case (2). In Case (d), h-functions for L_1 and L_2 are like Case (3). Thus we can compute (5-1) in these cases as well.

If the h-function corresponding to $\widehat{\mathrm{HFL}}(s_1, s_2)$ is not in these four cases, then $\widehat{\mathrm{HFL}}(s_1, s_2) = 0$, and at least one of $\widehat{\mathrm{HFK}}(L_1, s_1)$ and $\widehat{\mathrm{HFK}}(L_2, s_2)$ is zero. Thus the conclusion also holds. □

Proof of Theorem 1.8. Let $L = L_1 \cup L_2$ be an L-space link with vanishing Alexander polynomial. The linking number of L_1 and L_2 is 0 by (2-1). By Theorem 1.2, the Heegaard Floer link homology $\widehat{\mathrm{HFL}}(s_1, s_2)$ is determined by $\Delta_L(t_1, t_2)$, $\Delta_{L_1}(t)$ and $\Delta_{L_2}(t)$. So

$$\widehat{\mathrm{HFL}}(L, s_1, s_2) \cong \widehat{\mathrm{HFL}}(L_1 \sqcup L_2, s_1, s_2) \cong \widehat{\mathrm{HFK}}(L_1, s_1) \otimes \widehat{\mathrm{HFK}}(L_2, s_2) \otimes (\mathbb{F} \oplus \mathbb{F}_{(-1)})$$

for any $(s_1, s_2) \in \mathbb{H}$. □

Example 5.5 (the link Floer homology polytope for the split disjoint union of two right-handed trefoils). Let $L = L_1 \sqcup L_2$ be the split disjoint union of two right-handed trefoils. Recall that the right-handed trefoil is an L-space knot with Alexander polynomial $\Delta_{L_1}(t) = t - 1 + t^{-1}$, and

$$\sum_{s_1 \in \mathbb{Z}} \chi(\mathrm{HFK}^-(L_1, s_1)) t^{s_1} = \frac{\Delta_{L_1}}{1 - t^{-1}} = t + t^{-1} + t^{-2} + t^{-3} + t^{-4} + \cdots .$$

Observe the short exact sequence $0 \to A^-(s_1 - 1) \to A^-(s_1) \to \mathrm{CFK}^-(s_1) \to 0$. We have

$$\mathrm{HFK}^-(L_1, s_1) = H_*(A^-(s_1)/A^-(s_1 - 1)),$$
$$\chi(\mathrm{HFK}^-(L_1, s_1)) = h_1(s_1 - 1) - h_1(s_1),$$

which is also the coefficient of t^{s_1} in $\Delta_{L_1}(t)/(1 - t^{-1})$. Since L_1 is an L-space knot, $h_1(s_1) = 0$ for sufficiently large $s_1 \gg 0$. So the h-function $h_1(s_1)$ can be determined as follows:

$$\ldots, 7, 6, 5, 4, 3, 2, 1, 1, 0, 0, 0, 0, 0, \ldots,$$

where $h_1(0) = h_1(-1) = 1$, $h_1(s) = 0$ if $s \geq 1$, and $h_1(s) = -s$ if $s \leq -1$. Similarly, for another right-handed trefoil L_2, the h-function $h_2(s_2)$ is the same as $h_1(s_1)$. By Proposition 5.4, we can find all $(s_1, s_2) \in \mathbb{H}$ where $\widehat{\mathrm{HFL}}(L, s_1, s_2)$ are nonzero. So

$$\begin{aligned}
\widehat{\mathrm{HFL}}(L, 1, 1) &= \mathbb{F}[0] \oplus \mathbb{F}[-1],\\
\widehat{\mathrm{HFL}}(L, 0, 1) = \widehat{\mathrm{HFL}}(L, 1, 0) &= \mathbb{F}[-1] \oplus \mathbb{F}[-2],\\
\widehat{\mathrm{HFL}}(L, -1, 1) = \widehat{\mathrm{HFL}}(L, 0, 0) = \widehat{\mathrm{HFL}}(L, 1, -1) &= \mathbb{F}[-2] \oplus \mathbb{F}[-3],\\
\widehat{\mathrm{HFL}}(L, -1, 0) = \widehat{\mathrm{HFL}}(L, 0, -1) &= \mathbb{F}[-3] \oplus \mathbb{F}[-4],\\
\widehat{\mathrm{HFL}}(L, -1, -1) &= \mathbb{F}[-4] \oplus \mathbb{F}[-5].
\end{aligned}$$

For other lattice points $(s_1, s_2) \in \mathbb{H}$, $\widehat{\mathrm{HFL}}(L, s_1, s_2) = 0$. Thus the link Floer homology polytope is the shaded square in Figure 14.

Remark 5.6. In general, let $L = L_1 \sqcup L_2$ be the split union of any two L-space knots. The genus of a knot K is defined as

$$g(K) = \min\{\mathrm{genus}(F) \mid F \subseteq S^3 \text{ is an oriented, embedded surface with } \partial F = K\}.$$

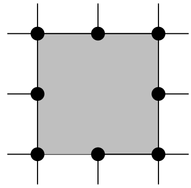


Figure 14. The link Floer homology polytope for L.

Observe that $g(L_i) = \max\{s \geq 0 \mid \widehat{\mathrm{HFK}}_*(L_i, s) \neq 0\}$ for $i = 1, 2$ [Ozsváth and Szabó 2004a, Theorem 1.2], and $\widehat{\mathrm{HFK}}(L_1, g(L_1)) \cong \mathbb{Z}$, $\widehat{\mathrm{HFK}}(L_2, g(L_2)) \cong \mathbb{Z}$, [Ozsváth and Szabó 2005, Theorem 1.2]. The link Floer homology polytope of L_i is the interval $[-g(L_i), g(L_i)]$, where $i = 1, 2$. By Proposition 5.4, the link Floer homology polytope for L is a rectangle with vertices $(g(L_1), g(L_2))$, $(g(L_1), -g(L_2))$, $(-g(L_1), g(L_2))$ and $(-g(L_1), -g(L_2))$ (see Figure 14).

Acknowledgements

I deeply appreciate Eugene Gorsky for introducing this interesting topic to me and his patient teaching on Heegaard Floer homology, and also for his constant guidance and discussions during the project. I am also grateful to Allison Moore, Yi Ni and Jacob Rasmussen for useful discussions on L-space links. I especially wish to thank the referee for helpful suggestions and corrections. The paper is inspired by the work of Yajing Liu, and the project is partially supported by NSF-1559338.

References

[Agol and Dunfield 2015] I. Agol and N. M. Dunfield, "Certifying the Thurston norm via SL(2, C)-twisted homology", 2015. To appear in Thurston memorial conference proceedings, Princeton Univ. Press. arXiv

[Borodzik and Gorsky 2016] M. Borodzik and E. Gorsky, "Immersed concordances of links and Heegaard Floer homology", preprint, 2016. arXiv

[Dawra 2015] N. Dawra, "On the link Floer homology of L-space links", preprint, 2015. arXiv

[Friedl and Kim 2008] S. Friedl and T. Kim, "Twisted Alexander norms give lower bounds on the Thurston norm", *Trans. Amer. Math. Soc.* **360**:9 (2008), 4597–4618. MR Zbl

[Friedl and Vidussi 2015] S. Friedl and S. Vidussi, "The Thurston norm and twisted Alexander polynomials", *J. Reine Angew. Math.* **707** (2015), 87–102. MR Zbl

[Gorsky and Némethi 2015] E. Gorsky and A. Némethi, "Lattice and Heegaard Floer homologies of algebraic links", *Int. Math. Res. Not.* **2015**:23 (2015), 12737–12780. MR Zbl

[Gorsky and Némethi 2016] E. Gorsky and A. Némethi, "Links of plane curve singularities are L-space links", *Algebr. Geom. Topol.* **16**:4 (2016), 1905–1912. MR Zbl

[Liu 2017] Y. Liu, "L-space surgeries on links", *Quantum Topol.* **8**:3 (2017), 505–570. MR Zbl

[Manolescu and Ozsváth 2010] C. Manolescu and P. Ozsváth, "Heegaard Floer homology and integer surgeries on links", preprint, 2010. arXiv

[McMullen 2002] C. T. McMullen, "The Alexander polynomial of a 3-manifold and the Thurston norm on cohomology", *Ann. Sci. École Norm. Sup.* (4) **35**:2 (2002), 153–171. MR Zbl

[Ozsváth and Szabó 2004a] P. Ozsváth and Z. Szabó, "Holomorphic disks and genus bounds", *Geom. Topol.* **8** (2004), 311–334. MR Zbl

[Ozsváth and Szabó 2004b] P. Ozsváth and Z. Szabó, "Holomorphic disks and topological invariants for closed three-manifolds", *Ann. of Math.* (2) **159**:3 (2004), 1027–1158. MR Zbl

[Ozsváth and Szabó 2005] P. Ozsváth and Z. Szabó, "On knot Floer homology and lens space surgeries", *Topology* **44**:6 (2005), 1281–1300. MR Zbl

[Ozsváth and Szabó 2006] P. Ozsváth and Z. Szabó, "Heegaard diagrams and Floer homology", pp. 1083–1099 in *International Congress of Mathematicians, II*, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl arXiv

[Ozsváth and Szabó 2008a] P. Ozsváth and Z. Szabó, "Holomorphic disks, link invariants and the multi-variable Alexander polynomial", *Algebr. Geom. Topol.* **8**:2 (2008), 615–692. MR Zbl

[Ozsváth and Szabó 2008b] P. Ozsváth and Z. Szabó, "Link Floer homology and the Thurston norm", *J. Amer. Math. Soc.* **21**:3 (2008), 671–709. MR Zbl

[Remigio-Juárez and Rieck 2012] J. Remigio-Juárez and Y. Rieck, "The link volumes of some prism manifolds", *Algebr. Geom. Topol.* **12**:3 (2012), 1649–1665. MR Zbl

[Thurston 1986] W. P. Thurston, "A norm for the homology of 3-manifolds", pp. 99–130 in Mem. Amer. Math. Soc. **339**, Amer. Math. Soc., Providence, RI, 1986. MR Zbl

[Vafaee 2015] F. Vafaee, "On the knot Floer homology of twisted torus knots", *Int. Math. Res. Not.* **2015**:15 (2015), 6516–6537. MR Zbl

Received May 9, 2017. Revised February 9, 2018.

BEIBEI LIU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
DAVIS, CA
UNITED STATES

bxliu@math.ucdavis.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.113

ON THE Σ-INVARIANTS OF WREATH PRODUCTS

LUIS AUGUSTO DE MENDONÇA

We present a full description of the Bieri–Neumann–Strebel invariant of restricted permutational wreath products of groups. We also give partial results about the 2-dimensional homotopical invariant of such groups. These results may be turned into a full picture of these invariants when the abelianization of the basis group is infinite. We apply these descriptions to the study of the Reidemeister number of automorphisms of wreath products in some specific cases.

1. Introduction

In this paper we study the so called Σ-invariants of restricted permutational wreath products of groups. The Σ-invariants of a group are some subsets of its character sphere and contain a lot of information on finiteness properties of its subgroups. Their definitions and most general results appeared in a series of papers by Bieri, Neumann, Strebel, Renz ([Bieri et al. 1987; Bieri and Renz 1988; Bieri and Strebel 1980]) and others.

Let Γ be a finitely generated group. The *character sphere* $S(\Gamma)$ is the set of nonzero homomorphisms $\chi : \Gamma \to \mathbb{R}$ (these homomorphisms are called *characters*) modulo the equivalence relation given by $\chi_1 \sim \chi_2$ if there is some $r \in \mathbb{R}_{>0}$ such that $\chi_2 = r\chi_1$. The class of χ will be denoted by $[\chi]$. The character sphere may be seen as the $(n-1)$-sphere in the vector space $\mathrm{Hom}(\Gamma, \mathbb{R}) \simeq \mathbb{R}^n$, where n is the torsion-free rank of the abelianization of Γ.

In this paper we deal with the homotopical invariants in low dimension, that is, those denoted by $\Sigma^1(\Gamma)$ and $\Sigma^2(\Gamma)$, the second defined when Γ is finitely presented. They are defined as certain subsets of $S(\Gamma)$; we leave the details to Section 2. Their most important feature is that they classify the properties of being finitely generated and being finitely presented for subgroups of Γ containing the derived subgroup $[\Gamma, \Gamma]$ (see Theorem 2.1).

Recall that a group Γ is of type F_n if there is a $K(\Gamma, 1)$-complex with compact n-skeleton. A group is of type F_1 (respectively, F_2) if and only if it is finitely

MSC2010: primary 20E22, 20F65; secondary 20E45.
Keywords: sigma theory, wreath product, twisted conjugacy.

generated (respectively, finitely presented). The homological version of the property F_n is the property FP_n: a group Γ is of type FP_n if the trivial $\mathbb{Z}\Gamma$-module $\mathbb{Z}$ admits a projective resolution

$$\mathscr{P} : \cdots \to P_n \to P_{n-1} \to \cdots \to P_1 \to P_0 \to \mathbb{Z} \to 0$$

with P_j finitely generated for all $j \leq n$. Again a group is of type FP_1 if and only if it is finitely generated, but the properties F_n are in general stronger than FP_n. In particular, FP_2 is strictly weaker than finite presentability ([Bestvina and Brady 1997; Bieri 1976]).

There are some higher homotopical invariants, denoted by $\Sigma^n(\Gamma)$, which are defined for groups of type F_n and fit in a decreasing sequence

$$S(\Gamma) \supseteq \Sigma^1(\Gamma) \supseteq \Sigma^2(\Gamma) \supseteq \cdots \supseteq \Sigma^n(\Gamma) \supseteq \cdots$$

whenever defined. They classify the property F_n for subgroups above the derived subgroup. Similarly, the homological invariants $\Sigma^n(\Gamma; \mathbb{Z})$ are defined for groups of type FP_n and classify this same property for subgroups containing the derived subgroup. In general $\Sigma^1(\Gamma) = \Sigma^1(\Gamma; \mathbb{Z})$ if Γ is finitely generated and $\Sigma^n(\Gamma) \subseteq \Sigma^n(\Gamma; \mathbb{Z})$ if Γ has type F_n ([Bieri and Renz 1988]).

All these invariants are in general hard to describe for specific groups, and this has been done only for a few classes of groups. For right-angled Artin groups, for example, the invariant Σ^1 was computed first by Meier and VanWyk [1995] and then generalized for higher dimensions (for both homotopical and homological versions) by the same authors and Meinert [1998]. This is connected with the existence of subgroups of these groups having a wide variety of finiteness properties, as shown by Bestvina and Brady [1997]. Another line of generalization was followed by Meinert [1995], who computed the invariants in dimension 1 for graph products.

Another interesting group for which the invariants are known is Thompson's group F. Both homological and homotopical invariants have been computed in all dimensions by Bieri, Geoghegan and Kochloukova [Bieri et al. 2010]. The Σ^2-invariants of the generalized Thompson groups $F_{n,\infty}$ were then computed by Kochloukova [2012] and recently Zaremsky [2017] extended it to higher dimensions.

We considered the homotopical invariants Σ^1 and Σ^2 of wreath products. Recall that given H and G groups and a G-set X, the wreath product $H \wr_X G$ is defined as the semidirect product $M \rtimes G$, where $M = \bigoplus_{x \in X} H_x$ is the direct sum (that is, the restricted direct product) of copies of H indexed by X and G acts by permuting these copies according to the action on X. We shall always assume that $X \neq \varnothing$ and $H \neq 1$, to avoid trivial cases. The finiteness properties for these groups were studied by de Cornulier [2006] (finite generation and finite presentability) and more recently by Bartholdi, de Cornulier and Kochloukova [Bartholdi et al. 2015] (properties FP_n and F_n).

Remark 1.1. By H_x we always mean the copy of H associated to the element $x \in X$. On the other hand, G_x denotes the stabilizer of $x \in X$ in the action of G. To avoid confusion, we will always denote by G the group that acts.

Our first result is the full description of Σ^1.

Theorem A. *Let $\Gamma = H \wr_X G$ be a finitely generated wreath product and let $\chi : \Gamma \to \mathbb{R}$ be a nontrivial character. We set $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$.*

(1) *If $\chi|_M = 0$, then $[\chi] \in \Sigma^1(\Gamma)$ if and only if $[\chi|_G] \in \Sigma^1(G)$ and $\chi|_{G_x} \neq 0$ for all $x \in X$.*

(2) *If $\chi|_M \neq 0$, then $[\chi] \in \Sigma^1(\Gamma)$ if and only if at least one of the following conditions holds*:

  (a) *There exist $x, y \in X$ with $x \neq y$, $\chi|_{H_x} \neq 0$ and $\chi|_{H_y} \neq 0$.*

  (b) *There exists $x \in X$ with $\chi|_{H_x} \neq 0$ and $[\chi|_{H_x}] \in \Sigma^1(H)$.*

  (c) $\chi|_G \neq 0$.

Part (1) of the above theorem generalizes Theorem 8.1 in [Bartholdi et al. 2015] in dimension 1, where H has infinite abelianization by hypothesis. For regular wreath products, that is, $\Gamma = H \wr_G G$, the action being by multiplication on the left, the Σ^1-invariant was already computed by Strebel [2012, Proposition C1.18].

For the invariant Σ^2 we consider two cases, the same as in the theorem above. For characters $\chi : H \wr_X G \to \mathbb{R}$ such that $\chi|_M \neq 0$ the criteria developed by Renz [1989] are especially powerful, and have allowed us to prove part (2) of Theorem A and a similar result for Σ^2.

Theorem B. *Let $\Gamma = H \wr_X G$ be a finitely presented wreath product and let $\chi : \Gamma \to \mathbb{R}$ be a nontrivial character. If the set*

$$T = \{x \in X \mid \chi|_{H_x} \neq 0\}$$

has at least 3 *elements, then* $[\chi] \in \Sigma^2(\Gamma)$.

The cases where T is nonempty but has less than 3 elements can be dealt with using the direct product formula (see Theorem 2.2) and the results on the Σ^1-invariant (see Theorem 6.5 and the comment right before it).

For the characters $\chi : \Gamma \to \mathbb{R}$ with $\chi|_M = 0$ we were not able to obtain a complete result, for lack of a general method to study necessary conditions for $[\chi] \in \Sigma^2(\Gamma)$. By the results of Bartholdi, de Cornulier and Kochloukova on homological invariants, the most general theorem we can enunciate is the following, where $G_{(x,y)}$ denotes the stabilizer subgroup associated to an element (x, y) of X^2, which is equipped with the diagonal G-action.

Theorem C. *Let $\Gamma = H \wr_X G$ be a finitely presented wreath product and let $\chi : \Gamma \to \mathbb{R}$ be a nonzero character such that $\chi|_M = 0$. Then $[\chi] \in \Sigma^2(\Gamma)$ if*

(1) $[\chi|_G] \in \Sigma^2(G)$,

(2) $[\chi|_{G_x}] \in \Sigma^1(G_x)$ *for all* $x \in X$, *and*

(3) $\chi|_{G_{(x,y)}} \neq 0$ *for all* $(x, y) \in X^2$.

In general, conditions (1) *and* (3) *are necessary for* $[\chi] \in \Sigma^2(\Gamma)$. *If we assume further that the abelianization of* H *is infinite, then condition* (2) *is necessary as well.*

Restrictions on the abelianization of the basis group H have been recurrent in the study of finiteness properties of wreath products and related constructions. Besides appearing in the work of Bartholdi, de Cornulier and Kochloukova [2015], they also pop up in the paper by Kropholler and Martino [2016], which deals with the wider class of *graph-wreath products* (see Section 5) from a more homotopical point of view.

Finally, we consider some applications to twisted conjugacy. Recall that given an automorphism φ of a group G, the *Reidemeister number* $R(\varphi)$ is defined as the number of orbits of the twisted conjugacy action, which is given by $g \cdot h := gh\varphi(g^{-1})$, for $g, h \in G$.

Exploring the connections between Σ-theory and Reidemeister numbers, as found out by Koban and Wong [2011] and Gonçalves and Kochloukova [2010], we obtain some results about the Reidemeister numbers of automorphisms contained in some subgroups of finite index of $\mathrm{Aut}(H \wr_X G)$, under some relatively strong restrictions. For precise statements, see Corollaries 9.3 and 9.5.

2. Background on the Σ-invariants

Let us start by recalling the definition of the invariant Σ^1. For a finitely generated group Γ and a finite generating set $\mathscr{X} \subseteq \Gamma$, we consider the Cayley graph $\mathrm{Cay}(\Gamma; \mathscr{X})$. Its vertex set is Γ and two vertices γ_1 and γ_2 are connected by an edge if and only if there is some $x \in \mathscr{X}^{\pm 1}$ such that $\gamma_2 = \gamma_1 x$ (therefore Γ acts on the left). This graph is always connected. Given a nonzero character $\chi : \Gamma \to \mathbb{R}$ we can define the submonoid

$$\Gamma_\chi = \{\gamma \in \Gamma \mid \chi(\gamma) \geq 0\}.$$

Notice that $\Gamma_{\chi_1} = \Gamma_{\chi_2}$ if and only if χ_1 and χ_2 represent the same class in the character sphere $S(\Gamma)$. The full subgraph spanned by Γ_χ, which we denote by $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$, may not be connected. We put:

$$\Sigma^1(\Gamma) = \{[\chi] \in S(\Gamma) \mid \mathrm{Cay}(\Gamma; \mathscr{X})_\chi \text{ is connected}\}.$$

It can be shown that this definition does not depend on the (finite) generating set $\mathscr{X}$. This invariant is known as the *Bieri–Neumann–Strebel* invariant (or simply BNS-invariant), in reference to the authors who studied it first [Bieri et al. 1987].

The invariant Σ^2 is defined similarly. If Γ is finitely presented and $\langle \mathscr{X} \mid \mathscr{R} \rangle$ is a finite presentation, we consider the Cayley complex $\mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)$. This complex is obtained from the Cayley graph by gluing 2-dimensional cells with boundary determined by the loops defined by the relations $r \in R$, for each base point in Γ. The resulting complex is always 1-connected. Again we define $\mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)_\chi$ to be the full subcomplex spanned by Γ_χ. The 1-connectedness of this complex depends on the choice of the presentation. We define $\Sigma^2(\Gamma)$ as the subset of $S(\Gamma)$ containing exactly all the classes $[\chi]$ of characters such that $\mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)_\chi$ is 1-connected for some finite presentation $\langle \mathscr{X} \mid \mathscr{R} \rangle$ of Γ. More details on these definitions may be found in [Meinert 1997].

The main feature of these invariants is that they classify the related finiteness properties for subgroups containing the derived subgroup. For the invariants Σ^1 and Σ^2, this can be stated as follows.

Theorem 2.1 [Bieri et al. 1987; Renz 1988]. *Suppose that Γ is finitely generated and let $N \subseteq \Gamma$ be a subgroup such that $[\Gamma, \Gamma] \subseteq N$. Then N is finitely generated if and only if*

$$\Sigma^1(\Gamma) \supseteq \{[\chi] \in S(\Gamma) \mid \chi|_N = 0\}.$$

Similarly, if Γ is further finitely presented then N is finitely presented if and only if

$$\Sigma^2(\Gamma) \supseteq \{[\chi] \in S(\Gamma) \mid \chi|_N = 0\}.$$

The homological invariants can be defined by means of the monoid ring $\mathbb{Z}\Gamma_\chi$. This is of course the subring of $\mathbb{Z}\Gamma$ containing exactly all elements $\sum a_\gamma \gamma \in \mathbb{Z}\Gamma$ such that $a_\gamma \neq 0$ only if $\gamma \in \Gamma_\chi$. We put

$$\Sigma^m(\Gamma; \mathbb{Z}) = \{[\chi] \in S(\Gamma) \mid \mathbb{Z} \text{ is of type } FP_m \text{ over } \mathbb{Z}\Gamma_\chi\}.$$

As observed by Bieri and Renz [1988] if $\Sigma^m(\Gamma; \mathbb{Z}) \neq \varnothing$ then Γ is of type FP_m. All we need about these homological invariants is that $\Sigma^2(\Gamma) \subseteq \Sigma^2(\Gamma; \mathbb{Z})$ whenever Γ is finitely presented. Details may be found in [Bieri and Renz 1988; Renz 1988].

Some of the general results we will need about these invariants concern direct products of groups, subgroups of finite index and retracts.

Theorem 2.2 (direct product formulas [Gehrke 1998]). *Let G_1 and G_2 be finitely generated groups and let $\chi = (\chi_1, \chi_2) : G_1 \times G_2 \to \mathbb{R}$ be a nonzero character. Then $[\chi] \in \Sigma^1(G_1 \times G_2)$ if and only if at least one of the following conditions holds*:

(1) $\chi_i \neq 0$ *for* $i = 1, 2$.

(2) $[\chi_i] \in \Sigma^1(G_i)$ *for some* $i \in \{1, 2\}$.

Similarly, if G_1 and G_2 are finitely presented, then $[\chi] \in \Sigma^2(G_1 \times G_2)$ if and only if at least one of the following conditions holds:

(1) $[\chi_1] \in \Sigma^1(G_1)$ *and* $\chi_2 \neq 0$.

(2) $[\chi_2] \in \Sigma^1(G_2)$ *and* $\chi_1 \neq 0$.

(3) $[\chi_i] \in \Sigma^2(G_i)$ *for some* $i \in \{1, 2\}$.

There was a conjecture suggesting how to compute the Σ-invariants of direct products in higher dimensions, but it turned out to be false. Counterexamples were found by Meier, Meinert and VanWyk [Meier et al. 2001] for the homotopical invariants and by Schütz [2008] in the homological case. For precise statements see [Bieri and Geoghegan 2010], which also brings a proof of the homological conjecture if coefficients are taken in a field (rather than $\mathbb{Z}$).

Theorem 2.3 (finite index subgroups [Meinert 1997]). *Let G be a finitely presented group and let $H \leqslant G$ be a subgroup of finite index. Let $\chi : G \to \mathbb{R}$ be a nonzero character and denote by χ_0 its restriction to H. Then $[\chi] \in \Sigma^2(G)$ if and only if $[\chi_0] \in \Sigma^2(H)$.*

Theorem 2.4 (retracts [Meinert 1997]). *Let G be a finitely presented group and suppose that H is a retract, that is, there are homomorphisms $p : G \to H$ and $j : H \to G$ such that $p \circ j = id_H$. Suppose that $\chi : H \to \mathbb{R}$ is a nonzero character. Then*

$$[\chi \circ p] \in \Sigma^2(G) \Rightarrow [\chi] \in \Sigma^2(H).$$

Theorem 2.5 [Kochloukova 2001, Theorem C]. *Suppose that G is a group of type FP_m (respectively, F_m) and N is a normal subgroup of G that is locally nilpotent-by-finite. Then*

$$\{[\chi] \in S(G) \mid \chi(N) \neq 0\} \subseteq \Sigma^m(G; \mathbb{Z}) \text{ (respectively, } \Sigma^m(G)).$$

As pointed out to me by D. Kochloukova, in [Kochloukova 2001] the result is stated for N locally polycyclic-by-finite, but actually the proof works for nilpotent-by-finite. We will use it with N being abelian. The case $m = 1$, with N abelian, can also be found as Lemma C1.20 in Strebel's notes [2012].

3. The Σ^1-invariant of wreath products

Let $\Gamma = H \wr_X G$ be a finitely generated wreath product. As shown by de Cornulier [2006], both G and H are finitely generated and G acts on X with finitely many orbits. Denote $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$. We start working with the characters $\chi : \Gamma \to \mathbb{R}$ such that $\chi|_M = 0$, for which there are some partial results by Bartholdi, de Cornulier and Kochloukova. We quote their result in its most general form, which deals with the higher homological invariants.

Theorem 3.1 [Bartholdi et al. 2015, Theorem 8.1]. *Let $\Gamma = H \wr_X G$ be a wreath product of type FP_m and let $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$. Let $\chi : \Gamma \to \mathbb{R}$ be a nonzero character such that $\chi|_M = 0$. The following conditions are sufficient for $[\chi] \in \Sigma^m(\Gamma; \mathbb{Z})$:*

(1) $[\chi|_G] \in \Sigma^m(G;\mathbb{Z})$.

(2) $[\chi|_{G_\alpha}] \in \Sigma^{m-i}(G_\alpha;\mathbb{Z})$ *for all stabilizers* G_α *of the diagonal action of* G *on* X^i *and for all* $1 \le i \le m$.

Moreover, if the abelianization of H *is infinite, such conditions are also necessary.*

Notice that item (2) contains a statement about invariants in dimension 0. For any finitely generated group V and $\chi : V \to \mathbb{R}$, the condition $[\chi] \in \Sigma^0(V;\mathbb{Z})$ amounts to saying that χ is a nonzero homomorphism.

Recall that the homological and homotopical invariants coincide in dimension 1, that is, $\Sigma^1(V;\mathbb{Z}) = \Sigma^1(V)$ whenever V is a finitely generated group (see [Strebel 2012, Corollary C1.5], for instance). It is worth mentioning that if we consider the original definitions of the invariants in [Bieri et al. 1987] and [Bieri and Renz 1988], we get that actually the sets $\Sigma^1(V)$ and $\Sigma^1(V;\mathbb{Z})$ are antipodal in $S(V)$, that is, $\Sigma^1(V;\mathbb{Z}) = -\Sigma^1(V)$. This happens because in [Bieri and Renz 1988] the authors chose to work with left group actions, while in [Bieri et al. 1987] the actions are on the right. The sign disappears if the choice is consistent.

We can now extract from Theorem 3.1 a set of sufficient conditions for $[\chi] \in \Sigma^1(\Gamma)$. Namely:

Proposition 3.2. *Let* $\Gamma = H \wr_X G$ *be a finitely generated wreath product and let* $\chi : \Gamma \to \mathbb{R}$ *be a nonzero character such that* $\chi|_M = 0$. *If* $[\chi|_G] \in \Sigma^1(G)$ *and if* $\chi|_{G_x} \neq 0$ *for all stabilizers* G_x *of the action of* G *on* X, *then* $[\chi] \in \Sigma^1(\Gamma)$.

Remark 3.3. These conditions could also be obtained by considering an action of Γ on a sufficiently nice complex. We shall apply this reasoning in the study of the invariant $\Sigma^2(H \wr_X G)$.

This set of conditions is in fact necessary. First, if $\chi : \Gamma \to \mathbb{R}$ and $M \subseteq \ker(\chi)$, then

$$[\chi] \in \Sigma^1(\Gamma) \Rightarrow [\chi|_G] \in \Sigma^1(G),$$

since $\chi|_G$ coincides with the character $\bar{\chi}$ induced on the quotient $\Gamma/M \simeq G$ (see [Strebel 2012, Propoisiton A4.5]).

It suffices then to analyze the restriction of χ to the stabilizer subgroups under the hypothesis that $[\chi] \in \Sigma^1(\Gamma)$.

Proposition 3.4. *If* $[\chi] \in \Sigma^1(\Gamma)$ *and* $\chi|_M = 0$, *then* $\chi|_{G_x} \neq 0$ *for all* $x \in X$.

Proof. Let $X = G \cdot x_1 \sqcup \cdots \sqcup G \cdot x_n$. We only need to show that $\chi|_{G_{x_i}} \neq 0$ for all i. By taking the quotient by $M' = \bigoplus_{x \in X \smallsetminus G \cdot x_i} H_x$, we may assume that $n = 1$, that is, we consider wreath products of the form $\Gamma = H \wr_X G$ with $X = G \cdot x_1$.

Let Y and Z be finite generating sets for H and G, respectively. Since $X = G \cdot x_1$ it is clear that $Y \cup Z$ is a finite generating set for Γ (we see Y as a subset of the copy H_{x_1}). Then $\mathrm{Cay}(\Gamma; Y \cup Z)_\chi$ must be connected, since $[\chi] \in \Sigma^1(\Gamma)$ by hypothesis.

First, we show that M can be generated by the left conjugates of elements of $Y^{\pm 1}$ by elements of G_χ. Indeed if $m \in M$ there is a path in $\mathrm{Cay}(\Gamma; Y \cup Z)_\chi$ connecting 1 to m, since $m \in M \subseteq \ker(\chi) \subseteq \Gamma_\chi$. Such a path has as its label a word with letters in $Y^{\pm 1} \cup Z^{\pm 1}$, so we can write

$$m = w_1 v_1 w_2 v_2 \cdots w_k v_k,$$

where each w_j is a word in $Y^{\pm 1}$ and each v_j is a word in $Z^{\pm 1}$ (possibly trivial). We rewrite:

$$m = w_1 ({}^{v_1} w_2)({}^{v_1 v_2} w_3) \cdots ({}^{v_1 \cdots v_{k-1}} w_k)(v_1 \cdots v_k).$$

Now, $w_1 ({}^{v_1} w_2)({}^{v_1 v_2} w_3) \cdots ({}^{v_1 \cdots v_{k-1}} w_k) \in M$ and $v_1 \cdots v_k \in G$. But $m \in M$ and $\Gamma = M \rtimes G$, so $v_1 \cdots v_k = 1_G$. Moreover, since $\chi|_M = 0$, it is clear that $\chi(v_1 \cdots v_j) \geq 0$ for all $1 \leq j \leq k$, so

$$m = w_1 ({}^{v_1} w_2)({}^{v_1 v_2} w_3) \cdots ({}^{v_1 \cdots v_{k-1}} w_k) \in \langle {}^{G_\chi}(Y^{\pm 1}) \rangle,$$

as we wanted.

But then

$$M = \langle {}^{G_\chi}(Y) \rangle \subseteq \langle {}^{G_\chi}(H_{x_1}) \rangle = \bigoplus_{x \in G_\chi \cdot x_1} H_{x_1},$$

that is, $X = G_\chi \cdot x_1$. Finally, as $\chi|_G \neq 0$ there is some $g_1 \in G$ such that $\chi(g_1) < 0$. On the other hand, there must be some $g_0 \in G_\chi$ such that $g_0 \cdot x_1 = g_1 \cdot x_1$. It follows that $g_1^{-1} g_0 \in G_{x_1}$, with $\chi(g_1^{-1} g_0) = -\chi(g_1) + \chi(g_0) > 0$, hence $\chi|_{G_{x_1}} \neq 0$. □

We obtain part (1) of Theorem A by combining Propositions 3.2 and 3.4.

4. The Σ^1-invariant and Renz's criterion

We shall use the results of Renz [1989] to consider the characters $\chi : H \wr_X G \to \mathbb{R}$ such that $\chi|_M \neq 0$. Let Γ be any finitely generated group and let $\mathscr{X} \subseteq \Gamma$ be a finite generating set. For a nonzero character $\chi : \Gamma \to \mathbb{R}$ and for any word $w = x_1 \cdots x_n$, with $x_i \in \mathscr{X}^{\pm 1}$, we denote

$$v_\chi(w) := \min\{\chi(x_1 \cdots x_j) \mid 1 \leq j \leq n\}.$$

Theorem 4.1 [Renz 1989, Theorem 1]. *With the notation above,* $[\chi] \in \Sigma^1(\Gamma)$ *if and only if there exists* $t \in \mathscr{X}^{\pm 1}$ *with* $\chi(t) > 0$ *and such that for all* $x \in \mathscr{X}^{\pm 1} \setminus \{t, t^{-1}\}$ *the conjugate* $t^{-1} x t$ *can be represented by a word* w_x *in* $\mathscr{X}^{\pm 1}$ *such that*

$$v_\chi(t^{-1} x t) < v_\chi(w_x).$$

Proposition 4.2. *Let* $\Gamma = H \wr_X G$ *be a finitely generated wreath product and let* $[\chi] \in S(\Gamma)$. *Suppose that there is some* $x_1 \in X$ *such that* $G \cdot x_1 \neq \{x_1\}$ *and* $\chi|_{H_{x_1}} \neq 0$. *Then* $[\chi] \in \Sigma^1(\Gamma)$.

Proof. Let Y and Z be finite generating sets for H and G, respectively, and choose $x_1, \ldots, x_n \in X$ such that $X = \bigsqcup_{j=1}^n G \cdot x_j$ (the element x_1 is already chosen to satisfy the hypotheses). For each $1 \le j \le n$ let Y_j be a copy of Y inside H_{x_j}. It is clear that Γ is generated by $Y_1 \cup \cdots \cup Y_n \cup Z$.

Now, since $G \cdot x_1 \neq \{x_1\}$ we can choose $g_1 \in G$ such that $g_1 \cdot x_1 \neq x_1$. Furthermore, since $\chi|_{H_{x_1}} \neq 0$, we can choose a generator $h \in Y_1$ such that $\chi(h) \neq 0$. We may assume without loss of generality that $\chi(h) > 0$. Define $t := {}^{g_1}h \in H_{g_1 \cdot x_1}$. We take $\mathscr{X} = Y_1 \cup \cdots \cup Y_n \cup Z \cup \{t\}$ as a generating set for Γ and we show that the conditions of Theorem 4.1 are satisfied.

If $y \in (Y_1 \cup \cdots \cup Y_n)^{\pm 1}$ then t and y commute in Γ, hence $w_y := y$ is word that represents $t^{-1}yt$. Also, $v_\chi(w_y) = \chi(y)$ and

$$v_\chi(t^{-1}yt) \le \chi(t^{-1}y) = \chi(y) - \chi(t) < \chi(y),$$

so $v_\chi(t^{-1}yt) < v_\chi(w_y)$.

If $z \in Z^{\pm 1}$, there are two cases: $z \in G_{g_1 \cdot x_1}$ or $z \notin G_{g_1 \cdot x_1}$. In the first case z and t commute in Γ, so we may proceed as in the previous paragraph: we take the word $w_z := z$, which represents $t^{-1}zt$ and satisfies $v_\chi(t^{-1}zt) < v_\chi(w_z)$. If $z \notin G_{g_1 \cdot x_1}$ notice that ${}^{z}t$ and t^{-1} lie in different copies of H in Γ, therefore they commute, so

$$t^{-1}zt = t^{-1}({}^{z}t)z = ({}^{z}t)t^{-1}z = ztz^{-1}t^{-1}z.$$

In this case, define $w_z := ztz^{-1}t^{-1}z$. Observe that $v_\chi(w_z) = \min\{0, \chi(z)\}$. If this minimum is 0 then $\chi(z) \ge 0$, and so $v_\chi(t^{-1}zt) = -\chi(t) < 0$. Otherwise $v_\chi(w_z) = \chi(z) < 0$ and so $v_\chi(t^{-1}zt) \le \chi(t^{-1}z) = \chi(z) - \chi(t) < \chi(z)$. In both cases, $v_\chi(t^{-1}zt) < v_\chi(w_z)$.

Thus $[\chi] \in \Sigma^1(\Gamma)$ by Theorem 4.1. □

In order to complete the proof of Theorem A, we only need to consider the cases where the restriction of χ to the copies of H is nonzero only for copies associated to orbits that are composed by only one element, and this is done by use of the direct product formula, as follows.

Theorem 4.3. *Let $\Gamma = H \wr_X G$ be a finitely generated wreath product and set $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$. Let $\chi : \Gamma \to \mathbb{R}$ be a nonzero character such that $\chi|_M \neq 0$. Then $[\chi] \in \Sigma^1(\Gamma)$ if and only if at least one the following conditions holds*:

(1) *The set $T = \{x \in X \mid \chi|_{H_x} \neq 0\}$ has at least two elements.*

(2) *$T = \{x_1\}$ and $\chi|_G \neq 0$.*

(3) *$T = \{x_1\}$ and $[\chi|_{H_{x_1}}] \in \Sigma^1(H)$.*

Proof. By Proposition 4.2 it is enough to consider the case where $G \cdot x = \{x\}$ for all $x \in T$. Notice that in this case T must be finite, since each of its elements is an entire orbit of G on X, and there are finitely many of those. Let $P = \prod_{x \in T} H_x$ and $X' = X \setminus T$. Then,

$$\Gamma = H \wr_X G \simeq P \times (H \wr_{X'} G).$$

If T has at least two elements, then $[\chi|_P] \in \Sigma^1(P)$ and hence $[\chi] \in \Sigma^1(\Gamma)$, by two applications of the direct product formula for Σ^1. If $T = \{x_1\}$, the formula gives us exactly that $[\chi] \in \Sigma^1(\Gamma)$ if and only if one of conditions (2) or (3) holds, since $\chi|_G \neq 0$ if and only if $\chi|_{H \wr_{X'} G} \neq 0$. □

5. Graph-wreath products

We now digress a bit and obtain a generalization of the results of Section 3 to a wider class of groups. Besides being interesting in its own right, this will be useful in the analysis of the Σ^2-invariants of wreath products.

Given two groups G and H, and K a G-graph, the *graph-wreath product* $H \wr_K G$ is defined by Kropholler and Martino [2016] as the semidirect product $H^{\langle K \rangle} \rtimes G$, where $H^{\langle K \rangle}$ is the graph product of H with respect to the graph K (that is, H is the group associated to every vertex of K). The action of G is given by permutation of the copies of H according to the G-action on the vertex set of K. When K is the complete graph, $H \wr_K G$ is simply $H \wr_X G$, where X is the vertex set of K.

Kropholler and Martino showed that $H \wr_K G$ is finitely generated if and only if G and H are finitely generated and G acts with finitely many orbits of vertices on K, that is, $H \wr_K G$ is finitely generated under the same conditions as $H \wr_X G$ is, where X is the vertex set of K.

In what follows we fix $\Gamma = H \wr_K G$ and $M = H^{\langle K \rangle} \subseteq \Gamma$. We assume that Γ is finitely generated and we decompose X in orbits as $X = G \cdot x_1 \sqcup \cdots \sqcup G \cdot x_n$. Moreover, we choose finite generating sets Z for G and Y_i for H_{x_i} for all $i = 1, \ldots, n$ and we denote $\mathscr{X} = \left(\bigcup_{i=1}^n Y_i\right) \cup Z$, which is seen as a generating set for Γ.

Theorem 5.1. *Let $\chi : H \wr_K G \to \mathbb{R}$ be a nonzero character such that $\chi|_M = 0$. Then $[\chi] \in \Sigma^1(H \wr_K G)$ if and only if $[\chi|_G] \in \Sigma^1(G)$ and $\chi|_{G_x} \neq 0$ for all $x \in X$.*

Proof. Let N_K be the kernel of the obvious homomorphism $M \twoheadrightarrow \bigoplus_{x \in X} H_x$. Note that $N_K \subseteq \ker(\chi)$ and that $\overline{\Gamma} := \Gamma / N_K \simeq H \wr_X G$. It follows that χ induces a character $\bar{\chi} : \overline{\Gamma} \to \mathbb{R}$. For an element $\gamma \in \Gamma$, we denote by $\bar{\gamma}$ its image in $\overline{\Gamma}$.

If $[\chi] \in \Sigma^1(\Gamma)$, then $[\bar{\chi}] \in \Sigma^1(\overline{\Gamma})$ (again by Proposition A4.5 in [Strebel 2012]). Thus $[\chi|_G] \in \Sigma^1(G)$ and $\chi|_{G_x} \neq 0$ for all $x \in X$ by Theorem A.

Conversely, suppose that $[\chi|_G] \in \Sigma^1(G)$ and that $\chi|_{G_x} \neq 0$ for all $x \in X$. Then $[\bar{\chi}] \in \Sigma^1(\overline{\Gamma})$. We will show that this implies that $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$ is connected.

We need to show that for all $\gamma \in \Gamma_\chi$, there is a path in $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$ connecting 1 and γ. Given such a γ, notice that $\bar{\gamma} \in \overline{\Gamma}_{\bar{\chi}}$, so there must be a path from 1 to $\bar{\gamma}$ in $\mathrm{Cay}(\overline{\Gamma}; \overline{\mathscr{X}})_{\bar{\chi}}$. Its obvious lift to $\mathrm{Cay}(\Gamma; \mathscr{X})$ with 1 as initial vertex is a path in $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$ that ends at an element of the form γn, with $n \in N_K$. If we can connect γ to γn inside $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$ we are done. For that it suffices to find a path in $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$ connecting 1 and n, and then act with γ on the left.

Since $N_K \subseteq M$, each $n \in N_K$ can be written as

$$n = ({}^{g_1}h_1)({}^{g_2}h_2)\cdots({}^{g_k}h_k), \tag{5-1}$$

with $h_j \in \bigcup_{i=1}^{n} Y_i^{\pm 1}$ and $g_j \in G$ for all j. Even more, we may assume that each $\chi(g_j) \geq 0$. Indeed, since $\chi|_{G_x} \neq 0$ for all x, we can always pick $t_j \in G$ such that $\chi(t_j) > 0$ and ${}^{t_j}h_j = h_j$. Then we may change g_j for $g_j t_j^{k_j}$ in (5-1), where k_j is some integer such that $k_j\chi(t_j) \geq -\chi(g_j)$.

But if $\chi(g_j) \geq 0$ then $g_j \in G_{\chi|_G}$, and since $[\chi|_G] \in \Sigma^1(G)$, we can choose words w_j in $Z^{\pm 1}$ representing g_j and such that $v_\chi(w_j) \geq 0$. Finally, the word

$$w = (w_1 h_1 w_1^{-1})(w_2 h_2 w_2^{-1})\cdots(w_k h_k w_k^{-1})$$

is the label for a path connecting 1 and n in $\mathrm{Cay}(\Gamma; \mathscr{X})_\chi$, by the choice of each w_j together with the fact that $\chi(h_j) = 0$ for all j by hypothesis. □

The above result will be needed only in a special case, namely when K is a graph without edges, so that $\Gamma \simeq (*_{x\in X} H_x) \rtimes G$.

6. The Σ^2-invariant

Renz's paper [1989] also brings a criterion for the invariant Σ^2. In order to state it, we need to introduce the concept of a diagram over a group presentation, for which we follow [Bridson 2002]. Fix an orientation on $\mathbb{R}^2$. Define a *diagram* to be a subset $M \subseteq \mathbb{R}^2$ endowed with the structure of a finite combinatorial 2-complex. Thus to each 1-cell of M correspond two opposite directed edges. If $\langle \mathscr{X} \mid \mathscr{R} \rangle$ is a presentation for a group Γ, a *labeled diagram over* $\langle \mathscr{X} \mid \mathscr{R} \rangle$ is a diagram M endowed with an edge labeling satisfying:

(1) The edges of M are labeled by elements of $\mathscr{X}^{\pm 1}$.

(2) If an edge e has label x, then its opposite edge has label x^{-1}.

(3) The boundary of each face of M, read as a word in $\mathscr{X}^{\pm 1}$, beginning at any vertex and proceeding with either orientation, is either a cyclic permutation of some $r \in \mathscr{R}^{\pm 1}$, or a word of the form $tt^{-1}t^{-1}t$ for some $t \in \mathscr{X}^{\pm 1}$.

A labeled diagram M is said to be *simple* if it is connected and simply connected.

Remark 6.1. This is a weakening of the definition of the usual *van Kampen diagrams*. In fact, a simple diagram M, with a vertex chosen as a base point, differs from a van Kampen diagram only by the fact that it can have what we call *trivial faces*, that is, those labeled by $tt^{-1}t^{-1}t$ for some $t \in \mathscr{X}^{\pm 1}$. This weakening has the effect of simplifying the drawing of some diagrams that we will consider in the sequence (see [Renz 1989, Section 3.3]).

Suppose that we are given a simple diagram M with a base point u (a vertex in the boundary of M) and an element $\gamma \in \Gamma$. Then to each vertex u' of M corresponds a unique element of Γ, given by $\gamma\eta$, where η is the image in Γ of the label of any path connecting u and u' inside M. In particular, the given group element γ corresponds to the base point u. For any character $\chi : \Gamma \to \mathbb{R}$ we define the χ-*valuation* of M with respect to u and γ, denoted by $v_\chi(M)$, to be the minimum value of $\chi(g)$ when g runs over the elements of Γ corresponding to the vertices of M.

Now, suppose that Γ is finitely presented (with $\langle \mathscr{X} \mid \mathscr{R} \rangle$ a finite presentation) and assume $[\chi] \in \Sigma^1(\Gamma)$. Then we can distinguish an element $t \in \mathscr{X}^{\pm 1}$ with $\chi(t) > 0$ with which we can apply Renz's criterion for Σ^1: for each $x \in \mathscr{X}^{\pm 1} \smallsetminus \{t, t^{-1}\}$ we can associate a word w_x in $\mathscr{X}^{\pm 1}$ that represents $t^{-1}xt$ and for which $v_\chi(t^{-1}xt) < v_\chi(w_x)$. Additionally, we put $w_t := t$ and $w_{t^{-1}} := t^{-1}$. If $r = x_1 \cdots x_n \in \mathscr{R}^{\pm 1}$, we define

$$\hat{r} := w_{x_1} \cdots w_{x_n}.$$

We are now ready to state the criterion for Σ^2.

Theorem 6.2 [Renz 1989, Theorem 3]. *Let Γ, $\mathscr{X}$ and t be as above. Suppose that the set $\mathscr{R}$ of defining relations contains some cyclic permutation of the words $t^{-1}xtw_x^{-1}$, for all $x \in \mathscr{X}^{\pm 1}$. Then $[\chi] \in \Sigma^2(\Gamma)$ if and only if for each $r \in \mathscr{R}^{\pm 1}$ there exist a simple diagram $M_{\hat{r}}$ and vertex u in its boundary, such that both the following conditions hold*:

(1) *The boundary path of $M_{\hat{r}}$, read from u, has as label the word $\hat{r}$.*

(2) $v_\chi(r) < v_\chi(M_{\hat{r}})$, *where the valuation of $M_{\hat{r}}$ is taken with respect to the base point u and the element $t \in \Gamma$.*

Now, recall that a wreath product $H \wr_X G$ is finitely presented if and only if G and H are finitely presented, G acts diagonally on X^2 with finitely many orbits and the stabilizers of the G-action on X are finitely generated. This is the result by de Cornulier [2006].

We will apply Theorem 6.2 to show that if $\Gamma = H \wr_X G$ is finitely presented and if $\chi : \Gamma \to \mathbb{R}$ is a character such that $\chi|_{H_{x_1}} \neq 0$ for some $x_1 \in X$ with $|G \cdot x_1| = \infty$, then $[\chi] \in \Sigma^2(\Gamma)$.

We start by assuming that G acts transitively on X, with $X = G \cdot x_1$. Let $\langle Y \mid R \rangle$ and $\langle Z \mid S \rangle$ be finite presentations for H and G, respectively. We may assume that Z contains a generating set E for the stabilizer subgroup G_{x_1} and a set J of representatives for the nontrivial double cosets of (G_{x_1}, G_{x_1}) in G, since both E and J can be taken to be finite by the proof of the main theorem in [de Cornulier 2006].

We think of $\Gamma = H \wr_X G$ with the presentation considered by de Cornulier. So Γ is generated by the set $Y \cup Z$, subject to the defining relations

(1) r, for all $r \in R$ (defining relations for H);

(2) s, for all $s \in S$ (defining relations for G);

(3) $[^g y_1, y_2]$, for $g \in J$, $y_1, y_2 \in Y$;

(4) $[e, y]$, for $e \in E$ and $y \in Y$.

Let us adapt this presentation a bit. We are under the hypothesis that $\chi|_{H_{x_1}} \neq 0$ and $|G \cdot x_1| = \infty$. We may assume without loss of generality that $\chi(h) > 0$ for some $h \in Y$. Choose $g_i \in Z$, for $1 \leq i \leq 5$, such that $\{x_1\} \cup \{g_i \cdot x_1 \mid 1 \leq i \leq 5\}$ is a set with exactly six elements (of course we may assume that Z contains elements g_i with this property). Define

$$t_i := {}^{g_i}h,$$

for $i = 1, \ldots, 5$. Then Γ is generated by $Y \cup Z \cup \{t_i \mid 1 \leq i \leq 5\}$, subject to the defining relations

(1) r, for all $r \in R$ (defining relations for H);

(2) s, for all $s \in S$ (defining relations for G);

(3) $[^g y_1, {}^{g'} y_2]$, for all $y_1, y_2 \in Y \cup \{t_i \mid 1 \leq i \leq 5\}$ and $g, g' \in Z \cup \{1\}$ whenever the commutator $[^g y_1, {}^{g'} y_2]$ is indeed a relation in Γ;

(4) $[e, y]$, for $e \in E$ and $y \in Y$ and $[z, t_1]$, for $z \in Z \cap G_{g_1 \cdot x_1}$;

(5) $g_i h g_i^{-1} t_i^{-1}$, for $1 \leq i \leq 5$.

Remark 6.3. We could write the conditions of item (3) in a more precise way, but it would require writing many cases. If $y_1 \in Y$ and $y_2 = t_1$, for example, then $[^g y_1, {}^{g'} y_2]$ is a defining relation if $g \cdot x_1 \neq (g' g_1) \cdot x_1$.

Note that we have added a few relations of types (3) and (4), but clearly they are consequences of the others. Furthermore, the set of relations is clearly still finite.

Set $t = t_1$. We will continue using the notation of Proposition 4.2. Thus for $y \in Y^{\pm 1}$ we have chosen $w_y = y$. If $z \in Z^{\pm 1}$, then $w_z = z$ if $z \in G_{g_1 \cdot x_1}$ and $w_z = ztz^{-1}t^{-1}z$ otherwise. Moreover, since t_i and t commute in Γ for all $1 \leq i \leq 5$, we can define $w_{t_i} := t_i$ and $w_{t_i^{-1}} := t_i^{-1}$.

Let us check that the chosen presentation satisfies the conditions of Theorem 6.2. First, the set of defining relations contains the relations $t^{-1}xtw_x^{-1}$. Indeed if $y \in Y^{\pm 1} \cup \{t_i \mid 1 < i \leq 5\}^{\pm 1}$ then $t^{-1}ytw_y^{-1}$ is a relation of type (3), since $w_y = y$. If $z \in Z^{\pm 1} \cap G_{g_1 \cdot x_1}$, then $w_z = z$ and $t^{-1}ztz^{-1}$ is a relation of type (4). Finally, if $z \in Z^{\pm 1}$ but $z \notin G_{g_1 \cdot x_1}$, then $w_z = ztz^{-1}t^{-1}z$ and

$$t^{-1}ztw_z^{-1} = t^{-1}ztz^{-1}tzt^{-1}z^{-1} = t^{-1}({}^z t)t({}^z t)^{-1},$$

which is a cyclic permutation of $[^z t, t]$, a relation of type (3).

According to Theorem 6.2, now we need to apply the transformation $r \mapsto \hat{r}$ to each defining relation and then find a simple diagram $M_{\hat{r}}$ satisfying the stated conditions. The following subsections are devoted to the verification of the existence

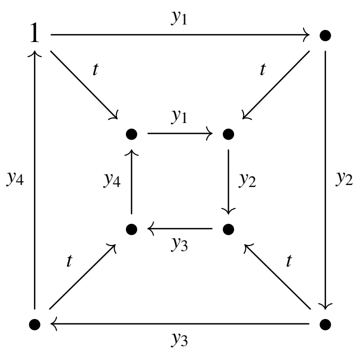


Figure 1. Diagram for a relation of type (1), $r = \hat{r} = y_1 y_2 y_3 y_4$

of these diagrams. Observe that we do not need to consider the inverses of defining relations, since any simple diagram for $\hat{r}$ is a simple diagram for the inverse of $\hat{r}$ if we read its boundary backwards.

***Relations of type* (1).** Note that the relations of type (1) involve only generators in $Y^{\pm 1}$. But $w_y = y$ for all $y \in Y^{\pm 1}$, so $\hat{r} = r$ whenever r is a relation of type (1). Thus the one-faced diagram M that represents the relation r, with base point corresponding to t, is already a choice for $M_{\hat{r}}$, since its χ-value is increased by $\chi(t) > 0$.

In Figure 1 we represent the diagram $M_{\hat{r}}$, for $r = \hat{r} = y_1 y_2 y_3 y_4$, as the internal square of the figure. The external boundary represents the path beginning at the base point $1 \in \Gamma$ and with label the original relation r. The edges labeled by t indicate that $\hat{r}$ is obtained from r by conjugation by t.

***Relations of type* (2).** Since $\chi(h) \neq 0$, the order of h in Γ is infinite. Consider the subgroup

$$\Gamma_0 := \langle h, G \rangle \leqslant \Gamma.$$

Notice $\Gamma_0 \simeq \mathbb{Z} \wr_X G$. Let χ_0 be the restriction of χ to Γ_0. The group Γ_0 is an extension of an abelian group $A = \bigoplus_{x \in X} \mathbb{Z}$ by G, so it follows from Theorem 2.5 that $[\chi_0] \in \Sigma^2(\Gamma)$, as $\chi_0|_A \neq 0$. Now choose a presentation for Γ_0 that is compatible with the chosen presentation for Γ: write the same presentation with $Y = \{h\}$ and discard the relations of type (1). Naturally, this presentation satisfies the hypothesis of Theorem 6.2.

Let $r = z_1 \cdots z_n$ be a relation of type (2). We can see r as a relation in Γ_0. By Theorem 6.2 there is a simple diagram $M_{\hat{r}}$, with respect to the presentation of Γ_0, whose base point corresponds to t and such that $\partial M_{\hat{r}} = \hat{r}$ and $v_\chi(r) < v_\chi(M_{\hat{r}})$. But all the relations in the chosen presentation of Γ_0 are also relations in the original presentation of Γ, after identifying the generating sets. Then $M_{\hat{r}}$, if seen as a diagram over the presentation of Γ, is the diagram we wanted.

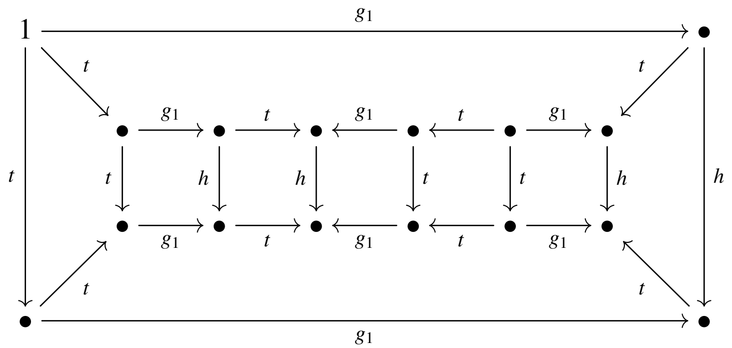


Figure 2. Diagram for relations of type (5)

***Relations of the types* (4) *or* (5).** All cases are similar: we can obtain simple diagrams whose only vertices are those of the boundary, that is, those that are defined by the word $\hat{r}$. In this case, the diagram automatically satisfies the hypothesis about its χ-value, exactly as in the case of the relations of type (1). See the diagram for the relation $g_1 h g_1^{-1} t^{-1}$ in Figure 2. As before, the external boundary represents $r = g_1 h g_1^{-1} t^{-1}$, and the diagram $M_{\hat{r}}$ itself is the internal diagram composed by the five squares. Again, the edges labeled by t and with origin at some point in the external path indicate conjugation by t and represent the growth of v_χ from r to $\hat{r}$.

Figure 2 is an illustration of the case when $g = g_1 \notin G_{g_1 \cdot x}$, when the word w_g is more complicated. If the letter representing an element of G is an element of $G_{g_1 \cdot x}$, the argument is simpler: all the letters involved in the relation commute with t, so $r = \hat{r}$ and the argument follows as in the case of the relations of type (1).

***Relations of type* (3).** Let $y \in Y \cup \{t_i \mid 1 \le i \le 5\}$ and $g \in Z$. Let $\eta_{g,y}$ be the word obtained from gyg^{-1} by applying the transformation that takes each letter α to w_α:

$$\eta_{g,y} = ({}^g t) t^{-1} ({}^g y) t ({}^g t)^{-1}, \tag{6-1}$$

if $g \notin G_{g_1 \cdot x_1}$, or

$$\eta_{g,y} = {}^g y, \tag{6-2}$$

if $g \in G_{g_1 \cdot x_1}$. If $g = 1$, put $\eta_{1,y} := y$. In all cases we see that $\eta_{g,y}$ is a product of subwords representing elements of at most 3 copies of H in Γ. Indeed, ${}^g t$ and $({}^g t)^{-1}$ are elements of $H_{gg_1 \cdot x_1}$, while t and t^{-1} are elements of $H_{g_1 \cdot x_1}$ and, finally, ${}^g y$ is an element of $H_{g \cdot x_1}$ or $H_{gg_i \cdot x_1}$ for some $1 \le i \le 5$, depending on y.

Consider $r = [{}^g y_1, {}^{g'} y_2]$, a relation of type (3). The word $\hat{r} = [\eta_{g,y_1}, \eta_{g',y_2}]$ is a relation in Γ, so we can always find a simple diagram M_1 with some base point corresponding to t and such that $\partial M_1 = \hat{r}$. If $v_\chi(M_1) > v_\chi(r)$ we are done. Otherwise there is a vertex p of M_1 such that $\chi(p) \le v_\chi(r)$. Notice that this vertex cannot lie on the boundary of M_1, since $v_\chi(\hat{r}) > v_\chi(t^{-1} r t)$.

Now, the commutator between the words η_{g,y_1} and η_{g',y_2} is a product of elements of the form ${}^z y$, with $z \in Z \cup \{1\}$ and $y \in Y^{\pm 1} \cup \{t_1, \ldots, t_5\}^{\pm 1}$. By the remarks above, these elements lie in at most five different copies of H, one of them being indexed by $g_1 \cdot x_1$ when five copies do pop up. It follows that for some $u \in \{h, t_2, t_3, t_4, t_5\}$, the words $[{}^z y, u]$ are defining relations for all the subwords ${}^z y$ appearing in $\hat{r} = [\eta_{g,y_1}, \eta_{g',y_2}]$ (we consider the subwords ${}^z y$ that appear when η_{g,y_1} and η_{g',y_2} are written exactly as in (6-1) or (6-2)). Observe that $\chi(u) = \chi(h) > 0$ in all cases. So we can build a diagram M_2 by surrounding M_1 with faces representing the commutators $[{}^z y, u]$, for all these subwords ${}^z y$.

Clearly the boundary of M_2 is also labeled by $\hat{r}$. If we set as base point the vertex on the new boundary corresponding to the base point of M_1 (that is, the one joined to it by an edge with label u), then the χ-value of the interior points (including p) is increased by $\chi(u) > 0$, so that $v_\chi(M_2) > v_\chi(M_1)$. Repeating this process finitely many times, we obtain a simple diagram M_n satisfying the conditions of the theorem.

We record what we have proved in the following proposition.

Proposition 6.4. *Let $\Gamma = H \wr_X G$ be a finitely presented wreath product and let $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$. Suppose that G acts transitively on the infinite set X. If $\chi : \Gamma \to \mathbb{R}$ is a character with $\chi|_M \neq 0$, then $[\chi] \in \Sigma^2(\Gamma)$.*

The arguments above essentially contain what we need when $G \cdot x$ is infinite for some $x \in X$ such that $\chi|_{H_x} \neq 0$ (but the G-action on X is not necessarily transitive), so in the proof of Theorem 6.5 we will only indicate how to deal with this case.

Recall that we denote by T the set of elements $x \in X$ such that $\chi|_{H_x} \neq 0$. Notice that if $T = \{x_1\}$, then Γ is a direct product

$$\Gamma \simeq H_{x_1} \times (H \wr_{X'} G),$$

where $X' = X \smallsetminus \{x_1\}$. Then the direct product formula and the results on the Σ^1-invariants of wreath products already contain all the information we need. The remaining cases are all part of the following theorem, which includes Theorem B.

Theorem 6.5. *Let $\Gamma = H \wr_X G$ be a finitely presented wreath product and let $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$. Suppose that the set*

$$T = \{x \in X \mid \chi|_{H_x} \neq 0\}$$

has at least two elements. Then $[\chi] \in \Sigma^2(\Gamma)$ if and only if at least one of the following conditions holds:

(1) $[\chi|_{H_x}] \in \Sigma^1(H)$ *for some $x \in T$.*

(2) $\chi|_G \neq 0$.

(3) *T has at least three elements.*

Proof. Suppose first that T is a finite set and consider the subgroup $B = \cap_{x\in T} G_x \leqslant G$. It is of finite index in G, so $\Gamma_1 = H \wr_X B$ is of finite index in Γ. Notice that

$$\Gamma_1 \simeq \left(\prod_{x\in T} H_x \right) \times (H \wr_{X'} B).$$

Denote $P = \prod_{x\in T} H_x$ and $Q = H \wr_{X'} B$. The fact that T has at least two elements implies $[\chi|_P] \in \Sigma^1(P)$, by the direct product formula. By applying the formula again, now to the product $\Gamma_1 = P \times Q$, we get that $[\chi|_{\Gamma_1}] \in \Sigma^2(\Gamma_1)$ if and only if $[\chi|_P] \in \Sigma^2(P)$ or $\chi|_Q \neq 0$. The former happens if and only if at least one of conditions (1) or (3) is satisfied (once again, by the direct product formula), while the latter clearly happens if and only if $\chi|_B \neq 0$, which in turn is equivalent to $\chi|_G \neq 0$, since B is a subgroup of finite index. Finally, since the index of Γ_1 in Γ is finite, we are done, by Theorem 2.3.

We are left with the case where T is infinite and we want to show that $[\chi] \in \Sigma^2(\Gamma)$. Since G acts on X with finitely many orbits, there must be some $x_1 \in T$ such that $|G \cdot x_1| = \infty$. We adapt the proof of Proposition 6.4 putting the orbit of x_1 in a distinguished position.

Choose $x_2, \ldots, x_n \in X$ such that $X = \bigsqcup_{j=1}^n G \cdot x_j$. For each j choose a finite generating set E_j for the stabilizer subgroup G_{x_j}. For each pair par (i, j), with $1 \leq i, j \leq n$, choose a finite set $J_{i,j}$ of representatives of the nontrivial double cosets of (G_{x_i}, G_{x_j}) in G. Finally, choose finite presentations $\langle Y \mid R\rangle$ and $\langle Z \mid S\rangle$ for H and G respectively. We may assume that Z contains E_j and $J_{i,j}$ for all $1 \leq i, j \leq n$.

A finite presentation for Γ, adapted from the presentation given by de Cornulier [2006], can be given as follows. For each $1 \leq i \leq n$ we associate a copy $\langle Y_i \mid R_i\rangle$ of the presentation for H and, as before, we define $t_i := {}^{g_i}h$ for some $g_i \in Z$ and $h \in Y_1$ with $\chi(h) > 0$ and $|\{x_1\} \cup \{g_i \cdot x_1 \mid 1 \leq i \leq 5\}| = 6$. We think of Γ as generated by $\left(\bigcup_{i=1}^n Y_i\right) \cup Z \cup \{t_i \mid 1 \leq i \leq 5\}$ and subject to the defining relations given by

(1) r, for all $r \in \bigcup_{i=1}^n R_i$ (defining relations for the copies of H);

(2) s, for all $s \in S$ (defining relations for G);

(3) $[{}^{g}y_1, {}^{g'}y_2]$, for $y_1, y_2 \in \left(\bigcup_{i=1}^n Y_i\right) \cup \{t_i \mid 1 \leq i \leq 5\}$ and $g, g' \in Z \cup \{1\}$ whenever $[{}^{g}y_1, {}^{g'}y_2]$ is indeed a relation in Γ;

(4) $[e_i, y_i]$, for all $e_i \in E_i,\ y_i \in Y_i$ and $1 \leq i \leq n$ and $[z, t_1]$, for all $z \in Z \cap G_{g_1 \cdot x_1}$;

(5) $g_i h g_i^{-1} t_i^{-1}$, for $1 \leq i \leq 5$.

Set $t = t_1$. We use again the notation of Proposition 4.2. So we use the same words w_z if $z \in Z^{\pm 1}$, and $w_y = y$ for all other generators y. It is clear that the set of defining relations above still satisfies the hypothesis of Theorem 6.2. The construction of the diagrams associated to each defining relation can be done exactly as in the case where the action is transitive, as we will argue below. The key fact is that the generators coming from copies of H associated to all other orbits of G (other than $G \cdot x_1$) commute with $t = t_1$.

First, notice that the construction of the diagrams associated to the relations of types (1) or (4) in the case of a transitive action depends only on the fact that $[t, y]$ is a defining relation for all $y \in Y = Y_1$. But t_1 commutes also with all elements of $Y_2 \cup \cdots \cup Y_n$, so the construction can be carried out in the same way. For the case of relations of type (3), it was only necessary that for any generators $g, g' \in Z \cup \{1\}$ and $y, y' \in Y_1$, we could find some $u \in \{h, t_2, \ldots, t_5\}$ that commutes with all the following elements: t, ${}^g t$, ${}^{g'} t$, ${}^g y$ and ${}^{g'} y'$. If we allow y to be an element of $Y_2 \cup \cdots \cup Y_n$, then any u that commutes with t, ${}^g t$, ${}^{g'} t$ and ${}^{g'} y'$ will do it, since ${}^g y$ commutes any choice of u. Thus the five options for u, coming from different copies of H, are enough to let us repeat the argument. Similar considerations cover the cases where either only y', or both y and y' are elements of $Y_2 \cup \cdots \cup Y_n$.

This is all we needed to check, since relations of types (2) and (5) do not involve any of the new generators. □

7. Some observations about Σ^2

Let Γ be a finitely presented group and let $[\chi] \in S(\Gamma)$. Let $\langle \mathscr{X} \mid \mathscr{R} \rangle$ be a finite presentation for Γ. Denote by $C = \mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)$ the associated Cayley complex and by C_χ the full subcomplex of C spanned by Γ_χ. The canonical action of Γ on C restricts to an action by the monoid Γ_χ on C_χ.

Remark 7.1. If a monoid K acts on some set X we still say that the sets $K \cdot x$ are *orbits*. By "K has finitely many orbits on X" we mean that there are elements $x_1, \ldots, x_n \in X$ such that $X = \bigcup_{j=1}^n K \cdot x_j$.

The following lemma can be found in Renz's thesis [1988].

Lemma 7.2. *C_χ has finitely many Γ_χ-orbits of k-cells for $k \leq 2$.*

Proof. Denote by D and D_χ the sets of k-cells of C and C_χ, respectively, (for a fixed $k \leq 2$). We know that Γ acts on D with finitely many orbits. Choose representatives $d_1, \ldots, d_n$ for these orbits so that $d_j \in D_\chi$ but $\gamma \cdot d_j \notin D_\chi$ for all j and for all $\gamma \in \Gamma$ with $\chi(\gamma) < 0$. For this it suffices to take any representatives $\tilde{d}_1, \ldots, \tilde{d}_n$ and then put $d_j := \gamma_j^{-1} \cdot \tilde{d}_j$, where $\gamma_j \in \Gamma$ is the vertex of $\tilde{d}_j$ with lowest χ-value. Thus if $d \in D_\chi$, then $d = \gamma \cdot d_j$ for some j and, by choice of d_j, we have that $\chi(\gamma) \geq 0$. So $D_\chi = \bigcup_{j=1}^n \Gamma_\chi \cdot d_j$. □

Denote by $F(\mathscr{X}, \chi)$ the submonoid of $F(\mathscr{X})$ consisting of the classes of reduced words w with $v_\chi(w) \geq 0$. Note that $F(\mathscr{X}, \chi)$ is indeed closed under the product, since $w_1, w_2 \in F(\mathscr{X}, \chi)$ implies $v_\chi(w_1 w_2) \geq 0$, and this property is preserved by elementary reductions (that is, canceling out terms of the form xx^{-1} or $x^{-1}x$). Let $R(\chi)$ be the subgroup of $F(\mathscr{X})$ consisting of the classes of reduced words w

that represent relations (that is, $w \in \langle \mathscr{R} \rangle^{F(\mathscr{X})}$) and such that $v_\chi(w) \geq 0$. Observe that $R(\chi) \subseteq F(\mathscr{X}, \chi)$ and notice that $R(\chi)$ is indeed a subgroup, since $v_\chi(w) \geq 0$ implies $v_\chi(w^{-1}) \geq 0$ whenever w is a relation. Finally, observe that $R(\chi)$ admits an action by the monoid $F(\mathscr{X}, \chi)$ via left conjugation.

Now, let r be a reduced word in $\mathscr{X}^{\pm 1}$ representing a relation in Γ, that is, $r \in \langle \mathscr{R} \rangle^{F(\mathscr{X})}$. Suppose that M is a van Kampen diagram over $\langle \mathscr{X} \mid \mathscr{R} \rangle$ whose boundary, read in some orientation from some base point p, is exactly r. Then it holds in $F(\mathscr{X})$ that

$$r = {}^{w_1}r_1 \cdots {}^{w_n}r_n, \tag{7-1}$$

where each r_i is a word read on the boundary of some face of M and w_i is the label for a path in M connecting p to a base point of the face associated to r_i. Both the facts that such a diagram exists and that r can be written as above are consequences of van Kampen's lemma (see Proposition 4.1.2 and Theorem 4.2.2 in [Bridson 2002], for instance).

Lemma 7.3. *If $\chi : \Gamma \to \mathbb{R}$ is a character such that $C_\chi = \mathrm{Cay}(G, \langle \mathscr{X} \mid \mathscr{R} \rangle)_\chi$ is* 1*-connected*, *then $R(\chi)$ is finitely generated over $F(\mathscr{X}, \chi)$.*

Remark 7.4. By "$R(\chi)$ is finitely generated over $F(\mathscr{X}, \chi)$" we mean that every element of $R(\chi)$ can be written as a product of elements of the form ${}^{w}s$, where $w \in F(\mathscr{X}, \chi)$ and $s \in S$ for some finite set $S \subseteq R(\chi)$.

Proof. Let $r \in R(\chi)$ and consider the path ρ in C beginning at 1 and with label r. Notice that this path runs inside C_χ, since $v_\chi(r) \geq 0$. Also, ρ is clearly a loop and it must be nullhomotopic in C_χ, since C_χ is 1-connected. A homotopy from ρ to the trivial path can then be realized by a van Kampen diagram M with $v_\chi(M) \geq 0$ (the valuation is taken with respect to 1, seen both as a base point in C and a group element). This is made precise by Theorem 2 in [Renz 1989].

Write r as in (7-1). Thus r is a product of relations corresponding to the faces of M conjugated on the left by elements of $F(\mathscr{X}, \chi)$. Since $v_\chi(M) \geq 0$, such faces are faces of C_χ, so by Lemma 7.2 and using that every element of Γ_χ can be written as a word in $F(\mathscr{X}, \chi)$, each ${}^{w_j}r_j$ can be rewritten as ${}^{u_j}s_j$ where $u_j \in F(\mathscr{X}, \chi)$ and each s_j is a word read on the boundary of a face in a finite set S of representatives of Γ_χ-orbits of faces of C_χ. It follows that S is a finite generating set for $R(\chi)$ modulo the action of $F(\mathscr{X}, \chi)$. □

8. Σ^2 for characters with $\chi|_M = 0$

We get back to a finitely presented wreath product $\Gamma = H \wr_X G = M \rtimes G$. We consider now the nonzero characters $\chi : \Gamma \to \mathbb{R}$ such that $\chi|_M = 0$.

In order to find sufficient conditions for $[\chi] \in \Sigma^2(\Gamma)$, we consider a nice action of Γ on a complex. We will briefly describe the construction in the proof of

Theorem A in [Kropholler and Martino 2016], with the simplifications allowed by the fact that our situation is less general than what is considered in that paper.

We are assuming that $\Gamma = H \wr_X G$ is finitely presented, so H is also finitely presented. Choose a $K(H, 1)$-complex Y, with base point $*$, having a single 0-cell and finitely many 1-cells and 2-cells. Let $Z = \bigoplus_{x \in X} Y_x$ be the *finitary product* of copies of Y indexed by X, that is, Z is the subset of the cartesian product $\prod_{x \in X} Y_x$ consisting on the families $(y_x)_{x \in X}$ such that y_x is not the base point $*$ only for finitely many indices $x \in X$. It follows by the results in [Davis 2012] that Z is an Eilenberg–MacLane space for $M = \bigoplus_{x \in X} H_x$. Notice that Z has a natural cell structure. There is a single 0-cell, given by the family $(y_x)_{x \in X}$ with $y_x = *$ for all x. For $n \geq 1$, an n-cell can be seen as a product $c_1 \times \cdots \times c_k$ of cells of Y, supported by some tuple $(x_1, \ldots, x_k) \in X^k$, such that $\dim(c_1) + \cdots + \dim(c_k) = n$.

There is an obvious action of G on Z. On the other hand, M acts freely on the universal cover E of Z. By putting together these two actions, we get an action of $\Gamma = M \rtimes G$ on E. Notice that, since we are assuming that Γ is finitely presented (in particular G acts on X^2 with finitely many orbits by de Cornulier's results), the 2-skeleton of E has finitely many Γ-orbits of cells. Moreover, since the action of M is free, the stabilizer subgroups are all conjugate to subgroups of G, and can be described as follows:

(1) The stabilizer subgroup of any 0-cell is a conjugate of G.

(2) For $n \geq 1$, the stabilizer subgroup of each n-cell contains a conjugate of the stabilizer $G_{(x_1,\ldots,x_n)}$ of some $(x_1, \ldots, x_n) \in X^n$ as a finite index subgroup.

We make stabilizers of n-cells correspond to stabilizers of n-tuples (rather than k-tuples, for $k \leq n$) by repeating some indices if necessary. Also, the reason why we need to pass to a finite index subgroup is that cells of Z written as products of cells of Y may contain some repetition. For instance, a cell of Z that arises as a product $c \times c$, supported by (x_1, x_2), is also fixed by elements of G that interchange x_1 and x_2. This will also happen in the Γ-action on the universal cover E.

For groups admitting sufficiently nice actions on complexes, there is a criterion for the Σ-invariants.

Theorem 8.1 [Meinert 1997, Theorem B]. *Let E' be a 2-dimensional 1-connected complex. Suppose that a group Γ acts on E' with finitely many orbits of cells. If $\chi : \Gamma \to \mathbb{R}$ is a character such that $[\chi|_{\Gamma_c}] \in \Sigma^{2-\dim(c)}(\Gamma_c)$ for all cells c in E', then $[\chi] \in \Sigma^2(\Gamma)$.*

We apply the theorem above with E' being the 2-skeleton of E. We obtain:

Proposition 8.2. *Suppose that $\Gamma = H \wr_X G$ is finitely presented and let $\chi : \Gamma \to \mathbb{R}$ be a nonzero character such that $\chi|_M = 0$. Suppose also that*

(1) $[\chi|_G] \in \Sigma^2(G)$,

(2) $[\chi|_{G_x}] \in \Sigma^1(G_x)$ *for all* $x \in X$, *and*

(3) $\chi|_{G_{(x,y)}} \neq 0$ *for all* $(x, y) \in X^2$.

Then $[\chi] \in \Sigma^2(\Gamma)$.

The fact that we can state the proposition above with reference only to the stabilizers contained in G follows from the invariance of the Σ^1-invariants under isomorphisms (Proposition B1.5 in [Strebel 2012]). It is also clear that item (3) is equivalent to requiring that the restriction of χ to the actual stabilizers is nonzero.

By Theorem 2.4, if $[\chi] \in \Sigma^2(\Gamma)$ and $\chi|_M = 0$, then $[\chi|_G] \in \Sigma^2(G)$. We can also show that condition (3) of Proposition 8.2 is necessary.

Lemma 8.3. *If* $\chi|_{G_{(x,y)}} = 0$, *then the monoid* G_χ *cannot have finitely many orbits on* $G \cdot (x, y)$.

Proof. Suppose that $G \cdot (x, y) = \bigcup_{j=1}^n G_\chi \cdot (x_j, y_j)$ and choose $g_1, \dots, g_n \in G$ such that $(x_j, y_j) = g_j \cdot (x, y)$. Choose $g \in G$ such that

$$\chi(g) < \min\{\chi(g_j) \mid 1 \leq j \leq n\}.$$

Since $g \cdot (x, y) \in \bigcup_{j=1}^n G_\chi \cdot (x_j, y_j)$, there must be some $g_0 \in G_\chi$ and $1 \leq j \leq n$ such that $g \cdot (x, y) = g_0 \cdot (x_j, y_j) = g_0 g_j (x, y)$. But then $g^{-1} g_0 g_j \in G_{(x,y)}$, with

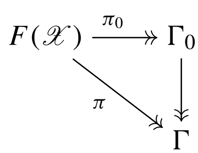
$$\chi(g^{-1} g_0 g_j) = \chi(g_0) + (\chi(g_j) - \chi(g)) > 0,$$

so $\chi|_{G_{(x,y)}} \neq 0$. □

Proposition 8.4. *Let* $\Gamma = H \wr_X G$ *be a finitely presented wreath product and let* $\chi : \Gamma \to \mathbb{R}$ *be a nonzero character. Let* $M = \bigoplus_{x \in X} H_x \subseteq \Gamma$ *and suppose that* $\chi|_M = 0$. *If* $\chi|_{G_{(x,y)}} = 0$ *for some* $(x, y) \in X^2$, *then* $[\chi] \notin \Sigma^2(\Gamma)$.

Proof. We may assume that $[\chi] \in \Sigma^1(\Gamma)$, otherwise there is nothing to do. Thus $[\chi|_G] \in \Sigma^1(G)$ and $\chi|_{G_x} \neq 0$ for all $x \in X$ by Proposition 3.4.

Let $\Gamma_0 = (*_{x \in X} H_x) \rtimes G$ and let $\mathscr{X} \subseteq \Gamma_0$ be a finite generating set. Note that Γ is a quotient of Γ_0, so we can consider the following diagram:



The homomorphism π defines presentations for Γ with generating set $\mathscr{X}$. We first show that for finite presentations of type $\Gamma = \langle \mathscr{X} \mid \mathscr{R} \rangle$ (with $ker(\pi) = \langle \mathscr{R} \rangle^{F(\mathscr{X})}$) the complex $\mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)_\chi$ cannot be 1-connected.

Fix $\langle \mathscr{X} \mid \mathscr{R} \rangle$ to be such a presentation. We use the notation $F(\mathscr{X}, \chi)$ and $R(\chi)$ defined in Section 7. We want to show that $R(\chi)$ is not finitely generated over

$F(\mathscr{X}, \chi)$, from which it follows that $\mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)_\chi$ is not 1-connected by Lemma 7.3.

If $\chi|_{G_{(x,y)}} = 0$ then by Lemma 8.3 we can build a strictly increasing sequence

$$I_1 \subsetneq I_2 \subsetneq \cdots \subsetneq I_j \subsetneq \cdots$$

of G_χ-invariant subsets of X^2 such that $X^2 = \bigcup_j I_j$.

Let N be the normal subgroup of $*_{x\in X} H_x$ such that $M = (*_{x\in X} H_x)/N$. Note that N admits an action by $(*_{x\in X} H_x) \rtimes G$ (which defines the wreath product $H \wr_X G$). Let N_j be the normal subgroup of $*_{x\in x} H_x$ generated by the commutators $[H_x, H_y]$ with $(x, y) \in I_j$. Note that $N_1 \subsetneq N_2 \subsetneq \ldots$, that $N = \bigcup_j N_j$ and that each N_j is $(*_{x\in X} H_x) \rtimes G_\chi$-invariant.

Put

$$\Gamma_{j,\chi} = \frac{*_{x\in X} H_x}{N_j} \rtimes G_\chi.$$

This is a well defined monoid under the operation we use to define the semidirect product. This defines a sequence $\{\Gamma_{j,\chi}\}_j$ of monoids that converges to Γ_χ.

Now, remember we have chosen $\mathscr{X}$ so that the projection $\pi_0 : F(\mathscr{X}) \twoheadrightarrow \Gamma_0$ is well defined. Passing to monoids, we obtain a homomorphism

$$p_0 : F(\mathscr{X}, \chi) \to (\Gamma_0)_{\chi_0},$$

where χ_0 is the obvious lift of χ to Γ_0. From p_0 we define

$$p_j : F(\mathscr{X}, \chi) \to \Gamma_{j,\chi}$$

for each $j \geq 1$. Let

$$R_j = p_j^{-1}(\{1\}).$$

Notice that $R_j \subsetneq R(\chi)$ for all j. Indeed, it is clear that χ_0 and χ restrict to the same homomorphisms on G and G_x, for all $x \in X$. Also, χ_0 restricts to zero on $*_{x\in X} H_x$ by construction. So it follows from Theorem 5.1 that $[\chi_0] \in \Sigma^1(\Gamma_0)$, since we are assuming that $[\chi] \in \Sigma^1(\Gamma)$. Thus p_0 is surjective and any $n \in N \setminus N_j$ defines an element in $R(\chi) \setminus R_j$. Observe further that each that R_j is actually a $F(\mathscr{X}, \chi)$-invariant subgroup of $R(\chi)$ and that $\bigcup_j R_j = R(\chi)$. The existence of the sequence $\{R_j\}_j$ implies that $R(\chi)$ cannot be finitely generated over $F(\mathscr{X}, \chi)$.

For the general case, let $\langle \mathscr{X} \mid \mathscr{R} \rangle$ be any finite presentation for Γ and suppose by contradiction that $\mathrm{Cay}(\Gamma; \langle \mathscr{X} \mid \mathscr{R} \rangle)_\chi$ is 1-connected. From $\langle \mathscr{X} \mid \mathscr{R} \rangle$ we build another finite presentation $\langle \mathscr{X}' \mid \mathscr{R}' \rangle$ for Γ with $\mathscr{X} \subseteq \mathscr{X}'$, $\mathscr{R} \subseteq \mathscr{R}'$ and satisfying the previous hypothesis (that is, $\mathscr{X}'$ is actually a generating set for Γ_0). For this, it suffices to add the necessary generators and include the relations that define them in Γ in terms of the previous generating set $\mathscr{X}$. It may be the case that $\mathrm{Cay}(\Gamma; \langle \mathscr{X}' | \mathscr{R}' \rangle)_\chi$ is not 1-connected anymore, but by Lemma 3 in [Renz 1989], we can always enlarge $\mathscr{R}'$ to a (still finite) set $\mathscr{R}''$ so that $\mathrm{Cay}(\Gamma; \langle \mathscr{X}' \mid \mathscr{R}'' \rangle)_\chi$ is indeed 1-connected. This is done by adding the relations of the form $t^{-1}xtw_x^{-1}$, as in Theorem 6.2. We arrive at a contradiction with the first part of the proof, since $\mathscr{X}'$ satisfies the previous hypothesis, that is, $\mathscr{X}'$ can be lifted to a generating set for Γ_0. □

The above proposition completes the proof of Theorem C as stated in the introduction, since its last assertion (when we assume that H has infinite abelianization) follows from Theorem 3.1.

9. Applications to twisted conjugacy

We now derive some consequences of the previous results to twisted conjugacy, more specifically to the study of Reidemeister numbers of automorphisms of wreath products. For this we start by considering the Koban invariant Ω^1.

Given a finitely generated group Γ, endow $\mathrm{Hom}(\Gamma, \mathbb{R})$ with an inner product structure, so that it makes sense to talk about angles in $S(\Gamma)$. Denote by $N_{\pi/2}([\chi])$ the open neighborhood of angle $\pi/2$ and centered at $[\chi] \in S(\Gamma)$. Following Koban [2006], we can define the invariant $\Omega^1(\Gamma)$ in terms of $\Sigma^1(\Gamma)$:

$$\Omega^1(\Gamma) = \{[\chi] \in S(\Gamma) \mid N_{\pi/2}([\chi]) \subseteq \Sigma^1(\Gamma)\}.$$

A proof of the fact that this does not depend on the inner product can be found in the above-mentioned paper, which contains the original definition of the invariant.

Let $\Gamma = H \wr_X G$ be a finitely generated wreath product. With some restrictions on the action by G on X, we can obtain nice descriptions of $\Omega^1(\Gamma)$. Notice that, since the invariant does not depend on the choice of inner product, we can assume that characters $[\chi], [\eta] \in S(\Gamma)$ such that $\chi|_G = 0$ and $\eta|_M = 0$ are always orthogonal, and this will be done in the proposition below.

Proposition 9.1. *Let $\Gamma = H \wr_X G$ be a finitely generated wreath product. Suppose*

$$\Sigma^1(\Gamma) = \{[\chi] \in S(\Gamma) \mid \chi|_M \neq 0\},$$

where $M = \bigoplus_{x\in x} H_x \subseteq \Gamma$. Then

$$\Omega^1(\Gamma) = \{[\chi] \in S(\Gamma) \mid \chi|_G = 0\}.$$

Proof. Let $[\chi] \in S(\Gamma)$ with $\chi|_G = 0$. Clearly $\chi|_M \neq 0$, so $[\chi] \in \Sigma^1(\Gamma)$. Furthermore, if $[\eta] \in N_{\pi/2}([\chi])$, then $\eta|_M \neq 0$, otherwise χ and η would be orthogonal. So $N_{\pi/2}([\chi]) \subseteq \Sigma^1(\Gamma)$ whenever $\chi|_G = 0$. On the other hand, if there were some $[\chi] \in \Omega^1(\Gamma)$ with $\chi|_G \neq 0$, then by taking $\eta : \Gamma \to \mathbb{R}$ defined by $\eta|_M = 0$ and $\eta|_G = \chi|_G$, we would have that $[\eta] \in N_{\pi/2}([\chi])$, but $[\eta] \notin \Sigma^1(\Gamma)$. □

For any group V, we denote by V^{ab} its abelianization. By Theorem A, if the G-action on X does not contain orbits composed by only one element, then many conditions imply the hypothesis on the description of $\Sigma^1(\Gamma)$, such as

(1) $(G_x)^{\mathrm{ab}}$ is finite for some $x \in X$, or
(2) the set $\{[\chi] \in \Sigma^1(G) \mid \chi|_{G_x} \neq 0\}$ is empty for some $x \in X$.

This includes the cases where the G-action is free (in particular the regular wreath products $\Gamma = H \wr G$) and the case where $\Sigma^1(G) = \varnothing$.

Recall that the *Reidemeister number* $R(\varphi)$, for a group isomorphism $\varphi : V \to V$, is defined as the number of orbits of the φ-twisted conjugacy action of V on itself. A connection between the invariant Ω^1 and Reidemeister numbers was studied by Koban and Wong [2011]. Recall that a character χ is *discrete* if its image is infinite cyclic.

Theorem 9.2 [Koban and Wong 2011, Theorem 4.3]. *Let G be a finitely generated group and suppose that $\Omega^1(G)$ contains only discrete characters.*

(1) *If $\Omega^1(G)$ contains only one element, then G is of type R_∞, that is, $R(\varphi) = \infty$ for all $\varphi \in Aut(G)$.*
(2) *If $\Omega^1(G)$ has exactly two elements, then there is a subgroup $N \subseteq Aut(G)$, with $[Aut(G) : N] = 2$, such that $R(\varphi) = \infty$ for all $\varphi \in N$.*

Corollary 9.3. *Let $\Gamma = H \wr_X G$ be a finitely generated wreath product and suppose that the G-action on X is transitive. Suppose further that $\Sigma^1(\Gamma)$ is as described in Proposition 9.1 and that H^{ab} has torsion-free rank 1. Then there is a subgroup $N \subseteq Aut(\Gamma)$, with $[Aut(\Gamma) : N] = 2$, such that $R(\varphi) = \infty$ for all $\varphi \in N$.*

Proof. By the hypothesis on H^{ab} we have that

$$\Omega^1(\Gamma) = \{[\nu_1], [\nu_2]\},$$

where $\nu_j(G) = 0$, $\nu_1(h) = 1$ and $\nu_2(h) = -1$ for some lift $h \in H$ of a generator for the infinite cyclic factor of H^{ab}. It suffices then to apply part (2) of Theorem 9.2. □

The applications that we keep in mind are the finitely generated regular wreath products of the form $\mathbb{Z} \wr G$.

Gonçalves and Kochloukova [2010] exhibited other connections between the Σ-theory and the property R_∞. Below we denote by $\Sigma^1(G)^c$ the complement of $\Sigma^1(G)$ in $S(G)$, that is, $\Sigma^1(G)^c = S(G) \smallsetminus \Sigma^1(G)$.

Theorem 9.4 [Gonçalves and Kochloukova 2010, Corollary 3.4]. *Let G be a finitely generated group and suppose that*

$$\Sigma^1(G)^c = \{[\chi_1], \ldots, [\chi_n]\},$$

where $n \geq 1$ and each χ_j is a discrete character. Then there is a subgroup of finite index $N \subseteq Aut(G)$ such that $R(\varphi) = \infty$ for all $\varphi \in N$.

Corollary 9.5. *Let $\Gamma = H \wr_X G$ be a finitely generated wreath product. Once again, suppose that $\Sigma^1(\Gamma)$ is as described in Proposition 9.1. Suppose further that G^{ab} has torsion-free rank* 1*. Then there is a subgroup of finite index $N \subseteq Aut(\Gamma)$ such that $R(\varphi) = \infty$ for all $\varphi \in N$.*

Proof. Under the hypothesis above, we have

$$\Sigma^1(\Gamma)^c = \{[\chi_1], [\chi_2]\},$$

where $\chi_j|_M = 0$ and $\chi_1(g) = 1$ and $\chi_2(g) = -1$ for some $g \in G$ whose image in G^{ab} is a generator of the infinite cyclic factor. Then Theorem 9.4 applies. □

This time we can take as an example the regular wreath product $\Gamma = H \wr \mathbb{Z}$.

We note that Gonçalves and Wong [2006] and Taback and Wong [2011] had already obtained some results about the property R_∞ for regular wreath products of the form $H \wr \mathbb{Z}$, with H abelian or finite. Our results complement theirs in the sense that they consider other basis groups H and nonregular actions, but here we were limited to talk about Reidemeister numbers of automorphisms contained in subgroups of finite index in the automorphism group. In the above-mentioned papers, on the other hand, the authors were able to determine positively the property R_∞ for some choices of H.

Acknowledgements

The author is supported by PhD grants 2015/22064-6 and 2016/24778-9, from São Paulo Research Foundation (FAPESP). The author would like to thank his advisor, Prof. Dessislava Kochloukova, for her guidance. He also wishes to thank P. H. Kropholler and A. Martino, for valuable discussions about their paper, and the anonymous referee, for many suggestions.

References

[Bartholdi et al. 2015] L. Bartholdi, Y. de Cornulier, and D. H. Kochloukova, "Homological finiteness properties of wreath products", *Q. J. Math.* **66**:2 (2015), 437–457. MR Zbl

[Bestvina and Brady 1997] M. Bestvina and N. Brady, "Morse theory and finiteness properties of groups", *Invent. Math.* **129**:3 (1997), 445–470. MR Zbl

[Bieri 1976] R. Bieri, *Homological dimension of discrete groups*, Queen Mary College, London, 1976. MR Zbl

[Bieri and Geoghegan 2010] R. Bieri and R. Geoghegan, "Sigma invariants of direct products of groups", *Groups Geom. Dyn.* **4**:2 (2010), 251–261. MR Zbl

[Bieri and Renz 1988] R. Bieri and B. Renz, "Valuations on free resolutions and higher geometric invariants of groups", *Comment. Math. Helv.* **63**:3 (1988), 464–497. MR Zbl

[Bieri and Strebel 1980] R. Bieri and R. Strebel, "Valuations and finitely presented metabelian groups", *Proc. London Math. Soc.* (3) **41**:3 (1980), 439–464. MR Zbl

[Bieri et al. 1987] R. Bieri, W. D. Neumann, and R. Strebel, "A geometric invariant of discrete groups", *Invent. Math.* **90**:3 (1987), 451–477. MR Zbl

[Bieri et al. 2010] R. Bieri, R. Geoghegan, and D. H. Kochloukova, "The sigma invariants of Thompson's group F", *Groups Geom. Dyn.* **4**:2 (2010), 263–273. MR Zbl

[Bridson 2002] M. R. Bridson, "The geometry of the word problem", pp. 29–91 in *Invitations to geometry and topology*, edited by M. R. Bridson and S. M. Salamon, Oxford Grad. Texts Math. **7**, Oxford Univ. Press, 2002. MR Zbl

[de Cornulier 2006] Y. de Cornulier, "Finitely presented wreath products and double coset decompositions", *Geom. Dedicata* **122** (2006), 89–108. MR Zbl

[Davis 2012] M. W. Davis, "Right-angularity, flag complexes, asphericity", *Geom. Dedicata* **159** (2012), 239–262. Correction by M. W. Davis and P. H. Kropholler in **179** (2015), 39–44. MR Zbl

[Gehrke 1998] R. Gehrke, "The higher geometric invariants for groups with sufficient commutativity", *Comm. Algebra* **26**:4 (1998), 1097–1115. MR Zbl

[Gonçalves and Kochloukova 2010] D. Gonçalves and D. H. Kochloukova, "Sigma theory and twisted conjugacy classes", *Pacific J. Math.* **247**:2 (2010), 335–352. MR Zbl

[Gonçalves and Wong 2006] D. Gonçalves and P. Wong, "Twisted conjugacy classes in wreath products", *Internat. J. Algebra Comput.* **16**:5 (2006), 875–886. MR Zbl

[Koban 2006] N. Koban, "Controlled topology invariants of translation actions", *Topology Appl.* **153**:12 (2006), 1975–1993. MR Zbl

[Koban and Wong 2011] N. Koban and P. Wong, "A relationship between twisted conjugacy classes and the geometric invariants Ω^n", *Geom. Dedicata* **151** (2011), 233–243. MR Zbl

[Kochloukova 2001] D. H. Kochloukova, "More about the geometric invariants $\Sigma^m(G)$ and $\Sigma^m(G, \mathbb{Z})$ for groups with normal locally polycyclic-by-finite subgroups", *Math. Proc. Cambridge Philos. Soc.* **130**:2 (2001), 295–306. MR Zbl

[Kochloukova 2012] D. H. Kochloukova, "On the Σ^2-invariants of the generalised R. Thompson groups of type F", *J. Algebra* **371** (2012), 430–456. MR Zbl

[Kropholler and Martino 2016] P. H. Kropholler and A. Martino, "Graph-wreath products and finiteness conditions", *J. Pure Appl. Algebra* **220**:1 (2016), 422–434. MR Zbl

[Meier and VanWyk 1995] J. Meier and L. VanWyk, "The Bieri–Neumann–Strebel invariants for graph groups", *Proc. London Math. Soc.* (3) **71**:2 (1995), 263–280. MR Zbl

[Meier et al. 1998] J. Meier, H. Meinert, and L. VanWyk, "Higher generation subgroup sets and the Σ-invariants of graph groups", *Comment. Math. Helv.* **73**:1 (1998), 22–44. MR Zbl

[Meier et al. 2001] J. Meier, H. Meinert, and L. VanWyk, "On the Σ-invariants of Artin groups", *Topology Appl.* **110**:1 (2001), 71–81. MR Zbl

[Meinert 1995] H. Meinert, "The Bieri–Neumann–Strebel invariant for graph products of groups", *J. Pure Appl. Algebra* **103**:2 (1995), 205–210. MR Zbl

[Meinert 1997] H. Meinert, "Actions on 2-complexes and the homotopical invariant Σ^2 of a group", *J. Pure Appl. Algebra* **119**:3 (1997), 297–317. MR Zbl

[Renz 1988] B. Renz, *Geometrische Invarianten und Endlichkeitseigenschaften von Gruppen*, Ph.D. thesis, Goethe-Universität Frankfurt am Main, 1988, Available at https://tinyurl.com/renzdis. Zbl

[Renz 1989] B. Renz, "Geometric invariants and HNN-extensions", pp. 465–484 in *Group theory* (Singapore, 1987), edited by K. N. Cheng and Y. K. Leong, de Gruyter, Berlin, 1989. MR Zbl

[Schütz 2008] D. Schütz, "On the direct product conjecture for sigma invariants", *Bull. Lond. Math. Soc.* **40**:4 (2008), 675–684. MR Zbl

[Strebel 2012] R. Strebel, "Notes on the Sigma invariants", preprint, 2012. arXiv

[Taback and Wong 2011] J. Taback and P. Wong, "The geometry of twisted conjugacy classes in wreath products", pp. 561–587 in *Geometry, rigidity, and group actions*, edited by B. Farb and D. Fisher, Univ. Chicago Press, 2011. MR Zbl arXiv

[Zaremsky 2017] M. C. B. Zaremsky, "On the Σ-invariants of generalized Thompson groups and Houghton groups", *Int. Math. Res. Not.* **2017**:19 (2017), 5861–5896. MR

Received September 15, 2017. Revised May 15, 2018.

LUIS AUGUSTO DE MENDONÇA
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CAMPINAS (UNICAMP)
CAMPINAS
BRAZIL

luismendonca@ime.unicamp.br

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.141

ENHANCED ADJOINT ACTIONS AND THEIR ORBITS FOR THE GENERAL LINEAR GROUP

KYO NISHIYAMA AND TAKUYA OHTA

We study an enhanced adjoint action of the general linear group on a product of its Lie algebra and a vector space consisting of several copies of defining representations and its duals. We determine regular semisimple orbits (i.e., closed orbits of maximal dimension) and the structure of enhanced null cone, including its irreducible components and their dimensions.

Introduction

Let G be a reductive algebraic group over the complex number field $\mathbb{C}$, and $\mathfrak{g}$ its Lie algebra. The adjoint action of G on $\mathfrak{g}$ is a basic tool for many aspects of representation theory, and is also useful for invariant theory, the theory of singularities, and so on.

Achar, Henderson and Johnson [Achar and Henderson 2008; Achar et al. 2011; Johnson 2010] considered an enhanced version of nilpotent varieties and classified the nilpotent orbits (there are only finitely many of them). Kato [2009] considered an "exotic" nilpotent cone and derived the Deligne–Langlands theory for those exotic nilpotent orbits. There are many related works based on algebraic geometry, combinatorial theory, and the theory of character sheaves [Travkin 2009; Finkelberg et al. 2009; Henderson and Trapa 2012; Fresse and Nishiyama 2016; Rosso 2012].

In these papers, enhancement of the nilpotent cone is only "one-sided" to get a criterion of finiteness of orbits. However, from the viewpoint of symmetric spaces and invariant theory, it seems better to enhance all the adjoint orbits in two-sided directions. In this respect, we already had two results that relate the orbit structure of two enhanced actions [Ohta 2008; Nishiyama 2014], but we did not know the explicit orbit structures of individual enhanced adjoint actions.

In this paper, we begin to study (two-sided) "enhanced adjoint action" of G for $G = \mathrm{GL}_n(\mathbb{C})$ (type A). The big difference from those one-sided enhanced (or exotic) ones is that there exist infinitely many nilpotent orbits. So the analysis becomes more difficult, but involves less combinatorics. In the easiest cases, we can

Nishiyama is supported by JSPS KAKENHI Grant #16K05070.
MSC2010: primary 14L30; secondary 15A72, 16W22.
Keywords: enhanced nilpotent cone, exotic nilpotent cone, adjoint quotient, classical invariant theory.

describe enhanced adjoint orbits fairly explicitly, but in general, we have obtained coarser structures, like regular orbits of maximal possible dimensions, the structure of invariants, irreducible components of nilpotent variety.

To state the main results more explicitly, let us introduce some notation. Let $V = \mathbb{C}^n$ be a vector space of dimension n. We consider a natural action of $G = \mathrm{GL}(V) = \mathrm{GL}_n(\mathbb{C})$ on

$$W = (\mathbb{C}^n)^{\oplus p} \oplus (\mathbb{C}^{*n})^{\oplus q} \oplus \mathrm{M}_n = \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n,$$

with the action of $g \in G$ given by

$$g \cdot (B, C, A) = (gB, Cg^{-1}, \mathrm{Ad}(g)A) \quad \text{for } (B, C, A) \in \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n.$$

Thus, the part M_n is considered to be $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{C})$ and the action is the adjoint action. For the other parts, $\mathrm{M}_{n,p}$ is a p-copy of natural representations and $\mathrm{M}_{q,n}$ is a q-copy of its dual, i.e., as a representation space we will study

$$W \simeq V^{\oplus p} \oplus (V \otimes V^*) \oplus (V^*)^{\oplus q}.$$

The space W is the *fully* enhanced adjoint representation as we explained. Here we note that, from the opposite view point, the space W is also considered as an extension of $V^{\oplus p} \oplus (V^*)^{\oplus q}$ by adding $V \otimes V^*$. Hence it is a generalization of what H. Weyl considered in the course of his study of classical invariant theory [1939].

There are obvious invariants for the action of $G = \mathrm{GL}_n(\mathbb{C})$ on W. We put

$$\begin{aligned} \tau_k &:= \operatorname{trace} A^k && (1 \le k \le n-1), \\ \gamma_{i,j}^{\ell} &:= (C A^{\ell} B)_{i,j} && (0 \le \ell \le n-1,\ 1 \le i \le q,\ 1 \le j \le p). \end{aligned}$$

These invariants are generators of the whole invariant ring $\mathbb{C}[W]^G$, and they seem to be known to experts in various forms, including in quiver theory (see Theorem 1.1). Thus, we can define a quotient map $\pi_W : W \to \mathbb{C}^n \times (\mathrm{M}_{q,p})^n$ using these invariants (see (2-3)).

If $p = 1$ or $q = 1$, the quotient map has a very good property. Namely, we get:

Theorem 0.1 (Theorem 2.1(2)). *If $p = 1$ or $q = 1$, the map $\pi_W : W \to \mathbb{C}^n \times (\mathrm{M}_{q,p})^n$ is an affine categorical quotient map (note that $\mathrm{M}_{q,p} = \mathbb{C}^p$ or $\mathbb{C}^q$). In particular, the quotient map π_W is coregular, and $\mathbb{C}[W]^G$ is a polynomial ring generated by the fundamental invariants listed above.*

For general $p \ge 1$ and $q \ge 1$, the following theorem gives a generic structure of enhanced adjoint orbits.

Theorem 0.2 (Theorem 2.1, Corollary 2.2). *The dimension of the image* $\dim \operatorname{Im} \pi_W$ *is equal to $n(p+q)$, and a general fiber of π_W is a single G-orbit of dimension n^2. This implies that general orbits for the enhanced adjoint action are closed of dimension n^2.*

These orbits are called *regular semisimple orbits*. Another extreme case are nilpotent orbits. We investigate the null cone $\mathfrak{N}(W) \subset W$ in Section 3, and get the following results.

Theorem 0.3 (Theorem 3.3). *The null cone $\mathfrak{N}(W)$ is reducible and it has $n+1$ irreducible components $C_k \subset \mathfrak{N}(W)$ $(0 \le k \le n)$ given in Lemma 3.2. The dimension of the null cone is $n^2 - n + n \cdot \max\{p, q\}$ and $\mathfrak{N}(W)$ is equidimensional if and only if $p = q$.*

Finally, we get the structure of general (enhanced) nilpotent orbits contained in each component C_k in Theorem 3.4.

1. Setting

Let $V = \mathbb{C}^n$ be a vector space of dimension n. We consider a natural action of $G = \mathrm{GL}(V)$ on

$$W = W(p, q; r) := V^{\oplus p} \oplus (V^*)^{\oplus q} \oplus (V \otimes V^*)^{\oplus r}$$

in the obvious manner. In explicit matrix form, we can identify

$$W = (\mathbb{C}^n)^{\oplus p} \oplus (\mathbb{C}^{*n})^{\oplus q} \oplus (\mathrm{M}_n)^{\oplus r} = \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n^r,$$

with the action of $g \in G$ on

$$(B, C, (A_1, \dots, A_r)) \in \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n^r$$

given by

$$g \cdot (B, C, (A_1, \dots, A_r)) = (gB, Cg^{-1}, (\mathrm{Ad}(g)A_i)_{i=1}^r).$$

There are obvious invariants, which we list below. For a multi-index

$$I = (i_1, i_2, \dots, i_\ell) \quad (1 \le i_k \le r),$$

let us write $A_I = A_{i_1} A_{i_2} \cdots A_{i_\ell}$. We denote $[n] = \{1, 2, \dots, n\}$ as usual; then the multi-index I above is an element in $[r]^\ell$. We put

$$\begin{aligned} \tau_I &:= \mathrm{trace}(A_I) \qquad (I \in [r]^\ell), \\ \gamma_{i,j}^K &:= (C A_K B)_{i,j} \quad (K \in [r]^\ell, 1 \le i \le q, 1 \le j \le p), \end{aligned}$$

where we allow $\ell = 0$ for K, which means $A_K = 1_n$ (the identity matrix). These invariants are generators of the whole invariant ring, which is essentially due to a more general result of Le Bruyn and Procesi [1990, § 3, Theorem 1] (see also [Le Bruyn and Procesi 1987; Itoh 2013]).

Theorem 1.1. *The invariant ring* $\mathbb{C}[W]^G$ *is generated by the elements* τ_I *with* $I \in [r]^\ell$ $(\ell \geq 0)$ *and the elements* $\gamma_{i,j}^K$ *with* $K \in [r]^\ell$ $(\ell \geq 0)$, $i \in [q]$, $j \in [p]$; *that is,*

$$\mathbb{C}[W]^G = \mathbb{C}\big[\tau_I, \gamma_{i,j}^K \mid I, K \in [r]^\ell \ (\ell \geq 0), i \in [q], j \in [p]\big].$$

Proof. We largely follow the notation of [Le Bruyn and Procesi 1990]. We denote a connected quiver by Q and by α its dimension vector. For a representation space $R(Q,\alpha)$ of Q, Theorem 1 in [loc. cit.] states that the invariant ring $\mathbb{C}[Q,\alpha]^{\mathrm{GL}(\alpha)}$ is generated by traces of oriented cycles. So we will consider a quiver Q of two vertices $Q_0 = \{1, 2\}$ with arrows

$$Q_1 = \{a_i \mid 1 \leq i \leq r\} \cup \{b_i \mid 1 \leq i \leq p\} \cup \{c_i \mid 1 \leq i \leq q\},$$

where the a_i are loops connecting 1 and itself (i.e., $h(a_i) = t(a_i) = 1$), the b_i are arrows from 2 to 1 ($h(b_i) = 2,\ t(b_i) = 1$), and the c_i are arrows from 1 to 2 ($h(c_i) = 1,\ t(c_i) = 2$). Take a dimension vector $\alpha = (\alpha(1), \alpha(2)) = (n, 1)$, so that $V(1) = \mathbb{C}^n$ and $V(2) = \mathbb{C}$. Then our $W = W(p,q,r)$ coincides with the representation space $R(Q,\alpha)$.

The invariants are considered with respect to the action of $G(\alpha) = \mathrm{GL}_n \times \mathrm{GL}_1$. However, the representation image of $G(\alpha)$ on $W = R(Q,\alpha)$ and that of GL_n are the same because the action of the torus GL_1 on $V(2) = \mathbb{C}$ can be recaptured by the center of GL_n. So both invariant rings for $G(\alpha)$ and GL_n are the same.

Let us consider any closed cycles. Since we take traces, we can start from any vertices contained in the cycle. If it only contains the vertex 1, the traces are τ_I's. If it contains the vertex 2, we will start from 2 which necessarily ends in 2. Decompose the cycle into several cycles which start from 2 and end in 2. Since $V(2) = \mathbb{C}$ is 1-dimensional, a decomposed cycle starting from 2 represents a scalar being equal to its trace. Thus the trace of the cycle which we are considering is a product of various $\gamma_{i,j}^K$'s. □

Let us denote $\pi = \pi_W : W \to W/\!/G$, an affine quotient map by the action above. As a set, the quotient $W/\!/G$ corresponds to the set of closed G-orbits in W. It is known that these closed orbits are precisely the set of equivalence classes of completely reducible representations of a quiver corresponding to W.

Let $\mathfrak{N}(W) = \pi_W^{-1}(\pi_W(0))$ be the nilpotent variety, which consists of the nilpotent elements x with the property $\overline{G \cdot x} \ni 0$. The nilpotent variety $\mathfrak{N}(W)$ is the "worst" fiber. So we are strongly interested in its structure. In particular, we are interested in $\dim \mathfrak{N}(W)$, its irreducible components, its orbit structure, and whether it is reduced or not. For the dimensions and irreducible components, we have a complete result, which is stated in Section 3 in detail. The problems of orbit structure and reducibility of $\mathfrak{N}(W)$ also seem very interesting but these are our future subjects.

On the other hand, general fibers are supposed to have "best" properties we can expect. This will be helpful for studying the quotient space (at least its smooth part), which we shall do in Section 2.

It would be too ambitious to expect to get a very explicit orbit structure of the whole space W. Also it seems to be a difficult problem to clarify the structure of the singularities of the quotient space.

2. Enhanced adjoint action

In the following, we restrict ourselves to the case $r = 1$, so $W = \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n$, on which $G = \mathrm{GL}_n$ acts. In matrix form, $g \in \mathrm{GL}_n$ acts on

$$(B, C, A) \in \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n$$

via $g \cdot (B, C, A) = (gB, Cg^{-1}, \mathrm{Ad}(g)A)$. We call this action the *enhanced adjoint action*.

Now Theorem 1.1 gives a set of generators of G-invariants:

$$\tau_k := \mathrm{trace}(A^k) \qquad (1 \le k \le n), \tag{2-1}$$

$$\gamma_{i,j}^k := (CA^kB)_{i,j} \quad (0 \le k \le n-1,\ 1 \le i \le q,\ 1 \le j \le p). \tag{2-2}$$

Note that A^n is a linear combination of A^k's $(0 \le k \le n-1)$ thanks to the Cayley–Hamilton formula, so we don't need higher powers of A in τ_k or $\gamma_{i,j}^k$. Let us denote the affine quotient map by

$$\begin{aligned} \pi_W &: W \to \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^n, \\ (A, B, C) &\mapsto \big((\tau_k)_{k=1}^n; ((\gamma_{i,j}^k)_{i,j})_{k=0}^{n-1}\big) = \big((\tau_k)_{k=1}^n; (CA^kB)_{k=0}^{n-1}\big). \end{aligned} \tag{2-3}$$

By the general theory of quotients, we know the image $\mathrm{Im}\,\pi_W$ is a closed subvariety of $\mathbb{C}^n \oplus (\mathrm{M}_{q,p})^n$. Let us denote by $\mathrm{Det}_r(\mathrm{M}_{q,p})$ the determinantal variety consisting of matrices in $\mathrm{M}_{q,p}$ of rank less than or equal to r. Clearly, if we put $m = \min\{p, q, n\}$, $\mathrm{Im}\,\pi_W$ is contained in $\mathbb{C}^n \times \mathrm{Det}_m(\mathrm{M}_{q,p})^n$. However, it is much smaller, as you can see from the theorem below.

Theorem 2.1. *Under the setting above, the image* $\mathrm{Im}\,\pi_W$ *is isomorphic to the affine quotient* $W/\!/G = \mathrm{Spec}\,(\mathbb{C}[W]^G)$. *Moreover*:

(1) *There is a dominant map*

$$\Psi : \mathbb{C}^n \times (\mathrm{Det}_1(\mathrm{M}_{q,p}))^n \to \mathrm{Im}\,\pi_W,$$

whose restriction to a dense open subset of $\mathbb{C}^n \times (\mathrm{Det}_1(\mathrm{M}_{q,p}))^n$ *gives an affine quotient map under the diagonal action of* S_n *(permuting both coordinates) to a dense open subset of* $\mathrm{Im}\,\pi_W$. *Consequently, we get* $\dim W/\!/G = \dim \mathrm{Im}\,\pi_W = n(p+q)$, *and a general fiber of* π_W *is of dimension* n^2.

(2) *If* $p = 1$ *or* $q = 1$, *the quotient map* π_W *is surjective, and*

$$\operatorname{Im}\pi_W = \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^{\oplus n}$$

is an affine space. In particular, the quotient map π_W *is coregular, and* $\mathbb{C}[W]^G$ *is a polynomial ring of the fundamental invariants listed in* (2-1) *and* (2-2).

Proof. Let us fix a generic diagonal matrix $A = t = \operatorname{diag}(t_1, \dots, t_n)$, where $t_i \neq t_j$ $(i \neq j)$. For $1 \le r \le n$, put

$$X^{(r)} = \begin{pmatrix} c_{1,r} \\ c_{2,r} \\ \vdots \\ c_{p,r} \end{pmatrix} (b_{r,1}, b_{r,2}, \dots, b_{r,q}) \in \operatorname{Det}_1(\mathrm{M}_{q,p}),$$

where $c_{i,j}$ denotes the (i, j)-element of the matrix $C \in \mathrm{M}_{q,n}$ and similarly $b_{i,j}$ for $B \in \mathrm{M}_{n,p}$. We get

$$CA^kB = (\gamma^k_{i,j})_{i,j} = \Big(\sum_{r=1}^{n} c_{i,r} t_r^k b_{r,j}\Big)_{i,j} = \sum_{r=1}^{n} t_r^k X^{(r)} =: \Gamma^{(k)}. \tag{2-4}$$

Thus, in the matrix expression,

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \\ \vdots & \vdots & \ddots & \vdots \\ t_1^{n-1} & t_2^{n-1} & \cdots & t_n^{n-1} \end{pmatrix} \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(n)} \end{pmatrix} = \begin{pmatrix} \Gamma^{(0)} \\ \Gamma^{(1)} \\ \vdots \\ \Gamma^{(n-1)} \end{pmatrix}, \tag{2-5}$$

hence

$$\begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(n)} \end{pmatrix} = D(t)^{-1} \begin{pmatrix} \Gamma^{(0)} \\ \Gamma^{(1)} \\ \vdots \\ \Gamma^{(n-1)} \end{pmatrix}, \tag{2-6}$$

where $D(t) = (t_j^{i-1})_{i,j}$ denotes the Vandermonde matrix in (2-5). Bearing this calculation in mind, we define a map $\Psi : \mathbb{C}^n \times (\operatorname{Det}_1(\mathrm{M}_{q,p}))^n \to \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^n$ by

$$\Psi(t; (X^{(k)})_{k=1}^n) = \bigg(\Big(\sum_{i=1}^{n} t_i^k\Big)_{k=1}^{n}; (\Gamma^{(k)})_{k=1}^n = D(t)(X^{(k)})_{k=1}^n\bigg). \tag{2-7}$$

We will show that $\operatorname{Im}\Psi \subset \operatorname{Im}\pi_W$ so that we get a map from $U := \mathbb{C}^n \times (\operatorname{Det}_1(\mathrm{M}_{q,p}))^n$ to $\operatorname{Im}\pi_W$, denoted by the same letter Ψ.

To see $\operatorname{Im}\Psi \subset \operatorname{Im}\pi_W$, take $(\tau; (\Gamma^{(k)})_k) \in \operatorname{Im}\pi_W$ for which τ is in an image of regular semisimple A. For this A, we can pick a diagonal matrix $t = \operatorname{diag}(t_1, \dots, t_n)$ in the same adjoint orbit of A, which implies $\tau_k(A) = \sum_{i=1}^{n} t_i^k$. Using this t, we can recover $X^{(k)}$'s via (2-6), since $t_i \neq t_j$ $(i \neq j)$. As we saw above, if $(\tau; (\Gamma^{(k)})_k) \in \operatorname{Im} \pi_W$ then $\operatorname{rank} X^{(k)} \leq 1$. Here we require that those $X^{(k)}$'s are all exactly rank one matrices. This is an open condition (in $\operatorname{Im} \pi_W$) and it does not depend on the choice of the diagonal representatives of A. We define an open dense set $(\operatorname{Im} \pi_W)' \subset \operatorname{Im} \pi_W$ consisting of $(\tau; (\Gamma^{(k)})_k) \in \operatorname{Im} \pi_W$ for which (i) τ is in an image of regular semisimple A; and (ii) $\operatorname{rank} X^{(k)} = 1$ for $1 \leq k \leq n$. Thus we conclude that Ψ is a surjective map from an open dense subset of U to $(\operatorname{Im} \pi_W)'$. Consequently, the image $\operatorname{Im} \Psi$ is contained in the closed subvariety $\operatorname{Im} \pi_W$, and we get a well defined map from $U = \mathbb{C}^n \times (\operatorname{Det}_1(\mathrm{M}_{q,p}))^n$ to $\operatorname{Im} \pi_W$ by

$$\Psi : U \to \operatorname{Im} \pi_W,$$
$$(t; (X^{(k)})_{k=1}^n) \mapsto \left(\left(\sum_{i=1}^{n} t_i^k \right)_{k=1}^{n} ; (\Gamma^{(k)})_{k=1}^n = D(t)(X^{(k)})_{k=1}^n \right). \tag{2-8}$$

The map Ψ is generically an $n!$-fold covering map, and it is invariant under S_n which acts on U by the diagonal coordinate permutation on both factors.[1]

Since Ψ is a dominant map with generically finite fibers, we conclude that

$$\dim \operatorname{Im} \pi_W = \dim U = n + n(p + q - 1) = n(p + q),$$

where we used $\dim \operatorname{Det}_1(\mathrm{M}_{q,p}) = p + q - 1$. Comparing the dimension, we know the dimension of a generic fiber of π_W is $n^2 = \dim W - \dim \operatorname{Im} \pi_W$.

Now let us assume $p = 1$ or $q = 1$. Then $\mathrm{M}_{q,p} = \mathbb{C}^q$ or $\mathbb{C}^p$, and

$$\dim(\mathbb{C}^n \oplus (\mathrm{M}_{q,p})^{\oplus n}) = n(p + q) = \dim \operatorname{Im} \pi_W$$

(the last equality follows from Theorem 2.1(1)). Since the image $\operatorname{Im} \pi_W$ is closed in $\mathbb{C}^n \oplus (\mathrm{M}_{q,p})^{\oplus n}$, we have a surjective quotient map $\pi_W : W \to \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^{\oplus n}$ so $W /\!/ G \simeq \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^{\oplus n}$, an affine space. This means the invariants are algebraically independent and $\mathbb{C}[W]^G$ is a polynomial ring. □

Corollary 2.2. *Let us denote the quotient map by* $\pi_W : W \to \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^n$ *as in* (2-3). *Assume that* $(\tau; \Gamma) = (\tau; (\Gamma^{(k)})_{k=1}^n) \in \mathbb{C}^n \oplus (\mathrm{M}_{q,p})^n$ *satisfies the following conditions*:

(i) *There exists a regular diagonal matrix* t *with* $\tau = (\tau_k(t))_{k=1}^n$, *i.e.*, $\tau \in \mathbb{C}^n$ *with the* k*-th coordinate being* $\tau_k = \sum_{i=1}^{n} t_i^k$, *where* $t_i \neq t_j$ *if* $i \neq j$.

(ii) $\Gamma^{(k)}$ $(0 \leq k \leq n-1)$ *corresponds to* $X^{(k)}$ *via* (2-6), *which are of rank* 1.

Then $(\tau; \Gamma)$ *is in the image* $\operatorname{Im} \pi_W$ *and* $\dim \pi_W^{-1}(\tau; \Gamma) = n^2$, *i.e.*, *the fiber of* $(\tau; \Gamma)$ *is generic and of dimension* n^2. *Moreover*, *it is a single closed* G*-orbit.*

[1] Unfortunately, Ψ may not be a quotient map. See Remark 2.4

Proof. By condition (i), we can choose a regular diagonal matrix t with $\tau = (\tau_k(t))_{k=1}^n$. Thus we can define $(X^{(k)}) = D(t)^{-1}\Gamma$ via (2-6). If $X^{(k)}$ is of rank 1, then we can write $X^{(k)} = c_k\, {}^t b_k$ for certain $c_k \in \mathbb{C}^q$ and $b_k \in \mathbb{C}^p$. From these vectors, we can restore ${}^t B = (b_1, \dots, b_n)$ and $C = (c_1, \dots, c_n)$. Thus

$$(\tau; \Gamma) = \pi_W(t, B, C) \in \operatorname{Im} \pi_W.$$

There is not so much choice for the fiber. We know the fiber over τ of the adjoint quotient is just the conjugation of t, which is of dimension $n^2 - n$. For B and C, since any column of B and C is nonzero, we can only multiply scalars column by column, which is of dimension n.

It is now clear that any element in the fiber can be obtained from (t, B, C) through the action of G. Since the stabilizer of the fiber (t, B, C) is trivial, we again get the right dimension n^2. □

Remark 2.3. Let us assume $p = 1$ or $q = 1$. In this case, the action of $G = \mathrm{GL}_n(\mathbb{C})$ on W is coregular, i.e., the quotient space is an affine space and the generators listed in (2-1) and (2-2) are algebraically independent.

However, if we consider an action of the simple group $\mathrm{SL}_n(\mathbb{C})$ instead of $\mathrm{GL}_n(\mathbb{C})$, this action is not coregular (coregular actions are classified for simple groups; see [Schwarz 1978; Adamovich and Golovina 1979]).

To see this, let us assume $p = q = 1$ for simplicity. Consider two invariants D_1 and D_2, with respect to the action of SL_n defined as follows. For

$$(u, v, A) \in V \oplus V^* \oplus \mathrm{M}_n$$

(we consider $V = \mathbb{C}^n$ as a column vector), we put

$$D_1(u, v, A) = \det \begin{pmatrix} v \\ vA \\ vA^2 \\ \vdots \\ vA^{n-1} \end{pmatrix}, \qquad D_2(u, v, A) = \det(u, Au, A^2u, \dots, A^{n-1}u).$$

Both D_1 and D_2 are clearly SL_n-invariants, and they are not GL_n-invariants so they cannot be expressible by using τ_k and γ^k above.[2] However, it is easy to see

$$D_1 \cdot D_2 = \det\bigl(vA^{i+j}u\bigr)_{i,j} = \det(\gamma^{i+j})_{i,j},$$

which gives a relation. This shows that the action of SL_n is not coregular.

When $p > 1$ or $q > 1$, similar arguments lead to the same conclusion.

However, even if it is not coregular, it seems the SL_n-orbit structure has good properties. We will discuss it in the future.

[2]Note that, since $p = q = 1$, we do not need subscripts i and j for $\gamma^k_{i,j}$

Remark 2.4. Let us consider a toy model for the map (2-8), as illustrated below. Assume that V is a vector space and S_n acts on $\mathbb{C}^n \times V^n$ as the diagonal coordinate permutation.

$$\begin{array}{ccc} \mathbb{C}^n \times V^n \ni (a_1, \dots, a_n; v_1, \dots, v_n) & & \\ \downarrow \psi & \searrow^{\pi} & \\ \mathbb{C}^n \times V^n \ni \Big((\sum_{i=1}^n a_i^k)_{k=1}^n ; (\sum_{i=1}^n a_i^k v_i)_{k=1}^n \Big) & \xleftarrow{\varphi} & (\mathbb{C}^n \times V^n)/S_n \end{array}$$

Consider a closed set $Z = \{(a; v) \mid a_i v_i = u \ (1 \le i \le n)\}$ for a fixed nonzero vector u, which is stable under the S_n-action. The image $\psi(Z)$ does not contain an element of the form $(0; w)$, however its closure contains $(0; (n\,u, 0, \dots, 0))$. Thus the image $\psi(Z)$ is not closed, hence ψ is not a quotient map.

Remark 2.5. Let us consider a semidirect sum $L = \mathfrak{gl}(V) \ltimes (V \oplus V^*)$ and the corresponding Lie group S. Then L admits a deformed universal enveloping algebra called "infinitesimal Cherednik algebra". The infinitesimal invariant ring $\mathbb{C}[L^*]^S$ is isomorphic to the center of the infinitesimal Cherednik algebra, which is a polynomial ring of n-variables ($n = \dim V$). Our invariant ring naturally contains it as a subalgebra if $p = q = 1$. For details, see [Tikaradze 2010; Panyushev 2007; Raïs 2009].

3. Structure of the null cone

We will study the structure of the null cone $\mathfrak{N}(W) = \pi_W^{-1}(\pi_W(0))$ in this section. For this, we follow the strategy of [Popov 2003] and [Kraft and Wallach 2006]. We briefly recall their theory.

3A. In this subsection, we consider a general situation so that the notation is independent of those in the former (sub)sections.

Let G be a connected reductive algebraic group G over $\mathbb{C}$, which acts on a vector space V linearly. Let $\pi : V \to V/\!/G$ be the quotient map, and

$$\mathcal{N}_V := \pi^{-1}(\pi(0)) = \{v \in V \mid \overline{Gv} \ni 0\}$$

be the null cone. For any one parameter subgroup (abbreviated as "1-PSG") $\lambda : \mathbb{C}^\times \to G$, we define $V(\lambda) := \{v \in V \mid \lim_{t\to 0} \lambda(t)v = 0\}$. Then $v \in V$ is in the null cone $\mathcal{N}_V$ if and only if $v \in V(\lambda)$ for a suitable 1-PSG λ (the Hilbert–Mumford criterion).

Let $T \subset G$ be a maximal torus. We fix T once and for all, and denote by $X^*(T)$ the character group of T. Then V has the weight space decomposition

$$V = \bigoplus_{\gamma \in X^*(T)} V_\gamma, \qquad V_\gamma := \{v \in V \mid tv = \gamma(t)v, \ t \in T\}.$$

We denote the set of 1-PSGs $\lambda : \mathbb{C}^\times \to T$ by $X_*(T)$. Then there is a natural pairing $\langle -, - \rangle : X_*(T) \times X^*(T) \to \mathbb{Z}$ determined as follows. For $(\lambda, \gamma) \in X_*(T) \times X^*(T)$, $m = \langle \lambda, \gamma \rangle$ if $\gamma(\lambda(t)) = t^m$ $(t \in \mathbb{C}^\times)$.

With these notations, for a 1-PSG $\lambda : \mathbb{C}^\times \to T \subset G$, we have

$$V(\lambda) = \bigoplus_{\langle \lambda, \gamma \rangle > 0} V_\gamma.$$

Since every 1-PSG of G is conjugate to a certain $\lambda \in X_*(T)$, we get

$$\mathcal{N}_V = \bigcup_{\lambda \in X_*(T)} G \cdot V(\lambda).$$

In this decomposition, there appear only finitely many different $V(\lambda) \neq 0$. Thus, a maximal $V(\lambda)$ may contribute to an irreducible component of $\mathcal{N}_V$ (but not always). We call such $U = V(\lambda)$ a maximal unstable subspace, and put

$$\mathcal{X}_U := \{\gamma \in X^*(T) \mid V_\gamma \subset U\} = \{\gamma \mid \langle \lambda, \gamma \rangle > 0\},$$

a maximal unstable subset of weights. Let $\mathcal{X}_1, \dots, \mathcal{X}_s$ be a complete set of representatives of maximal unstable subsets of weights up to the conjugation of the Weyl group $W_G(T)$, and $U_i = \bigoplus_{\gamma \in \mathcal{X}_i} V_\gamma$ $(1 \leq i \leq s)$ be the corresponding maximal unstable subspace.

For a 1-PSG λ, put

$$P(\lambda) := \{g \in G \mid \text{the limit } \lim_{t \to 0} \operatorname{Ad}(\lambda(t))\, g \text{ exists}\}.$$

Then $P(\lambda)$ is a parabolic subgroup which leaves $V(\lambda)$ stable; see Kempf [1978]. If $U = V(\lambda)$ is a maximal unstable subspace, then the stabilizer $\operatorname{Stab}_G(U)$ contains $P(\lambda)$ and hence it is a parabolic subgroup.

Define $P_i := \operatorname{Stab}_G(U_i)$ for each $1 \leq i \leq s$. Thus, we get a natural multiplication map $G \times_{P_i} U_i \to C_i \subset \mathcal{N}_V$, where $C_i = G \cdot U_i$. Since G/P_i is projective, the image C_i is closed and irreducible. Thus we can choose $C_1, \dots, C_r$ which give irreducible components of $\mathcal{N}_V$, after renumbering if necessary. In this way, we can determine the irreducible decomposition of $\mathcal{N}_V$:

$$\mathcal{N}_V = \bigcup_{k=1}^{r} C_k. \tag{3-1}$$

Let us apply this theory to our situation of the enhanced adjoint representation.

3B. Now let us return to our original notation, so $G = \mathrm{GL}_n(\mathbb{C})$ which acts on $W = \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n$ as before. It is easy to see that the set of weights of W is given by

$$\Lambda = \Lambda(W) := \{0\} \cup \Delta_n \cup \{\pm \varepsilon_i \mid 1 \leq i \leq n\},$$
$$\Delta_n = \{\varepsilon_i - \varepsilon_j \mid 1 \leq i \neq j \leq n\}.$$

Here, Δ_n denotes the set of roots of type A_{n-1} and ε_i denotes the standard basis in $\mathfrak{t}^*$, where $\mathfrak{t}$ is the Lie algebra of the diagonal torus $T \subset G$. The multiplicity of $\alpha \in \Delta_n$ is 1, while the multiplicity of $\alpha = 0$ is n; that of ε_i is p and that of $-\varepsilon_i$ is q. We describe a family of maximal unstable subsets of weights up to the Weyl group conjugation. Take a standard positive system $\Delta_n^+ = \{\varepsilon_i - \varepsilon_j \mid 1 \le i < j \le n\}$ of Δ_n.

Lemma 3.1. *For $0 \le k \le n$, put $X_k := \Delta_n^+ \cup \{\varepsilon_i \mid 1 \le i \le k\} \cup \{-\varepsilon_j \mid k < j \le n\}$. Then $X_0, X_1, \ldots, X_n$ gives a complete system of representatives of maximal unstable subset of weights up to the conjugation of the Weyl group $W_G(T) = S_n$.*

Proof. Let X be a maximal unstable subset corresponding to a 1-PSG λ. Taking conjugation of λ by S_n, we can assume $\lambda = (\lambda_1, \ldots, \lambda_n)$, with $\lambda_1 > \lambda_2 > \cdots > \lambda_n$. Note that, if an equality appears among λ_i's or one of λ_i's is equal to zero, the corresponding unstable subset is not maximal. If $\lambda_k > 0 > \lambda_{k+1}$, X is given by X_k. □

Let $U_k \subset W$ be the maximal unstable subspace corresponding to X_k, so that

$$(3\text{-}2)\quad U_k = \bigoplus_{\alpha \in X_k} W_\alpha = \big\{(\xi, \eta, v) \in \mathrm{M}_{n,p} \oplus \mathrm{M}_{q,n} \oplus \mathrm{M}_n \mid \xi_{i,j} = 0\ (i > k), \eta_{i',j'} = 0\ (j' \le k), v \in \mathfrak{n}^+\big\},$$

where $\mathfrak{n}^+$ denotes a maximal nilpotent subalgebra consisting of upper triangular matrices with 0's on the diagonal. It is the Lie algebra of the unipotent radical of a Borel subgroup B of upper triangular matrices in $G = \mathrm{GL}_n$. Note that

$$\xi = \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} \quad (\xi_1 \in \mathrm{M}_{k,p}), \qquad \text{while} \qquad \eta = (0, \eta_2) \quad (\eta_2 \in \mathrm{M}_{q,n-k}).$$

Lemma 3.2. *Let U_k $(0 \le k \le n)$ be a maximal unstable subspace as above. Then the stabilizer $P_k = \mathrm{Stab}_G(U_k)$ of U_k is the Borel subgroup B for any k and $\psi_k : G \times_B U_k \to C_k \subset \mathfrak{N}(W)$ is a resolution of singularity. In particular, C_k is an irreducible closed subvariety in $\mathfrak{N}(W)$ of dimension $(n^2 - n) + pk + q(n-k)$.*

Proof. Since P_k stabilizes $\mathfrak{n}^+$, it is contained in B. On the other hand, clearly B stabilizes U_k, hence $P_k = B$.

Let us show that a generic fiber of the map ψ_k is a one-point set. Since $C_k \supset U_k$, we will examine the fiber of $(\xi, \eta, v) \in U_k$, where $v \in \mathfrak{n}^+$ is a principal nilpotent element. Take an element $[g, (\xi', \eta', u)] \in \psi_k^{-1}((\xi, \eta, v))$. Then $(\xi, \eta, v) = \psi_k([g, (\xi', \eta', u)]) = (g\xi', \eta' g^{-1}, \mathrm{Ad}(g)u)$. In particular, $v = \mathrm{Ad}(g)u \in \mathrm{Ad}(g)\mathfrak{b} =: \mathfrak{b}^g$. It is well known that a principal element belongs to a unique Borel subalgebra. Since $v \in \mathfrak{b}$, we conclude $\mathfrak{b} = \mathfrak{b}^g$, hence $g \in B$. Now we know $[g, (\xi', \eta', u)] \sim [1_n, (\xi, \eta, v)]$, which means the element in the fiber is uniquely determined.

The set of elements $\{(\xi, \eta, v) \in C_k \mid v \text{ is principal nilpotent}\}$ is open dense in C_k, so the map ψ_k is generically one-to-one, hence it is birational. Since $G \times_B U_k$ is a vector bundle over a projective variety, the map ψ_k is proper and it is a resolution. □

Theorem 3.3. *Let $\mathfrak{N}(W)$ be the null cone, and let $C_k \subset \mathfrak{N}(W)$ $(0 \le k \le n)$ be as in Lemma 3.2.*

1. *$\mathfrak{N}(W) = \bigcup_{k=0}^{n} C_k$ gives the irreducible decomposition. So the null cone has $(n+1)$ components, the number of which is independent of $p \ge 1$ and $q \ge 1$. The dimension of $\mathfrak{N}(W)$ is $n^2 - n + n \cdot \max\{p, q\}$.*
2. *The null cone $\mathfrak{N}(W)$ is equidimensional if and only if $p = q$. In this case, the dimension of $\mathfrak{N}(W)$ is $n^2 - n + pn$.*
3. *The dimension of $\mathfrak{N}(W)$ is n^2 if and only if $p = q = 1$. If this is the case, any fiber $\pi_W^{-1}((\tau; \Gamma))$ of $(\tau; \Gamma) \in \operatorname{Im} \pi_W$ is of dimension n^2.*

Proof. From Lemma 3.2, the subvariety C_k is closed and irreducible. The general theory described in Section 3A gives the irreducible decomposition of $\mathfrak{N}(W)$ (cf. (3-1)). Since $\dim C_k = (n^2 - n) + pk + q(n-k)$,

$$\dim \mathfrak{N}(W) = \max_{0 \le k \le n} \bigl\{(n^2 - n) + pk + q(n-k)\bigr\} = n^2 - n + n \cdot \max\{p, q\}.$$

This proves (1). The claim (2) follows immediately from (1).

Let us prove (3). For any $(\tau; \Gamma) \in \operatorname{Im} \pi_W$, the dimension of the fiber $\pi_W^{-1}((\tau; \Gamma))$ is greater than or equal to that of a general fiber, which is n^2 by Theorem 2.1. On the other hand, the dimension of the null cone is the greatest among those of the fibers (see [Popov and Vinberg 1994]). □

3C. ***Orbits in the null cone.*** Let us investigate orbits in an irreducible component $C_k = G \cdot U_k \subset \mathfrak{N}(W)$ (cf. (3-2)). So pick $w = (\xi, \eta, v) \in U_k$, where $v \in \mathfrak{n}^+$ is a principal nilpotent element. We denote the G orbit through w by $\mathbb{O}(w)$.

We compute the stabilizer $Z_G(w)$ of w. Up to G conjugacy, we can assume

$$v = e := \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

By direct calculation, we get

$$Z_G(e) = \exp\Bigl(\Bigl\{\sum_{i=0}^{n-1} \sigma_i e^i \mid \sigma_i \in \mathbb{R}\Bigr\}\Bigr) \ni \sum_{j=1}^{n} x_j e^{j-1} =: g. \tag{3-3}$$

Assume that $k \ge n - k$, and denote $\xi \in \mathrm{M}_{n,p}$ and $\eta \in \mathrm{M}_{q,n}$ as

$$\xi = \begin{pmatrix} \xi_1 \\ 0 \end{pmatrix} \quad (\xi_1 \in \mathrm{M}_{k,p}) \quad \text{and} \quad \eta = (\, 0 \mid \eta_1 \,) \quad (\eta_1 \in \mathrm{M}_{q,n-k}). \tag{3-4}$$

Here we take

$$\xi_1 = (\boldsymbol{e}_k, \xi_1') \qquad (\xi_1' \in \mathrm{M}_{k,p-1}), \tag{3-5}$$

where $\boldsymbol{e}_k \in \mathbb{C}^k$ is the k-th elementary vector whose k-th coordinate is 1 and whose other coordinates are 0. Then, the element g in (3-3) stabilizes ξ and η if and only if $x_1 = 1,\ x_2 = \cdots = x_k = 0$. Thus we get

$$Z_G(w) = \Big\{ 1_n + \sum_{j=k+1}^{n} x_j e^{j-1} \Big\}.$$

In particular, we know $\operatorname{codim} \mathbb{O}(w) = n-k$. For the orbit $\mathbb{O}(w)$, we can take ξ_1' in (3-5) and η_1 in (3-4) freely, and they are uniquely determined by the orbit. So there is a fibration of orbits $\mathbb{O}(w)$ with the base space $\mathrm{M}_{k,p-1} \times \mathrm{M}_{q,n-k}$ of dimension

$$\begin{aligned} \dim \mathbb{O}(w) + \dim \mathrm{M}_{k,p-1} \times \mathrm{M}_{q,n-k} &= n^2 - (n-k) + k(p-1) + q(n-k) \\ &= n^2 - n + kp + (n-k)q = \dim C_k. \end{aligned}$$

This means the family of orbits $\{\mathbb{O}(w)\}$ makes up an open dense subset of the irreducible component C_k. Since the orbits of the largest possible dimension constitute an open set, $\dim \mathbb{O}(w) = n^2 - n + k$ is the largest among the orbits in C_k. For the family parametrized by $\mathrm{M}_{k,p-1} \times \mathrm{M}_{q,n-k}$, there is no reason to specialize the first column of ξ. So, if the k-th row of ξ does not vanish, we can follow the same arguments.

This construction also applies to the case of $k \le n-k$, if we take η instead of ξ.

Let us summarize what we have proven here.

Theorem 3.4. *Let $C_k \subset \mathfrak{N}(W)$ $(0 \le k \le n)$ be an irreducible component of the null cone $\mathfrak{N}(W)$ (see Lemma 3.2). The largest dimension of the nilpotent orbits in C_k is $n^2 - \min\{k, n-k\}$. Moreover, there exists an open dense subset of C_k which is fibered over an affine space of dimension $kp + q(n-k) - \max\{k, n-k\}$ with the fiber of isomorphic nilpotent orbits $\mathbb{O}$ of the largest dimension.*

In particular, an irreducible component C_k contains a nilpotent orbit of dimension n^2 if and only if $k = 0$ or n.

Remark 3.5. Let us consider $w = (\xi, \eta, v) \in U_k$ as above. Even if v is not principal, a G-orbit $\mathbb{O}(w)$ through w can attain the largest possible dimension in the irreducible component C_k. This seems difficult to describe when an orbit $\mathbb{O}(w)$ has the largest dimension.

Acknowledgements

We thank Minoru Itoh for useful discussions on invariants. We thank Prof. Mustapha Raïs for pointing out relevant references.

References

[Achar and Henderson 2008] P. N. Achar and A. Henderson, "Orbit closures in the enhanced nilpotent cone", *Adv. Math.* **219**:1 (2008), 27–62. Correction in **228**:5 (2011), 2984–2988. Zbl

[Achar et al. 2011] P. N. Achar, A. Henderson, and B. F. Jones, "Normality of orbit closures in the enhanced nilpotent cone", *Nagoya Math. J.* **203** (2011), 1–45. MR Zbl

[Adamovich and Golovina 1979] O. M. Adamovich and E. O. Golovina, "Simple linear Lie groups having a free algebra of invariants", *Vopr. Teor. Grupp Gomologicheskoj Algebry* **2** (1979), 3–41. In Russian; translated in *Selecta Math. Soviet.* **3**:2 (1983/84), 183–220. Zbl

[Finkelberg et al. 2009] M. Finkelberg, V. Ginzburg, and R. Travkin, "Mirabolic affine Grassmannian and character sheaves", *Selecta Math.* (*N.S.*) **14**:3-4 (2009), 607–628. MR Zbl

[Fresse and Nishiyama 2016] L. Fresse and K. Nishiyama, "On the exotic Grassmannian and its nilpotent variety", *Represent. Theory* **20** (2016), 451–481. MR Zbl

[Henderson and Trapa 2012] A. Henderson and P. E. Trapa, "The exotic Robinson–Schensted correspondence", *J. Algebra* **370** (2012), 32–45. MR Zbl

[Itoh 2013] M. Itoh, "Cayley–Hamilton type theorems and invariant theory", pp. art. id. 10, 20 pp. in *Proceedings of the Symposium on Representation Theory* (Miura, 2013), 2013. In Japanese.

[Johnson 2010] C. P. Johnson, *Enhanced nilpotent representations of a cyclic quiver*, Ph.D. thesis, University of Utah, 2010, Available at https://search.proquest.com/docview/748818453.

[Kato 2009] S. Kato, "An exotic Deligne–Langlands correspondence for symplectic groups", *Duke Math. J.* **148**:2 (2009), 305–371. MR Zbl

[Kempf 1978] G. R. Kempf, "Instability in invariant theory", *Ann. of Math.* (2) **108**:2 (1978), 299–316. MR Zbl

[Kraft and Wallach 2006] H. Kraft and N. R. Wallach, "On the nullcone of representations of reductive groups", *Pacific J. Math.* **224**:1 (2006), 119–139. MR Zbl

[Le Bruyn and Procesi 1987] L. Le Bruyn and C. Procesi, "Étale local structure of matrix invariants and concomitants", pp. 143–175 in *Algebraic groups* (Utrecht, 1986), edited by A. M. Cohen et al., Lecture Notes in Math. **1271**, Springer, 1987. MR Zbl

[Le Bruyn and Procesi 1990] L. Le Bruyn and C. Procesi, "Semisimple representations of quivers", *Trans. Amer. Math. Soc.* **317**:2 (1990), 585–598. MR Zbl

[Nishiyama 2014] K. Nishiyama, "Enhanced orbit embedding", *Comment. Math. Univ. St. Pauli* **63**:1-2 (2014), 223–232. MR Zbl arXiv

[Ohta 2008] T. Ohta, "An inclusion between sets of orbits and surjectivity of the restriction map of rings of invariants", *Hokkaido Math. J.* **37**:3 (2008), 437–454. MR Zbl

[Panyushev 2007] D. I. Panyushev, "On the coadjoint representation of $\mathbb{Z}_2$-contractions of reductive Lie algebras", *Adv. Math.* **213**:1 (2007), 380–404. MR Zbl

[Popov 2003] V. L. Popov, "The cone of Hilbert null forms", *Tr. Mat. Inst. Steklova* **241** (2003), 192–209. In Russian; translation in *Proc. Steklov Inst. Math.* **241**:2 (2003), 177–194. MR

[Popov and Vinberg 1994] V. L. Popov and E. B. Vinberg, "Invariant theory", pp. 123–278 in *Algebraic geometry, IV*, edited by A. N. Parshin and I. R. Shafarevich, Encyc. Math. Sciences **55**, Springer, 1994.

[Raïs 2009] M. Raïs, "Les invariants polynômes de la représentation coadjointe de groupes inhomogènes", preprint, 2009. arXiv

[Rosso 2012] D. Rosso, "Classic and mirabolic Robinson–Schensted–Knuth correspondence for partial flags", *Canad. J. Math.* **64**:5 (2012), 1090–1121. MR Zbl

[Schwarz 1978] G. W. Schwarz, "Representations of simple Lie groups with regular rings of invariants", *Invent. Math.* **49**:2 (1978), 167–191. MR Zbl

[Tikaradze 2010] A. Tikaradze, "Center of infinitesimal Cherednik algebras of $\mathfrak{gl}_n$", *Represent. Theory* **14** (2010), 1–8. MR Zbl

[Travkin 2009] R. Travkin, "Mirabolic Robinson–Schensted–Knuth correspondence", *Selecta Math.* (*N.S.*) **14**:3-4 (2009), 727–758. MR Zbl

[Weyl 1939] H. Weyl, *The classical groups: their invariants and representations*, Princeton Univ. Press, 1939. MR Zbl

Received March 19, 2018.

KYO NISHIYAMA
DEPARTMENT OF PHYSICS AND MATHEMATICS
AOYAMA GAKUIN UNIVERSITY
SAGAMIHARA
JAPAN
kyo@gem.aoyama.ac.jp

TAKUYA OHTA
DEPERTMENT OF MATHEMATICS
TOKYO DENKI UNIVERSITY
ADACHI-KU
JAPAN
ohta@cck.dendai.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.157

REVISITING THE SADDLE-POINT METHOD OF PERRON

CORMAC O'SULLIVAN

Perron's saddle-point method gives a way to find the complete asymptotic expansion of certain integrals that depend on a parameter going to infinity. We give two proofs of the key result. The first is a reworking of Perron's original proof, showing the clarity and simplicity that has been lost in some subsequent treatments. The second proof extends the approach of Olver which is based on Laplace's method. New results include more precise error terms and bounds for the expansion coefficients. We also treat Perron's original examples in greater detail and give a new application to the asymptotics of Sylvester waves.

1. Introduction

The main problem under consideration here is the accurate estimation of

$$\int_{\mathcal{C}} e^{N\cdot p(z)} q(z)\, dz \tag{1-1}$$

as $N \to \infty$, where p and q are holomorphic functions and integration is along a contour $\mathcal{C}$. If the contour can be moved to pass through a saddle-point of $p(z)$ so that $\operatorname{Re} p(z)$ achieves its maximum on $\mathcal{C}$ there, then the complete asymptotic expansion of (1-1) may be given quite explicitly. This was established one hundred years ago by Perron in a groundbreaking paper [1917].

Unfortunately, this paper is now difficult to obtain. There seem to be two detailed accounts of the method that are more recent. Wong [1989, Part II, Section 5] refers to *Perron's method* and gives a statement and proof based on [Wyman 1964]. These include an extra condition that does not appear in [Perron 1917]. The second account, by Olver [1974, Theorem 6.1, p. 125], refers only to the *saddle-point method* and does not include this extra condition. However it also does not include Perron's formula for the asymptotic expansion coefficients, nor give Perron's clear description of how the result is affected by the behavior of the contour $\mathcal{C}$ near the

Support for this project was provided by a PSC-CUNY award, jointly funded by The Professional Staff Congress and The City University of New York.
MSC2010: primary 41A60; secondary 11P82.
Keywords: asymptotics, saddle-point method, Sylvester waves.

saddle-point. Olver refers to [Wyman 1964] but his proof is different and more similar to Laplace's method.

To resolve these discrepancies, our first aim is to produce a clear proof of the asymptotic expansion of (1-1) based closely on Perron's original ideas. We see that the result may be stated simply and is easy to apply. We also give a second proof that extends the work of Olver mentioned above. In two innovations, the dependence of the error on $q(z)$ is made explicit, as required by our new application to the asymptotics of Sylvester waves in Section 9, and we show a bound for the expansion coefficients with Proposition 7.3.

As a simple example of the asymptotics that Perron's method produces, we see in Section 8A that

$$\int_{1/2}^{3/2} e^{N(-z+\log z)}\,dz = \frac{\sqrt{2\pi}}{N^{1/2}e^N}\left(1+\frac{1}{12N}+\frac{1}{288N^2}-\frac{139}{51840N^3}+O\left(\frac{1}{N^4}\right)\right)$$

as $N \to \infty$. Perron's original motivation was in finding the asymptotics of the integral

$$\int_{-\pi}^{\pi} \frac{e^{Ni(z-\varepsilon\sin z)}}{1-\varepsilon\cos z}\,dz, \tag{1-2}$$

which occurs in Kepler's theory when relating the true anomaly to the mean anomaly for a body orbiting in an ellipse with eccentricity ε. As described in [Burkhardt 1914], the initial terms of the asymptotic expansion of (1-2) had already been found by Jacobi, Cauchy and Debye, for example, with difficult methods. Burkhardt [1914] outlined a simpler approach and Perron was able to extend Burkhardt's ideas and make them rigorous. In [Perron 1917, Section 5] it is shown how to calculate as many terms as one wishes in the expansion of (1-2) and several related integrals. We complete these examples in Section 8 by giving explicit formulas for all their expansion coefficients.

Perron's method has many other applications, for example to the asymptotics of special functions used in pure and applied mathematics [Copson 1965; Olver 1974, Chapter 4; López et al. 2009; López and Pagola 2011], to statistics and probability [Small 2010, Chapter 7], and to results in combinatorics and number theory [de Bruijn 1958, Chapter 6; Flajolet and Sedgewick 2009, Section VIII]. The author's interest in this area began with [O'Sullivan 2015; 2016], where Perron's method was key in obtaining the asymptotics of Rademacher's coefficients and disproving Rademacher's conjecture about them. The results described in Section 9 on Sylvester waves are an extension of the work in [O'Sullivan 2016].

1A. ***Main results.*** The usual convention that the principal branch of log has arguments in $(-\pi, \pi]$ is used. As in (1-7), powers of nonzero complex numbers take the corresponding principal value $z^\tau := e^{\tau \log z}$ for $\tau \in \mathbb{C}$. This convention will be

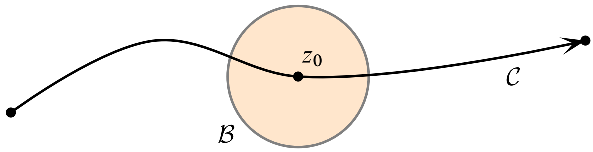


Figure 1. Neighborhood $\mathcal{B}$ and path of integration $\mathcal{C}$.

in place throughout the paper, however in some cases we will specify different branches of the power.

Our contours of integration $\mathcal{C}$ will lie in a bounded region of $\mathbb{C}$ and be parametrized by a continuous function $c : [0, 1] \to \mathbb{C}$ that has a continuous derivative except at a finite number of points. For any appropriate f, integration along the corresponding contour $\mathcal{C}$ is defined as $\int_{\mathcal{C}} f(z)\,dz := \int_0^1 f(c(t))c'(t)\,dt$ in the normal way.

The notation $f(z) = O(g(z))$, or equivalently $f(z) \ll g(z)$, means that there exists a C such that $|f(z)| \leqslant C \cdot g(z)$ for all z in a specified range. The number C is called the *implied constant*.

In our main results we make the following assumptions and definitions.

Assumptions 1.1. We have $\mathcal{B}$ a neighborhood of $z_0 \in \mathbb{C}$. Let $\mathcal{C}$ be a contour as described above, with z_0 a point on it (see Figure 1 for an example). Suppose $p(z)$ and $q(z)$ are holomorphic functions on a domain containing $\mathcal{B} \cup \mathcal{C}$. We assume $p(z)$ is not constant and hence there must exist $\mu \in \mathbb{Z}_{\geqslant 1}$ and $p_0 \in \mathbb{C}_{\neq 0}$ so that

$$p(z) = p(z_0) - p_0(z - z_0)^{\mu}(1 - \phi(z)) \qquad (z \in \mathcal{B}) \tag{1-3}$$

with ϕ holomorphic on $\mathcal{B}$ and $\phi(z_0) = 0$. Let $\omega_0 := \arg(p_0)$ and we will need the *steepest-descent angles*

$$\theta_\ell := -\frac{\omega_0}{\mu} + \frac{2\pi\ell}{\mu} \qquad (\ell \in \mathbb{Z}). \tag{1-4}$$

For later results we require $a \in \mathbb{C}$. We also assume that $\mathcal{B},\ \mathcal{C},\ p(z),\ q(z),\ z_0$ and a are independent of $N > 0$. Finally, let K_q be a bound for $|q(z)|$ on $\mathcal{B} \cup \mathcal{C}$.

The following is a slight restatement of Perron's key result in [Perron 1917, p. 202]. It may be compared with [Wong 1989, Theorem 4, p. 105] and [Olver 1974, Theorem 6.1, p. 125].

Theorem 1.2 (Perron's method for a holomorphic integrand with contour starting at a maximum). *Suppose that Assumptions 1.1 hold, with $\mathcal{C}$ a contour from z_0 to z_1 in $\mathbb{C}$, where $z_0 \neq z_1$. Suppose that*

$$\operatorname{Re} p(z) < \operatorname{Re} p(z_0) \quad \textit{for all } z \in \mathcal{C},\ z \neq z_0. \tag{1-5}$$

We may choose $k \in \mathbb{Z}$ so that the initial part of $\mathcal{C}$ lies in the sector of angular width

$2\pi/\mu$ about z_0 with bisecting angle θ_k. Then for every $S \in \mathbb{Z}_{\geqslant 0}$, we have

$$\text{(1-6)} \quad \int_{\mathcal{C}} e^{N \cdot p(z)} q(z)\, dz = e^{N \cdot p(z_0)} \left(\sum_{s=0}^{S-1} \Gamma\left(\frac{s+1}{\mu}\right) \frac{\alpha_s \cdot e^{2\pi i k(s+1)/\mu}}{N^{(s+1)/\mu}} + O\left(\frac{K_q}{N^{(S+1)/\mu}}\right)\right)$$

as $N \to \infty$, where the implied constant in (1-6) is independent of N and q. The numbers α_s are given by

$$\text{(1-7)} \qquad \alpha_s = \frac{1}{\mu \cdot s!} p_0^{-(s+1)/\mu} \frac{d^s}{dz^s} \{q(z) \cdot (1 - \phi(z))^{-(s+1)/\mu}\}_{z=z_0}.$$

To understand the geometry of the condition (1-5) we first write

$$\text{(1-8)} \qquad p(z) - p(z_0) = -\sum_{s=0}^{\infty} p_s (z - z_0)^{\mu+s} \qquad (z \in \mathcal{B}).$$

By Taylor's theorem, for each S there exists $K_{p,S}$ such that

$$\text{(1-9)} \qquad \left| p(z) - p(z_0) + \sum_{s=0}^{S-1} p_s (z - z_0)^{\mu+s} \right| \leqslant K_{p,S} |z - z_0|^{\mu+S}$$

for all $z \in \mathcal{B}$. Write

$$\text{(1-10)} \qquad p_s = |p_s| e^{i\omega_s} \quad \text{and} \quad z = z_0 + r \cdot e^{i\theta}$$

so that we obtain

$$\text{(1-11)} \qquad \operatorname{Re}(p(z) - p(z_0)) = -r^\mu \sum_{s=0}^{\infty} |p_s| r^s \cos(\omega_s + (\mu+s)\theta).$$

Then (1-9) and (1-11) imply that $\operatorname{Re}(p(z) - p(z_0)) \approx -r^\mu |p_0| \cos(\omega_0 + \mu\theta)$ for small r. Thus, in a small neighborhood of z_0, the regions where $\operatorname{Re}(p(z) - p(z_0)) < 0$ correspond approximately to μ sectors of angular width π/μ. These "valleys" alternate with μ "hill" sectors, of the same size, where $\operatorname{Re}(p(z) - p(z_0)) > 0$. The exact boundaries where $\operatorname{Re}(p(z) - p(z_0)) = 0$ will be differentiable curves, as we see in Section 2. See Figure 2 for an example with $\mu = 3$. In Proposition 2.1 we show it is possible to choose $R_p > 0$ and small enough so that these boundary curves behave nicely in the disk of radius R_p about z_0, approximating 2μ regularly spaced spokes in a wheel.

The bisecting lines of the valley sectors are clearly given by $z_0 + re^{i\theta}$ for $r \geqslant 0$ and θ satisfying $\cos(\omega_0 + \mu\theta) = 1$. These bisecting angles are the θ_ℓ defined in (1-4) and correspond to the directions of greatest decrease (steepest descent) of $\operatorname{Re}(p(z) - p(z_0))$.

The condition (1-5) means that the initial part of $\mathcal{C}$ must lie in one of the valley regions. To specify which one, we use the fact that the part of this region within

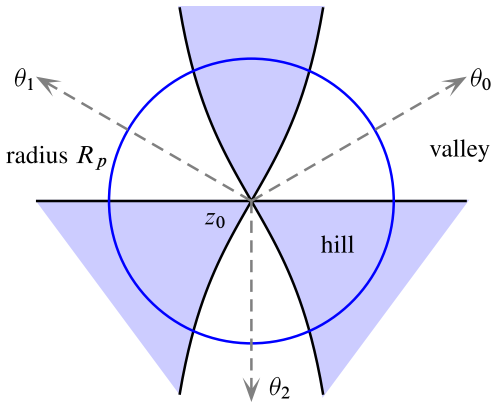


Figure 2. Hills and valleys near $z_0 = 0$ for $p(z) = i(z - \sin z)$.

a distance R_p from z_0 must lie inside the sector of angular width $2\pi/\mu$ about z_0 with bisecting angle θ_k for some $k \in \mathbb{Z}$. For the details of this see Section 2.

The proofs of Theorem 1.2 we give in Sections 3 and 4 rely on the important simplification of Perron stated next and proved in Section 2.

Proposition 1.3. *Suppose all the assumptions of Theorem 1.2 are true. Let b be the point on the bisecting line with angle θ_k that is a distance R_p from z_0. Then there exists $\varepsilon > 0$ so that*

$$\int_{\mathcal{C}} e^{N\cdot p(z)} q(z)\,dz = e^{N\cdot p(z_0)}\left(\int_{z_0}^{b} e^{N(p(z)-p(z_0))} q(z)\,dz + O(K_q e^{-\varepsilon N})\right) \tag{1-12}$$

as $N \to \infty$, where ε and the implied constant in (1-12) are independent of N and q.

The point b is shown in Figure 3. It is clear from Proposition 1.3 that most details of the contour $\mathcal{C}$ are irrelevant for our asymptotic results; we only need to know which sector the contour starts off in.

As a simple corollary to Theorem 1.2, the next result is obtained by breaking the contour of integration into $\int_{\mathcal{C}} = \int_{z_0}^{z_2} - \int_{z_0}^{z_1}$. This may also be compared with Theorem 1 of [López et al. 2009].

Corollary 1.4 (Perron's method for a holomorphic integrand with contour passing through a maximum). *Suppose Assumptions 1.1 hold. Let $\mathcal{C}$ be a contour starting at z_1, passing through z_0 and ending at z_2, with these three points all distinct. Suppose that*

$$\operatorname{Re} p(z) < \operatorname{Re} p(z_0) \quad \textit{for all } z \in \mathcal{C},\ z \neq z_0. \tag{1-13}$$

Let $\mathcal{C}$ approach z_0 in the sector of angular width $2\pi/\mu$ about z_0 with bisecting angle θ_{k_1} and leave z_0 in a sector of the same size with bisecting angle θ_{k_2}. Then for *every $S \in \mathbb{Z}_{\geqslant 0}$, we have*

$$\int_{\mathcal{C}} e^{N\cdot p(z)} q(z)\, dz = e^{N\cdot p(z_0)}\left(\sum_{s=0}^{S-1} \Gamma\left(\frac{s+1}{\mu}\right) \frac{\alpha_s (e^{2\pi i k_2 (s+1)/\mu} - e^{2\pi i k_1 (s+1)/\mu})}{N^{(s+1)/\mu}} + O\left(\frac{K_q}{N^{(S+1)/\mu}}\right)\right) \tag{1-14}$$

as $N \to \infty$, where the implied constant is independent of N and q. The numbers α_s are given by (1-7).

We will see generalizations of these results in Section 6. In Section 7, more explicit formulas for the numbers α_s are given.

Prior to [Burkhardt 1914] and [Perron 1917], different techniques to estimate integrals by moving the path of integration to a saddle-point were pioneered by Cauchy, Stokes, Riemann, Nekrasov, Kelvin and Debye. See, for example, [Olver 1970; 1974, pp. 104–105; Petrova and Solov'ev 1997; Temme 2013] where their contributions are described. These techniques include the method of steepest descent, and an advantage of Corollary 1.4 is that it does not require computing steepest descent paths.

1B. ***Burkhardt's heuristic.*** Before proving the above results, we give Burkhardt's heuristic and show how the form of (1-14) arises. Suppose $p'(z_0) = 0$ and $p''(z_0) < 0$. For simplicity we take $\mathcal{C} = [-1, 1]$ and $z_0 = 0$. Expanding $p(z)$ as in (1-8) with $p(z) = p(0) - p_0 z^2 - p_1 z^3 - \cdots$ and $q(z) = q_0 + q_1 z + \cdots$ yields

$$\int_{-1}^{1} e^{N\cdot p(z)} q(z)\, dz = e^{N\cdot p(0)} \int_{-1}^{1} e^{-N p_0 z^2} e^{-N z^2 (p_1 z + p_2 z^2 + \cdots)} (q_0 + q_1 z + \cdots)\, dz,$$

where we may write

$$e^{-N z^2 (p_1 z + p_2 z^2 + \cdots)} = 1 - N z^2 (p_1 z + p_2 z^2 + \cdots) + \frac{(N z^2)^2}{2} (p_1 z + p_2 z^2 + \cdots)^2 + \cdots .$$

Since $p_0 > 0$ and $N > 0$, the term $e^{-N p_0 z^2}$ will have exponential decay and so extending the path of integration to $\mathbb{R}$ should not affect the result. Let $w = N p_0 z^2$ to obtain

$$e^{N\cdot p(0)} \int_{-\infty}^{\infty} e^{-w} \left(1 - \frac{w}{p_0}(p_1 z + p_2 z^2 + \cdots) + \frac{w^2}{2 p_0^2}(p_1 z + p_2 z^2 + \cdots)^2 + \cdots\right) \times (q_0 + q_1 z + \cdots)\, dz. \tag{1-15}$$

By symmetry, the contributions from the odd powers of z will cancel. From the z^0 term of (1-15) we get the first term of the asymptotic expansion:

$$2 e^{N\cdot p(0)} \int_{0}^{\infty} e^{-w} q_0 \frac{dw}{2 (N p_0)^{1/2} w^{1/2}} = e^{N\cdot p(0)} \frac{\Gamma(1/2) q_0}{(N p_0)^{1/2}}. \tag{1-16}$$

From the z^2 term of (1-15) we get the next term of the expansion:

$$2e^{N\cdot p(0)}\int_0^\infty e^{-w}\left(q_2-\frac{w(p_1q_1+p_2q_0)}{p_0}+\frac{w^2p_1^2q_0}{2p_0^2}\right)z^2\frac{dw}{2(Np_0)^{1/2}w^{1/2}} \tag{1-17}$$
$$=\frac{e^{N\cdot p(0)}}{(Np_0)^{3/2}}\int_0^\infty e^{-w}\left(wq_2-\frac{w^2(p_1q_1+p_2q_0)}{p_0}+\frac{w^3p_1^2q_0}{2p_0^2}\right)\frac{dw}{w^{1/2}}$$
$$=\frac{e^{N\cdot p(0)}}{(Np_0)^{3/2}}\left(\Gamma(3/2)q_2-\frac{\Gamma(5/2)(p_1q_1+p_2q_0)}{p_0}+\frac{\Gamma(7/2)p_1^2q_0}{2p_0^2}\right).$$

The formulas (1-16) and (1-17) will reappear in Section 7.

2. Preliminary results

This section is an elaboration of the paragraph in [Perron 1917] before equation (11) and gives a detailed description of $p(z)$ for z near z_0.

Proposition 2.1. *Suppose $p(z)$ is holomorphic in a neighborhood $\mathcal{B}$ of z_0. As in Assumptions 1.1, we assume $p(z)$ is not constant and hence there must exist $\mu\in\mathbb{Z}_{\geqslant 1}$ and $p_0\in\mathbb{C}_{\neq 0}$ so that*

$$p(z)=p(z_0)-p_0(z-z_0)^\mu(1-\phi(z))\qquad(z\in\mathcal{B}) \tag{2-1}$$

with ϕ holomorphic on $\mathcal{B}$ and $\phi(z_0)=0$. Then there exists $R_p>0$ so that the closed disk centered at z_0 of radius R_p is contained in $\mathcal{B}$ and we have the following additional properties.

(i) *All solutions to $\operatorname{Re}(p(z_0+re^{i\theta})-p(z_0))/r^\mu=0$ for $r\in[0,R_p]$ have the form $(r,\theta)=(r,f_\ell(r))$ for functions $f_\ell(r)$ with $\ell\in\mathbb{Z}$.*

(ii) *These functions $f_\ell(r)$ are all defined on an interval containing $[0,R_p]$ and are differentiable.*

(iii) *We have*

$$f_\ell(0)=\delta_\ell\quad\text{for }\ \delta_\ell:=-\frac{\omega_0}{\mu}+\frac{\pi(\ell+1/2)}{\mu}. \tag{2-2}$$

(iv) *Also $|f_\ell(r)-\delta_\ell|\leqslant\pi/(4\mu)$ for $r\in[0,R_p]$.*

Proof. Set $H(r,\theta):=-\operatorname{Re}\big(p(z_0+re^{i\theta})-p(z_0)\big)/r^\mu$. By (1-11)

$$H(r,\theta)=\sum_{s=0}^\infty|p_s|r^s\cos(\omega_s+(\mu+s)\theta)$$

and so $H(0,\theta)=|p_0|\cos(\omega_0+\mu\theta)$. Then the solutions to $H(0,\theta)=0$ are $\theta=\delta_\ell$ for $\ell\in\mathbb{Z}$ with δ_ℓ defined in (2-2).

For (r,θ) in a neighborhood of $(0,\delta_\ell)$ the partial derivatives of $H(r,\theta)$ exist and are continuous. Also

$$\frac{\partial H}{\partial \theta}\bigg|_{(r,\theta)=(0,\delta_\ell)} = -|p_0|\mu\sin(\omega_0+\mu\delta_\ell) = (-1)^{\ell+1}|p_0|\mu \neq 0.$$

Therefore, by the implicit function theorem, all the solutions to $H(r,\theta)=0$ for (r,θ) in some neighborhood of $(0,\delta_\ell)$ take the form $(r,\theta)=(r, f_\ell(r))$ for differentiable functions f_ℓ. Note that $H(r,\theta+2\pi)=H(r,\theta)$ so that, for all $\ell\in\mathbb{Z}$,

$$f_{\ell+2\mu}(r) = f_\ell(r)+2\pi. \tag{2-3}$$

We choose $R_p>0$ small enough so that the interval $[0, R_p]$ is contained in the above neighborhoods for all $\ell\in\mathbb{Z}$. By (2-3), this choice involves only 2μ conditions. We have proved parts (i), (ii) and (iii).

Suppose $\varepsilon>0$ is given. Since $f_\ell(r)$ is continuous at $r=0$ we may decrease R_p again, if necessary, to ensure that $|f_\ell(r)-f_\ell(0)|\leqslant\varepsilon$ for $r\in[0,R_p]$. We do this for each $\ell \bmod 2\mu$ and with $\varepsilon=\pi/(4\mu)$. This proves part (iv). □

Corollary 2.2. *Suppose all the assumptions of Proposition 2.1 hold. Then*

$$f_{2\ell-1}(r) < \theta_\ell < f_{2\ell}(r) \quad \textit{for all } r\in[0,R_p],\ \ell\in\mathbb{Z}. \tag{2-4}$$

Also

$$\mathrm{Re}(p(z_0+re^{i\theta_\ell})-p(z_0)) < 0 \quad \textit{for all } r\in(0,R_p],\ \ell\in\mathbb{Z}. \tag{2-5}$$

Inequalities (2-4) and (2-5) are special cases of the following. For every $r\in(0,R_p]$ we have

$$\mathrm{Re}(p(z_0+re^{i\theta})-p(z_0)) < 0 \tag{2-6}$$

if and only if θ satisfies $f_{2\ell-1}(r)<\theta<f_{2\ell}(r)$ for some $\ell\in\mathbb{Z}$.

Proof. By Proposition 2.1 part (iii), we have

$$f_{2\ell-1}(0)+\frac{\pi}{2\mu} = \theta_\ell = f_{2\ell}(0)-\frac{\pi}{2\mu}.$$

Hence, with part (iv), it is clear that (2-4) holds. Therefore $\mathrm{Re}(p(z_0+re^{i\theta_\ell})-p(z_0))$ does not change sign for $r\in(0,R_p]$. Since

$$\mathrm{Re}(p(z_0+re^{i\theta_\ell})-p(z_0)) \approx -r^\mu|p_0|\cos(\omega_0+\mu\theta_\ell) = -r^\mu|p_0|<0$$

for small r we obtain (2-5). Similarly, along the directions of steepest ascent,

$$\mathrm{Re}(p(z_0+re^{i(\theta_\ell+\pi/\mu)})-p(z_0)) > 0 \quad \text{for all } r\in(0,R_p],\ \ell\in\mathbb{Z}. \tag{2-7}$$

For fixed $r\in(0,R_p]$, consider $\mathrm{Re}(p(z_0+re^{i\theta})-p(z_0))$ as a continuous function of θ with zeros only at $\theta=f_\ell(r)$ for $\ell\in\mathbb{Z}$. Thus $\mathrm{Re}(p(z_0+re^{i\theta})-p(z_0))$ is always

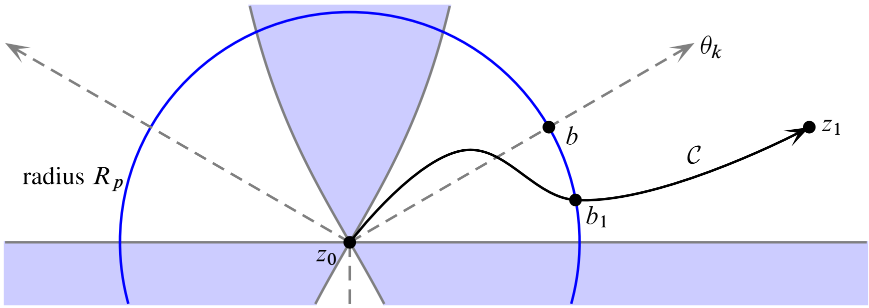


Figure 3. Replacing $\mathcal{C}$ by the line from z_0 to b.

positive or always negative for $f_{2\ell-1}(r) < \theta < f_{2\ell}(r)$. By (2-4) and (2-5) it must be negative. Similarly, with (2-7), it must be positive for $f_{2\ell}(r) < \theta < f_{2\ell+1}(r)$. □

Proof of Proposition 1.3. If the contour $\mathcal{C}$ is not contained in the disk of radius R_p about z_0 then let b_1 be the first point of $\mathcal{C}$ that is a distance R_p from z_0, as shown in Figure 3. Let $\mathcal{C}'$ be the contour from b to z_1 that follows the circular arc about z_0 from b to b_1. From b_1 the contour now follows $\mathcal{C}$ to z_1. (If $\mathcal{C}$ is contained in the disk of radius R_p about z_0 then $\mathcal{C}'$ could move from b to a point b_0 on the line between z_0 and b that is the same distance as z_1 from z_0. It then follows the circular arc about z_0 from b_0 to z_1.)

Since the integrand is holomorphic, Cauchy's theorem tells us that

$$\int_{\mathcal{C}} e^{N\cdot p(z)} q(z)\, dz = \left(\int_{z_0}^{b} + \int_{\mathcal{C}'}\right) e^{N\cdot p(z)} q(z)\, dz.$$

It is clear from Corollary 2.2 and (1-5) that $\operatorname{Re}(p(z) - p(z_0)) < 0$ for $z \in \mathcal{C}'$. Hence there exists $\varepsilon > 0$, depending only on $\mathcal{C}$, $p(z)$ and R_p, such that $\operatorname{Re}(p(z)-p(z_0)) \leqslant -\varepsilon$ for all $z \in \mathcal{C}'$. Therefore

$$\left|\int_{\mathcal{C}'} e^{N(p(z)-p(z_0))} q(z)\, dz\right| \leqslant e^{-\varepsilon N} \int_{\mathcal{C}'} |q(z)|\, |dz| \leqslant K_q |\mathcal{C}'| e^{-\varepsilon N}, \tag{2-8}$$

where $|\mathcal{C}'|$ is the length of $\mathcal{C}'$ which is less than $R_p + R_p(\pi/\mu) + |\mathcal{C}|$. This completes the proof of Proposition 1.3. □

Therefore Perron shows us that in finding the asymptotic expansion of (1-1), we may replace $\mathcal{C}$ by the line from z_0 to b as shown in Figure 3. This important step is emphasized in [López et al. 2009]. Theorem 4 on page 105 of [Wong 1989] (based on the corresponding result of [Wyman 1964]) is similar to Theorem 1.2 but has the extra condition that there exists $\delta > 0$ so that $|\arg(p(z_0) - p(z))| \leqslant \pi/2 - \delta$ for all $z \in \mathcal{C}$. This condition seems to be caused by missing the step of Proposition 1.3. Olver [1970] also comments that this condition is unnecessary. (There are two further unnecessary conditions in [Wong 1989]: that the initial part of $\mathcal{C}$ may be deformed into a straight line and that the path $\mathcal{C}$ leaves z_0 at a well-defined angle.)

3. First proof of Theorem 1.2

This proof of Theorem 1.2 is based closely on the original in [Perron 1917] though including more detail. We follow Wyman [1964] in bounding $P_s(w)$ in Lemma 3.1 using Cauchy's inequality. We also depart from Perron by bounding $Q_s(z)$ in Lemma 3.2 using the integral form of the remainder from Taylor's theorem.

Proof of Theorem 1.2. Let

$$\mathcal{D} := \{z \in \mathbb{C} \,:\, |z - z_0| \leqslant \rho\}$$

for $\rho = R_p$ initially. Since $\phi(z_0) = 0$, there exists $K_\phi > 0$ such that

$$|\phi(z)| \leqslant K_\phi |z - z_0| \quad \text{for all } z \in \mathcal{D}. \tag{3-1}$$

Looking ahead to Lemma 3.2, we decrease ρ, if necessary, to ensure that

$$0 < \rho \leqslant 1/(2K_\phi). \tag{3-2}$$

By Proposition 1.3 we only need to estimate the integral

$$\int_{z_0}^{b} e^{N(p(z)-p(z_0))} q(z)\, dz,$$

where b is on the bisecting line with angle θ_k and a distance R_p from z_0. It is convenient to change the end point to b', on the same bisecting line and a distance $\rho/2$ from z_0. See Figure 4. By (2-5) there exists $\varepsilon' > 0$ such that $\operatorname{Re}(p(z) - p(z_0)) \leqslant -\varepsilon'$ for z on the line between b' and b. Hence

$$\int_{z_0}^{b'} e^{N(p(z)-p(z_0))} q(z)\, dz = \int_{z_0}^{b} e^{N(p(z)-p(z_0))} q(z)\, dz + O(K_q e^{-\varepsilon' N}). \tag{3-3}$$

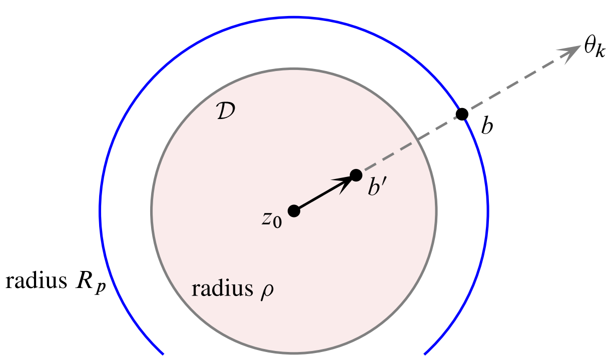


Figure 4. The line from z_0 to b' is the new path of integration.

For any $w \in \mathbb{C}$ we have the Taylor expansion

$$q(z)e^{w\phi(z)} = \sum_{s=0}^{\infty} P_s(w)(z-z_0)^s \qquad (z \in \mathcal{D}).$$

Since

$$q(z)e^{w\phi(z)} = q(z)(1 + w\phi(z) + w^2\phi(z)^2/2! + \cdots)$$

and $\phi(z_0) = 0$, it follows that $P_s(w)$ is a polynomial and

$$P_s(w) = \sum_{\ell=0}^{s} c_{s,\ell} \cdot w^\ell, \tag{3-4}$$

where $c_{s,\ell}$ is the coefficient of $(z-z_0)^s$ in the Taylor expansion of $q(z)\phi(z)^\ell/\ell!$ about z_0. The following bound for $P_s(w)$ will be needed for the proof of Proposition 3.5.

Lemma 3.1. *For all $w \in \mathbb{C}$,*

$$|P_s(w)| \leqslant K_q e^{K_\phi}(|w|^s + \rho^{-s}).$$

Proof. Starting with Cauchy's inequality [Ahlfors 1978, p. 120], we find that for every r with $0 < r \leqslant \rho$,

$$\begin{aligned} |P_s(w)| &\leqslant r^{-s} \max_{|z-z_0|=r} |q(z)e^{w\phi(z)}| \\ &\leqslant K_q r^{-s} \max_{|z-z_0|=r} e^{\mathrm{Re}(w\phi(z))} \\ &\leqslant K_q r^{-s} e^{K_\phi |w| r}. \end{aligned} \tag{3-5}$$

If $|w| \leqslant 1/\rho$ then letting $r = \rho$ in (3-5) shows $|P_s(w)| \leqslant K_q e^{K_\phi}\rho^{-s}$. If $|w| \geqslant 1/\rho$ then letting $r = 1/|w|$ in (3-5) shows $|P_s(w)| \leqslant K_q e^{K_\phi}|w|^s$. □

Now we take

$$w = N p_0 (z-z_0)^\mu. \tag{3-6}$$

It is an easy exercise to check that $w \geqslant 0$ when z is on the line between z_0 and b'. For these z values,

$$w^{1/\mu} = (N|p_0|)^{1/\mu}|z-z_0|. \tag{3-7}$$

Lemma 3.2. *With w given by (3-6), and z on the line between z_0 and b', we have*

$$e^{N(p(z)-p(z_0))}q(z) = \sum_{s=0}^{S-1} e^{-w} P_s(w)(z-z_0)^s + Q_S(z), \tag{3-8}$$

where

$$|Q_S(z)| \leqslant \frac{2K_q}{\rho^S}|z-z_0|^S e^{-w/2}. \tag{3-9}$$

Proof. By Taylor's theorem, see [Ahlfors 1978, pp. 125–126],

$$q(z)e^{w\phi(z)} = \sum_{s=0}^{S-1} P_s(w)(z-z_0)^s + \frac{(z-z_0)^S}{2\pi i}\int_\gamma \frac{q(\tau)e^{w\phi(\tau)}}{(\tau-z_0)^S(\tau-z)}\,d\tau,$$

where γ is the positively oriented circle of radius ρ about z_0. For $\tau \in \gamma$ we have

$$|q(\tau)e^{w\phi(\tau)}| \leqslant K_q e^{K_\phi w\rho}.$$

Also $|\tau - z| \geqslant \rho/2$ since $|z-z_0| \leqslant \rho/2$ by our choice of b'. The identity

$$e^{N(p(z)-p(z_0))}q(z) = e^{-w}\cdot q(z)e^{w\phi(z)}$$

proves (3-8) with

$$|Q_S(z)| \leqslant \frac{2K_q}{\rho^S}|z-z_0|^S e^{-w+K_\phi w\rho}.$$

The inequality (3-2) implies $\exp(-w+K_\phi w\rho) \leqslant \exp(-w/2)$ and we get (3-9). □

With Proposition 1.3, (3-3) and Lemma 3.2 we may write

$$\int_C e^{N\cdot p(z)}q(z)\,dz = e^{N\cdot p(z_0)}\biggl(\sum_{s=0}^{S-1} I_s(N) + \int_{z_0}^{b'} Q_S(z)\,dz + O(K_q e^{-\varepsilon N})\biggr) \tag{3-10}$$

for

$$I_s(N) := \int_{z_0}^{b'} e^{-w}P_s(w)(z-z_0)^s\,dz \tag{3-11}$$

(with w given by (3-6)) and where $\varepsilon > 0$ is independent of N and q.

Lemma 3.3. *We have*

$$\int_{z_0}^{b'} Q_S(z)\,dz = O\biggl(\frac{K_q}{N^{(S+1)/\mu}}\biggr).$$

Proof. The absolute value of the left side is

$$\begin{aligned}\left|\int_0^{\rho/2} Q_S(te^{i\theta_k}+z_0)e^{i\theta_k}\,dt\right| &\leqslant \int_0^{\rho/2} |Q_S(te^{i\theta_k}+z_0)|\,dt \\ &\leqslant \frac{2K_q}{\rho^S}\int_0^{\rho/2}\exp(-N|p_0|t^\mu/2)\cdot t^S\,dt.\end{aligned} \tag{3-12}$$

We used inequality (3-9) in (3-12) and that $w = N|p_0|t^\mu$ when $z = te^{i\theta_k}+z_0$. With the change of variables $u = N|p_0|t^\mu/2$ and extending the range of integration to ∞ we obtain

$$\left|\int_{z_0}^{b'} Q_S(z)\,dz\right| \leqslant \left[\frac{2\Gamma((S+1)/\mu)}{\mu\cdot\rho^S(|p_0|/2)^{(S+1)/\mu}}\right]\frac{K_q}{N^{(S+1)/\mu}}.$$ □

Combining the errors from (3-10) and Lemma 3.3 shows

$$\int_{\mathcal{C}} e^{N\cdot p(z)} q(z)\,dz = e^{N\cdot p(z_0)}\left(\sum_{s=0}^{S-1} I_s(N) + O\left(\frac{K_q}{N^{(S+1)/\mu}}\right)\right) \tag{3-13}$$

for an implied constant independent of N and q.

Lemma 3.4. *We have*

$$I_s(N) = \frac{e^{2\pi i k(s+1)/\mu}}{\mu\cdot(Np_0)^{(s+1)/\mu}} \int_0^{N|p_0|(\rho/2)^\mu} e^{-w} P_s(w) w^{(s+1)/\mu-1}\,dw.$$

Proof. Recall (3-6). First we claim that

$$z - z_0 = w^{1/\mu}(Np_0)^{-1/\mu} e^{2\pi i k/\mu} \tag{3-14}$$

for z on the line between z_0 and b'. This follows from the definitions

$$p_0 = |p_0| e^{i\omega_0}, \qquad \theta_k = -\frac{\omega_0}{\mu} + \frac{2\pi k}{\mu}, \qquad z - z_0 = |z - z_0| e^{i\theta_k}$$

and the relation (3-7). The proof is completed by using (3-14) in (3-11) to change the variable of integration to w. □

Proposition 3.5. *There exists $\varepsilon'' > 0$ so that*

$$\int_0^{N|p_0|(\rho/2)^\mu} e^{-w} P_s(w) w^{(s+1)/\mu-1}\,dw = \frac{\Gamma((s+1)/\mu)}{s!}\frac{d^s}{dz^s}\{q(z)(1-\phi(z))^{-(s+1)/\mu)}\}_{z=z_0} + O(K_q e^{-\varepsilon'' N}). \tag{3-15}$$

Proof. Put $T := N|p_0|(\rho/2)^\mu$ and write the integral in (3-15) as $\int_0^T = \int_0^\infty - \int_T^\infty$. Employing Lemma 3.1, we find

$$\begin{aligned}\left|\int_T^\infty e^{-w} P_s(w) w^{(s+1)/\mu-1}\,dw\right| &\leqslant K_q e^{K_\phi} \int_T^\infty e^{-w}(w^s + \rho^{-s}) w^{(s+1)/\mu-1}\,dw \\ &= K_q e^{K_\phi}\left(\int_T^\infty e^{-w} w^{d-1}\,dw + \rho^{-s}\int_T^\infty e^{-w} w^{d'-1}\,dw\right)\end{aligned} \tag{3-16}$$

for $d := (s+1)/\mu + s$ and $d' := (s+1)/\mu$. The estimate

$$\int_T^\infty e^{-w} w^{d-1}\,dw \leqslant 2^d \Gamma(d) e^{-T/2} \qquad (T \geqslant 0, d > 0) \tag{3-17}$$

follows from bounding e^{-w} in the integrand by $e^{-T/2}e^{-w/2}$. (More accurate estimates of the incomplete gamma function are possible; see for example [Olver 1974, Equation (2.02), p. 110].) Hence (3-16) is bounded by

$$K_q e^{K_\phi}(2^d\Gamma(d)+\rho^{-s}2^{d'}\Gamma(d'))e^{-T/2}.$$

We have shown that

$$\left|\int_T^\infty\right| = O(K_q e^{-\varepsilon'' N})$$

for $\varepsilon'' = |p_0|(\rho/2)^\mu/2$ and an implied constant independent of N and q.

Lastly, we calculate $\int_0^\infty$. Recalling (3-4),

$$\int_0^\infty e^{-w}P_s(w)w^{(s+1)/\mu-1}\,dw = \sum_{\ell=0}^{s} c_{s,\ell}\Gamma\left(\frac{s+1}{\mu}+\ell\right), \tag{3-18}$$

where $c_{s,\ell}$ is the coefficient of $(z-z_0)^s$ in the Taylor expansion of $q(z)\phi(z)^\ell/\ell!$. Therefore (3-18) is the coefficient of $(z-z_0)^s$ in

$$\begin{aligned} q(z)\sum_{\ell=0}^{s}\Gamma\left(\frac{s+1}{\mu}+\ell\right)\frac{\phi(z)^\ell}{\ell!} &= q(z)\Gamma\left(\frac{s+1}{\mu}\right)\sum_{\ell=0}^{s}\frac{\Gamma\left(\frac{s+1}{\mu}+\ell\right)}{\Gamma\left(\frac{s+1}{\mu}\right)\ell!}(-1)^\ell(-\phi(z))^\ell \\ &= q(z)\Gamma\left(\frac{s+1}{\mu}\right)\sum_{\ell=0}^{s}\binom{-(s+1)/\mu}{\ell}(-\phi(z))^\ell. \end{aligned} \tag{3-19}$$

Extending this sum to infinity will not affect the coefficient of $(z-z_0)^s$ and so we may replace (3-19) by

$$q(z)\Gamma\left(\frac{s+1}{\mu}\right)\sum_{\ell=0}^{\infty}\binom{-(s+1)/\mu}{\ell}(-\phi(z))^\ell = q(z)\Gamma\left(\frac{s+1}{\mu}\right)(1-\phi(z))^{-(s+1)/\mu}.$$

This completes the proof of Proposition 3.5. □

Theorem 1.2 now follows from (3-13), Lemma 3.4 and Proposition 3.5. □

4. Second proof of Theorem 1.2

This proof of Theorem 1.2 is based on [Olver 1970, Theorem I] or, equivalently, [Olver 1974, Theorem 6.1, p. 125]. Instead of employing the substitution $w = Np_0(z-z_0)^\mu$, Olver uses $v = p(z_0) - p(z)$ as in the usual proofs of Laplace's method (see Section 6C).

To get the result to match the statement of Theorem 1.2, we have to treat the branch factor $e^{2\pi ik/\mu}$ more explicitly than in [Olver 1974, Theorem 6.1, p. 125].

The coefficients α_s naturally appear in a power series in this proof and we use a method inspired by the application of Cauchy's differentiation formula in [Campbell et al. 1987] to obtain Perron's expression for them.

Second proof of Theorem 1.2. Let

$$\mathcal{D} := \{z \in \mathbb{C} \; : \; |z - z_0| \leqslant \rho\}$$

for $\rho = R_p$, initially. As in (3-1) and (3-2) we may decrease ρ to ensure that

$$|\phi(z)| \leqslant 1/2 \quad \text{for all } \; z \in \mathcal{D}. \tag{4-1}$$

By Proposition 1.3 we only need to estimate the integral

$$\int_{z_0}^{b} e^{N(p(z)-p(z_0))} q(z)\, dz, \tag{4-2}$$

where b is on the bisecting line with angle θ_k and a distance R_p from z_0. We will use the change of variables $v := p(z_0) - p(z)$ and, to prepare for this, set

$$\tau = \tau(z) := p_0^{1/\mu}(z - z_0)(1 - \phi(z))^{1/\mu} \tag{4-3}$$

with all roots principal. By (1-3) it is clear that τ is some μ-th root of $p(z_0) - p(z)$. We also see by (4-1) that τ is a holomorphic function of z for z in $\mathcal{D}$. We have $\frac{d\tau}{dz}\big|_{z=z_0} = p_0^{1/\mu} \neq 0$ and consequently, by the inverse function theorem for holomorphic functions, there exists a neighborhood $\mathcal{D}_\tau$ of 0 so that z is a holomorphic function of τ there:

$$z - z_0 = g(\tau) := \sum_{s=1}^{\infty} c_s \tau^s \qquad (\tau \in \mathcal{D}_\tau).$$

Choose $\mathcal{D}_\tau$ to be a disk centered at 0 and small enough that the image $\mathcal{D}_z := z_0 + g(\mathcal{D}_\tau)$ is contained in $\mathcal{D}$. See Figure 5 ($\mathcal{D}_z$ may not be a disk). From the

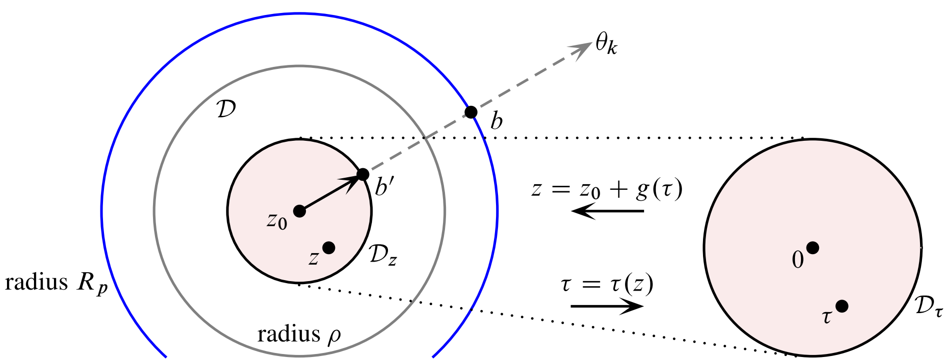


Figure 5. The regions $\mathcal{D}_z$ and $\mathcal{D}_\tau$.

equality

$$p'(z) = -\sum_{s=0}^{\infty}(s+\mu)p_s(z-z_0)^{s+\mu-1},$$

we obtain

$$-\frac{q(z)}{p'(z)} = -\frac{q(z_0+g(\tau))}{p'(z_0+g(\tau))} =: F(\tau) = \sum_{s=0}^{\infty}\beta_s \tau^{s-\mu+1} \qquad (z\in\mathcal{D}_z,\ \tau\in\mathcal{D}_\tau). \tag{4-4}$$

Shrink $\mathcal{D}_\tau$ (and correspondingly $\mathcal{D}_z$) if necessary so that $\tau^{\mu-1}F(\tau)$ is holomorphic on $\mathcal{D}_\tau$; we are avoiding any zeros of $p'(z)$ away from $z=z_0$. Taylor's theorem implies there exist constants $K_{F,S}$ such that

$$\left|\tau^{\mu-1}F(\tau) - \sum_{s=0}^{S-1}\beta_s\tau^s\right| \leqslant K_{F,S}|\tau|^S \qquad (\tau\in\mathcal{D}_\tau,\ S\in\mathbb{Z}_{\geqslant 0}). \tag{4-5}$$

To understand the dependence of $K_{F,S}$ on q we may write the remainder term explicitly as

$$\tau^{\mu-1}F(\tau) = \sum_{s=0}^{S-1}\beta_s\tau^s + \frac{\tau^S}{2\pi i}\int_{C_0}\frac{w^{\mu-1}F(w)}{w^S(w-\tau)}\,dw$$

with C_0 the boundary of $\mathcal{D}_\tau$, oriented positively. Since

$$\frac{1}{2\pi i}\int_{C_0}\frac{w^{\mu-1}F(w)}{w^S(w-\tau)}\,dw = -\frac{1}{2\pi i}\int_{C_0}\frac{q(z_0+g(w))}{p'(z_0+g(w))\cdot w^{S-\mu+1}(w-\tau)}\,dw \tag{4-6}$$

and $|q(z_0+g(w))| \leqslant K_q$ on the right of (4-6), we may write $K_{F,S} = K^*_{F,S}\cdot K_q$ with $K^*_{F,S}$ independent of q. For these estimates we have shrunk $\mathcal{D}_\tau$ (and $\mathcal{D}_z$) again, for example to half their size, so that $w-\tau$ in (4-6) is bounded away from zero for $w\in C_0$ and $\tau\in\mathcal{D}_\tau$.

Lemma 4.1. *For all $z\in\mathcal{D}_z$ with z also on the line between z_0 and b, we have*

$$\tau(z) = e^{2\pi i k/\mu}\big(p(z_0)-p(z)\big)^{1/\mu}. \tag{4-7}$$

Proof. Recall that $\arg(p_0)=\omega_0$ and $\arg(z-z_0)=\theta_k$. Hence $\arg(p_0(z-z_0)^\mu)=0$ and so

$$\begin{aligned}
\tau(z) &:= p_0^{1/\mu}(z-z_0)(1-\phi(z))^{1/\mu}\\
&= p_0^{1/\mu}(z-z_0)(p_0(z-z_0)^\mu)^{-1/\mu}(p(z_0)-p(z))^{1/\mu}\\
&= e^{i\omega_0/\mu}\cdot e^{i\theta_k}\left|p_0^{1/\mu}(z-z_0)(p_0(z-z_0)^\mu)^{-1/\mu}\right|(p(z_0)-p(z))^{1/\mu}\\
&= e^{2\pi i k/\mu}(p(z_0)-p(z))^{1/\mu}
\end{aligned}$$

as desired. □

Fix b' on the line between z_0 and b so that the segment from z_0 to b' is contained in $\mathcal{D}_z$. Hence Lemma 4.1 shows that we have

$$v := p(z_0) - p(z), \quad \tau = e^{2\pi ik/\mu} v^{1/\mu}, \quad v = \tau^\mu$$

for all z and $\tau(z)$ where z is on the line between z_0 and b'.

To estimate (4-2) we see first that $\int_{b'}^{b}$ is $O(K_q e^{-\varepsilon' N})$ as in (3-3). Using the equality $dv/dz = -p'(z)$ and (4-4) we find

$$\begin{aligned}
\text{(4-8)} \qquad \int_{z_0}^{b'} e^{N(p(z)-p(z_0))} q(z)\, dz &= \int_0^{p(z_0)-p(b')} e^{-Nv} q(z) \cdot \frac{dz}{dv}\, dv \\
&= -\int_0^{p(z_0)-p(b')} e^{-Nv} \frac{q(z)}{p'(z)}\, dv \\
&= \int_0^{p(z_0)-p(b')} e^{-Nv} F(e^{2\pi ik/\mu} v^{1/\mu})\, dv.
\end{aligned}$$

The contour of integration in (4-8) is the image of the line between z_0 and b' in the v-plane. Except for the starting point, this contour is contained in the half-plane with positive real part by (1-5). The principal root $v^{1/\mu}$ is holomorphic in this half-plane and therefore the integrand in (4-8) is holomorphic there too. Set $w := p(z_0) - p(b')$. By Cauchy's theorem we may change the contour of integration to the straight line from 0 to w. (The integrand may have a singularity at $v = 0$, but it is $\ll |v|^{1/\mu - 1}$ for $|v|$ small, and so moving the path of integration near 0 may be justified.) Employing (4-5) yields

$$\text{(4-9)} \quad \int_0^w e^{-Nv} F(e^{2\pi ik/\mu} v^{1/\mu})\, dv = \sum_{s=0}^{S-1} \beta_s e^{2\pi ik(s+1)/\mu} \int_0^w e^{-Nv} v^{(s+1)/\mu - 1}\, dv + E_S$$

with

$$\begin{aligned}
\text{(4-10)} \qquad |E_S| &\leqslant K_{F,S} \int_0^w |e^{-Nv}| |v|^{(S+1)/\mu - 1}\, |dv| \\
&= K_{F,S} \int_0^{|w|} e^{-N\,\mathrm{Re}(w/|w|)t} t^{(S+1)/\mu - 1}\, dt \\
&\leqslant K_{F,S} \Gamma\left(\frac{S+1}{\mu}\right) \left(N\,\mathrm{Re}\,\frac{w}{|w|}\right)^{-(S+1)/\mu}
\end{aligned}$$

upon extending the limit of integration to infinity. The next lemma estimates the integral in (4-9).

Lemma 4.2. *Suppose $N, r, \varepsilon > 0$ and $\mathrm{Re}\, w \geqslant \varepsilon$. For an implied constant depending only on r and w we have*

$$\int_0^w e^{-Nv} v^{r-1}\, dv = N^{-r}(\Gamma(r) + O(e^{-\varepsilon N/2})).$$

Proof. Continue the line of integration to $w\infty$ and write $\int_0^w = \int_0^{w\infty} - \int_w^{w\infty}$. The integral $\int_0^{w\infty}$ is computed by rotating the line of integration to $\mathbb{R}_{\geqslant 0}$ which is straightforward to justify:

$$\int_0^{w\infty} e^{-Nv} v^{r-1}\, dv = N^{-r}\Gamma(r).$$

The absolute value of $\int_w^{w\infty}$ is bounded by

$$\begin{aligned}|w|^r \left| \int_1^\infty e^{-Nwt} t^{r-1}\, dt \right| &\leqslant |w|^r \int_1^\infty e^{-N\varepsilon t} t^{r-1}\, dt \\ &= \frac{|w|^r}{(N\varepsilon)^r} \int_{N\varepsilon}^\infty e^{-u} u^{r-1}\, du \\ &\leqslant \left(\frac{2|w|}{\varepsilon}\right)^r \Gamma(r) N^{-r} e^{-\varepsilon N/2},\end{aligned}$$

where the last line used (3-17). □

We have shown so far, with (4-8), (4-9), (4-10) and with Lemma 4.2 applied to (4-9), that

$$\int_{z_0}^{b'} e^{N(p(z)-p(z_0))} q(z)\, dz = \sum_{s=0}^{S-1} \beta_s e^{2\pi i k(s+1)/\mu} \frac{\Gamma((s+1)/\mu)}{N^{(s+1)/\mu}} + E_S^*,$$

where

$$E_S^* \ll \frac{K_{F,S}}{N^{(S+1)/\mu}} + \left(\sum_{s=0}^{S-1} |\beta_s|\right) e^{-\varepsilon N/2} \tag{4-11}$$

for an implied constant independent of N and q. A similar argument to the one after (4-6), showing that $K_{F,S}/K_q$ may be bounded independently of q, shows that $|\beta_s|/K_q$ is also independent of q since

$$\begin{aligned}\beta_s &= \frac{1}{2\pi i} \int_{C_0} \frac{\tau^{\mu-1} F(\tau)}{\tau^{s+1}}\, d\tau \\ &= -\frac{1}{2\pi i} \int_{C_0} \frac{q(z_0+g(\tau))}{p'(z_0+g(\tau)) \cdot \tau^{s-\mu+2}}\, d\tau.\end{aligned} \tag{4-12}$$

We have already seen that integral $\int_{b'}^{b}$ has exponential decay in N, and so may be included in the error term (4-11). Consequently

$$\begin{aligned}&\int_{\mathcal{C}} e^{N\cdot p(z)} q(z)\, dz \\ &\qquad = e^{N\cdot p(z_0)} \left(\sum_{s=0}^{S-1} \Gamma\left(\frac{s+1}{\mu}\right) \frac{\beta_s \cdot e^{2\pi i k(s+1)/\mu}}{N^{(s+1)/\mu}} + O\left(\frac{K_q}{N^{(S+1)/\mu}}\right)\right)\end{aligned} \tag{4-13}$$

as desired.

It only remains to compute the numbers β_s. A change of variables in (4-12) shows

$$\beta_s = -\frac{1}{2\pi i}\int_{C_{z_0}} \frac{q(z)}{p'(z)\cdot \tau^{s-\mu+2}}\frac{d\tau}{dz}\,dz$$

for $C_{z_0} \subset \mathcal{D}_z$ a positively oriented circle centered at z_0. Use (1-3) and (4-3) to show

$$\frac{d\tau}{dz} = -\frac{1}{\mu}\tau^{1-\mu}p'(z).$$

Hence

$$\beta_s = \frac{1}{2\pi i\cdot\mu}\int_{C_{z_0}} \frac{q(z)}{\tau^{s+1}}\,dz$$

$$(4\text{-}14)\qquad = \frac{1}{2\pi i\cdot\mu}p_0^{-(s+1)/\mu}\int_{C_{z_0}} \frac{q(z)\cdot(1-\phi(z))^{-(s+1)/\mu}}{(z-z_0)^{s+1}}\,dz$$

$$(4\text{-}15)\qquad = \frac{1}{\mu\cdot s!}p_0^{-(s+1)/\mu}\frac{d^s}{dz^s}\{q(z)(1-\phi(z))^{-(s+1)/\mu)}\}_{z=z_0},$$

where (4-14) is related to (4-15) by Cauchy's differentiation formula. Thus β_s is recognized as α_s from (1-7). Combining (4-13) and (4-15) completes the second proof of Theorem 1.2. □

5. An important case

A case of Corollary 1.4 that often arises is when $\mathcal{C}$ passes through the saddle-point z_0 in a straight line or in a curve with a well-defined tangent at z_0. If μ is even then these paths will pass through opposite valley sectors, for example with $\theta_{k_2} = \theta_{k_1} \pm \pi$. In this case the terms in (1-14) with s odd vanish:

Corollary 5.1 (Perron's method for a holomorphic integrand with contour passing through a maximum between opposite sectors). *Suppose Assumptions 1.1 hold and μ is even. Let $\mathcal{C}$ be a contour beginning at z_1, passing through z_0 and ending at z_2, with these points all distinct. Suppose*

$$\operatorname{Re} p(z) < \operatorname{Re} p(z_0) \quad \textit{for all}\ z\in\mathcal{C},\ z\neq z_0.$$

Let $\mathcal{C}$ approach z_0 in a sector of angular width $2\pi/\mu$ about z_0 with bisecting angle $\theta_k + (2n+1)\pi$ for some $n \in \mathbb{Z}$, and initially leave z_0 in a sector of the same size with bisecting angle θ_k. Then for every $M \in \mathbb{Z}_{\geqslant 0}$,

$$\int_{\mathcal{C}} e^{N\cdot p(z)}q(z)\,dz = e^{N\cdot p(z_0)}\left(\sum_{m=0}^{M-1}\Gamma\left(\frac{2m+1}{\mu}\right)\frac{2\alpha_{2m}\cdot e^{2\pi i k(2m+1)/\mu}}{N^{(2m+1)/\mu}} + O\left(\frac{K_q}{N^{(2M+1)/\mu}}\right)\right)$$

as $N\to\infty$, where the implied constant is independent of N and q. The numbers α_s are given by (1-7).

Proof. Apply Corollary 1.4 with $k_2 = k$ and $k_1 = k + (2n+1)\mu/2$. Then the difference of exponentials in (1-14) is

$$e^{2\pi ik(s+1)/\mu} - e^{2\pi i(k+(2n+1)\mu/2)(s+1)/\mu} = e^{2\pi ik(s+1)/\mu}(1 - e^{2\pi i(\mu/2)(s+1)/\mu})$$
$$= e^{2\pi ik(s+1)/\mu}(1 - (-1)^{s+1})$$

and the corollary follows on writing $s = 2m$. □

The above result corresponds to [Olver 1974, Theorem 7.1, p. 127] when $\mu = 2$, giving a clearer description of how the result depends on $\mathcal{C}$ near z_0. Olver does not give the formula (1-7) for the coefficients and perhaps he was not aware of Perron's paper [1917]. It does not appear in the references of [Olver 1974], though [Burkhardt 1914] is listed. Perron's paper [1917] is not cited by the classic works [de Bruijn 1958; Copson 1965; Dingle 1973] either. It is briefly mentioned, along with [Burkhardt 1914], in Section 2.4 of Erdélyi's book [1956], though in a way which seems to imply that Perron only gives the main term of the asymptotic expansions.

6. Generalizations

6A. ***Including a factor*** $(z - z_0)^{a-1}$ ***with*** $\operatorname{Re} a > 0$. The results in [Perron 1917] cover a more general situation where we have $(z - z_0)^{a-1}q(z)$ in the integrand, instead of just $q(z)$. Unlike Perron, we do not assume that $q(z_0) \neq 0$. The number a is in $\mathbb{C}$ and so we must pay attention to which branch of $(z - z_0)^{a-1}$ is meant. For example, if z is on the bisecting line with angle θ_k (recall (1-4)) then possible branches are

$$(z - z_0)^{a-1} = |z - z_0|^{a-1} \cdot e^{i\theta_\ell(a-1)} \tag{6-1}$$

for $\ell \in \mathbb{Z}$ with $\ell \equiv k \bmod \mu$. The principal value of the power (6-1) has the unique such ℓ for which θ_ℓ is in $(-\pi, \pi]$.

The standard method for integrating a multivalued function such as (6-1) along a contour $\mathcal{C}$ is to begin with a specified branch, and as z moves along $\mathcal{C}$ the branch is determined by continuity. In particular, if $z - z_0$ crosses the negative real axis then $(z - z_0)^{a-1}$ enters another branch.

Theorem 6.1 (Perron's method for an integrand containing a factor $(z - z_0)^{a-1}$ for $\operatorname{Re} a > 0$ and with contour starting at a maximum). *Suppose Assumptions 1.1 hold. Let $\mathcal{C}$ be a contour from z_0 to z_1, with $z_0 \neq z_1$, that initially runs along the bisecting line with angle θ_k for some $k \in \mathbb{Z}$. Suppose* $\operatorname{Re} a > 0$ *and that*

$$\operatorname{Re} p(z) < \operatorname{Re} p(z_0) \quad \text{for all } z \in \mathcal{C},\ z \neq z_0. \tag{6-2}$$

On the initial part of $\mathcal{C}$ we take

$$(z - z_0)^{a-1} = |z - z_0|^{a-1} \cdot e^{i\theta_k(a-1)}. \tag{6-3}$$

Then for any $S \in \mathbb{Z}_{\geqslant 0}$,

$$\text{(6-4)} \quad \int_{\mathcal{C}} e^{N\cdot p(z)}(z-z_0)^{a-1}q(z)\,dz = e^{N\cdot p(z_0)}\left(\sum_{s=0}^{S-1}\Gamma\left(\frac{s+a}{\mu}\right)\frac{\alpha_s\cdot e^{2\pi i k(s+a)/\mu}}{N^{(s+a)/\mu}} + O\left(\frac{K_q}{N^{(S+\operatorname{Re} a)/\mu}}\right)\right),$$

where the implied constant in (6-4) *is independent of* N *and* q. *The numbers* α_s *are given by*

$$\text{(6-5)} \qquad \alpha_s = \frac{1}{\mu\cdot s!}p_0^{-(s+a)/\mu}\frac{d^s}{dz^s}\{q(z)\cdot(1-\phi(z))^{-(s+a)/\mu}\}_{z=z_0}.$$

The condition in Theorem 6.1 that $\mathcal{C}$ initially runs along the bisecting line with angle θ_k is not really necessary and just included for convenience. The theorem is true if $\mathcal{C}$ begins in the sector of angular width $2\pi/\mu$ about z_0 with this bisecting line, and the branch of $(z-z_0)^{a-1}$ is consistent with (6-3). The $a=1$ case of Theorem 6.1 is Theorem 1.2 and, in particular, (6-5) reduces to (1-7) when $a=1$.

Proof of Theorem 6.1. We may use a straightforward extension of the first proof of Theorem 1.2 given in Section 3. The key step is in Lemma 3.4, where we need to express $(z-z_0)^{a+n}$ in terms of w for any $n\in\mathbb{Z}$ and z on the bisecting line with angle θ_k. Here, $(z-z_0)^{a+n}=(z-z_0)^{a-1}\cdot(z-z_0)^{1+n}$ where $(z-z_0)^{a-1}$ is given by (6-3) and $(z-z_0)^{1+n}$ is unambiguous. Then

$$\begin{aligned}\text{(6-6)} \qquad (z-z_0)^{a+n} &= |z-z_0|^{a+n}\cdot e^{i\theta_k(a+n)}\\ &= \left(\frac{w}{N|p_0|}\right)^{(a+n)/\mu} e^{-i\omega_0(a+n)/\mu}e^{2\pi i k(a+n)/\mu}\\ &= w^{(a+n)/\mu}(Np_0)^{-(a+n)/\mu}e^{2\pi i k(a+n)/\mu},\end{aligned}$$

with the powers in (6-6) taking the principal values. Therefore

$$\begin{aligned} I_s(N) &:= \int_{z_0}^{b'} e^{-w}P_s(w)(z-z_0)^{s+a-1}\,dz\\ &= \frac{e^{2\pi i k(s+a)/\mu}}{\mu\cdot(Np_0)^{(s+a)/\mu}}\int_0^{N|p_0|(\rho/2)^\mu} e^{-w}P_s(w)w^{(s+a)/\mu-1}\,dw.\end{aligned}$$

The rest of the proof continues as in Section 3 to obtain the result. □

The second proof given in Section 4 may also be adapted to Theorem 6.1. The series $F(\tau)$ has a more complicated construction as described next. Define τ as in (4-3) and choose the branch of τ^{a-1} so that

$$\text{(6-7)} \qquad \tau^{a-1} = \tau(z)^{a-1} := p_0^{(a-1)/\mu}(z-z_0)^{a-1}(1-\phi(z))^{(a-1)/\mu},$$

where $(z-z_0)^{a-1}$ is consistent with (6-3) and the two other powers in (6-7) are principal. Then

$$-\frac{(z-z_0)^{a-1}}{\tau^{a-1}}q(z)\frac{\tau^{\mu-1}}{p'(z)} = -\frac{q(z)}{p_0^{(a-1)/\mu}(1-\phi(z))^{(a-1)/\mu}}\frac{\tau^{\mu-1}}{p'(z)} = h(z)$$

for $h(z)$ holomorphic on $\mathcal{D}_z$. As in (4-4) we may write

$$h(z_0+g(\tau)) =: \tau^{\mu-1}F(\tau) = \sum_{s=0}^{\infty}\beta_s\tau^s \qquad (z\in\mathcal{D}_z,\ \tau\in\mathcal{D}_\tau),$$

implying the identity

$$-(z-z_0)^{a-1}\frac{q(z)}{p'(z)} = \tau^{a-1}F(\tau) \qquad (z\in\mathcal{D}_z,\ \tau\in\mathcal{D}_\tau).$$

A calculation similar to Lemma 4.1 shows that

$$\tau^{a-1} = e^{2\pi ik(a-1)/\mu}v^{(a-1)/\mu}$$

when z in $\mathcal{D}_z$ is on the line from z_0 to b.

With the above results in place, the rest of the proof of Section 4 goes through easily. Of particular interest is the computation of β_s, as in the equations leading to (4-12) and (4-15):

(6-8)
$$\begin{aligned}\beta_s &= \frac{1}{2\pi i}\int_{C_0}\frac{\tau^{\mu-1}F(\tau)}{\tau^{s+1}}\,d\tau\\ &= -\frac{1}{2\pi i}\int_{C_0}\frac{q(z_0+g(\tau))}{p_0^{(a-1)/\mu}(1-\phi(z_0+g(\tau)))^{(a-1)/\mu}\cdot p'(z_0+g(\tau))\cdot\tau^{s-\mu+2}}\,d\tau\\ &= -\frac{1}{2\pi i\cdot\mu\cdot p_0^{(a-1)/\mu}}\int_{C_{z_0}}\frac{q(z)}{(1-\phi(z))^{(a-1)/\mu}\cdot\tau^{s+1}}\,dz\\ &= -\frac{1}{2\pi i\cdot\mu\cdot p_0^{(s+a)/\mu}}\int_{C_{z_0}}\frac{q(z)(1-\phi(z))^{-(s+a)/\mu}}{(z-z_0)^{s+1}}\,dz = \alpha_s.\end{aligned}$$

Formula (6-8) will be used in Proposition 7.3. When $a=1$, (6-8) reduces to (4-12).

6B. ***Including a factor*** $(z-z_0)^{a-1}$ ***with arbitrary*** $a\in\mathbb{C}$. Two applications of Theorem 6.1 give the following corollary.

Corollary 6.2 (Perron's method for an integrand containing a factor $(z-z_0)^{a-1}$ for $\operatorname{Re} a>0$ and with contour passing through a maximum). *Suppose Assumptions 1.1 hold. Let $\mathcal{C}$ be a contour starting at z_1, passing through z_0 and ending at z_2, with these three points all distinct. Suppose there are $k_1, k_2\in\mathbb{Z}$ so that, in a neighborhood* *of z_0, $\mathcal{C}$ runs along the bisecting line with angle θ_{k_1} as $\mathcal{C}$ approaches z_0 and $\mathcal{C}$ runs along the bisecting line with angle θ_{k_2} leaving z_0. Assume* $\operatorname{Re} a > 0$ *and*

$$\operatorname{Re} p(z) < \operatorname{Re} p(z_0) \quad \textit{for all } z \in \mathcal{C},\ z \neq z_0. \tag{6-9}$$

On the part of $\mathcal{C}$ approaching z_0 we take

$$(z - z_0)^{a-1} = |z - z_0|^{a-1} \cdot e^{i\theta_{k_1}(a-1)} \tag{6-10}$$

and on the part of $\mathcal{C}$ leaving z_0,

$$(z - z_0)^{a-1} = |z - z_0|^{a-1} \cdot e^{i\theta_{k_2}(a-1)}. \tag{6-11}$$

Then for any $S \in \mathbb{Z}_{\geqslant 0}$,

$$\int_{\mathcal{C}} e^{N\cdot p(z)}(z-z_0)^{a-1} q(z)\,dz = e^{N\cdot p(z_0)}\left(\sum_{s=0}^{S-1} \Gamma\left(\frac{s+a}{\mu}\right)\frac{\alpha_s(e^{2\pi i k_2(s+a)/\mu} - e^{2\pi i k_1(s+a)/\mu})}{N^{(s+a)/\mu}} + O\left(\frac{K_q}{N^{(S+\operatorname{Re} a)/\mu}}\right)\right), \tag{6-12}$$

where the implied constant in (6-12) is independent of N and q. The numbers α_s are given by (6-5).

The next result is an elegant extension of Corollary 6.2, where Perron shows that the condition $\operatorname{Re} a > 0$ may be dropped provided that the contour of integration is adjusted to make sure it avoids z_0. We will need this extension for the examples in Sections 8C and 8D.

Theorem 6.3 (Perron's method for an integrand containing a factor $(z - z_0)^{a-1}$ for arbitrary $a \in \mathbb{C}$). *Suppose Assumptions 1.1 hold. Let $\mathcal{C}$ be the following contour. Starting at z_1 it runs to the point z_1' which is a distance R_p from z_0 and on the bisecting line with angle θ_{k_1}. Then the contour circles z_0 to arrive at the point z_2' which is a distance R_p from z_0 and on the bisecting line with angle θ_{k_2}. Finally, the contour ends at z_2. The integers k_1 and k_2 keep track of how $\mathcal{C}$ rotates about z_0 between z_1' and z_2'; the angle of rotation is $2\pi(k_2 - k_1)/\mu$.*

Suppose that $\operatorname{Re} p(z) < \operatorname{Re} p(z_0)$ *for all z in the segments of $\mathcal{C}$ between z_1 and z_1' and between z_2' and z_2 (including endpoints). Let $a \in \mathbb{C}$. For $z \in \mathcal{C}$, the branch of $(z - z_0)^{a-1}$ is specified by requiring*

$$(z_1' - z_0)^{a-1} = |z_1' - z_0|^{a-1} \cdot e^{i\theta_{k_1}(a-1)} \tag{6-13}$$

when $z = z_1'$ and by continuity at the other points of $\mathcal{C}$. Then for any $S \in \mathbb{Z}_{\geqslant 0}$, (6-12) holds with an implied constant independent of N and q. If $(s + a)/\mu \in \mathbb{Z}_{\leqslant 0}$ then

$$\Gamma((s+a)/\mu)(e^{2\pi i k_2(s+a)/\mu} - e^{2\pi i k_1(s+a)/\mu})$$

in (6-12) is not defined and must be replaced by

$$2\pi i(k_2-k_1)(-1)^{(s+a)/\mu}/|(s+a)/\mu|!.$$

Proof. We will follow [Perron 1917, Section 4] and the first proof of Theorem 1.2 given in Section 3. It is convenient to move z_1', z_2' and the circular path of integration to the smaller radius $\rho/2$ with ρ satisfying (3-2). The points z_1' and z_2' are kept on their bisecting lines.

There exists $\varepsilon>0$ so that $\operatorname{Re}(p(z)-p(z_0))\leqslant -\varepsilon$ for all z in the segment of $\mathcal{C}$ between z_1 and z_1' (using (2-5) for the new part). It also follows that on this segment z is bounded away from z_0. Hence

$$\left|\int_{z_1}^{z_1'} e^{N(p(z)-p(z_0))}(z-z_0)^{a-1}q(z)\,dz\right|\leqslant K_q e^{-\varepsilon N}\int_{z_1}^{z_1'}|(z-z_0)^{a-1}|\,|dz|\ll K_q e^{-\varepsilon N}.$$

We obtain a similar bound for the integral between z_2' and z_2. The integral around the circular path from z_1' to z_2' remains to be estimated.

Following Lemma 3.2, write the integrand in the form

$$\text{(6-14)}\quad e^{N(p(z)-p(z_0))}(z-z_0)^{a-1}q(z) = \sum_{s=0}^{S-1} e^{-w}P_s(w)(z-z_0)^{s+a-1}+(z-z_0)^{a-1}Q_S(z)$$

with $w=Np_0(z-z_0)^{\mu}$ as in (3-6). The integer S should satisfy $S\geqslant 0$ and $S+\operatorname{Re}a>0$.

Lemma 6.4. *With this choice of* S,

$$\int_{z_1'}^{z_2'}(z-z_0)^{a-1}Q_S(z)\,dz = O\left(\frac{K_q}{N^{(S+\operatorname{Re}\,a)/\mu}}\right).$$

Proof. We may change the path of integration, moving the circular part closer to z_0 as follows. From z_1' the new path follows the bisecting line with angle θ_{k_1} to a point ζ_1 close to z_0. Then it circles z_0 until reaching ζ_2 on the bisecting line with angle θ_{k_2}. This bisecting line is followed to z_2'.

As in Lemma 3.2,

$$Q_S(z)=\frac{(z-z_0)^S}{2\pi i}\int_{\gamma}\frac{q(\tau)e^{-w+w\phi(\tau)}}{(\tau-z_0)^S(\tau-z)}\,d\tau,$$

where γ is the positively oriented circle of radius ρ about z_0. Note

$$\operatorname{Re}(-w+w\phi(\tau))\leqslant |w|(1+|\phi(\tau)|)\leqslant 3|w|/2\leqslant 2N|p_0|\,|z-z_0|^{\mu}.$$

Hence, for z with $|z-z_0|\leqslant\rho/2$,

$$(z-z_0)^{a-1}Q_S(z)\ll K_q|z-z_0|^{S+\operatorname{Re}\,a-1}e^{2N|p_0|\,|z-z_0|^{\mu}}.$$

Suppose ζ_1, ζ_2 and the circular path of integration between them are at a distance r from z_0. Then

$$\int_{\zeta_1}^{\zeta_2} (z-z_0)^{a-1} Q_S(z)\, dz = O(K_q \cdot r^{S+\mathrm{Re}\ a} e^{2N|p_0|r^\mu}). \tag{6-15}$$

Choosing any $r \leqslant N^{-1/\mu}$ shows that (6-15) satisfies the lemma's bound. The remaining integrals along the bisecting lines may now be bounded using (3-9) as in Lemma 3.3, completing the proof. □

Our work so far has shown

$$\int_{\mathcal{C}} e^{N\cdot p(z)} (z-z_0)^{a-1} q(z)\, dz = e^{N\cdot p(z_0)} \left(\sum_{s=0}^{S-1} I_s^*(N) + O\left(\frac{K_q}{N^{(S+\mathrm{Re}\ a)/\mu}} \right) \right) \tag{6-16}$$

for

$$I_s^*(N) := \int_{z_1'}^{z_2'} e^{-w} P_s(w) (z-z_0)^{s+a-1}\, dz. \tag{6-17}$$

As in Lemma 3.4, and using (6-6), we change variables to w in (6-17) to produce

$$I_s^*(N) = \frac{e^{2\pi i k_1 (s+a)/\mu}}{\mu \cdot (N p_0)^{(s+a)/\mu}} \int_{N|p_0|(\rho/2)^\mu}^{N|p_0|(\rho/2)^\mu} e^{-w} P_s(w) w^{(s+a)/\mu - 1}\, dw. \tag{6-18}$$

The path of integration in (6-18) starts and ends at the positive real number $T := N|p_0|(\rho/2)^\mu$, circling the origin $k_2 - k_1$ times. The value of $w^{(s+a)/\mu-1}$ in (6-18) is the principal power value at the beginning of the integration path and this value times $\exp(2\pi i (k_2-k_1)(s+a)/\mu)$ at the end of the integration path.

Lemma 6.5. *If* $(s+a)/\mu \in \mathbb{Z}_{\leqslant 0}$ *then*

$$I_s^*(N) = 2\pi i (k_2-k_1) \frac{(-1)^{(s+a)/\mu}}{|(s+a)/\mu|!} \frac{\alpha_s}{N^{(s+a)/\mu}}.$$

Proof. Letting $m := (s+a)/\mu$, the integral in (6-18) is

$$\int_T^T e^{-w} P_s(w) w^{m-1}\, dw. \tag{6-19}$$

When $m \in \mathbb{Z}_{\leqslant 0}$, the integrand has a pole with residue

$$\sum_{\ell=0}^{|m|} c_{s,\ell} \frac{(-1)^{|m|-\ell}}{(|m|-\ell)!},$$

where $c_{s,\ell}$ is the coefficient of $(z-z_0)^s$ in the Taylor expansion of $q(z)\phi(z)^\ell/\ell!$

about z_0 as in (3-4). Therefore (6-19) equals the coefficient of $(z-z_0)^s$ in

$$
\begin{aligned}
(6\text{-}20)\quad & 2\pi i(k_2-k_1)q(z)\sum_{\ell=0}^{|m|}\frac{(-1)^{|m|-\ell}}{(|m|-\ell)!}\frac{\phi(z)^\ell}{\ell!} \\
&= 2\pi i(k_2-k_1)\frac{(-1)^{|m|}}{|m|!}q(z)\sum_{\ell=0}^{|m|}\binom{|m|}{\ell}(-\phi(z))^\ell \\
&= 2\pi i(k_2-k_1)\frac{(-1)^m}{|m|!}q(z)(1-\phi(z))^{-m}.
\end{aligned}
$$

Putting this value into (6-18) and comparing with (6-5) completes the proof. □

Lemma 6.6. *If* $(s+a)/\mu \notin \mathbb{Z}_{\leqslant 0}$ *then, for* $\varepsilon'' > 0$,

$$
I_s^*(N) = \Gamma\left(\frac{s+a}{\mu}\right)\frac{\alpha_s(e^{2\pi i k_2(s+a)/\mu}-e^{2\pi i k_1(s+a)/\mu})}{N^{(s+a)/\mu}} + O(K_q e^{-\varepsilon'' N}).
$$

Proof. Let $\mathcal{H}_T$ be the path that starts at infinity, follows the positive real line to T, circles the origin k_2-k_1 times and then returns from T to its starting point at infinity. We need the simple extension of (3-17) given by

$$
\int_T^\infty e^{-w}w^{d-1}\,dw \leqslant e^{-T/2}\times\begin{cases}2^d\Gamma(d) & \text{if } d\geqslant 1,\\ 2T^{d-1} & \text{if } d\leqslant 1,\end{cases} \tag{6-21}
$$

for $T>0$. Then arguing as at the start of Proposition 3.5 shows that the integral in (6-18) satisfies

$$
\int_T^T e^{-w}P_s(w)w^{(s+a)/\mu-1}\,dw = \int_{\mathcal{H}_T} e^{-w}P_s(w)w^{(s+a)/\mu-1}\,dw + O(K_q e^{-\varepsilon'' N})
$$

for $T = N|p_0|(\rho/2)^\mu$ and $\varepsilon'' = |p_0|(\rho/2)^\mu/2$.

Now we claim that

$$
\begin{aligned}
(6\text{-}22)\quad & \frac{e^{2\pi i k_1(s+a)/\mu}}{\mu\cdot(Np_0)^{(s+a)/\mu}}\int_{\mathcal{H}_T} e^{-w}P_s(w)w^{(s+a)/\mu-1}\,dw \\
&= \Gamma\left(\frac{s+a}{\mu}\right)\frac{\alpha_s(e^{2\pi i k_2(s+a)/\mu}-e^{2\pi i k_1(s+a)/\mu})}{N^{(s+a)/\mu}}
\end{aligned}
$$

for all $T>0$ and for all $a\in\mathbb{C}$ with $(s+a)/\mu\notin\mathbb{Z}_{\leqslant 0}$. If $s+\operatorname{Re} a>0$ then we may let $T\to 0$ and evaluate the integrals along $\mathbb{R}_{\geqslant 0}$ as in the second half of Proposition 3.5. This proves (6-22) for a in a right half plane. However, the left side of (6-22) is a holomorphic function of a for all $a\in\mathbb{C}$. The right side of (6-22) is also holomorphic for all $a\in\mathbb{C}$ except that the Γ function has poles at the nonpositive integers. Hence, the holomorphic functions on each side (6-22) must agree for all $a\in\mathbb{C}$, except for the nonpositive integers, and the lemma follows. □

We note that, since the left side of (6-22) is holomorphic in a, taking a limit in a so that $(s+a)/\mu$ approaches a nonpositive integer on the right side of (6-22) can also be used to prove Lemma 6.5.

With (6-16) and Lemmas 6.5 and 6.6, we have proved Theorem 6.3 at least for S sufficiently large to satisfy the conditions before Lemma 6.4. The terms $\alpha_s/N^{(s+a)/\mu}$ are $O(K_q/N^{(s+\text{Re }a)/\mu})$, (see Proposition 7.3 below), and so we obtain the theorem for all $S \in \mathbb{Z}_{\geqslant 0}$. □

6C. *Further generalizations.* The main results of Theorems 1.2, 6.1 and 6.3 may be extended in different directions:

- The case where the contour of integration $\mathcal{C}$ has an endpoint at infinity can easily be handled if the part of the integral near infinity has a bound such as $O(e^{-\varepsilon N})$.
- It is possible to let μ in (1-3) be a positive real number instead of just a positive integer — see, for example, [Olver 1974, Theorem 6.1, p. 125]. Of course $p(z)$ will no longer be holomorphic in a neighborhood of z_0 if μ is not an integer.
- With extra conditions, as described in [Wong 1989, Theorem 4, p. 105] or [Olver 1974, Theorem 6.1, p. 125], we may allow N to approach infinity in a sector in $\mathbb{C}$
- Laplace's method, originating with Laplace in the 18th century, gives the main term of the asymptotics of (1-1) where $\mathcal{C}$ is an interval on the real line and $p(z)$ and $q(z)$ are real-valued. It is assumed that there exists a unique maximum of $p(z)$ on $\mathcal{C}$ (at $z = z_0$, say) along with the weak conditions that $p(z)$ is differentiable with $p'(z)$ and $q(z)$ continuous; see, for example, [Olver 1974, Theorem 7.1, p. 81] for the precise statement. When $p(z)$ and $q(z)$ have series expansions in a neighborhood of z_0 then as in [Olver 1974, Theorem 8.1, p. 86], the full asymptotic expansion of (1-1) can be given. If $p(z)$ and $q(z)$ are restrictions of holomorphic functions on a domain containing $\mathcal{C}$, then Perron's method may be applied to obtain the same result since z_0 is necessarily a saddle-point with steepest descent angles lying on the real line.
- In Section VIII of [Flajolet and Sedgewick 2009], a general type of saddle-point algorithm is provided to attempt to find the asymptotics as $N \to \infty$ of integrals $\int_{\mathcal{C}} F(z)\, dz$ where $F(z)$ depends in some way on N.

7. More formulas for α_s

If we know the order of vanishing of $q(z)$ at $z = z_0$ then we can say which of the first numbers $\alpha_0, \alpha_1, \ldots$ in Theorems 1.2 or 6.1 are zero.

Proposition 7.1. *Let the order of vanishing of $q(z)$ at $z = z_0$ be m and write $q(z) = (z - z_0)^m \psi(z)$, where $\psi(z_0) \neq 0$. Then we have $\alpha_0 = \alpha_1 = \cdots = \alpha_{m-1} = 0$*

and $\alpha_m \neq 0$. *Also, for* $s \in \mathbb{Z}_{\geqslant 0}$,

$$\alpha_{s+m} = \frac{1}{\mu \cdot s!} p_0^{-(s+m+a)/\mu} \frac{d^s}{dz^s} \{\psi(z) \cdot (1-\phi(z))^{-(s+m+a)/\mu)}\}_{z=z_0}. \tag{7-1}$$

Proof. Replace $q(z)$ by $(z-z_0)^m \psi(z)$ in (6-5), and evaluate the derivative with Leibniz's rule and the fact that

$$\frac{d^s}{dz^s}(z-z_0)^m \bigg|_{z=z_0} = \begin{cases} m! & \text{if } s=m, \\ 0 & \text{if } s \neq m. \end{cases}$$

It follows easily that $\alpha_s = 0$ for $s \leqslant m-1$ and that (7-1) holds. Also (7-1) implies that α_m takes the nonzero value $p_0^{-(m+a)/\mu} \psi(z_0)/\mu$. □

Therefore, in Theorems 1.2 and 6.1 where $\mathcal{C}$ starts at z_0, the main term of the asymptotic expansion has $s=m$, where m is the order of vanishing of $q(z)$.

In Corollaries 1.4 and 6.2 and Theorem 6.3 where $\mathcal{C}$ passes through z_0, the main term of the asymptotic expansion may not be $s=m$, since the factor $e^{2\pi i k_2(s+a)/\mu} - e^{2\pi i k_1(s+a)/\mu}$ vanishes when $(k_2-k_1)(s+a)/\mu \in \mathbb{Z}$, and a calculation is required to find the first nonzero term. In some cases, the terms $\alpha_s(e^{2\pi i k_2(s+a)/\mu} - e^{2\pi i k_1(s+a)/\mu})$ vanish for all s and we do not obtain exact asymptotics with these results. This happens for example when $\mu = 1$ and $a \in \mathbb{Z}$, or when $q(z) = p'(z)$.

As before, write

$$p(z) - p(z_0) = -\sum_{s=0}^{\infty} p_s (z-z_0)^{s+\mu}, \qquad q(z) = \sum_{s=0}^{\infty} q_s (z-z_0)^s.$$

The next result is due to Campbell, Fröman and Walles [Campbell et al. 1987, pp. 157–158] and expresses α_s in terms of the coefficients p_s and q_s. It requires the *partial ordinary Bell polynomials* which may be defined with the generating function

$$(p_1 x + p_2 x^2 + p_3 x^3 + \cdots)^j = \sum_{i=j}^{\infty} \hat{B}_{i,j}(p_1, p_2, p_3, \ldots) x^i. \tag{7-2}$$

It is straightforward to see they may also be given as

$$\hat{B}_{i,j}(p_1, p_2, p_3, \ldots) = \sum_{\substack{1\ell_1 + 2\ell_2 + 3\ell_3 + \cdots = i \\ \ell_1 + \ell_2 + \ell_3 + \cdots = j}} \frac{j!}{\ell_1! \ell_2! \ell_3! \ldots} p_1^{\ell_1} p_2^{\ell_2} p_3^{\ell_3} \cdots \tag{7-3}$$

from [Comtet 1974, Section 3.3] where the sum is over all possible $\ell_1, \ell_2, \ldots \in \mathbb{Z}_{\geqslant 0}$, or as

$$\hat{B}_{i,j}(p_1, p_2, p_3, \ldots) = \sum_{n_1 + n_2 + \cdots + n_j = i} p_{n_1} p_{n_2} \cdots p_{n_j} \tag{7-4}$$

for $j \geqslant 1$ from [Campbell et al. 1987, p. 156] where the sum is over all possible

$n_1, n_2, \ldots \in \mathbb{Z}_{\geqslant 1}$. See [Comtet 1974, Section 3.3] for more information on Bell polynomials, including their recurrence relations.

Proposition 7.2. *For α_s defined in (6-5),*

$$\alpha_s = \frac{1}{\mu} p_0^{-(s+a)/\mu} \sum_{i=0}^{s} q_{s-i} \sum_{j=0}^{i} \binom{-(s+a)/\mu}{j} \hat{B}_{i,j}\left(\frac{p_1}{p_0}, \frac{p_2}{p_0}, \ldots\right). \tag{7-5}$$

Proof. We have

$$\begin{aligned}(1-\phi(z))^{-(s+a)/\mu} &= \left(1+\sum_{s=1}^{\infty} \frac{p_s}{p_0}(z-z_0)^s\right)^{-(s+a)/\mu} \\ &= \sum_{j=0}^{\infty} \binom{-(s+a)/\mu}{j} \left(\sum_{s=1}^{\infty} \frac{p_s}{p_0}(z-z_0)^s\right)^j \\ &= \sum_{j=0}^{\infty} \binom{-(s+a)/\mu}{j} \sum_{i=j}^{\infty} \hat{B}_{i,j}\left(\frac{p_1}{p_0}, \frac{p_2}{p_0}, \ldots\right)(z-z_0)^i.\end{aligned} \tag{7-6}$$

Therefore the coefficient of $(z-z_0)^s$ in $q(z)(1-\phi(z))^{-(s+a)/\mu}$ is

$$\sum_{i=0}^{s} q_{s-i} \sum_{j=0}^{i} \binom{-(s+a)/\mu}{j} \hat{B}_{i,j}\left(\frac{p_1}{p_0}, \frac{p_2}{p_0}, \ldots\right)$$

and the result follows. □

With $a = 1$, the first cases are

$$\begin{aligned}\alpha_0 &= p_0^{-1/\mu} q_0/\mu, \\ \alpha_1 &= \frac{1}{\mu} p_0^{-2/\mu}\left(-\frac{2p_1q_0}{\mu p_0} + q_1\right), \\ \alpha_2 &= \frac{1}{\mu} p_0^{-3/\mu}\left(\frac{3(1+3/\mu)p_1^2 q_0}{2\mu p_0^2} - \frac{3(p_2q_0 + p_1q_1)}{\mu p_0} + q_2\right).\end{aligned}$$

Moving p_0 out of the sum in (7-6) gives the slightly different formulation

$$\alpha_s = \frac{1}{\mu} p_0^{-(s+a)/\mu} \sum_{i=0}^{s} q_{s-i} \sum_{j=0}^{i} p_0^{-j} \binom{-(s+a)/\mu}{j} \hat{B}_{i,j}(p_1, p_2, \ldots). \tag{7-7}$$

Wojdylo [2006] rediscovered the formula (7-5) in the context of Laplace's method, though his proof seems incomplete; the form of [Wojdylo 2006, Equation (2.34)] needs to be justified. A comparison of the schemes to give α_s explicitly in [Perron 1917; de Bruijn 1958; Dingle 1973; Campbell et al. 1987; Wojdylo 2006] is discussed in the appendix of [López and Pagola 2011]. See also [Nemes 2013].

We finish this section with a new bound for these expansion coefficients.

Proposition 7.3. *With Assumptions 1.1 and α_s defined in (6-5),*

$$\alpha_s = O\big(K_q^* \cdot C^s\big) \quad \textit{for } s \in \mathbb{Z}_{\geqslant 0}, \tag{7-8}$$

where K_q^* is a bound for $|q(z)|$ on $\mathcal{B}$. The positive constant C and the implied constant in (7-8) are both independent of q and s.

Proof. The result follows from (6-8) with C taken as the reciprocal of the radius of $\mathcal{D}_\tau$. □

8. Applications

The next examples illustrate how to apply Perron's method. Given an integral depending on a parameter N going to infinity, the first task is to try to get it into the form (1-1), perhaps with a change of variables. We are free to move the path of integration $\mathcal{C}$ continuously wherever the integrand is holomorphic. If we can ensure that Re $p(z)$ is maximized at an endpoint then Theorems 1.2 or 6.1 may be applied. Otherwise we move $\mathcal{C}$ to pass through saddle-points and employ Corollaries 1.4, 5.1, 6.2 or Theorem 6.3.

8A. ***Gamma function asymptotics.*** The standard example, see, e.g., [Perron 1917, Section 5], is the important gamma function. For $N > 0$ we have

$$\Gamma(N+1) = \int_0^\infty e^{-t} t^N \, dt = N^{N+1} \int_0^\infty e^{N(-z+\log z)} \, dz$$

with the change of variables $t = Nz$. Fitting the last integral into (1-1) and Assumptions 1.1, write $q(z) = 1$ and $p(z) = -z + \log z$ with $p'(z) = -1 + 1/z$. This shows there is a saddle-point at $z_0 = 1$. Close to $z = 1$ we have the expansion

$$-\log z = (1-z) + (1-z)^2/2 + (1-z)^3/3 + \cdots,$$

so the range of integration can be restricted to $[1/2, 3/2]$, say, and it is easy to see that the remaining integral will be too small to affect the result.

Hence, for $|z-1| \leqslant 1/2$, $p(z)$ equals

$$p(1) - p_0(z-1)^\mu(1-\phi(z)) = -1 - \tfrac{1}{2}(z-1)^2\big(1 - \tfrac{2}{3}(z-1) + \tfrac{2}{4}(z-1)^2 + \cdots\big)$$

so that $p_0 = 1/2$, $\omega_0 = 0$, $\mu = 2$ and $p_s/p_0 = (-1)^s 2/(s+2)$. The steepest descent angles are $\theta_\ell = \pi\ell$. The assumptions of Corollary 5.1 hold (with $k = 0$) and on simplifying it shows

$$\Gamma(N+1) = \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \left(1 + \frac{\gamma_1}{N} + \frac{\gamma_2}{N^2} + \cdots + \frac{\gamma_{k-1}}{N^{k-1}} + O\left(\frac{1}{N^k}\right)\right)$$

for, by Proposition 7.2 (with $a = 1$),

$$\gamma_m = \frac{(2m)!}{m!2^m} \sum_{j=0}^{2m} \binom{-m-1/2}{j} \hat{B}_{2m,j}\left(-\tfrac{2}{3}, \tfrac{2}{4}, -\tfrac{2}{5}, \tfrac{2}{6}, -\tfrac{2}{7}, \dots\right). \tag{8-1}$$

The first coefficients are $\gamma_1 = 1/12$, $\gamma_2 = 1/288$, $\gamma_3 = -139/51840$ as Laplace already knew. See [Nemes 2013, Example 1] for different treatments of (8-1). Approximations to the gamma function are still an interesting and active area of research as shown in [Chen 2013].

8B. ***The equation of the center.*** In Kepler's theory of motion, the planets orbit the sun in ellipses of eccentricity ε with the sun at one focus. The true anomaly ν is the angle made from this focus and may be compared with the angle M (the mean anomaly) made if the planet were in uniform circular motion, with the same period, about the mid point of the foci. These quantities are related by Kepler's equations

$$\cos \nu = \frac{\cos E - \varepsilon}{1 - \varepsilon \cos E}, \qquad M = E - \varepsilon \sin E$$

for an intermediate quantity E, called the eccentric anomaly. The *equation of the center* refers to different ways to relate ν to M directly. An important way is through the Fourier expansion

$$\nu - M = \sum_{n=1}^{\infty} C_n \sin(nM) \qquad \text{for } C_n = \frac{\sqrt{1-\varepsilon^2}}{\pi n} \int_{-\pi}^{\pi} \frac{e^{ni(z-\varepsilon \sin z)}}{1 - \varepsilon \cos z}\, dz, \tag{8-2}$$

as derived in [Battin 1999, pp. 210–212], for example. The integral appearing in (8-2) is the one from the introduction, (1-2). Before working on the asymptotics of (1-2) we take a simpler case.

The integral

$$\int_{-\pi}^{\pi} e^{Ni(z-\sin z)}\, dz \qquad (N \in \mathbb{Z}_{\geqslant 1}) \tag{8-3}$$

is studied in [Burkhardt 1914] and [Perron 1917]. Fitting it to the assumptions of Corollary 1.4, we have $q(z) = 1$ and $p(z) = i(z - \sin z)$ with $p'(z) = i(1 - \cos z)$. This shows there is a saddle-point at $z_0 = 0$ and writing

$$p(z) = 0 - p_0 z^{\mu}(1 - \phi(z)) = -\left(\frac{-i}{3!}\right) z^3 \left(1 - \frac{3!}{5!} z^2 + \cdots\right)$$

means that $p_0 = -i/6$, $\omega_0 = -\pi/2$, $\mu = 3$ and $1 - \phi(z) = 6(z - \sin z)/z^3$. The steepest descent angles are

$$\theta_\ell = \pi/6 + 2\pi\ell/3$$

as shown in Figure 2. We change the path of integration to

$$\int_{-\pi}^{-\pi+\pi i/\sqrt{3}} + \int_{-\pi+\pi i/\sqrt{3}}^{0} + \int_{0}^{\pi+\pi i/\sqrt{3}} + \int_{\pi+\pi i/\sqrt{3}}^{\pi} \tag{8-4}$$

so that 0 is approached along the line with angle θ_{k_1} for $k_1 = 1$ and, on leaving 0, the line with angle θ_{k_2} for $k_2 = 0$ is followed. The integrals along the vertical lines cancel since the integrand has period 2π. We have

$$\operatorname{Re} p(te^{i\theta_0}) = \operatorname{Re} p(te^{i\theta_1}) = f(t) \quad \text{for} \;\; f(t) := \cos(\sqrt{3}t/2)\sinh(t/2) - t/2$$

with $f(0) = 0$. To confirm condition (1-13) we need to show that $f(t) < 0$ for $0 < t \leqslant 2\pi/\sqrt{3}$. One approach is to first note that

$$f'''(t) = -\cos(\sqrt{3}t/2)\cosh(t/2).$$

Hence $f''(t)$ is decreasing on $[0, \pi/\sqrt{3})$ and increasing on $[\pi/\sqrt{3}, 2\pi/\sqrt{3}]$. As $f''(0) = 0$ and $f''(2\pi/\sqrt{3})$ is positive, this means that $f''(t)$ is negative on an interval $(0, c)$ and positive on $(c, 2\pi/\sqrt{3}]$ for some c. We see that $f'(t)$ decreases from $f'(0) = 0$ and then increases from $t = c$ to $f'(2\pi/\sqrt{3})$ which is < 0. Therefore $f'(t)$ is negative on $(0, 2\pi/\sqrt{3}]$ and so $f(t)$ is decreasing in this range as we wanted.

Write

$$\alpha_s = \frac{e^{\pi i(s+1)/6} \cdot 6^{(s+1)/3} \cdot d(s)}{3} \quad \text{for} \;\; d(s) = \frac{1}{s!}\frac{d^s}{dz^s}\left\{\left(\frac{6(z-\sin z)}{z^3}\right)^{-(s+1)/3}\right\}_{z=0}.$$

Also, by Proposition 7.2,

$$d(s) = \sum_{j=0}^{s} \binom{-(s+1)/3}{j} \hat{B}_{s,j}\left(0, -\frac{3!}{5!}, 0, \frac{3!}{7!}, 0, -\frac{3!}{9!}, \dots\right) \tag{8-5}$$

and computations yield, for example,

$$d(0) = 1, \quad d(2) = \tfrac{1}{20}, \quad d(4) = \tfrac{1}{280}, \quad d(6) = \tfrac{1}{3600}, \quad d(8) = \tfrac{387}{17248000}$$

with $d(s) = 0$ for s odd. Then by Corollary 1.4, for an implied constant depending only on S,

$$\begin{aligned} \int_{-\pi}^{\pi} e^{Ni(z-\sin z)}\,dz &= \sum_{s=0}^{S-1} \Gamma\left(\frac{s+1}{3}\right) \frac{\alpha_s(1-e^{2\pi i(s+1)/3})}{N^{(s+1)/3}} + O\left(\frac{1}{N^{(S+1)/3}}\right) \\ &= \frac{2}{3}\sum_{s=0}^{S-1} \cos\left(\frac{\pi(s+1)}{6}\right)\Gamma\left(\frac{s+1}{3}\right) d(s)\left(\frac{6}{N}\right)^{(s+1)/3} + O\left(\frac{1}{N^{(S+1)/3}}\right). \end{aligned} \tag{8-6}$$

We can obtain nonzero terms in the sum only for $s \equiv 0, 4 \bmod 6$. Formulas (8-5) and (8-6) give the complete asymptotic expansion of the integral (8-3). With $S = 10$ for example,

$$\int_{-\pi}^{\pi} e^{Ni(z-\sin z)}\,dz = \frac{2}{3}\cos\Bigl(\frac{\pi}{6}\Bigr)\Gamma\Bigl(\frac{1}{3}\Bigr)\Bigl(\frac{6}{N}\Bigr)^{1/3} + \frac{1}{420}\cos\Bigl(\frac{5\pi}{6}\Bigr)\Gamma\Bigl(\frac{5}{3}\Bigr)\Bigl(\frac{6}{N}\Bigr)^{5/3} + \frac{1}{5400}\cos\Bigl(\frac{7\pi}{6}\Bigr)\Gamma\Bigl(\frac{7}{3}\Bigr)\Bigl(\frac{6}{N}\Bigr)^{7/3} + O\Bigl(\frac{1}{N^{11/3}}\Bigr) \tag{8-7}$$

which is equivalent to [Perron 1917, Equation (53)]. When $N = 50$, for instance, the integral in (8-7) is approximately $\underline{0.762835382}546$ with the underlined digits indicating the agreement with the right side of (8-7). All the numerical calculations in this paper were carried out using Mathematica.

8C. ***Asymptotics of the true anomaly Fourier coefficient C_N.*** We now turn to the integral

$$\int_{-\pi}^{\pi} \frac{e^{Ni(z-\varepsilon\sin z)}}{1-\varepsilon\cos z}\,dz \qquad (N \in \mathbb{Z}_{\geqslant 1},\ 0 < \varepsilon < 1), \tag{8-8}$$

appearing in the equation of the center (8-2), and the main motivation of the papers [Burkhardt 1914; Perron 1917]. We initially follow [Perron 1917, pp. 210–214] and then go more deeply into the combinatorics of the expansion coefficients.

Set $p(z) = i(z - \varepsilon\sin z)$ and so $p'(z) = i(1 - \varepsilon\cos z)$. It is convenient to define

$$\gamma := \frac{1+\sqrt{1-\varepsilon^2}}{\varepsilon} > 1.$$

Then $p'(z) = 0$ for z taking the two values $\pm i\log\gamma$ and we choose z_0 to be $i\log\gamma$. Our computations will show that this is the correct choice. Expanding about this saddle-point gives

$$\begin{aligned} \varepsilon\sin(z+z_0) &= \sin z + i\sqrt{1-\varepsilon^2}\cos z,\\ \varepsilon\cos(z+z_0) &= \cos z - i\sqrt{1-\varepsilon^2}\sin z, \end{aligned} \tag{8-9}$$

as in [Perron 1917, p. 212]. Hence

$$\begin{aligned} p(z)-p(z_0) &= i(z-z_0) - i\sin(z-z_0) + \sqrt{1-\varepsilon^2}(-1+\cos(z-z_0))\\ &= -\frac{\sqrt{1-\varepsilon^2}}{2}(z-z_0)^2\Bigl(1 - \frac{2i}{3!\sqrt{1-\varepsilon^2}}(z-z_0) - \frac{2}{4!}(z-z_0)^2 + \cdots\Bigr) \end{aligned} \tag{8-10}$$

which implies that $p_0 = \sqrt{1-\varepsilon^2}/2$, $\mu = 2$ and the steepest descent angles are $\theta_\ell = \pi\ell$. Clearly, $1/(1-\varepsilon\cos z)$ has a simple pole at $z = z_0$, so we let $q(z) = (z-z_0)/(1-\varepsilon\cos z)$ with $a = 0$ and will be applying Theorem 6.3.

The contour of integration should therefore be moved from the real line and go vertically from $-\pi$ to $-\pi + i\log\gamma$. The path then approaches z_0 along the steepest descent angle θ_{k_1} for $k_1 = 1$, circles below z_0 and leaves along the angle θ_{k_2} for $k_2 = 2$. After reaching $\pi + i \log \gamma$ it then moves vertically to π. The integrals on the vertical paths cancel since the integrand has period 2π. For $t \in \mathbb{R}$ we see by (8-10) that $\operatorname{Re}(p(t+z_0)-p(z_0)) = \sqrt{1-\varepsilon^2}(-1+\cos t)$ and so $\operatorname{Re} p(z) < \operatorname{Re} p(z_0)$ for z on the horizontal part of the contour and the conditions for Theorem 6.3 are satisfied. (The other saddle-point, $-i \log \gamma$, has vertical steepest descent lines and so we cannot use it in a similar treatment.)

Writing w for $z - z_0$ we obtain, by (8-9),

$$q(z) = \frac{w}{1-\cos w + i\sqrt{1-\varepsilon^2}\sin w} = \frac{2\gamma}{\varepsilon} \cdot \frac{w}{(e^{iw}-1)} \cdot \frac{1}{(\gamma^2 e^{-iw}-1)}.$$

We have the expansions

$$\text{(8-11)} \qquad \frac{z}{e^z-1} = \sum_{m=0}^{\infty} \frac{B_m}{m!} z^m, \qquad \frac{1}{\xi e^z - 1} = \sum_{m=0}^{\infty} \frac{\beta_{m+1}(\xi)}{(m+1)!} z^m \qquad (\xi \neq 1)$$

for the Bernoulli numbers B_m and the coefficients

$$\text{(8-12)} \qquad \beta_m(\xi) = (-1)^{m-1} m \sum_{j=1}^{m} \left\{ {m \atop j} \right\} \frac{(j-1)!}{(\xi-1)^j} \qquad (\xi \neq 1),$$

where $\left\{ {m \atop j} \right\}$ is the Stirling number, denoting the number of ways to partition a set of size m into j nonempty subsets. See [O'Sullivan 2015, Proposition 3.2] for the formula (8-12) which is similar to a result of Glaisher. Then

$$q(z) = \sum_{s=0}^{\infty} q_s w^s = \frac{-2\gamma \cdot i}{\varepsilon} \left(\sum_{m=0}^{\infty} \frac{B_m}{m!} (iw)^m \right) \left(\sum_{n=0}^{\infty} \frac{\beta_{n+1}(\gamma^2)}{(n+1)!} (-iw)^n \right)$$

and we obtain the expression

$$\text{(8-13)} \qquad q_s = \frac{-2\gamma \cdot i^{s+1}}{\varepsilon} \sum_{n=0}^{s} (-1)^n \frac{\beta_{n+1}(\gamma^2) B_{s-n}}{(n+1)!(s-n)!}.$$

With Proposition 7.2 we may write $\alpha_s = d(s)/(2p_0^{s/2})$ for

$$\text{(8-14)} \quad d(s) = \sum_{i=0}^{s} q_{s-i} \sum_{j=0}^{i} \binom{-s/2}{j} \hat{B}_{i,j} \left(-\frac{2i}{3!\sqrt{1-\varepsilon^2}}, -\frac{2}{4!}, \frac{2i}{5!\sqrt{1-\varepsilon^2}}, \frac{2}{6!}, \dots \right),$$

where the arguments in the above Bell polynomial are

$$\frac{p_s}{p_0} = \frac{2 \cdot i^s}{(s+2)!} \times \begin{cases} 1 & \text{if } s \text{ is even,} \\ -1/\sqrt{1-\varepsilon^2} & \text{if } s \text{ is odd.} \end{cases}$$

A short calculation with (8-9) shows

$$e^{p(z_0)} = e^{\sqrt{1-\varepsilon^2}}/\gamma < 1.$$

Putting everything together, and using the last line in the statement of Theorem 6.3 for the $s = 0$ term, we obtain

$$\int_{-\pi}^{\pi} \frac{e^{Ni(z-\varepsilon\sin z)}}{1-\varepsilon\cos z}\,dz = \left(\frac{e^{\sqrt{1-\varepsilon^2}}}{\gamma}\right)^N \left(\frac{\pi}{\sqrt{1-\varepsilon^2}} + \sum_{1\leqslant s\leqslant S-1,\ s \text{ odd}} \Gamma(s/2)\cdot d(s)\left(\frac{2}{\sqrt{1-\varepsilon^2}N}\right)^{s/2} + O\left(\frac{1}{N^{S/2}}\right)\right) \tag{8-15}$$

which, along with (8-13) and (8-14), gives the complete asymptotic expansion. Computing the first values of $d(s)$, for s odd, we observe that they take the form $f_s(\varepsilon^2)/(1-\varepsilon^2)^{(s+1)/2}$ for f_s a polynomial with rational coefficients and degree $(s-1)/2$. For instance

$$f_1(x) = 2/3, \quad f_3(x) = -(46+189x)/540, \quad f_5(x) = (92+6228x+4887x^2)/36288.$$

It would be interesting to prove that this form always holds. With $S = 5$ we find

$$\int_{-\pi}^{\pi} \frac{e^{Ni(z-\varepsilon\sin z)}}{1-\varepsilon\cos z}\,dz = \left(\frac{e^{\sqrt{1-\varepsilon^2}}}{\gamma}\right)^N \left(\frac{\pi}{\sqrt{1-\varepsilon^2}} + \Gamma(1/2)\frac{2}{3(1-\varepsilon^2)}\left(\frac{2}{\sqrt{1-\varepsilon^2}N}\right)^{1/2} - \Gamma(3/2)\frac{46+189\varepsilon^2}{540(1-\varepsilon^2)^2}\left(\frac{2}{\sqrt{1-\varepsilon^2}N}\right)^{3/2} + O\left(\frac{1}{N^{5/2}}\right)\right) \tag{8-16}$$

which is equivalent to [Perron 1917, Equation (45)]. When $N = 50$ and $\varepsilon = 2/5$, for example, the integral in (8-16) is $\approx \underline{2.8171}413884 \times 10^{-14}$ with the underlined digits indicating the agreement with the right side of (8-16). Taking $S = 13$, i.e., using the first seven terms in the expansion (8-15), yields the agreement $\underline{2.817141388}4 \times 10^{-14}$.

As a referee noted, the method of steepest descent for this example requires moving the contour of integration to a more complicated path near z_0 than the horizontal line above. It requires part of the path described by the equation $\cosh(y) = x/(\varepsilon\sin(x))$ for $z = x+iy$. This is where $\mathrm{Im}(p(z)-p(z_0)) = 0$.

8D. *The case $\varepsilon = 1$.* Taking $\varepsilon = 1$ in (8-8) produces the integral

$$\int_{-\pi}^{\pi} \frac{e^{Ni(z-\sin z)}}{1-\cos z}\,dz \qquad (N \in \mathbb{Z}_{\geqslant 1}) \tag{8-17}$$

which is studied in Example 4 of [Perron 1917]. This would correspond to a parabolic orbit if (8-2) were valid for $\varepsilon = 1$. The path of integration in (8-17) must avoid the double pole at $z = 0$ in order to converge. The expansion of the integrand at $z = 0$ begins $2/z^2 + 1/6 + (Niz)/3 + z^2/120 + \cdots$, implying the residue at $z = 0$ is zero. Since the integrand has period 2π, all the residues are zero and so the integral is completely independent of any pole-avoiding path of integration from $-\pi$ to π.

The function $p(z)$ is the same as in Section 8B, but now $q(z) = z^2/(1-\cos z)$ and $a = -1$. We will use Theorem 6.3 and so the path of integration (8-4) must be adjusted to circle at a small radius about the pole at $z_0 = 0$. Then

$$\alpha_s = \frac{1}{3} e^{\pi i(s-1)/6} \cdot 6^{(s-1)/3} \cdot d^*(s)$$
$$\text{for } d^*(s) = \frac{1}{s!} \frac{d^s}{dz^s} \left\{ \frac{z^2}{1-\cos z} \left(\frac{6(z-\sin z)}{z^3} \right)^{-(s-1)/3} \right\}_{z=0}.$$

We have

$$\sum_{s=0}^{\infty} q_s z^s = \frac{z^2}{1-\cos z} = 2 \frac{iz}{e^{iz}-1} \frac{-iz}{e^{-iz}-1} = 2\left(\sum_{m=0}^{\infty} \frac{B_m}{m!}(iz)^m \right)\left(\sum_{n=0}^{\infty} \frac{B_n}{n!}(-iz)^n \right).$$

It follows that q_s is 0 for odd s, and for s even,

$$q_s = 2(-1)^{s/2} \sum_{n=0}^{s} (-1)^n \frac{B_n B_{s-n}}{n!(s-n)!}. \tag{8-18}$$

Proposition 7.2 tells us

$$d^*(s) = \sum_{i=0}^{s} q_{s-i} \sum_{j=0}^{i} \binom{-(s-1)/3}{j} \hat{B}_{i,j}\left(0, -\frac{3!}{5!}, 0, \frac{3!}{7!}, 0, -\frac{3!}{9!}, \dots\right) \tag{8-19}$$

and computations yield, for example,

$$d^*(0) = 2, \quad d^*(2) = \tfrac{1}{5}, \quad d^*(4) = \tfrac{27}{1400}, \quad d^*(6) = \tfrac{23}{12600}, \quad d^*(8) = \tfrac{947}{5544000}$$

with $d^*(s) = 0$ for s odd. Then, for an implied constant depending only on S,

$$\int_{-\pi}^{\pi} \frac{e^{Ni(z-\sin z)}}{1-\cos z}\, dz = \frac{2}{3} \sum_{s=0}^{S-1} \cos\left(\frac{\pi(s-1)}{6}\right) \Gamma\left(\frac{s-1}{3}\right) d^*(s) \left(\frac{6}{N}\right)^{(s-1)/3} + O\left(\frac{1}{N^{(S-1)/3}}\right). \tag{8-20}$$

We can obtain nonzero terms in the sum only for $s \equiv 0, 2 \bmod 6$. The term with $s = 1$ needs the formula from the last line of the statement of Theorem 6.3, but in any case vanishes since $d^*(1) = 0$. Formulas (8-18), (8-19) and (8-20) give the complete asymptotic expansion of the integral (8-17). Taking $S=8$ for example,

$$\text{(8-21)}\quad \int_{-\pi}^{\pi} \frac{e^{Ni(z-\sin z)}}{1-\cos z}\,dz = \frac{4}{3}\cos\Big(\frac{-\pi}{6}\Big)\Gamma\Big(\frac{-1}{3}\Big)\Big(\frac{N}{6}\Big)^{1/3} + \frac{2}{15}\cos\Big(\frac{\pi}{6}\Big)\Gamma\Big(\frac{1}{3}\Big)\Big(\frac{6}{N}\Big)^{1/3} + \frac{23}{18900}\cos\Big(\frac{5\pi}{6}\Big)\Gamma\Big(\frac{5}{3}\Big)\Big(\frac{6}{N}\Big)^{5/3} + O\Big(\frac{1}{N^{7/3}}\Big)$$

with the first two terms of this expansion given in [Perron 1917, Equation (50)]. When $N=50$ the integral in (8-21) is $\approx \underline{-9.357585}773084$ and the underlined digits show the agreement with the right-hand side.

9. The asymptotics of Sylvester waves

In this section we give an application of Perron's method to number theory. Let $p(n)$ be the number of partitions of the positive integer n. This is the number of ways to write n as a sum of nonincreasing positive integers. Also let $p_N(n)$ count the partitions of n with at most N summands. Since the work of Cayley and Sylvester in the 19th century, we know that

$$p_N(n) = \sum_{k=1}^{N} W_k(N,n),$$

where each $W_k(N,n)$ may be expressed in terms of a sequence of k polynomials $w_{k,m}(N,x) \in \mathbb{Q}[x]$ for $0 \leqslant m \leqslant k-1$. Write

$$\text{(9-1)}\qquad W_k(N,n) = [w_{k,0}(N,n),\ w_{k,1}(N,n),\ \dots,\ w_{k,k-1}(N,n)],$$

where the notation in (9-1) indicates that the value of $W_k(N,n)$ is given by one of the polynomials on the right and we select $w_{k,j}(N,n)$ when $n \equiv j \mod k$. The degrees of the polynomials on the right of (9-1) are at most $\lfloor N/k \rfloor - 1$.

For example, with $N=3$ we have $p_3(n) = W_1(3,n) + W_2(3,n) + W_3(3,n)$, where

$$W_1(3,n) = [6n^2+36n+47]/72,$$
$$W_2(3,n) = [1,\ -1]/8,$$
$$W_3(3,n) = [2,\ -1,\ -1]/9.$$

Sylvester called $W_k(N,n)$ the *k-th wave* and provided the formula

$$\text{(9-2)}\qquad W_k(N,n) = \operatorname*{Res}_{z=0} \sum_{\rho} \frac{\rho^n e^{nz}}{(1-\rho^{-1}e^{-z})(1-\rho^{-2}e^{-2z})\cdots(1-\rho^{-N}e^{-Nz})}$$

in [Sylvester 1882], where $\mathrm{Res}_{z=0}$ indicates the coefficient of $1/z$ in the Laurent expansion about 0, and the sum is over all primitive k-th roots of unity ρ. For a more detailed discussion of the above results with references, see Sections 1 and 2 of [O'Sullivan 2018].

When $N=3$ it is clear that the first wave $W_1(3,n)$ will make the largest contribution to $p_3(n)$ for large n. Similarly, $p_N(n) \sim W_1(N,n)$ for any fixed N as $n\to\infty$. A more difficult question, which we answer for the first time in [O'Sullivan 2018], is how the first waves $W_1(N,n)+W_2(N,n)+\cdots$ compare with $p_N(n)$ as N and n both go to ∞. The answer, perhaps surprisingly, is that when N and n grow at approximately the same rate, the first waves quickly become much larger than $p_N(n)$ (in absolute value, since these waves also oscillate like a sine with period ≈ 31.963 in N).

The asymptotics of the first 100 waves is given in [O'Sullivan 2018] as follows, in terms of two uniquely defined complex numbers with approximations $w_0 \approx 0.916198-0.182459i$ and $z_0 \approx 1.181475+0.255528i$.

Theorem 9.1. *Let λ^+ be a positive real number. Suppose $N\in\mathbb{Z}_{\geqslant 1}$ and $\lambda N\in\mathbb{Z}$ for λ satisfying $|\lambda|\leqslant\lambda^+$. Then there are explicit coefficients $a_0(\lambda), a_1(\lambda),\dots$ so that*

$$\sum_{k=1}^{100} W_k(N,\lambda N)=\operatorname{Re}\left[\frac{w_0^{-N}}{N^2}\left(a_0(\lambda)+\frac{a_1(\lambda)}{N}+\cdots+\frac{a_{m-1}(\lambda)}{N^{m-1}}\right)\right]+O\left(\frac{|w_0|^{-N}}{N^{m+2}}\right) \tag{9-3}$$

as $N\to\infty$, where $a_0(\lambda)=2z_0e^{-\pi i z_0(1+2\lambda)}$ and the implied constant depends only on λ^+ and m.

In the rest of this section we briefly sketch the proof of Theorem 9.1, highlighting the role of Perron's method in the form of Corollary 5.1. We require the dilogarithm, which is initially defined as

$$\mathrm{Li}_2(z):=\sum_{n=1}^{\infty}\frac{z^n}{n^2} \quad \text{for } |z|\leqslant 1,$$

with an analytic continuation given by $-\int_0^z \log(1-u)/u\,du$.

Sketch of proof of Theorem 9.1. In [O'Sullivan 2018, Equation (3.6)], it is shown that the left side of (9-3) may be expressed as a sum of three parts. As in the proof of [O'Sullivan 2018, Theorem 1.2], two of these parts are $O(e^{0.055N})$. The third part may be expressed as an integral (see [O'Sullivan 2018, Equation (5.13)]) to obtain

$$\sum_{k=1}^{100} W_k(N,\lambda N)=\frac{2}{N^{3/2}}\operatorname{Im}\int_{1.01}^{1.49} e^{N\cdot p(z)} f_\lambda(z)\cdot\exp(v(z;N))\,dz+O(e^{0.055N}), \tag{9-4}$$

for an implied constant depending only on λ^+, where

$$p(z):=\frac{\mathrm{Li}_2(e^{2\pi i z})-\mathrm{Li}_2(1)}{2\pi i z},$$

$$f_\lambda(z):=\left(\frac{z}{2\sin(\pi(z-1))}\right)^{1/2}\exp\bigl(-\pi i z(2\lambda+1/2)\bigr).$$

(In [O'Sullivan 2018], the function $p(z)$ is used with the opposite sign.) To describe a useful approximation to the function $\exp(v(z;N))$, we first define

$$g_\ell(z) := -\frac{B_{2\ell}}{(2\ell)!}(\pi z)^{2\ell-1}\cot^{(2\ell-2)}(\pi z),$$
$$u_j(z) := \sum_{m_1+3m_2+5m_3+\cdots=j} \frac{g_1(z)^{m_1}}{m_1!}\frac{g_2(z)^{m_2}}{m_2!}\cdots\frac{g_j(z)^{m_j}}{m_j!}$$

with $u_0 := 1$. Also define the box

$$\mathbb{B}_1 := \{z \in \mathbb{C} : 1.01 \leqslant \operatorname{Re} z \leqslant 1.49,\ -1 \leqslant \operatorname{Im} z \leqslant 1\}.$$

Then there are functions $u_j(z)$ (defined above) and $\zeta_d(z;N)$ which are holomorphic on a domain containing the box $\mathbb{B}_1$ and have the following property. For all $z \in \mathbb{B}_1$,

$$\exp(v(z;N)) = \sum_{j=0}^{d-1} \frac{u_j(z)}{N^j} + \zeta_d(z;N) \quad \text{for } \zeta_d(z;N) = O\Big(\frac{1}{N^d}\Big) \tag{9-5}$$

with an implied constant depending only on d, where $1 \leqslant d \leqslant 2L-1$ and $L = \lfloor 0.006\pi e \cdot N \rfloor$.

Since $|\exp(-2\pi i\lambda z)| \leqslant \exp(\lambda^+ 2\pi|z|)$ it follows that

$$f_\lambda(z) \ll 1 \quad \text{for } z \in \mathbb{B}_1, \tag{9-6}$$

with an implied constant depending only on λ^+.

To apply Corollary 5.1 we need the relevant saddle-point of $p(z)$ and this turns out to be $z_0 := 1 + \log(1-w_0)/(2\pi i)$, where w_0 is the unique solution to $\operatorname{Li}_2(w) - 2\pi i \log w = 0$. Both z_0 and w_0 may be found to any precision and their approximations were given before Theorem 9.1. (It is straightforward to compute the size of the error introduced into (9-3) by using approximations to z_0 and w_0.) We find $\mu = 2$, $p_0 \approx 0.504 - 0.241i$ and the steepest-descent angles are $\theta_0 \approx 0.223$ and $\theta_1 = \pi + \theta_0$.

Let $c := 1 + i\operatorname{Im}(z_0)/\operatorname{Re}(z_0)$. We move the path of integration in (9-4) to the path $\mathcal{P}$ through z_0 consisting of the straight line segments joining the points 1.01, $1.01c$, $1.49c$ and 1.49. Since the integrand in (9-4) is holomorphic on a domain containing $\mathbb{B}_1$, Cauchy's theorem ensures that the integral remains the same under this change of path. It is proved in [O'Sullivan 2016, Theorem 5.2] that

$$\operatorname{Re}(p(z) - p(z_0)) < 0 \quad \text{for all } z \in \mathcal{P},\ z \neq z_0. \tag{9-7}$$

We also need from [O'Sullivan 2018, Equation (5.16)] that

$$e^{p(z_0)} = w_0^{-1} \quad \text{and} \quad e^{\operatorname{Re} p(z_0)} = |w_0|^{-1} \approx e^{0.068}.$$

Using (9-5) in (9-4) implies

$$\text{(9-8)}\quad \sum_{k=1}^{100} W_k(N,\lambda N)= \\ \operatorname{Im}\left[\sum_{j=0}^{d-1}\frac{2}{N^{3/2+j}}\int_{\mathcal{P}} e^{N\cdot p(z)}\cdot f_\lambda(z)\cdot u_j(z)\,dz+\frac{2}{N^{3/2}}\int_{\mathcal{P}} e^{N\cdot p(z)}\cdot f_\lambda(z)\cdot \zeta_d(z;N)\,dz\right] \\ +O(e^{0.055N}),$$

where, by (9-5), (9-6) and (9-7), the last term in parentheses in (9-8) is

$$\ll \frac{1}{N^{3/2}}\int_{\mathcal{P}} |e^{N\cdot p(z)}|\cdot 1\cdot \frac{1}{N^d}\,dz \ll \frac{1}{N^{d+3/2}} e^{N\operatorname{Re} p(z_0)} = \frac{|w_0|^{-N}}{N^{d+3/2}},$$

for an implied constant depending only on λ^+. Applying Corollary 5.1 to each integral in the first part of (9-8) we obtain, since $k=0$,

$$\text{(9-9)}\quad \int_{\mathcal{P}} e^{N\cdot p(z)}\cdot f_\lambda(z)\cdot u_j(z)\,dz \\ = e^{N\cdot p(z_0)}\left(\sum_{m=0}^{M-1}\Gamma\left(m+\frac{1}{2}\right)\frac{2\alpha_{2m}(f_\lambda\cdot u_j)}{N^{m+1/2}}+O\left(\frac{K(f_\lambda\cdot u_j)}{N^{M+1/2}}\right)\right).$$

We have written $\alpha_{2m}(q)$, to show the dependence of α_{2m} on $q=f_\lambda\cdot u_j$, and also $K(q)$ instead of K_q. The error term in (9-9) corresponds to an error in (9-8) of size $O(|w_0|^{-N}/N^{M+j+2})$. Choose $M=d$ so that this error is less than $O(|w_0|^{-N}/N^{d+3/2})$ for all $j\geqslant 0$. Therefore

$$\sum_{k=1}^{100} W_k(N,\lambda N) \\ =\operatorname{Im}\left[\sum_{j=0}^{d-1}\frac{4}{N^{j+3/2}}e^{N\cdot p(z_0)}\sum_{m=0}^{d-1}\Gamma\left(m+\frac{1}{2}\right)\frac{\alpha_{2m}(f_\lambda\cdot u_j)}{N^{m+1/2}}\right]+O\left(\frac{|w_0|^{-N}}{N^{d+3/2}}\right)$$

$$\text{(9-10)}\quad =\operatorname{Im}\left[w_0^{-N}\sum_{t=0}^{2d-2}\frac{4}{N^{t+2}}\sum_{m=\max(0,t-d+1)}^{\min(t,d-1)}\Gamma\left(m+\frac{1}{2}\right)\alpha_{2m}(f_\lambda\cdot u_{t-m})\right] \\ +O\left(\frac{|w_0|^{-N}}{N^{d+3/2}}\right)$$

$$\text{(9-11)}\quad =\operatorname{Re}\left[w_0^{-N}\sum_{t=0}^{d-2}\frac{-4i}{N^{t+2}}\sum_{m=0}^{t}\Gamma\left(m+\frac{1}{2}\right)\alpha_{2m}(f_\lambda\cdot u_{t-m})\right]+O\left(\frac{|w_0|^{-N}}{N^{d+1}}\right)$$

for implied constants depending only on λ^+ and d. (In going from (9-10) to (9-11) we used that $|\alpha_{2m}(f_\lambda\cdot u_j)|$ has a bound depending only on λ^+ and d, by

Proposition 7.3, when $m, j \leqslant d-1$.) Hence, with

$$a_t(\lambda) := -4i \sum_{m=0}^{t} \Gamma\Big(m+\frac{1}{2}\Big)\alpha_{2m}(f_\lambda \cdot u_{t-m}), \tag{9-12}$$

we obtain (9-3) in the statement of the theorem.

The first coefficient is

$$a_0(\lambda) = -4i\Gamma(1/2)\alpha_0(f_\lambda \cdot u_{0,0}) = -4i\sqrt{\pi}\alpha_0(f_\lambda) = -2i\sqrt{\pi}\, p_0^{1/2} f_\lambda(z_0), \tag{9-13}$$

using our formula for α_0 from Section 7. The calculations [O'Sullivan 2018, Equations (5.24) and (5.26)] show

$$p_0^{1/2} = -\frac{\sqrt{\pi}e^{-\pi i/4}e^{\pi i z_0}}{z_0^{1/2} w_0^{1/2}}, \qquad f_\lambda(z_0) = -\frac{e^{\pi i/4} z_0^{1/2}}{w_0^{1/2}} e^{-2\pi i \lambda z_0}. \tag{9-14}$$

The formula for $a_0(\lambda)$ in the theorem's statement follows from (9-13) and (9-14). □

We may take $N = 2000$ and $\lambda = 1$ as an example of Theorem 9.1. The first wave $W_1(N, N)$ is $\approx 4.37 \times 10^{53}$ with the next waves much smaller: $W_2(N, N) \approx 4.98 \times 10^{23}$, $W_3(N, N) \approx -8.22 \times 10^{13}$, etc. We find that the main term on the right of (9-3) is $\approx 4.56 \times 10^{53}$. Taking the first three terms on the right of (9-3) gives the more accurate 4.37×10^{53}. By comparison, the corresponding partition number $p(N)$ $(= p_N(N))$ is a lot smaller and approximately 4.72×10^{45}.

See [O'Sullivan 2018] for the detailed proof of Theorem 9.1 as well as more extensive discussion and numerical work. We expect, as in [O'Sullivan 2018, Conjecture 9.1], that Theorem 9.1 is true with the sum of the first 100 waves on the left of (9-3) replaced by just the first wave $W_1(N, \lambda N)$.

References

[Ahlfors 1978] L. V. Ahlfors, *Complex analysis*, 3rd ed., McGraw-Hill, New York, 1978. MR Zbl

[Battin 1999] R. H. Battin, *An introduction to the mathematics and methods of astrodynamics*, revised ed., Amer. Inst. Aeronautics Astronautics, Reston, VA, 1999. MR Zbl

[de Bruijn 1958] N. G. de Bruijn, *Asymptotic methods in analysis*, Bibliotheca Mathematica **4**, North-Holland, Amsterdam, 1958. MR Zbl

[Burkhardt 1914] H. Burkhardt, "Über Funktionen großer Zahlen, insbesondere über die näherungsweise Bestimmung entfernter Glieder in den Reihenentwicklungen der Theorie der Keplerschen Bewegung", *Sitzungsberichte Bayer. Akad. Wissensch.* **1914**:1 (1914), 1–11. Zbl

[Campbell et al. 1987] J. A. Campbell, P. O. Fröman, and E. Walles, "Explicit series formulae for the evaluation of integrals by the method of steepest descents", *Stud. Appl. Math.* **77**:2 (1987), 151–172. MR Zbl

[Chen 2013] C.-P. Chen, "Unified treatment of several asymptotic formulas for the gamma function", *Numer. Algorithms* **64**:2 (2013), 311–319. MR Zbl

[Comtet 1974] L. Comtet, *Advanced combinatorics*, revised ed., Reidel, Dordrecht, 1974. MR Zbl

[Copson 1965] E. T. Copson, *Asymptotic expansions*, Cambridge Tracts in Math. and Mathematical Phys. **55**, Cambridge Univ. Press, 1965. MR Zbl

[Dingle 1973] R. B. Dingle, *Asymptotic expansions: their derivation and interpretation*, Academic Press, London, 1973. MR Zbl

[Erdélyi 1956] A. Erdélyi, *Asymptotic expansions*, Dover, New York, 1956. MR Zbl

[Flajolet and Sedgewick 2009] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge Univ. Press, 2009. MR Zbl

[López and Pagola 2011] J. L. López and P. J. Pagola, "An explicit formula for the coefficients of the saddle point method", *Constr. Approx.* **33**:2 (2011), 145–162. MR Zbl

[López et al. 2009] J. L. López, P. Pagola, and E. Pérez Sinusía, "A systematization of the saddle point method: application to the Airy and Hankel functions", *J. Math. Anal. Appl.* **354**:1 (2009), 347–359. MR Zbl

[Nemes 2013] G. Nemes, "An explicit formula for the coefficients in Laplace's method", *Constr. Approx.* **38**:3 (2013), 471–487. MR Zbl

[Olver 1970] F. W. J. Olver, "Why steepest descents?", *SIAM Rev.* **12** (1970), 228–247. MR Zbl

[Olver 1974] F. W. J. Olver, *Asymptotics and special functions*, Academic Press, New York, 1974. MR Zbl

[O'Sullivan 2015] C. O'Sullivan, "On the partial fraction decomposition of the restricted partition generating function", *Forum Math.* **27**:2 (2015), 735–766. MR Zbl

[O'Sullivan 2016] C. O'Sullivan, "Asymptotics for the partial fractions of the restricted partition generating function, I", *Int. J. Number Theory* **12**:6 (2016), 1421–1474. MR Zbl

[O'Sullivan 2018] C. O'Sullivan, "Partitions and Sylvester waves", *Ramanujan J.* **47**:2 (2018), 339–381. MR Zbl

[Perron 1917] O. Perron, "Über die näherungsweise Berechnung von Funktionen großer Zahlen", *Sitzungsberichte Bayer. Akad. Wissensch.* **1917**:11 (1917), 191–220. Zbl

[Petrova and Solov'ev 1997] S. S. Petrova and A. D. Solov'ev, "The origin of the method of steepest descent", *Historia Math.* **24**:4 (1997), 361–375. MR Zbl

[Small 2010] C. G. Small, *Expansions and asymptotics for statistics*, Monographs Stat. Appl. Probab. **115**, CRC Press, Boca Raton, FL, 2010. MR Zbl

[Sylvester 1882] J. J. Sylvester, "On subvariants; i.e., semi-invariants to binary quantics of an unlimited order", *Amer. J. Math.* **5**:1 (1882), 79–136. MR Zbl

[Temme 2013] N. M. Temme, "Uniform asymptotic methods for integrals", *Indag. Math.* (*N.S.*) **24**:4 (2013), 739–765. MR Zbl

[Wojdylo 2006] J. Wojdylo, "On the coefficients that arise from Laplace's method", *J. Comput. Appl. Math.* **196**:1 (2006), 241–266. MR Zbl

[Wong 1989] R. Wong, *Asymptotic approximations of integrals*, Academic Press, Boston, 1989. MR Zbl

[Wyman 1964] M. Wyman, "The method of Laplace", *Trans. Roy. Soc. Canada* (4) **2** (1964), 227–256. MR

Received June 14, 2017. Revised February 23, 2018.

Cormac O'Sullivan
Department of Mathematics
CUNY - Graduate Center
New York, NY
United States
cosullivan@gc.cuny.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.201

THE GAUSS–BONNET–CHERN MASS OF HIGHER-CODIMENSION GRAPHS

ALEXANDRE DE SOUSA AND FREDERICO GIRÃO

We give an explicit formula for the Gauss–Bonnet–Chern mass of an asymptotically flat graph of arbitrary codimension and use it to prove the positive mass theorem and the Penrose inequality for graphs with flat normal bundle.

1. Introduction

A complete Riemannian manifold (M^n, g), $n \geq 3$, is said to be asymptotically flat of order τ (with one end) if there exists a compact subset K of M and a diffeomorphism $\Psi : M \setminus K \to \mathbb{R}^n \setminus \bar{B}_1(0)$, introducing coordinates in $M \setminus K$, say $x = (x_1, \dots, x_n)$, such that, in these coordinates,

$$g_{ij} = \delta_{ij} + \sigma_{ij} \tag{1}$$

and

$$|\sigma_{ij}| + |x|\,|\sigma_{ij,k}| + |x|^2\,|\sigma_{ij,kl}| = O(|x|^{-\tau}), \tag{2}$$

where the σ_{ij}'s are the coefficients of σ with respect to x, $\sigma_{ij,k} = \partial\sigma_{ij}/\partial x_k$, $\sigma_{ij,kl} = \partial^2\sigma_{ij}/\partial x_k \partial x_l$, and $|\cdot|$ is the standard Euclidean norm. The ADM mass of (M, g), introduced by Arnowitt, Deser and Misner in [Arnowitt et al. 1961] (see also [Jaramillo and Gourgoulhon 2011]) is defined by

$$\mathrm{m}_{\mathrm{ADM}} = \frac{1}{2(n-1)\omega_{n-1}} \lim_{r\to\infty} \int_{S_r} (g_{ij,i} - g_{ii,j})\nu^j \, dS_r, \tag{3}$$

where ω_{n-1} is the volume of the $(n-1)$-dimensional unit sphere, S_r is the Euclidean coordinate sphere of radius r, dS_r is the volume form of S_r induced by the Euclidean metric, and $\nu = r^{-1}x$ is the outward unit normal to S_r (with respect to the Euclidean metric).

Alexandre de Sousa was a CAPES/Brazil doctoral fellow at Universidade Federal do Ceará. Frederico Girão was partially supported by CNPq/Brazil, grants number 483844/2013-6 and 306196/2016-6.
MSC2010: 53C40, 53C80, 83C99.
Keywords: Gauss–Bonnet–Chern mass, higher-codimension asymptotically flat graphs, positive mass theorem, Penrose inequality.

It is known that if $\tau > (n-2)/2$ and the scalar curvature of (M, g) is integrable, then the limit (3) exists, is finite, and is a geometric invariant, that is, two coordinate systems satisfying (1) and (2) yield the same value for it [Bartnik 1986; Chruściel 1986].

One of the most important conjectures in mathematical general relativity is the famous positive mass conjecture (PMC):

Conjecture 1. *If (M^n, g), $n \geq 3$, is an asymptotically flat Riemannian manifold of order $\tau > (n-2)/2$ whose scalar curvature is nonnegative and integrable, then the ADM mass of (M, g) is nonnegative. Moreover, if the mass is zero, then (M, g) is isometric to the Euclidean space $(\mathbb{R}^n, \delta)$.*

The PMC was settled for $n \leq 7$ in [Schoen and Yau 1979], for (M, g) conformally flat in [Schoen and Yau 1988], and for M spin in [Witten 1981] (see also [Parker and Taubes 1982] and [Choquet-Bruhat 1984]). Very elegant proofs for the case when (M, g) is a Euclidean graph were given in [Lam 2011] (see also [de Lima and Girão 2015]) for graphs of codimension one and in [Mirandola and Vitório 2015] for graphs of arbitrary codimension with flat normal bundle (notice that the case of graphs also follows from Witten's argument, since a Euclidean graph is spin). The case of Euclidean hypersurfaces (not necessarily graphs), including the rigidity statement, was treated in [Huang and Wu 2013], under appropriate decay conditions.

The Penrose inequality (PI) is a conjectured sharpening of the PMC when (M, g) has a compact boundary Γ which is an outermost minimal hypersurface.

Conjecture 2. *If (M^n, g), $n \geq 3$, is an asymptotically flat Riemannian manifold of order $\tau > (n-2)/2$ whose scalar curvature is nonnegative (and integrable), and Γ is a (possibly disconnected) outermost minimal hypersurface of area A, then*

$$\mathrm{m}_{\mathrm{ADM}} \geq \frac{1}{2}\left(\frac{A}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}}.$$

Furthermore, if the equality holds, then (M, g) is isometric to the Riemannian Schwarzschild manifold.

The PI was proved by Huisken and Ilmanen [2001] for $n = 3$ and Γ connected, and by Bray [2001] for $n = 3$ and general Γ. Bray and Lee [2009] established the conjecture for $n \leq 7$, with the extra requirement that M be spin for the rigidity statement. The case of Euclidean graphs of codimension one was treated by Lam [2011] (see also [de Lima and Girão 2015]) and generalized by Mirandola and Vitório for graphs of arbitrary codimension with flat normal bundle [2015]. The equality case for graphs of codimension one was treated in [Huang and Wu 2015a].

In [Ge et al. 2014b], a new mass for asymptotically flat Riemannian manifolds, named Gauss–Bonnet–Chern mass, was introduced. For a positive integer $q < n/2$, consider the q-th Gauss–Bonnet curvature, denoted $L_{(q)}$, and defined by

$$L_{(q)} = \frac{1}{2^q}\delta^{a_1a_2\cdots a_{2q}}_{b_1b_2\cdots b_{2q}}\left(\prod_{s=1}^{q} R_{a_{2s-1}a_{2s}}{}^{b_{2s-1}b_{2s}}\right) = P^{ijkl}_{(q)}R_{ijkl}, \tag{4}$$

where R is the Riemann curvature tensor of (M, g) and $P_{(q)}$, which has the same symmetries of the Riemann tensor (see [Ge et al. 2014b, Section 3]), is given by

$$P^{ijkl}_{(q)} = \frac{1}{2^q}\delta^{a_1a_2\cdots a_{2q-3}a_{2q-2}ij}_{b_1b_2\cdots b_{2q-3}b_{2q-2}b_{2q-1}b_{2q}}\left(\prod_{s=1}^{q-1} R_{a_{2s-1}a_{2s}}{}^{b_{2s-1}b_{2s}}\right)g^{b_{2q-1}k}g^{b_{2q}l}. \tag{5}$$

Remark 3. One can considerably simplify this complicated tensorial expression by rewriting it in the language of double forms, which are a special type of vector valued forms (see [Labbi 2007], for example).

The q-th Gauss–Bonnet–Chern mass (GBC mass) of (M, g) is defined by

$$\mathrm{m}_q = c_q(n)\lim_{r\to\infty}\int_{S_r} P^{ijkl}_{(q)}g_{jk,l}\nu_i\,dS_r, \tag{6}$$

where

$$c_q(n) = \frac{(n-2q)!}{2^{q-1}(n-1)!\,\omega_{n-1}} \tag{7}$$

and S_r, dS_r, ν and ω_{n-1} are as in the definition of the ADM mass.

As observed in [Ge et al. 2014b], m_1 coincides with the ADM mass. In the same article, the authors show that, if $\tau > (n-2q)/(q+1)$ and $L_{(q)}$ is integrable, then the limit (6) exists, is finite, and is a geometric invariant. Next, we state versions of the PMC and PI for the GBC mass. We start with the version of the PMC.

Conjecture 4. *Let n and q be integers such that $n \geq 3$ and $1 \leq q < n/2$. If (M^n, g) is an asymptotically flat Riemannian manifold of order $\tau > (n-2q)/(q+1)$ whose q-th Gauss–Bonnet curvature $L_{(q)}$ is nonnegative and integrable, then the q-th GBC mass of (M, g) is nonnegative. Moreover, if the mass is zero, then (M, g) is isometric to the Euclidean space $(\mathbb{R}^n, \delta)$.*

Before we state the analogue of the PI, we recall the Riemannian manifold known as the q-th Riemannian Schwarzschild [Ge et al. 2014b, Section 6], which is $(\mathbb{R}\times\mathbb{S}^{n-1}, g^q_{\mathrm{Sch}})$ with

$$g^q_{\mathrm{Sch}} = \left(1 + \frac{m}{2r^{\frac{n}{q}-2}}\right)^{\frac{4q}{n-2q}}(dr^2 + r^2d\theta^2),$$

where $d\theta^2$ is the round metric on $\mathbb{S}^{n-1}$ and $m \in \mathbb{R}$ is the mass parameter. Let $r_0 = (2m)^{q/(n-2q)}$. The hypersurface $r = r_0$ is an outermost minimal hypersurface of area $A = \omega_{n-1}r_0^{n-1}$, and the q-th GBC mass of $(\mathbb{R}\times\mathbb{S}^{n-1}, g^q_{\mathrm{Sch}})$ is $\mathrm{m}_q = m^q$.

Thus, for the q-th Riemannian Schwarzschild manifold, one has

$$\mathrm{m}_q = \frac{1}{2^q}\Big(\frac{A}{\omega_{n-1}}\Big)^{\frac{n-2q}{n-1}}.$$

We can now state the version of the PI for the GBC mass.

Conjecture 5. *Let n and q be integers such that $n \geq 3$ and $1 \leq q < n/2$. If (M^n, g) is an asymptotically flat Riemannian manifold of order $\tau > (n-2q)/(q+1)$ whose q-th Gauss–Bonnet curvature $L_{(q)}$ is nonnegative and integrable, and Γ is a (possibly disconnected) outermost minimal hypersurface of area A, then*

$$\mathrm{m}_q \geq \frac{1}{2^q}\Big(\frac{A}{\omega_{n-1}}\Big)^{\frac{n-2q}{n-1}},$$

where m_q is the q-th GBC mass. Moreover, if the equality holds, then (M, g) is isometric to the q-th Riemannian Schwarzschild manifold.

We now turn to the special case of graphs. Let Ω be a (possibly empty) bounded open subset of $\mathbb{R}^n$ such that $\Sigma = \partial\Omega$ is the union of finitely many smooth connected hypersurfaces. Let $f : \mathbb{R}^n \setminus \Omega \to \mathbb{R}^m$ be a continuous map such that its restriction to $\mathbb{R}^n \setminus \bar{\Omega}$ is smooth. Let f^α, $1 \leq \alpha \leq m$, be the components of f and let f_i^α, f_{ij}^α and f_{ijk}^α denote the first, second and third partial derivatives of f^α on $\mathbb{R}^n \setminus \bar{\Omega}$, where $1 \leq i, j, k \leq n$. The map f is said to be asymptotically flat of order τ if

$$|f_i^\alpha(x)| + |f_{ij}^\alpha(x)|\,|x| + |f_{ijk}^\alpha(x)|\,|x|^2 = O(|x|^{-\tau/2}), \tag{8}$$

for each $\alpha \in \{1, \dots, m\}$ and each triple (i, j, k) with $1 \leq i, j, k \leq n$.

We assume throughout the paper that

$$M = \{(x, f(x)) : x \in \mathbb{R}^n \setminus \Omega,\ f(x) \in \mathbb{R}^m\},$$

the graph of f, is a smooth submanifold with (possibly empty) boundary and that $g_{ij} = \delta_{ij} + f_i^\alpha f_j^\alpha$, the metric induced by the Euclidean metric on $\mathbb{R}^{n+m}$, extends to a smooth metric on M. Notice that if f is asymptotically flat of order τ, then from (8) we get that (M, g) is asymptotically flat of order τ.

Conjectures 4 and 5 have been proved for graphs of codimension one [Ge et al. 2014b]. When $q = 2$, Li, Wei and Xiong proved these conjectures for graphs of higher codimension with flat normal bundle [Li et al. 2014]. Conjecture 4 is also known to be true for conformally flat manifolds [Ge et al. 2014a].

The purpose of the present article is to prove Conjectures 4 and 5 for a family of higher-codimension Euclidean graphs (without the rigidity statements). This family includes the graphs with flat normal bundle. The exposition follows closely the ones given in [Lam 2011; Mirandola and Vitório 2015; Ge et al. 2014b; Li et al. 2014]. Before stating our main results, we need to introduce some notation.

Denote by $\{e_i\}_{i=1}^n$ the standard basis of $\mathbb{R}^n$ and by $\{e_\alpha\}_{\alpha=1}^m$ the standard basis of $\mathbb{R}^m$. The coordinate vector fields on M are given by $\partial_i = (e_i, f_i^\alpha e_\alpha)$, and the vector fields $\eta_\alpha = (-Df^\alpha, e_\alpha)$, where Df^α denotes the Euclidean gradient of f^α, give us a (global) frame field for the normal bundle of M. We denote by B the second fundamental form of M, by B_α its α-th component with respect to the frame $\{\eta_\alpha\}_{\alpha=1}^m$, and by A_α the shape operator with respect to η_α. Also, let $U = (U_{\alpha\beta})$ be the metric on the normal bundle induced by the Euclidean metric $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{n+m}$. The components of U are given by

$$U_{\alpha\beta} = \delta_{\alpha\beta} + \langle Df^\alpha, Df^\beta \rangle.$$

The inverse of U is denoted by $(U^{\alpha\beta})$.

Recall the Gauss and the Ricci equations, which are, respectively, given by

$$R_{ijkl} = \langle B_{ik}, B_{jl} \rangle - \langle B_{il}, B_{jk} \rangle \tag{9}$$

and

$$\langle R^\perp_{\alpha\beta}(X), Y \rangle = \langle [A_\beta, A_\alpha](X), Y \rangle, \tag{10}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean metric on $\mathbb{R}^{n+m}$, $R^\perp$ is the normal curvature operator and

$$[A_\beta, A_\alpha] = A_\alpha \circ A_\beta - A_\beta \circ A_\alpha.$$

We denote by $T_{(2q-1)}$ the Newton tensor of order $(2q-1)$ and denote by $T_{(2q-1)\alpha}$ its α-th component with respect to the frame $\{\eta_\alpha\}_{\alpha=1}^m$ (see [Grosjean 2002] and [Cao and Li 2007]). The expression for $T_{(2q-1)}$ in coordinates is

$$T_{(2q-1)i}{}^{j} = \frac{1}{(2q-1)!} \delta^{a_1 \cdots a_{2q-1} j}_{b_1 \cdots b_{2q-1} i} \left(\prod_{s=1}^{q-1} \langle B^{b_{2s-1}}_{a_{2s-1}}, B^{b_{2s}}_{a_{2s}} \rangle \right) B^{b_{2q-1}}_{a_{2q-1}}, \tag{11}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean metric on $\mathbb{R}^{n+m}$. As we will see in Section 2, if M has flat normal bundle, then $T_{(2q-1)\alpha}$ commutes with A_β, for $1 \le \alpha, \beta \le m$.

We can now state the main results of the article. The first of them is this:

Theorem 6. *Let n and q be integers such that $n \ge 3$ and $1 \le q < n/2$, and let (M, g) be the graph of an asymptotically flat map $f : \mathbb{R}^n \to \mathbb{R}^m$ of order $\tau > (n-2q)/(q+1)$. If the q-th Gauss–Bonnet curvature $L_{(q)}$ of (M, g) is integrable, then the q-th Gauss–Bonnet–Chern mass m_q satisfies*

$$\mathrm{m}_q = \frac{1}{2} c_q(n) \int_M \left(L_{(q)} + (2q-1)! \langle [T_{(2q-1)\alpha}, A_\beta] \cdot e_\alpha^\top, e_\beta^\top \rangle \right) \frac{1}{\sqrt{G}} \, dM, \tag{12}$$

where $c_q(n)$ is the constant (7), G is the determinant of (g_{ij}),

$$[T_{(2q-1)\alpha}, A_\beta] = T_{(2q-1)\alpha} \circ A_\beta - A_\beta \circ T_{(2q-1)\alpha}$$

is the commutator of the operators $T_{(2q-1)\alpha}$ and A_β, and $e_\alpha^\top$ is the tangent part (*along the graph* M) *of the canonical lift to* $\mathbb{R}^{n+m} \equiv \mathbb{R}^n \times \mathbb{R}^m$ *of the standard frame field on* $\mathbb{R}^m$. *Moreover*, *if* M *has flat normal bundle and* $L_{(q)}$ *is nonnegative*, *then* m_q *is nonnegative.*

Remark 7. Notice that, since the graph structure is used in the definition of the vector fields η_α, the tensor $[T_{(2q-1)\alpha}, A_\beta]$ is defined only for graphs. It is desirable to find an expression similar to (12) that holds for any asymptotically flat submanifold (not necessarily a graph), but we were unable to do it. One strategy in order to do this is to rewrite $[T_{(2q-1)\alpha}, A_\beta]$ in such a way that it also makes sense for submanifolds which are not necessarily graphs and try proving that (12) also holds in this case. Another strategy is to find a similar expression for the mass by considering, instead of the vector field (25) used to get (12), one which is defined for any submanifold (compare, for example, the vector fields considered in [Lam 2011] and [de Lima and Girão 2015]).

Let $\Sigma \subset \mathbb{R}^n$ be an orientable hypersurface and let ξ be a unit normal vector field along Σ (chosen to point outwards, whenever this makes sense). The r-th mean curvature of Σ is defined as the r-th elementary symmetric function on the principal curvatures of Σ. Alternatively, if K is the second fundamental form of $\Sigma \subset \mathbb{R}^n$, then

$$H_r = \frac{1}{r!}\delta^{a_1\cdots a_r}_{b_1\cdots b_r}\prod_{s=1}^{r} K^{b_s}_{a_s}. \tag{13}$$

The hypersurface $\Sigma \subset \mathbb{R}^n$ is called *strictly* p*-mean convex*, $1 \le p \le n-1$, if $H_r > 0$ for all $1 \le r \le p$. Our second main result is the following:

Theorem 8. *Let* n *and* q *be integers such that* $n \ge 3$ *and* $1 \le q < n/2$. *Let* Ω *be a bounded and open subset of* $\mathbb{R}^n$ *such that* $\Sigma = \partial\Omega$ *is the union of finitely many smooth connected hypersurfaces. Let* $f : \mathbb{R}^n \setminus \Omega \to \mathbb{R}^m$ *be an asymptotically flat map of order* $\tau > (n-2q)/(q+1)$, *and let* (M, g) *be the graph of* f. *Assume that* f *extends smoothly to an open set containing* $\mathbb{R}^n \setminus \Omega$ *and that* f *is constant along each connected component of* Σ. *If the* q*-th Gauss–Bonnet curvature* $L_{(q)}$ *is integrable*, *then the* q*-th Gauss–Bonnet–Chern mass* m_q *satisfies*

$$\mathrm{m}_q = \frac{1}{2}c_q(n)\int_M \big(L_{(q)} + (2q-1)!\langle\big[T_{(2q-1)\alpha}, A_\beta\big]\cdot e_\alpha^\top, e_\beta^\top\rangle\big)\frac{1}{\sqrt{G}}\,dM \\ + \frac{1}{2}(2q-1)!c_q(n)\int_\Sigma \left(\frac{|Df|^2}{1+|Df|^2}\right)^q H_{(2q-1)}\,d\Sigma,$$

where

$$|Df|^2 = \sum_{\alpha=1}^{m}|Df^\alpha|^2$$

and $H_{(2q-1)}$ *is the* $(2q-1)$*-th mean curvature of* Σ.

Our third main result is the following:

Theorem 9. *Let n and q be integers such that $n \geq 3$ and $1 \leq q < n/2$. Let Ω be a bounded and open subset of $\mathbb{R}^n$ such that $\Sigma = \partial\Omega$ is the union of finitely many smooth hypersurfaces. Let*

$$f : \mathbb{R}^n \setminus \Omega \to \mathbb{R}^m$$

be an asymptotically flat map of order

$$\tau > (n-2q)/(q+1),$$

and let (M, g) be the graph of f. Assume that f is constant along each connected component of Σ and that

$$|Df| \to \infty \quad as \quad x \to \Sigma.$$

If the q-th Gauss–Bonnet curvature $L_{(q)}$ is integrable, then the q-th Gauss–Bonnet–Chern mass m_q satisfies

$$\mathrm{m}_q = \frac{1}{2}c_q(n)\int_M \big(L_{(q)} + (2q-1)!\langle\!\langle [T_{(2q-1)\alpha}, A_\beta] \cdot e_\alpha^\top, e_\beta^\top \rangle\!\rangle\big)\frac{1}{\sqrt{G}}\, dM + \frac{1}{2}(2q-1)!c_q(n)\int_\Sigma H_{(2q-1)}\, d\Sigma, \tag{14}$$

where $H_{(2q-1)}$ is the $(2q-1)$-th mean curvature of Σ. Furthermore, if M has flat normal bundle, $L_{(q)}$ is nonnegative and each component of Σ is star-shaped and strictly $(2q-1)$-mean convex, then

$$\mathrm{m}_q \geq \frac{1}{2^q}\left(\frac{A}{\omega_{n-1}}\right)^{\frac{n-2q}{n-1}}. \tag{15}$$

Remark 10. As explained in [Ge et al. 2014b, Remark 5.1], when $\Sigma \subset \mathbb{R}^n$ is strictly mean convex, the condition

$$|Df| \to \infty \quad \text{as} \quad x \to \partial\Omega \tag{16}$$

holds if and only if $\Gamma = \partial M$ is an outermost minimal hypersurface. Therefore, this is a natural assumption.

Remark 11. Geometrically, condition (16) is equivalent to saying that along each connected component of ∂M, the graph M meets orthogonally the hyperplane that contains that component (see [de Sousa 2016]).

2. Auxiliary results

Let n and q be positive integers such that $n \geq 3$ and $1 \leq q < n/2$. Throughout this section, the tensors $P_{(q)}$ and $T_{(2q-1)}$ are defined by (5) and (11), respectively. Also, unless stated otherwise, we will follow the notation introduced in Section 1.

If Ω is not empty, we assume, throughout this section, that f extends smoothly to an open set containing $\mathbb{R}^n \setminus \Omega$.

Lemma 12. *Under the notation introduced above, the following identities hold*:

$$g_{jk,l} = f^\alpha_{jl} f^\alpha_k + f^\alpha_j f^\alpha_{kl} \tag{17}$$

$$e^\top_\alpha = \nabla f^\alpha = g^{ij} f^\alpha_j \partial_i = U^{\alpha\beta} f^\beta_i \partial_i \tag{18}$$

$$A_\alpha \partial_i = f^\alpha_{ik} g^{kj} \partial_j \tag{19}$$

$$B(\partial_i, \partial_j) = f^\alpha_{ij} U^{\alpha\beta} \eta^\beta \tag{20}$$

$$(B_\alpha)_{ij} = f^\alpha_{ij} \tag{21}$$

$$\Gamma^k_{ij} = g^{kl} f^\alpha_l f^\alpha_{ij} \tag{22}$$

$$\nabla e^\top_\alpha = \langle B, e_\alpha \rangle \tag{23}$$

Proof. Identities (17) to (22) are proven in [Mirandola and Vitório 2015] and [Li et al. 2014] and identity (23) is proved in [Palais and Terng 1988, Proposition 4.1.1]. □

On an open set that contains $\mathbb{R}^n \setminus \Omega$, consider the vector field $X_{(q)}$ given by

$$X_{(q)} = X^i_{(q)} \partial_i = P^{ijkl}_{(q)} g_{jk,l} \partial_i. \tag{24}$$

Proposition 13. *It holds that*

$$X_{(q)} = \frac{1}{2}(2q-1)!\, T_{(2q-1)\alpha} \cdot e^\top_\alpha. \tag{25}$$

Proof. By (24) and (17) we have

$$X^i_{(q)} = P^{ijkl}_{(q)} (f^\alpha_{jl} f^\alpha_k + f^\alpha_j f^\alpha_{kl}).$$

Using the antisymmetry of $P^{ijkl}_{(q)}$ with respect to the indices k and l, we have

$$X^i_{(q)} = P^{ijkl}_{(q)} f^\alpha_{jl} f^\alpha_k.$$

Combining this identity with (5), (19) and (18), we find

$$\begin{aligned}
X^i_{(q)} &= \frac{1}{2^q} \delta^{a_1 a_2 \cdots a_{2q-3} a_{2q-2} ij}_{b_1 b_2 \cdots b_{2q-3} b_{2q-2} cd} \left(\prod_{s=1}^{q-1} R_{a_{2s-1} a_{2s}}{}^{b_{2s-1} b_{2s}} \right) g^{ck} g^{dl} f^\alpha_{jl} f^\alpha_k \\
&= \frac{1}{2^q} \delta^{a_1 a_2 \cdots a_{2q-3} a_{2q-2} ij}_{b_1 b_2 \cdots b_{2q-3} b_{2q-2} cd} \left(\prod_{s=1}^{q-1} R_{a_{2s-1} a_{2s}}{}^{b_{2s-1} b_{2s}} \right) g^{dl} f^\alpha_{jl} g^{ck} f^\alpha_k \\
&= \frac{1}{2^q} \delta^{a_1 a_2 \cdots a_{2q-3} a_{2q-2} ij}_{b_1 b_2 \cdots b_{2q-3} b_{2q-2} cd} \left(\prod_{s=1}^{q-1} R_{a_{2s-1} a_{2s}}{}^{b_{2s-1} b_{2s}} \right) (A_\alpha)^d_j g^{ck} f^\alpha_k \\
&= \frac{1}{2^q} \delta^{a_1 a_2 \cdots a_{2q-3} a_{2q-2} ij}_{b_1 b_2 \cdots b_{2q-3} b_{2q-2} cd} \left(\prod_{s=1}^{q-1} R_{a_{2s-1} a_{2s}}{}^{b_{2s-1} b_{2s}} \right) (A_\alpha)^d_j (\nabla f^\alpha)^c.
\end{aligned}$$

Hence, using (9), (11), (18) and switching i with j and c with d, we find

$$\begin{aligned}X^i_{(q)} &= \frac{2^{q-1}}{2^q}\delta^{a_1a_2\cdots a_{2q-3}a_{2q-2}ij}_{b_1b_2\cdots b_{2q-3}b_{2q-2}cd}\left(\prod_{s=1}^{q-1}\langle B^{b_{2s-1}}_{a_{2s-1}}, B^{b_{2s}}_{a_{2s}}\rangle\right)(A_\alpha)^d_j(\nabla f^\alpha)^c \\ &= \frac{1}{2}\delta^{a_1a_2\cdots a_{2q-3}a_{2q-2}ji}_{b_1b_2\cdots b_{2q-3}b_{2q-2}dc}\left(\prod_{s=1}^{q-1}\langle B^{b_{2s-1}}_{a_{2s-1}}, B^{b_{2s}}_{a_{2s}}\rangle\right)(A_\alpha)^d_j(\nabla f^\alpha)^c \\ &= \frac{1}{2}(2q-1)!(T_{(2q-1)\alpha})^i_c(\nabla f^\alpha)^c \\ &= \frac{1}{2}(2q-1)!(T_{(2q-1)\alpha}\cdot\nabla f^\alpha)^i \\ &= \frac{1}{2}(2q-1)!(T_{(2q-1)\alpha}\cdot e^\top_\alpha)^i. \qquad\square\end{aligned}$$

The next identity is a higher-codimensional version of Proposition 3.5(b) in [Reilly 1973] (see also [Alías et al. 2006, Section 8]).

Proposition 14. *It holds that*

$$\operatorname{div}_e X = \tfrac{1}{2}L_{(q)} + \tfrac{1}{2}(2q-1)!\big\langle\big[T_{(2q-1)\alpha}, A_\beta\big]\cdot e^\top_\alpha, e^\top_\beta\big\rangle,$$

where div_e *denotes the Euclidean divergence.*

Proof. Using the identity

$$\operatorname{div}_e X_{(q)} = \partial_i X^i_{(q)}$$

and the identities (18), (21) and (22), we have

$$\begin{aligned}\operatorname{div}_g X_{(q)} = \nabla_i X^i_{(q)} &= \partial_i X^i_{(q)} + \Gamma^i_{ij}X^j_{(q)} \\ &= \operatorname{div}_e X_{(q)} + (e^\top_\beta)^i(B_\beta)_{ij}X^j_{(q)} \\ &= \operatorname{div}_e X_{(q)} + \langle A_\beta\cdot X_{(q)}, e^\top_\beta\rangle \\ &= \operatorname{div}_e X_{(q)} + \tfrac{1}{2}(2q-1)!\langle(A_\beta\circ T_{(2q-1)\alpha})\cdot e^\top_\alpha, e^\top_\beta\rangle.\end{aligned}$$

By the expression for the vector field $X_{(q)}$ established in the previous proposition, it follows that

$$\begin{aligned}\operatorname{div}_g X_{(q)} &= \tfrac{1}{2}(2q-1)!\ \operatorname{div}_g(T_{(2q-1)\beta}\cdot e^\top_\beta) \\ &= \tfrac{1}{2}(2q-1)!\big[\operatorname{div}_g(T_{(2q-1)\beta})\cdot e^\top_\beta + T_{(2q-1)\beta}\cdot\nabla e^\top_\beta\big].\end{aligned}$$

By (18) and (23), the identities

$$\nabla e^\top_\beta = \langle B, e_\beta\rangle = U^{\gamma\alpha}\langle\eta_\gamma, e_\beta\rangle B_\alpha = U^{\beta\alpha}B_\alpha$$

hold. Therefore, the Gauss equation together with identities (11) and (4) give

$$T_{(2q-1)\beta}\cdot\nabla e^\top_\beta = U^{\beta\alpha}T_{(2q-1)\beta}\cdot B_\alpha = \frac{1}{(2q-1)!}L_{(q)}.$$

Thus,

$$\operatorname{div}_e X_{(q)} = \tfrac{1}{2} L_{(q)} + \tfrac{1}{2}(2q-1)!\big[\operatorname{div}_g(T_{(2q-1)\beta}) \cdot e_\beta^\top - \big\langle (A_\beta \circ T_{(2q-1)\alpha}) \cdot e_\alpha^\top, e_\beta^\top \big\rangle\big].$$

Recall that the Newton tensors of a submanifold of Euclidean space are divergence free (see, for example, Lemmata 3.1 and 3.2 of [Cao and Li 2007]) and that each of the fields in the normal frame is given by the expression $\eta_\beta = (-Df^\beta, e_\beta)$. Hence, using identity (18), we have

$$\begin{aligned}
(\operatorname{div}_g T_{(2q-1)\beta})_j &= \nabla_i (T_{(2q-1)\beta})^i{}_j = \nabla_i \big\langle (T_{(2q-1)})^i{}_j, \eta_\beta \big\rangle \\
&= \big\langle \nabla_i^\perp (T_{(2q-1)})^i{}_j, \eta_\beta \big\rangle + \big\langle (T_{(2q-1)})^i{}_j, \nabla_i^\perp \eta_\beta \big\rangle \\
&= \big\langle (\operatorname{div}\ T_{(2q-1)})_j, \eta_\beta \big\rangle + U^{\gamma\alpha} (T_{(2q-1)\alpha})^i{}_j \big\langle \eta_\gamma, \bar{D}_i \eta_\beta \big\rangle \\
&= (T_{(2q-1)\alpha})^i{}_j U^{\gamma\alpha} f_k^\gamma f_{ik}^\beta = (T_{(2q-1)\alpha})^i{}_j (e_\alpha^\top)^k (B_\beta)_{ik} \\
&= \big((T_{(2q-1)\alpha} \circ A_\beta) \cdot e_\alpha^\top \big)_j,
\end{aligned}$$

where $\bar{D}$ is the Levi-Civita connection of the ambient space $\mathbb{R}^{n+m} \equiv \mathbb{R}^n \times \mathbb{R}^m$. Therefore,

$$\begin{aligned}
\operatorname{div}_e X_{(q)} &= \tfrac{1}{2} L_{(q)} + \tfrac{1}{2}(2q-1)! \big\langle \big(T_{(2q-1)\alpha} \circ A_\beta - A_\beta \circ T_{(2q-1)\alpha} \big) \cdot e_\alpha^\top, e_\beta^\top \big\rangle \\
&= \tfrac{1}{2} L_{(q)} + \tfrac{1}{2}(2q-1)! \big\langle [T_{(2q-1)\alpha}, A_\beta] \cdot e_\alpha^\top, e_\beta^\top \big\rangle. \qquad \square
\end{aligned}$$

Proposition 15. *For a level set $\Sigma \subset \mathbb{R}^n$ in the domain of a Euclidean graph, the identity*

$$\langle X_{(q)}, \xi \rangle = -\frac{1}{2}(2q-1)! \left(\frac{|Df|^2}{1+|Df|^2} \right)^q H_{(2q-1)}$$

holds, where ξ denotes a unit normal vector field along Σ (chosen to point outwards, whenever this makes sense).

Proof. We have

$$\langle X_{(q)}, \xi \rangle = \tfrac{1}{2}(2q-1)! (T_{(2q-1)\alpha} \cdot e_\alpha^\top)^i \xi_i.$$

Let $x \in \Sigma$. Rotate the coordinates such that, at x, $e_1 = \xi$ and $\{e_A\}_{A=2}^n$ is an orthonormal frame for the tangent space of Σ at x. With respect to this new frame $\{e_i\}_{i=1}^n$ on $\mathbb{R}^n$,

$$\xi_i = \delta_i^1,$$

for $i = 1, \dots, n$. Thus,

$$\langle X_{(q)}, \xi \rangle = \tfrac{1}{2}(2q-1)! (T_{(2q-1)\alpha} \cdot e_\alpha^\top)^1.$$

As in [Li et al. 2014, Section 4], we find that the inverse of g is given by

$$\begin{aligned}
g^{11} &= \frac{1}{1+|Df|^2}, \\
g^{A1} &= 0, \\
g^{AB} &= \delta^{AB}.
\end{aligned}$$

It follows that

$$e_\alpha^\top = \frac{f_1^\alpha}{1+|Df|^2}\partial_1 = \frac{\langle Df^\alpha, \xi\rangle}{1+|Df|^2}\partial_1.$$

Therefore,

$$\langle X_{(q)}, \xi\rangle = \frac{1}{2}(2q-1)!\frac{\langle Df^\alpha, \xi\rangle}{1+|Df|^2}(T_{(2q-1)\alpha})_1^1.$$

Since, by (9) and (11),

$$(T_{(2q-1)\alpha})_1^1 = \frac{1}{2^{q-1}}\frac{1}{(2q-1)!}\delta^{a_1\cdots a_{2q-1}1}_{b_1\cdots b_{2q-1}1}\Bigg(\prod_{s=1}^{q-1} R_{a_{2s-1}a_{2s}}{}^{b_{2s-1}b_{2s}}\Bigg)(A_\alpha)^{b_{2q-1}}_{a_{2q-1}},$$

using the antisymmetry of $\delta^{a_1\cdots a_{2q-1}1}_{b_1\cdots b_{2q-1}1}$ we find that

$$(T_{(2q-1)\alpha})_1^1 = \frac{1}{2^{q-1}}\frac{1}{(2q-1)!}\delta^{A_1\cdots A_{2q-1}1}_{B_1\cdots B_{2q-1}1}\Bigg(\prod_{s=1}^{q-1} R_{A_{2s-1}A_{2s}}{}^{B_{2s-1}B_{2s}}\Bigg)(A_\alpha)^{B_{2q-1}}_{A_{2q-1}}.$$

Recall that the generalized Kronecker delta is a determinant. Using the $2q$-th column to expand it, we find

$$\delta^{A_1\cdots A_{2q-1}1}_{B_1\cdots B_{2q-1}1} = \delta^{A_1\cdots A_{2q-1}}_{B_1\cdots B_{2q-1}}.$$

Hence,

$$(T_{(2q-1)\alpha})_1^1 = \frac{1}{2^{q-1}}\frac{1}{(2q-1)!}\delta^{A_1\cdots A_{2q-1}}_{B_1\cdots B_{2q-1}}\Bigg(\prod_{s=1}^{q-1} R_{A_{2s-1}A_{2s}}{}^{B_{2s-1}B_{2s}}\Bigg)(A_\alpha)^{B_{2q-1}}_{A_{2q-1}}.$$

Let $\hat{R}$ denote the Riemann curvature tensor of Σ, and denote by K and $\tilde{K}$, respectively, the second fundamental form of Σ as a hypersurface of $\mathbb{R}^n$ and the second fundamental form of $f(\Sigma)$ as a hypersurface of (M, g). By equations (4.3) and (4.4) of [Li et al. 2014], we have

$$\tilde{K} = \frac{K}{\sqrt{1+|Df|^2}}$$

and

$$R_{AB}{}^{CD} = \frac{|Df|^2}{1+|Df|^2}\hat{R}_{AB}{}^{CD}.$$

Plugging this into the expression for $(T_{(2q-1)\alpha})_1^1$, we find

$$\begin{aligned}\big(T_{(2q-1)\alpha}\big)_1^1 = {}& \frac{1}{2^{q-1}}\frac{1}{(2q-1)!}\bigg(\frac{|Df|^2}{1+|Df|^2}\bigg)^{q-1} \\ & \times \delta^{A_1\cdots A_{2q-1}}_{B_1\cdots B_{2q-1}}\Bigg(\prod_{s=1}^{q-1} \hat{R}_{A_{2s-1}A_{2s}}{}^{B_{2s-1}B_{2s}}\Bigg)(A_\alpha)^{B_{2q-1}}_{A_{2q-1}}.\end{aligned}$$

From

$$\eta^\alpha = e_\alpha - Df^\alpha = e_\alpha - \langle Df^\alpha, \xi\rangle\xi,$$

it follows that

$$(A_\alpha)_A^B = -\langle Df^\alpha, \xi\rangle K_A^B.$$

Also, the Gauss equation applied to $\Sigma \subset \mathbb{R}^n$ yields

$$\hat{R}_{AB}{}^{CD} = K_A^C K_B^D - K_A^D K_B^C.$$

We then conclude that

$$\begin{aligned}
(T_{(2q-1)\alpha})_1^1 &= -\frac{1}{2^{q-1}}\frac{1}{(2q-1)!}\langle Df^\alpha, \xi\rangle\left(\frac{|Df|^2}{1+|Df|^2}\right)^{q-1} \\
&\qquad\times \delta^{A_1\cdots A_{2q-1}}_{B_1\cdots B_{2q-1}}\left(\prod_{s=1}^{q-1}\hat{R}_{A_{2s-1}A_{2s}}{}^{B_{2s-1}B_{2s}}\right)(K)^{B_{2q-1}}_{A_{2q-1}} \\
&= -\frac{1}{(2q-1)!}\langle Df^\alpha, \xi\rangle\left(\frac{|Df|^2}{1+|Df|^2}\right)^{q-1} \\
&\qquad\times \delta^{A_1\cdots A_{2q-1}}_{B_1\cdots B_{2q-1}}\left(\prod_{s=1}^{q-1}K^{B_{2s-1}}_{A_{2s-1}}K^{B_{2s}}_{A_{2s}}\right)K^{B_{2q-1}}_{A_{2q-1}} \\
&= -\langle Df^\alpha, \xi\rangle\left(\frac{|Df|^2}{1+|Df|^2}\right)^{q-1} H_{(2q-1)},
\end{aligned}$$

where we have used the expression (13) to obtain the last equality. It follows that

$$\begin{aligned}
\langle X_{(q)}, \xi\rangle &= \frac{1}{2}(2q-1)!\frac{\langle Df^\alpha, \xi\rangle}{1+|Df|^2}(T_{(2q-1)\alpha})_1^1 \\
&= -\frac{1}{2}(2q-1)!\frac{\langle Df^\alpha, \xi\rangle^2}{1+|Df|^2}\left(\frac{|Df|^2}{1+|Df|^2}\right)^{q-1} H_{(2q-1)} \\
&= -\frac{1}{2}(2q-1)!\left(\frac{|Df|^2}{1+|Df|^2}\right)^{q} H_{(2q-1)},
\end{aligned}$$

where, to obtain the last equality, we have used that

$$Df^\alpha = \langle Df^\alpha, \xi\rangle\xi$$

implies

$$|Df|^2 = \sum_\alpha \langle Df^\alpha, \xi\rangle^2. \qquad \square$$

Remark 16. In Proposition 15, the expression $|Df|^2/(1+|Df|^2)$ is the cosine of the angle between the graph and the hyperplane containing its boundary (see [de Sousa 2016]).

3. Proof of the theorems

Suppose first that M has no boundary. Let S_r be an Euclidean coordinate sphere of radius r. By (6), (24) and the divergence theorem, we have

$$\begin{aligned} \mathrm{m}_q &= c_q(n) \lim_{r\to\infty} \int_{S_r} X^i_{(q)} \xi_i \, dS_r \\ &= c_q(n) \int_{\mathbb{R}^n} \mathrm{div}_e \, X_{(q)} \, dV, \end{aligned}$$

where dV denotes the Euclidean volume form. Thus, invoking Proposition 14 and using that

$$dV = \frac{1}{\sqrt{G}} dM, \tag{26}$$

we find

$$\mathrm{m}_q = \frac{1}{2} c_q(n) \int_M \big(L_{(q)} + (2q-1)! \langle \big[T_{(2q-1)\alpha}, A_\beta \big] \cdot e_\alpha^\top, e_\beta^\top \rangle \big) \frac{1}{\sqrt{G}} \, dM,$$

which is exactly the first part of Theorem 6.

To prove the second part of Theorem 6, notice that, from equations (3) and (6) of [Andrzejewski et al. 2016], the tensor $T_{(2q-1)\alpha}$ can be written as a polynomial on the A_α's. Also, if M has flat normal bundle, then the Ricci equation (10) yields

$$[A_\alpha, A_\beta] = 0, \tag{27}$$

for all $\alpha, \beta \in \{1, \ldots, m\}$. Thus, using (27) several times, we find

$$[T_{(2q-1)\alpha}, A_\beta] = 0,$$

for all $\alpha, \beta \in \{1, \ldots, m\}$. Hence, (12) becomes

$$\mathrm{m}_q = \frac{1}{2} c_q(n) \int_M L_{(q)} \frac{1}{\sqrt{G}} \, dM.$$

Therefore, if $L_{(q)}$ is nonnegative, then m_q is nonnegative. This finishes the proof of Theorem 6.

Suppose now that ∂M is not empty and that f can be extended to a smooth map on some open set containing $\mathbb{R}^n \setminus \Omega$. This assumption allows us to use the results of Section 2. Equations (6), (24) and the divergence theorem yield

$$\begin{aligned} \mathrm{m}_q &= c_q(n) \lim_{r\to\infty} \int_{S_r} X^i_{(q)} \nu_i \, dS_r \\ &= c_q(n) \int_{\mathbb{R}^n \setminus \Omega} \mathrm{div}_e \, X_{(q)} \, dV - c_q(n) \int_\Sigma \langle X_{(q)}, \xi \rangle \, d\Sigma. \end{aligned}$$

Invoking Propositions 14 and 15 and (26), we get

$$\mathrm{m}_q = \frac{1}{2}c_q(n)\int_M \big(L_{(q)} + (2q-1)!\langle\big[T_{(2q-1)\alpha}, A_\beta\big]\cdot e_\alpha^\top, e_\beta^\top\rangle\big)\frac{1}{\sqrt{G}}\,dM \\ + \frac{1}{2}(2q-1)!c_q(n)\int_\Sigma \left(\frac{|Df|^2}{1+|Df|^2}\right)^q H_{(2q-1)}\,d\Sigma. \tag{28}$$

This finishes the proof of Theorem 8.

Let us now prove Theorem 9. We cannot use (28) directly, since, by hypothesis,

$$|Df| \to \infty \ \text{ as } \ x \to \Sigma,$$

and hence, it is not possible to extend f to a smooth function on some open set containing $\mathbb{R}^n \setminus \Omega$. To circumvent this problem, we proceed as in the last section of [Mirandola and Vitório 2015]. Namely, we consider an approximating sequence

$$F^k = (f^{1,k}, \ldots, f^{m,k}) : \mathbb{R}^n \setminus \Omega \to \mathbb{R}^m,$$

$k \in \mathbb{N}$, of smooth maps such that each F^k extends to a smooth map on some open set containing $\mathbb{R}^n \setminus \Omega$. We then apply (28) to each F^k and take the limit as $k \to \infty$, reaching (14).

It remains to prove inequality (15). If Σ has only one component then, by a result of Guan and Li [2009], it holds that

$$\frac{1}{2}(2q-1)!c_q(n)\int_\Sigma H_{(2q-1)}\,d\Sigma \geq \frac{1}{2^q}\left(\frac{|\Sigma|}{\omega_{n-1}}\right)^{\frac{n-2q}{n-1}}, \tag{29}$$

with equality holding if and only if Σ is a round sphere.

Suppose now that Σ has more than one component. Recall that if $x_1, \ldots, x_j$ are nonnegative real numbers and $0 \leq s < 1$ then

$$\sum_{i=1}^{j} x_i^s \geq \left(\sum_{i=1}^{j} x_i\right)^s, \tag{30}$$

with equality holding if and only if at most one of the x_i's is positive (see [Huang and Wu 2015b, Proposition 5.2]). Inequality (15) then follows by combining inequalities (29) and (30).

Remark 17. Unfortunately our methods are not suitable to deal with the equality cases, that is, to prove the rigidity statements contained in Conjectures 4 and 5. If equality holds in Theorem 6, then we can only conclude that the Gauss–Bonnet curvature $L_{(q)}$ is identically zero. If equality holds in Theorem 9, then we can only conclude that $L_{(q)}$ is zero on M and that Σ has only one component which is a round sphere.

References

[Alías et al. 2006] L. J. Alías, J. H. S. de Lira, and J. M. Malacarne, "Constant higher-order mean curvature hypersurfaces in Riemannian spaces", *J. Inst. Math. Jussieu* **5**:4 (2006), 527–562. MR Zbl

[Andrzejewski et al. 2016] K. Andrzejewski, W. Kozłowski, and K. Niedziałomski, "Generalized Newton transformation and its applications to extrinsic geometry", *Asian J. Math.* **20**:2 (2016), 293–322. MR Zbl

[Arnowitt et al. 1961] R. Arnowitt, S. Deser, and C. W. Misner, "Coordinate invariance and energy expressions in general relativity", *Phys. Rev.* (2) **122** (1961), 997–1006. MR Zbl

[Bartnik 1986] R. Bartnik, "The mass of an asymptotically flat manifold", *Comm. Pure Appl. Math.* **39**:5 (1986), 661–693. MR Zbl

[Bray 2001] H. L. Bray, "Proof of the Riemannian Penrose inequality using the positive mass theorem", *J. Differential Geom.* **59**:2 (2001), 177–267. MR Zbl

[Bray and Lee 2009] H. L. Bray and D. A. Lee, "On the Riemannian Penrose inequality in dimensions less than eight", *Duke Math. J.* **148**:1 (2009), 81–106. MR Zbl

[Cao and Li 2007] L. Cao and H. Li, "r-minimal submanifolds in space forms", *Ann. Global Anal. Geom.* **32**:4 (2007), 311–341. MR Zbl

[Choquet-Bruhat 1984] Y. Choquet-Bruhat, "Positive-energy theorems", pp. 739–785 in *Relativity, groups and topology, II* (Les Houches, France, 1983), edited by B. S. DeWitt and R. Stora, North-Holland, Amsterdam, 1984. MR Zbl

[Chruściel 1986] P. Chruściel, "Boundary conditions at spatial infinity from a Hamiltonian point of view", pp. 49–59 in *Topological properties and global structure of space-time* (Erice, 1985), edited by P. G. Bergmann and V. De Sabbata, NATO Adv. Sci. Inst. Ser. B Phys. **138**, Plenum, New York, 1986. MR Zbl

[Ge et al. 2014a] Y. Ge, G. Wang, and J. Wu, "The Gauss–Bonnet–Chern mass of conformally flat manifolds", *Int. Math. Res. Not.* **2014**:17 (2014), 4855–4878. MR Zbl

[Ge et al. 2014b] Y. Ge, G. Wang, and J. Wu, "A new mass for asymptotically flat manifolds", *Adv. Math.* **266** (2014), 84–119. MR Zbl

[Grosjean 2002] J.-F. Grosjean, "Upper bounds for the first eigenvalue of the Laplacian on compact submanifolds", *Pacific J. Math.* **206**:1 (2002), 93–112. MR Zbl

[Guan and Li 2009] P. Guan and J. Li, "The quermassintegral inequalities for k-convex starshaped domains", *Adv. Math.* **221**:5 (2009), 1725–1732. MR Zbl

[Huang and Wu 2013] L.-H. Huang and D. Wu, "Hypersurfaces with nonnegative scalar curvature", *J. Differential Geom.* **95**:2 (2013), 249–278. MR Zbl

[Huang and Wu 2015a] L.-H. Huang and D. Wu, "The equality case of the Penrose inequality for asymptotically flat graphs", *Trans. Amer. Math. Soc.* **367**:1 (2015), 31–47. MR Zbl

[Huang and Wu 2015b] L.-H. Huang and D. Wu, "The equality case of the Penrose inequality for asymptotically flat graphs", *Trans. Amer. Math. Soc.* **367**:1 (2015), 31–47. MR Zbl

[Huisken and Ilmanen 2001] G. Huisken and T. Ilmanen, "The inverse mean curvature flow and the Riemannian Penrose inequality", *J. Differential Geom.* **59**:3 (2001), 353–437. MR Zbl

[Jaramillo and Gourgoulhon 2011] J. L. Jaramillo and E. Gourgoulhon, "Mass and angular momentum in general relativity", pp. 87–124 in *Mass and motion in general relativity* (Orléans, France, 2008), edited by L. Blanchet et al., Fundamental Theories of Physics **162**, Springer, 2011. Zbl

[Labbi 2007] M. L. Labbi, "On Gauss–Bonnet curvatures", *Symmetry Integrability Geom. Methods Appl.* **3** (2007), art. id. 118. MR Zbl

[Lam 2011] M.-K. G. Lam, *The graph cases of the Riemannian positive mass and Penrose inequalities in all dimensions*, Ph.D. thesis, Duke University, 2011, Available at https://search.proquest.com/docview/868529318.

[Li et al. 2014] H. Li, Y. Wei, and C. Xiong, "The Gauss–Bonnet–Chern mass for graphic manifolds", *Ann. Global Anal. Geom.* **45**:4 (2014), 251–266. MR Zbl

[de Lima and Girão 2015] L. L. de Lima and F. Girão, "The ADM mass of asymptotically flat hypersurfaces", *Trans. Amer. Math. Soc.* **367**:9 (2015), 6247–6266. MR Zbl

[Mirandola and Vitório 2015] H. Mirandola and F. Vitório, "The positive mass theorem and Penrose inequality for graphical manifolds", *Comm. Anal. Geom.* **23**:2 (2015), 273–292. MR Zbl

[Palais and Terng 1988] R. S. Palais and C.-L. Terng, *Critical point theory and submanifold geometry*, Lecture Notes in Math. **1353**, Springer, 1988. MR Zbl

[Parker and Taubes 1982] T. Parker and C. H. Taubes, "On Witten's proof of the positive energy theorem", *Comm. Math. Phys.* **84**:2 (1982), 223–238. MR Zbl

[Reilly 1973] R. C. Reilly, "On the Hessian of a function and the curvatures of its graph", *Michigan Math. J.* **20** (1973), 373–383. MR Zbl

[Schoen and Yau 1979] R. Schoen and S. T. Yau, "On the proof of the positive mass conjecture in general relativity", *Comm. Math. Phys.* **65**:1 (1979), 45–76. MR Zbl

[Schoen and Yau 1988] R. Schoen and S.-T. Yau, "Conformally flat manifolds, Kleinian groups and scalar curvature", *Invent. Math.* **92**:1 (1988), 47–71. MR Zbl

[de Sousa 2016] A. de Sousa, *Um breve estudo sobre a massa de Gauss–Bonnet–Chern dos gráficos euclidianos*, Ph.D. thesis, Universidade Federal do Ceará, 2016.

[Witten 1981] E. Witten, "A new proof of the positive energy theorem", *Comm. Math. Phys.* **80**:3 (1981), 381–402. MR Zbl

Received April 4, 2017. Revised January 16, 2018.

ALEXANDRE DE SOUSA
INSTITUTO DE MATEMÁTICA
UNIVERSIDADE FEDERAL DE ALAGOAS
MACEIÓ
BRAZIL
a.asm@protonmail.com

FREDERICO GIRÃO
DEPARTAMENTO DE MATEMÁTICA
UNIVERSIDADE FEDERAL DO CEARÁ
FORTALEZA
BRAZIL
fred@mat.ufc.br

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.217

THE ASYMPTOTIC BOUNDS OF VISCOSITY SOLUTIONS OF THE CAUCHY PROBLEM FOR HAMILTON–JACOBI EQUATIONS

KAIZHI WANG

We study the Cauchy problem for time-periodic Hamilton–Jacobi equations with Tonelli Hamiltonians. It is well known that the Cauchy problem admits a unique bounded viscosity solution. We provide a more precise description of the boundedness of the viscosity solution. We introduce the notion of asymptotic bounds of the viscosity solution of the Cauchy problem. An asymptotic bound is a 1-periodic viscosity solution of the Hamilton–Jacobi equation. We show how to obtain the optimal asymptotic bounds, i.e., minimal asymptotic upper bound and maximal asymptotic lower bound. Our method relies upon Mather theory and weak KAM theory on Lagrangian dynamics.

1. Introduction and main result

Consider the Hamilton–Jacobi equation

$$u_t + H(x, u_x, t) = c(H), \quad x \in M,\ t \in [0, +\infty), \tag{1-1}$$

where H is a Tonelli Hamiltonian, the constant $c(H)$ is the Mañé critical value of H [Mañé 1997], and M is a compact and connected smooth manifold without boundary. We choose, once and for all, a C^∞ Riemannian metric g on M. A C^2 function $H : T^*M \times \mathbb{R}^1 \to \mathbb{R}^1$ is called a Tonelli Hamiltonian if:

- (periodicity) H is 1-periodic in t.
- (strict convexity) For each $(x, p, t) \in T^*M \times \mathbb{R}^1$, the second partial derivative $\partial^2 H/\partial p^2(x, p, t)$ is positive definite.
- (superlinear growth) $\lim_{\|p\|_x \to +\infty} H(x, p, t)/\|p\|_x = +\infty$ uniformly on $x \in M$, $t \in \mathbb{R}^1$, where $\|\cdot\|_x$ denotes the norm on T_x^*M induced by g.
- (completeness of the Hamiltonian vector field) Each integral curve of the Hamiltonian vector field is defined on all of $\mathbb{R}^1$.

MSC2010: 35B40, 35F25, 37J99.
Keywords: Hamilton–Jacobi equations, viscosity solutions, Cauchy problem, asymptotic bounds, weak KAM theory.

For the Hamiltonian H, we can define the associated Lagrangian as its Fenchel–Legendre transform:

$$L : TM \times \mathbb{R}^1 \to \mathbb{R}^1, \quad (x, v, t) \mapsto \sup_{p \in T_x^* M} \{\langle p, v\rangle_x - H(x, p, t)\},$$

where $\langle \cdot, \cdot \rangle_x$ represents the canonical pairing between the tangent and cotangent space. Since H is a Tonelli Hamiltonian, one can easily prove that L is finite everywhere, a C^2 function, 1-periodic in t, superlinear and strictly convex in v, and that the Euler–Lagrange flow is complete. Such a Lagrangian will be called a Tonelli Lagrangian.

The Cauchy problem for (1-1) with Tonelli Hamiltonian H is well posed in the viscosity sense: given a continuous function $\varphi : M \to \mathbb{R}^1$, (1-1) admits a unique viscosity solution $u : M \times [0, +\infty) \to \mathbb{R}^1$ with $u|_{t=0} = \varphi$; see, e.g., [Lions 1982]. The notion of viscosity solutions was introduced by Crandall and Lions [1983] in the study of Hamilton–Jacobi equations.

Let $\tilde{H} = H - c(H)$. Then (1-1) can be rewritten as

$$u_t + \tilde{H}(x, u_x, t) = 0.$$

Since $\tilde{H}$ is still a Tonelli Hamiltonian and the Mañé critical value of $\tilde{H}$ is 0 (see Section 2A), then in the following we always assume that $c(H) = 0$ and consider the Hamilton–Jacobi equation

$$u_t + H(x, u_x, t) = 0, \quad x \in M, \ t \in [0, +\infty). \tag{1-2}$$

There exist viscosity solutions of (1-2) which are 1-periodic in time; see, e.g., [Wang and Yan 2012]. More precisely, Wang and Yan [2012] introduced a new kind of Lax–Oleinik type operators in the context of weak KAM theory [Fathi 2005]. The family of the new operators with an arbitrary continuous function φ on M as initial condition converges to a 1-periodic viscosity solution of (1-2). Moreover, using this method one can obtain all the 1-periodic viscosity solutions of (1-2). There is a nice representation formula for 1-periodic viscosity solutions: $u(x, t) = \inf_{y \in M}(\varphi(y) + h_{0,[t]}(y, x))$, where h is the Peierls barrier and $[t] = t \bmod 1$; see Section 2 for details. In addition, Wang and Yan also showed [2012] that weak KAM solutions and 1-periodic viscosity solutions of (1-2) are the same.

All the viscosity solutions of (1-2) are bounded; see Proposition 2.1. In general, it is not true that the viscosity solution converges, as $t \to +\infty$, to a 1-periodic viscosity solution; see [Barles and Souganidis 2000; Fathi and Mather 2000]. Roquejoffre [2001] and Bernard and Roquejoffre [2004] proved that the viscosity solution converges to a T-periodic viscosity solution in several nontrivial special cases, where T may be greater than 1.

In the present paper we aim to give a more precise description of the boundedness of the viscosity solution of (1-2). Denote by $u_\varphi(x, t)$ the unique viscosity solution of the Cauchy problem

$$\begin{cases} u_t + H(x, u_x, t) = 0 & \text{in } M \times (0, +\infty), \\ u(x, 0) = \varphi(x) & \text{on } M, \end{cases} \tag{1-3}$$

where $\varphi : M \to \mathbb{R}^1$ is a continuous function.

Definition 1.1 (asymptotic bounds). (i) We say that a 1-periodic viscosity solution v of (1-2) is an asymptotic upper bound (AUB) of u_φ, if for each $\epsilon > 0$ there exists $T > 0$ such that

$$u_\varphi(x, t) \le v(x, t) + \epsilon, \quad \forall (x, t) \in M \times [T, +\infty). \tag{1-4}$$

Furthermore, a function $\bar{v} : M \times [0, +\infty) \to \mathbb{R}^1$ is called the minimal asymptotic upper bound (min AUB) of u_φ, if it is an AUB and for each AUB v, we have

$$\bar{v}(x, t) \le v(x, t), \quad \forall (x, t) \in M \times [0, +\infty).$$

(ii) We say that a 1-periodic viscosity solution w of (1-2) is an asymptotic lower bound (ALB) of u_φ, if for each $\epsilon > 0$ there exists $T > 0$ such that

$$u_\varphi(x, t) \ge w(x, t) - \epsilon, \quad \forall (x, t) \in M \times [T, +\infty). \tag{1-5}$$

Furthermore, a function $\underline{w} : M \times [0, +\infty) \to \mathbb{R}^1$ is called the maximal asymptotic lower bound (max ALB) of u_φ, if it is an ALB and for each ALB w, we have

$$\underline{w}(x, t) \ge w(x, t), \quad \forall (x, t) \in M \times [0, +\infty).$$

Remark 1.1. (i) The existence of the AUB and the ALB of u_φ follows immediately from the existence of 1-periodic viscosity solutions of (1-2) and the boundedness of u_φ.

(ii) We assert that if the min AUB $\bar{v}$ and the max ALB $\underline{w}$ of u_φ exist, then for each $\epsilon > 0$ and each $T > 0$, there exist $(\bar{x}, \bar{t})$, $(\underline{x}, \underline{t}) \in M \times [T, +\infty)$ such that

$$u_\varphi(\bar{x}, \bar{t}) > \bar{v}(\bar{x}, \bar{t}) - \epsilon, \quad u_\varphi(\underline{x}, \underline{t}) < \underline{\mathrm{w}}(\underline{x}, \underline{t}) + \epsilon.$$

To show the first assertion, we argue by contradiction. For, otherwise, there would be $\epsilon_0 > 0$ and $T_0 > 0$ such that

$$u_\varphi(x, t) \le \bar{v}(x, t) - \epsilon_0 =: v'(x, t), \quad \forall (x, t) \in M \times [T_0, +\infty).$$

Note that v' is a 1-periodic viscosity solution of (1-2) and that for each $\epsilon > 0$,

$$u_\varphi(x, t) \le v'(x, t) \le v'(x, t) + \epsilon, \quad \forall (x, t) \in M \times [T_0, +\infty).$$

By definition, v' is an AUB of u_φ. Recall that $v'(x, t) := \bar{v}(x, t) - \epsilon_0$. Thus, we obtain a contradiction to the assumption that $\bar{v}$ is the min AUB of u_φ. By a similar argument, one can prove the second assertion.

The major result of the paper is as follows:

Theorem 1.1. *For every continuous function $\varphi : M \to \mathbb{R}^1$, both the min AUB and the max ALB of u_φ exist. Let $\bar{\varphi}$ and $\underline{\varphi}$ denote the min AUB and the max ALB respectively. Then there is a constant $C > 0$ such that*

$$|\bar{\varphi}(x,t) - \underline{\varphi}(x,t)| \le C, \quad \forall M \times [0,+\infty), \tag{1-6}$$

where C is independent of φ.

For a given $\varphi \in C(M, \mathbb{R}^1)$, we will show how to obtain $\bar{\varphi}$ and $\underline{\varphi}$ in Section 3. An outline of this paper is as follows. Section 2 includes some basic definitions and preliminary results. Section 3 is devoted to the proof of Theorem 1.1.

In recent years, many convergence results on the asymptotic behavior of viscosity solutions of Hamilton–Jacobi equations with the Hamiltonian independent of t have been obtained by various authors since the pioneering work of Lions [1982] and Barles [1985]. Among them, it is worth mentioning in particular that dynamical techniques were used first by Fathi [1998] and Roquejoffre [2001] to attack such problems. See [Ishii 2006] for more details.

2. Preliminaries

In this section we introduce the notation used in the sequel and review some definitions and results of Mather theory and weak KAM theory [Fathi 2005; Mañé 1997; Mather 1991; 1993]. We view $\mathbb{S}^1$ as a fundamental domain in $\mathbb{R}^1 : \bar{I} = [0, 1]$ with the two endpoints identified. The standard universal covering projection $\pi : \mathbb{R}^1 \to \mathbb{S}^1$ takes the form $\pi(t) = [t]$, where $[t] = t \mod 1$ denotes the fractional part of t.

The L-action of a continuous and piecewise C^1 curve $\gamma : [a,b] \to M$ is defined by

$$A_L(\gamma) = \int_a^b L(d\gamma(\sigma), \sigma)\, d\sigma,$$

where $d\gamma : [a,b] \to TM$ denotes the differential of γ.

2A. ***Mañé critical value.*** The notion of the critical value of autonomous Tonelli Hamiltonians (or Lagrangians) was introduced by Mañé [1997]; see also [Contreras et al. 1997]. We can define the critical value of time-periodic Tonelli Hamiltonians (or Lagrangians) in a similar way. Contreras et al. [2013] gave the following property of the critical value for time-periodic case, which can also be regarded as equivalent definitions of the critical value.

$$\begin{aligned} c(H) &:= \inf\{k \in \mathbb{R}^1 : A_{L+k}(\gamma) \ge 0 \text{ for all absolutely continuous closed curves } \gamma\} \\ &= \sup\{k \in \mathbb{R}^1 : A_{L+k}(\gamma) < 0 \text{ for some absolutely continuous closed curve } \gamma\}, \end{aligned}$$

where $L(x, v, t) = \sup_{p \in T_x^* M}\{\langle p, v\rangle_x - H(x, p, t)\}$ and a curve $\gamma : [a,b] \to M$

will be called closed if $\gamma(a) = \gamma(b)$ and $b - a$ is an integer. It is straightforward to verify that the critical value of $H - c(H)$ is 0.

2B. ***Lax–Oleinik semigroup and viscosity solutions.*** For each $t \geq 0$ and each $\varphi \in C(M, \mathbb{R}^1)$, let

$$T_t\varphi(x) = \inf_{y \in M}\big\{\varphi(y) + \inf_{\gamma} A_L(\gamma)\big\}$$

for all $x \in M$, where the second infimum is taken among the continuous and piecewise C^1 paths $\gamma : [0, t] \to M$ with $\gamma(0) = y$ and $\gamma(t) = x$. For each $t \geq 0$, T_t is an operator from $C(M, \mathbb{R}^1)$ to itself. Since L is time-periodic, $\{T_n\}_{n \in \mathbb{N}}$ is a one-parameter semigroup of operators, called the Lax–Oleinik semigroup associated with L, where $\mathbb{N} = \{0, 1, 2, \ldots\}$. By the definition of T_t, one can easily verify the following properties:

(i) $T_{n+t}\varphi(x) = T_t \circ T_n\varphi(x), \quad \forall n \in \mathbb{N},\ \forall t \geq 0.$

(ii) The function $(x, t) \mapsto T_t\varphi(x)$ is continuous on $M \times [0, +\infty)$, for each $\varphi \in C(M, \mathbb{R}^1)$.

(iii) For each $\varphi_1, \varphi_2 \in C(M, \mathbb{R}^1)$ and each $t \geq 0$, we have

$$\varphi_1 \leq \varphi_2 \Rightarrow T_t\varphi_1 \leq T_t\varphi_2. \qquad \text{(monotonicity)}$$

(iv) For each $\varphi_1, \varphi_2 \in C(M, \mathbb{R}^1)$ and each $t \geq 0$, we have

$$\|T_t\varphi_1 - T_t\varphi_2\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty, \qquad \text{(nonexpansiveness)}$$

where $\|\cdot\|_\infty$ denotes the supremum norm in the space $C(M, \mathbb{R}^1)$.

As mentioned in Section 1, the Cauchy problem (1-3) is well posed in the viscosity sense. Furthermore, $u_\varphi(x, t) = T_t\varphi(x)$, for all $(x, t) \in M \times [0, +\infty)$, which means that $T_\cdot\varphi(\,\cdot\,)$ is the unique viscosity solution of (1-3); see, e.g., [Fathi and Mather 2000].

2C. ***Peierls barrier.*** As in [Mather 1993], it is convenient to introduce, for all $t < t' \in \mathbb{R}^1$ and $x, x' \in M$, the quantity

$$F_{t,t'}(x, x') = \inf_{\gamma} A_L(\gamma),$$

where the infimum is taken over the continuous and piecewise C^1 paths $\gamma : [t, t'] \to M$ such that $\gamma(t) = x$ and $\gamma(t') = x'$. For all $t < t' \in \mathbb{R}$ and all $x, x' \in M$, A_L takes a finite minimum value over the set of continuous and piecewise C^1 paths $\gamma : [t, t'] \to M$ such that $\gamma(t) = x$ and $\gamma(t') = x'$ [Mather 1991]. For each $t \geq 0$, each $\varphi \in C(M, \mathbb{R}^1)$ and each $x \in M$, it is easy to see that

$$T_t\varphi(x) = \inf_{y \in M}\big\{\varphi(y) + F_{0,t}(y, x)\big\}. \tag{2-1}$$

The following lemma [Bernard 2002] will be useful later. Recall that the critical value of the Lagrangian is 0. Such a Lagrangian is called critical in [Bernard 2002].

Lemma 2.1. *The function*

$$F : \mathbb{R}^1 \times \mathbb{R}^1 \times M \times M \to \mathbb{R}^1, \quad (t, t', x, x') \mapsto F_{t,t'}(x, x')$$

is Lipschitz and bounded on $\{t' \geq t + 1\}$.

In light of (ii) in Section 2B, (2-1) and Lemma 2.1, we have the following proposition.

Proposition 2.1. *All the viscosity solutions of* (1-2) *are bounded on* $M \times [0, +\infty)$.

Recall the notion of Peierls barrier introduced in [Mather 1993], which is the main ingredient in Mather's approach. Define the Peierls barrier as

$$h_{s,s'}(x, x') = \liminf_{t'-t\to+\infty} F_{t,t'}(x, x'), \qquad s, s' \in \mathbb{S}^1,\ x, x' \in M,$$

where the lim inf is restricted to the set of $(t, t') \in \mathbb{R}^2$ such that $s = [t],\ s' = [t']$. In view of Lemma 2.1, the lim inf in the definition exists. It is clear that

$$h_{s,s'}(x, x') = \liminf_{n\to+\infty} F_{s,s'+n}(x, x') = \lim_{n\to+\infty} \inf_{k\geq n} F_{s,s'+k}(x, x').$$

Again by Lemma 2.1, the family of functions $\{\inf_{k\geq n} F_{s,s'+k}(x, x')\}_n$ is equi-Lipschitz and thus

$$h_{s,s'}(x, x') = \lim_{n\to+\infty} \inf_{k\geq n} F_{s,s'+k}(x, x') \tag{2-2}$$

uniformly on $\mathbb{S}^1 \times \mathbb{S}^1 \times M \times M$. An important property of the Peierls barrier is that it is Lipschitz; see [Contreras et al. 2013].

Lemma 2.2. *Given any* $s \in \mathbb{S}^1,\ t \in \mathbb{R}^1$ *with* $[t] = s$ *and any* $x, y \in M$,

(i) $h_{0,s}(x, y) = \inf_{z\in M}(h_{0,0}(x, z) + F_{0,t}(z, y))$,

(ii) $h_{0,s}(x, y) = \inf_{z\in M}(F_{0,t}(x, z) + h_{0,0}(z, y))$.

Proof. (i) The inequality

$$\Big|\inf_{z\in M}\Big(\inf_{k\geq n} F_{0,k}(x, z) + F_{0,t}(z, y)\Big) - \inf_{z\in M}(h_{0,0}(x, z) + F_{0,t}(z, y))\Big| \leq \sup_{z\in M}\Big|\inf_{k\geq n} F_{0,k}(x, z) - h_{0,0}(x, z)\Big|,$$

together with (2-2) implies that

$$\lim_{n\to+\infty} \inf_{z\in M}\Big(\inf_{k\geq n} F_{0,k}(x, z) + F_{0,t}(z, y)\Big) = \inf_{z\in M}(h_{0,0}(x, z) + F_{0,t}(z, y)).$$

This equality can be rewritten as

$$\begin{aligned}\inf_{z\in M}(h_{0,0}(x,z)+F_{0,t}(z,y)) &= \lim_{n\to+\infty}\inf_{z\in M}\Bigl(\inf_{k\geq n}F_{0,k}(x,z)+F_{0,t}(z,y)\Bigr)\\ &= \lim_{n\to+\infty}\inf_{z\in M}\Bigl(\inf_{k\geq n}(F_{0,k}(x,z)+F_{k,k+t}(z,y))\Bigr)\\ &= \lim_{n\to+\infty}\inf_{k\geq n}\Bigl(\inf_{z\in M}(F_{0,k}(x,z)+F_{k,k+t}(z,y))\Bigr)\\ &= \lim_{n\to+\infty}\inf_{k\geq n}F_{0,k+t}(x,y)\\ &= h_{0,s}(x,y),\end{aligned}$$

and therefore (i) holds.

The proof of (ii) follows in a similar manner. □

2D. *Mañé potential and 1-periodic viscosity subsolutions.* For each $(s,s')\in\mathbb{S}^1\times\mathbb{S}^1$, let

$$\Phi_{s,s'}(x,x')=\inf F_{t,t'}(x,x')$$

for all $(x,x')\in M\times M$, where the infimum is taken on the set of $(t,t')\in\mathbb{R}^2$ such that $s=[t],\ s'=[t']$ and $t'\geq t+1$. This quantity is commonly called Mañé potential [Mañé 1997].

Lemma 2.3. *A continuous function $u:M\times\mathbb{S}^1\to\mathbb{R}^1$ is a viscosity subsolution of (1-2) only if*

$$u(x',s')-u(x,s)\leq\Phi_{s,s'}(x,x'),\quad \forall(x,s),\ (x',s')\in M\times\mathbb{S}^1.$$

See, e.g., [Fathi 2005] for a proof.

2E. *Weak KAM solutions and 1-periodic viscosity solutions.* A function $u:M\times\mathbb{S}^1\to\mathbb{R}^1$ is called a weak KAM solution of (1-2) if u is a viscosity subsolution of (1-2) and if, for every $(x,s)\in M\times\mathbb{S}^1$ there exists a curve $\gamma:(-\infty,s]\to M$ with $\gamma(s)=x$ such that

$$w(x,s)-w(\gamma(t),[t])=\int_t^s L(d\gamma(\sigma),\sigma)\,d\sigma,\quad \forall t\in(-\infty,s].$$

Denote by $\mathcal{S}$ the set of weak KAM solutions.

Let us recall two elementary results [Contreras et al. 2013] about weak KAM solutions.

(i) Given $(x_0,s_0)\in M\times\mathbb{S}^1$, define $u^*(x,s):=h_{s_0,s}(x_0,x)$. Then $u^*\in\mathcal{S}$.

(ii) If $\mathcal{U}\subset\mathcal{S}$, let $u_*(x,s):=\inf_{u\in\mathcal{U}}u(x,s)$, then either $u_*\equiv-\infty$ or $u_*\in\mathcal{S}$.

In view of (i) and (ii), it is clear that for each $\varphi\in C(M,\mathbb{R}^1)$,

$$\underline{\varphi}(x,s):=\inf_{y\in M}(\varphi(y)+h_{0,s}(y,x))\in\mathcal{S}. \tag{2-3}$$

In Section 3, we will show that $\underline{\varphi}$ is the max ALB of u_φ.

The following result was proved in [Wang and Yan 2012].

Proposition 2.2. *Weak KAM solutions and* 1*-periodic viscosity solutions of* (1-2) *are the same.*

2F. *Projected Aubry sets and weak KAM solutions.* Recall the definition of the projected Aubry set $\mathcal{A}_0$:

$$\mathcal{A}_0 := \{(x, s) \in M \times \mathbb{S}^1 \mid h_{s,s}(x, x) = 0\}.$$

Define an equivalence relation on $\mathcal{A}_0$ by saying that (x, s) and (x', s') are equivalent if and only if

$$h_{s,s'}(x, x') + h_{s',s}(x', x) = 0.$$

The equivalent classes of this relation are called static classes. Let A be the set of static classes. For each static class $\Gamma \in$ A choose a point $(x, 0) \in \Gamma$ and let $\mathbb{A}_0$ be the set of such points.

The following result in [Contreras et al. 2013] characterizes weak KAM solutions of (1-2) in terms of their values at each static class and the Peierls barrier.

Proposition 2.3. *Let u be a weak KAM solution of* (1-2)*. Then we have*

$$u(x, [t]) = \min_{(p,0)\in\mathbb{A}_0} (u(p, 0) + h_{0,[t]}(p, x)), \quad \forall (x, t) \in M \times [0, +\infty). \tag{2-4}$$

3. Proof of Theorem 1.1

To prove the main result, we need some more auxiliary results.

Proposition 3.1. *Given a continuous function $\varphi : M \to \mathbb{R}^1$, we have*

(i) $\liminf_{n\to+\infty} T_{n+t}\varphi(x) = \inf_{y\in M}(\varphi(y)+h_{0,[t]}(y, x)) = \underline{\varphi}(x, [t])$, *for all* $(x, t) \in M \times [0, +\infty)$, *where* $\underline{\varphi}$ *denotes the function we have defined in* (2-3)*.*

(ii) $\liminf_{n\to+\infty} T_{n+t}\varphi(x) = T_t(\liminf_{n\to+\infty} T_n\varphi)(x)$, *for all* $(x, t) \in M \times [0, +\infty)$.

Proof. (i) Note that

$$\begin{aligned}\Big|\inf_{k\ge n} T_{k+t}\varphi(x) - \inf_{y\in M}(\varphi(y) + h_{0,[t]}(y, x))\Big| \\ &= \Big|\inf_{k\ge n}\inf_{y\in M}(\varphi(y) + F_{0,k+t}(y, x)) - \inf_{y\in M}(\varphi(y) + h_{0,[t]}(y, x))\Big| \\ &= \Big|\inf_{y\in M}(\varphi(y) + \inf_{k\ge n} F_{0,k+t}(y, x)) - \inf_{y\in M}(\varphi(y) + h_{0,[t]}(y, x))\Big| \\ &\le \sup_{y\in M}\Big|\inf_{k\ge n} F_{0,k+t}(y, x) - h_{0,[t]}(y, x)\Big|.\end{aligned}$$

Taking (2-2) into consideration, we have

$$\liminf_{n\to+\infty} T_{n+t}\varphi(x) = \lim_{n\to+\infty}\inf_{k\ge n} T_{k+t}\varphi(x) = \inf_{y\in M}(\varphi(y) + h_{0,[t]}(y, x)).$$

(ii) Let

$$\psi(x) = \underline{\varphi}(x,0), \quad \forall x \in M.$$

Then by (i), we have

$$\psi(x) = \inf_{y\in M}(\varphi(y) + h_{0,0}(y,x)) = \liminf_{n\to+\infty} T_n\varphi(x),$$

and thus it suffices to show that $T_t\psi(x) = \inf_{y\in M}(\varphi(y)+h_{0,[t]}(y,x))$, for all $x \in M$, for all $t \geq 0$. It is clear that

$$\begin{aligned} T_t\psi(x) &= \inf_{y\in M}(\psi(y) + F_{0,t}(y,x)) \\ &= \inf_{y\in M}\big(\inf_{z\in M}(\varphi(z) + h_{0,0}(z,y)) + F_{0,t}(y,x)\big) \\ &= \inf_{z\in M}\big(\varphi(z) + \inf_{y\in M}(h_{0,0}(z,y) + F_{0,t}(y,x))\big). \end{aligned} \tag{3-1}$$

Combining (3-1) and (i) of Lemma 2.2, we get

$$T_t\psi(x) = \inf_{z\in M}(\varphi(z) + h_{0,[t]}(z,x)).$$

□

Proposition 3.2. *Given a continuous function* $\varphi : M \to \mathbb{R}^1$, $\limsup_{n\to+\infty} T_{n+t}\varphi(x)$ *exists for all* $(x,t) \in M \times [0,+\infty)$. *Let*

$$\tilde{\varphi}(x,t) = \limsup_{n\to+\infty} T_{n+t}\varphi(x), \quad \forall (x,t) \in M \times [0,+\infty).$$

Then $\tilde{\varphi}$ *is a* 1*-periodic viscosity subsolution of* (1-2).

Proof. It is apparent from Lemma 2.1 that $\limsup_{n\to+\infty} T_{n+t}\varphi(x)$ exists for all $(x,t) \in M \times [0,+\infty)$. Since

$$\begin{aligned} \tilde{\varphi}(x,t+1) &= \lim_{n\to+\infty}\sup_{k\geq n} T_{k+t+1}\varphi(x) = \lim_{n\to+\infty}\sup_{k\geq n+1} T_{k+t}\varphi(x) \\ &= \lim_{n\to+\infty}\sup_{k\geq n} T_{k+t}\varphi(x) = \tilde{\varphi}(x,t) \end{aligned}$$

for all $(x,t) \in M \times [0,+\infty)$, then $\tilde{\varphi}$ is 1-periodic in t. Therefore, in order to complete the proof, by Lemma 2.3 we only need to show that

$$\tilde{\varphi}(x',s') - \tilde{\varphi}(x,s) \leq \Phi_{s,s'}(x,x'), \quad \forall (x,s),\ (x',s') \in M \times \mathbb{S}^1.$$

For any positive integer m, we have

$$\begin{aligned} &\sup_{k\geq n}\inf_{y\in M}(\varphi(y) + F_{0,k+m+s'}(y,x')) - \sup_{k\geq n}\inf_{y\in M}(\varphi(y) + F_{0,k+s}(y,x)) \\ &\qquad \leq \sup_{k\geq n}\big(\inf_{y\in M}(\varphi(y) + F_{0,k+m+s'}(y,x')) - \inf_{y\in M}(\varphi(y) + F_{0,k+s}(y,x))\big) \\ &\qquad \leq \sup_{k\geq n}\sup_{y\in M}(F_{0,k+m+s'}(y,x') - F_{0,k+s}(y,x)) \\ &\qquad \leq \sup_{k\geq n} F_{k+s,k+m+s'}(x,x') = F_{s,m+s'}(x,x'). \end{aligned}$$

By taking the limit for $n \to +\infty$, we find

$$\tilde{\varphi}(x', m+s') - \tilde{\varphi}(x, s) \le F_{s,m+s'}(x, x').$$

Since m is an arbitrary positive integer and $\tilde{\varphi}$ is 1-periodic in t, by the definition of Mañé potential, we have

$$\tilde{\varphi}(x', s') - \tilde{\varphi}(x, s) \le \Phi_{s,s'}(x, x'). \qquad \square$$

Proposition 3.3. *Let $\varphi \in C(M, \mathbb{R}^1)$ and $\hat{\varphi}(x) = \limsup_{n\to+\infty} T_n\varphi(x)$, for all $x \in M$. Then*

$$\hat{\varphi}(x) \le T_1\hat{\varphi}(x) \le \cdots \le T_n\hat{\varphi}(x) \le \cdots, \quad \forall x \in M, \tag{3-2}$$

and the uniform limit $\lim_{n\to+\infty} T_n\hat{\varphi}(x)$ exists. Let $\varphi_\infty(x) = \lim_{n\to+\infty} T_n\hat{\varphi}(x)$ and $\bar{\varphi}(x, t) = T_t\varphi_\infty(x)$, for all $x \in M$, for all $t \ge 0$. Then φ_∞ is a fixed point of T_1 and $\bar{\varphi}$ is a 1-periodic viscosity solution of (1-2).

Remark 3.1. It is easy to check that

$$\bar{\varphi}(x, t) = \lim_{n\to+\infty} T_{n+t}\hat{\varphi}(x), \quad \forall x \in M, \ \forall t \ge 0.$$

Proof of Proposition 3.3. By Proposition 3.2, $\hat{\varphi}$ is well defined. In view of the definition of $\hat{\varphi}$, for any $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that

$$T_n\varphi(x) \le \hat{\varphi}(x) + \epsilon,$$

for all $x \in M$ and all $n \ge N_0$. Using the monotonicity of the Lax–Oleinik operator, we see that

$$T_1 \circ T_n(x) \le T_1\hat{\varphi}(x) + \epsilon,$$

for all $x \in M$ and all $n \ge N_0$. Combining the above inequality, (i) in Section 2B and the definition of $\hat{\varphi}$, we have

$$\hat{\varphi}(x) = \lim_{n\to+\infty} \sup_{k\ge n+1} T_k\varphi(x) = \limsup_{n\to+\infty} T_{n+1}\varphi(x) \le T_1\hat{\varphi}(x) + \epsilon, \quad \forall x \in M.$$

It follows that

$$\hat{\varphi}(x) \le T_1\hat{\varphi}(x), \quad \forall x \in M,$$

since ϵ may be taken arbitrarily small. Again by the monotonicity of the Lax–Oleinik operator, we get that

$$\hat{\varphi}(x) \le T_1\hat{\varphi}(x) \le \cdots \le T_n\hat{\varphi}(x) \le \cdots, \quad \forall x \in M.$$

From Lemma 2.1, it is easy to see that $\{T_n\hat{\varphi}\}_n$ is uniformly bounded and equi-Lipschitz. Therefore, the uniform limit $\lim_{n\to+\infty} T_n\hat{\varphi}(x)$ exists, i.e.,

$$\varphi_\infty(x) := \lim_{n\to+\infty} T_n\hat{\varphi}(x), \tag{3-3}$$

uniformly on $x \in M$.

The following inequality which comes from the nonexpansiveness of the Lax–Oleinik semigroup;

$$\|T_n\hat{\varphi} - T_1\varphi_\infty\|_\infty \le \|T_{n-1}\hat{\varphi} - \varphi_\infty\|_\infty,$$

together with (3-3), implies that

$$\varphi_\infty(x) = \lim_{n\to+\infty} T_n\hat{\varphi}(x) = T_1\varphi_\infty(x), \quad \forall x \in M, \tag{3-4}$$

namely φ_∞ is a fixed point of T_1.

Finally we prove that $\bar{\varphi}(x,t) := T_t\varphi_\infty(x)$ is a 1-periodic viscosity solution of (1-2). As mentioned in Section 2B, $T_t\varphi_\infty(x)$ is a viscosity solution of (1-2). Thus, it suffices to show that $\bar{\varphi}(x,t)$ is 1-periodic in t. By (i) in Section 2B and (3-4),

$$\bar{\varphi}(x,t+1) = T_{t+1}\varphi_\infty(x) = T_t \circ T_1\varphi_\infty(x) = T_t\varphi_\infty(x) = \bar{\varphi}(x,t),\ \forall x \in M,\ \forall t \ge 0.$$

The proof of the proposition is now complete. □

Lemma 3.1. *For $\bar{\varphi}$ and $\hat{\varphi}$ defined in Proposition 3.3, we have $\bar{\varphi}(x,0) = \hat{\varphi}(x)$, for all $(x,0) \in \mathbb{A}_0$.*

Proof. By (3-2), we have $\hat{\varphi}(x) \le T_n\hat{\varphi}(x)$ for all $x \in M$ and all $n \in \mathbb{N}$, which implies

$$\hat{\varphi}(x) \le \lim_{n\to+\infty} T_n\hat{\varphi}(x) = \bar{\varphi}(x,0), \quad \forall x \in M.$$

On the other hand, for each $(x,0) \in \mathbb{A}_0$, we have

$$\begin{aligned}\bar{\varphi}(x,0) &= \lim_{n\to+\infty} T_n\hat{\varphi}(x)\\ &= \lim_{n\to+\infty}\inf_{y\in M}(\hat{\varphi}(y) + F_{0,n}(y,x)) \le \hat{\varphi}(x) + \liminf_{n\to+\infty} F_{0,n}(x,x)\\ &= \hat{\varphi}(x) + h_{0,0}(x,x) = \hat{\varphi}(x).\end{aligned}$$ □

Proof of Theorem 1.1. We divide our proof in three steps. First, we show that $\bar{\varphi}$ in Proposition 3.3 is an AUB of u_φ and that $\underline{\varphi}$ in (2-3) is an ALB of u_φ. Since we have shown that $\bar{\varphi}$ and $\underline{\varphi}$ are 1-periodic viscosity solutions of (1-2), it suffices to prove that $\bar{\varphi}$ and $\underline{\varphi}$ satisfy (1-4) and (1-5), respectively. Next, we prove that $\bar{\varphi}$ is the min AUB of u_φ and that $\underline{\varphi}$ is the max ALB of u_φ. Finally, we need to show that (1-6) holds for some constant $C > 0$ which depends only on L.

Step 1. Our task now is to verify that $\bar{\varphi}$ and $\underline{\varphi}$ are AUB and ALB of u_φ, respectively.

First, we show that $\bar{\varphi}$ satisfies (1-4), which implies that $\bar{\varphi}$ is an AUB of u_φ. From the definition of $\hat{\varphi}$, for every $\epsilon > 0$, there exists $N_1 \in \mathbb{N}$ such that

$$T_n\varphi(x) - \hat{\varphi}(x) \le \sup_{k\ge n} T_k\varphi(x) - \hat{\varphi}(x) \le \epsilon, \quad \forall x \in M,\ \forall n \ge N_1. \tag{3-5}$$

Combining (3-2) and (3-5), we have

$$T_n\varphi(x) - \varphi_\infty(x) \le \epsilon, \quad \forall x \in M,\ \forall n \ge N_1.$$

Again by the monotonicity of the Lax–Oleinik operator, we get

$$T_\tau \circ T_n\varphi(x) \le T_\tau\varphi_\infty(x) + \epsilon, \quad \forall x \in M,\ \forall \tau \in [0, 1],\ \forall n \ge N_1.$$

In view of (i) in Section 2B and the definition of $\bar{\varphi}$, the above inequality implies

$$T_{n+\tau}\varphi(x) \le \bar{\varphi}(x, \tau) + \epsilon, \quad \forall x \in M,\ \forall \tau \in [0, 1],\ \forall n \ge N_1.$$

Since $u_\varphi(x, t) = T_t\varphi(x)$ and $\bar{\varphi}(x, t)$ is 1-periodic in t, we have

$$u_\varphi(x, n+\tau) \le \bar{\varphi}(x, n+\tau) + \epsilon, \quad \forall x \in M,\ \forall \tau \in [0, 1],\ \forall n \ge N_1,$$

i.e.,

$$u_\varphi(x, t) \le \bar{\varphi}(x, t) + \epsilon, \quad \forall (x, t) \in M \times [N_1, +\infty).$$

Hence, $\bar{\varphi}$ is an AUB of u_φ.

Then, we show that $\underline{\varphi}$ satisfies (1-5), which implies that $\underline{\varphi}$ is an ALB of u_φ. For each $\epsilon > 0$, by (2-2) there exists $N_2 \in \mathbb{N}$ such that

$$\inf_{k\ge n} F_{0,k+\tau}(y, x) - h_{0,[\tau]}(y, x) \ge -\epsilon, \quad \forall x,\ y \in M,\ \forall \tau \in [0, 1], \tag{3-6}$$

if $n \ge N_2$. Since

$$\inf_{y\in M}(\varphi(y) + h_{0,[\tau]}(y, x)) - \inf_{y\in M}(\varphi(y) + F_{0,n+\tau}(y, x)) \le \sup_{y\in M}(h_{0,[\tau]}(y, x) - F_{0,n+\tau}(y, x)),$$

then by (3-6), we have

$$\inf_{y\in M}(\varphi(y) + h_{0,[\tau]}(y, x)) - \inf_{y\in M}(\varphi(y) + F_{0,n+\tau}(y, x)) \le \epsilon, \quad \forall x \in M,\ \forall \tau \in [0, 1],\ \forall n \ge N_2. \tag{3-7}$$

From the definition of $\underline{\varphi}$, (2-1) and $u_\varphi(x, t) = T_t\varphi(x)$, (3-7) becomes

$$\underline{\varphi}(x, \tau) - u_\varphi(x, n+\tau) \le \epsilon, \quad \forall x \in M,\ \forall \tau \in [0, 1],\ \forall n \ge N_2.$$

Since $\underline{\varphi}$ is 1-periodic in t, we get

$$\underline{\varphi}(x, n+\tau) - u_\varphi(x, n+\tau) \le \epsilon, \quad \forall x \in M,\ \forall \tau \in [0, 1],\ \forall n \ge N_2,$$

i.e.,

$$u_\varphi(x, t) \ge \underline{\varphi}(x, t) - \epsilon, \quad \forall (x, t) \in M \times [N_2, +\infty).$$

Hence, $\underline{\varphi}$ is an ALB of u_φ.

Step 2. We are now in a position to show that $\bar{\varphi}$ is the min AUB of u_φ and that $\underline{\varphi}$ is the max ALB of u_φ.

First, we prove that $\bar{\varphi}$ is the min AUB of u_φ, by contradiction. Otherwise, there would be an AUB v and a point $(x_0, t_0) \in M \times [0, +\infty)$ such that

$$v(x_0, t_0) < \bar{\varphi}(x_0, t_0). \tag{3-8}$$

Note that v and $\bar{\varphi}$ are both 1-periodic viscosity solutions of (1-2). In view of Propositions 2.2 and 2.3, we have

$$\begin{aligned} v(x,t) &= \min_{(p,0)\in\mathbb{A}_0}(v(p,0)+h_{0,[t]}(p,x)),\\ \bar{\varphi}(x,t) &= \min_{(p,0)\in\mathbb{A}_0}(\bar{\varphi}(p,0)+h_{0,[t]}(p,x)), \end{aligned} \tag{3-9}$$

for all $(x,t)\in M\times[0,+\infty)$. We assert that there exists a point $(p_0,0)\in\mathbb{A}_0$ such that

$$v(p_0,0)<\bar{\varphi}(p_0,0). \tag{3-10}$$

Suppose otherwise. Then

$$v(p,0)\geq\bar{\varphi}(p,0),\quad \forall (p,0)\in\mathbb{A}_0.$$

Therefore, we have

$$\min_{(p,0)\in\mathbb{A}_0}(v(p,0)+h_{0,[t]}(p,x))\geq \min_{(p,0)\in\mathbb{A}_0}(\bar{\varphi}(p,0)+h_{0,[t]}(p,x))$$

for all $(x,t)\in M\times[0,+\infty)$. The above inequality and (3-9) imply that

$$v(x,t)\geq\bar{\varphi}(x,t),\quad \forall (x,t)\in M\times[0,+\infty),$$

which contradicts (3-8). Hence (3-10) holds. Let $\delta_0=\bar{\varphi}(p_0,0)-v(p_0,0)$. Then $\delta_0>0$ and by Lemma 3.1, we have

$$v(p_0,0)=\hat{\varphi}(p_0)-\delta_0. \tag{3-11}$$

Since $\hat{\varphi}(p_0)=\limsup_{n\to+\infty}T_n\varphi(p_0)$, then for the above δ_0, there exists $N_3\in\mathbb{N}$ such that $\sup_{k\geq n}T_k\varphi(p_0)>\hat{\varphi}(p_0)-\frac{\delta_0}{2}$, if $n\geq N_3$, which implies that there exists $k_n\geq n$ such that

$$T_{k_n}\varphi(p_0)>\hat{\varphi}(p_0)-\frac{\delta_0}{2}.$$

From the above inequality and (3-11), we deduce that

$$T_{k_n}\varphi(p_0)>\hat{\varphi}(p_0)-\frac{\delta_0}{2}=v(p_0,0)+\frac{\delta_0}{2}.$$

It follows that there exist $\{k_n\}_n\subset\mathbb{N}$ with $k_n\to+\infty$ as $n\to+\infty$ such that

$$u_\varphi(p_0,k_n)>v(p_0,k_n)+\frac{\delta_0}{2},$$

which contradicts the assumption that v is an AUB of u_φ.

Next, we show that $\underline{\varphi}$ is the max ALB of u_φ. Suppose not. There exist an ALB w of u_φ and a point $(x_1,t_1)\in M\times[0,+\infty)$ such that

$$w(x_1,t_1)>\underline{\varphi}(x_1,t_1).$$

Let $\delta_1 = w(x_1, t_1) - \underline{\varphi}(x_1, t_1)$. From Proposition 3.1, we have

$$\underline{\varphi}(x_1, t_1) = \liminf_{n\to+\infty} T_{n+t_1}\varphi(x_1).$$

Thus, there exists $N_4 \in \mathbb{N}$ such that $\inf_{j\geq n} T_{j+t_1}\varphi(x_1) < \underline{\varphi}(x_1, t_1) + \frac{\delta_1}{2}$ if $n \geq N_4$. It follows that there exists $j_n \geq n$ such that

$$T_{j_n+t_1}\varphi(x_1) < \underline{\varphi}(x_1, t_1) + \frac{\delta_1}{2}.$$

This inequality and the definition of δ_1 imply that

$$T_{j_n+t_1}\varphi(x_1) < \underline{\varphi}(x_1, t_1) + \frac{\delta_1}{2} = w(x_1, t_1) - \frac{\delta_1}{2},$$

which means that there exist $\{j_n\}_n \subset \mathbb{N}$ with $j_n \to +\infty$ as $n \to +\infty$ such that

$$u_\varphi(x_1, j_n + t_1) < w(x_1, j_n + t_1) - \frac{\delta_1}{2},$$

and we have a contradiction to the assumption that w is an ALB of u_φ.

Step 3. It remains to show that (1-6) holds. Note that

$$\begin{aligned}
\text{(3-12)}\quad & \underline{\varphi}(x, 0) - \bar{\varphi}(x, 0) \\
&= \inf_{y\in M}(\varphi(y) + h_{0,0}(y, x)) - \lim_{n\to+\infty} T_n\big(\limsup_{n\to+\infty} T_n\varphi\big)(x) \\
&\leq \inf_{y\in M}(\varphi(y) + h_{0,0}(y, x)) - \limsup_{n\to+\infty} T_n\varphi(x) \\
&= \inf_{y\in M}(\varphi(y) + h_{0,0}(y, x)) - \lim_{n\to+\infty} \sup_{k\geq n} \inf_{y\in M}(\varphi(y) + F_{0,k}(y, x)) \\
&= \lim_{n\to+\infty} \sup_{k\geq n}\big(\inf_{y\in M}(\varphi(y) + h_{0,0}(y, x)) - \inf_{y\in M}(\varphi(y) + F_{0,k}(y, x))\big) \\
&\leq \lim_{n\to+\infty} \sup_{k\geq n} \sup_{y\in M}(h_{0,0}(y, x) - F_{0,k}(y, x))
\end{aligned}$$

for all $x \in M$, where for the first inequality we have used (3-2). In view of (2-2),

$$\lim_{n\to+\infty} \inf_{k\geq n} F_{0,k}(y, x) = h_{0,0}(y, x), \quad \text{uniformly on } (y, x) \in M \times M.$$

Therefore, for any $\epsilon > 0$, there exists $N_5 \in \mathbb{N}$ such that whenever $n \geq N_5$ it follows that

$$\inf_{k\geq n} F_{0,k}(y, x) \geq h_{0,0}(y, x) - \epsilon,$$

for all $x, y \in M$, which implies that

$$\epsilon \geq h_{0,0}(y, x) - F_{0,k}(y, x), \tag{3-13}$$

for all $k \geq N_5$ and all $x, y \in M$. Combining (3-12) and (3-13), we have

$$\underline{\varphi}(x, 0) - \bar{\varphi}(x, 0) \leq 0, \quad \forall x \in M, \tag{3-14}$$

since ϵ may be taken arbitrarily small.

On the other hand, we have

$$
\begin{aligned}
\bar{\varphi}(x,0)-\underline{\varphi}(x,0) &= \lim_{n\to+\infty} T_n\big(\limsup_{n\to+\infty} T_n\varphi\big)(x)-\inf_{y\in M}(\varphi(y)+h_{0,0}(y,x))\\
&= \lim_{n\to+\infty} T_n\hat{\varphi}(x)-\inf_{y\in M}(\varphi(y)+h_{0,0}(y,x))\\
&= \lim_{n\to+\infty}\big(\inf_{y\in M}(\hat{\varphi}(y)+F_{0,n}(y,x))-\inf_{y\in M}(\varphi(y)+h_{0,0}(y,x))\big)\\
&\le \limsup_{n\to+\infty}\sup_{y\in M}(\hat{\varphi}(y)-\varphi(y)+F_{0,n}(y,x)-h_{0,0}(y,x))\\
&= \limsup_{n\to+\infty}\sup_{y\in M}\big(\limsup_{m\to+\infty} T_m\varphi(y)-\varphi(y)+F_{0,n}(y,x)-h_{0,0}(y,x)\big)\\
&= \limsup_{n\to+\infty}\sup_{y\in M}\big(\lim_{m\to+\infty}\sup_{k\ge m} T_k\varphi(y)-\varphi(y)+F_{0,n}(y,x)-h_{0,0}(y,x)\big)\\
&= \limsup_{n\to+\infty}\sup_{y\in M}\big(\lim_{m\to+\infty}\sup_{k\ge m}\inf_{z\in M}(\varphi(z)+F_{0,k}(z,y))\\
&\qquad\qquad -\varphi(y)+F_{0,n}(y,x)-h_{0,0}(y,x)\big)\\
&\le \limsup_{n\to+\infty}\sup_{y\in M}\big(\lim_{m\to+\infty}\sup_{k\ge m} F_{0,k}(y,y)+F_{0,n}(y,x)-h_{0,0}(y,x)\big)
\end{aligned}
$$

for all $x\in M$. From Lemma 2.1, we get

$$\bar{\varphi}(x,0)-\underline{\varphi}(x,0)\le C,\quad \forall x\in M, \tag{3-15}$$

where $C>0$ depends only on L.

Combining (3-14) and (3-15), we obtain

$$\underline{\varphi}(x,0)\le\bar{\varphi}(x,0)\le\underline{\varphi}(x,0)+C,\quad \forall x\in M.$$

By the monotonicity of the Lax–Oleinik operator, we have

$$T_t\underline{\varphi}(x,0)\le T_t\bar{\varphi}(x,0)\le T_t\underline{\varphi}(x,0)+C,\quad \forall x\in M,\ \forall t\ge 0.$$

In view of Propositions 3.1 and 3.3, we have

$$|\bar{\varphi}(x,t)-\underline{\varphi}(x,t)|\le C,\quad \forall (x,t)\in M\times[0,+\infty),$$

which means that (1-6) holds true, and the proof is complete. □

Acknowledgements

The author would like to thank the referees for their valuable comments and suggestions, which helped to improve the presentation of this paper. The author is supported by NSFC Grant No. 11771283.

References

[Barles 1985] G. Barles, "Asymptotic behavior of viscosity solutions of first-order Hamilton–Jacobi equations", *Ricerche Mat.* **34**:2 (1985), 227–260. MR Zbl

[Barles and Souganidis 2000] G. Barles and P. E. Souganidis, "Some counterexamples on the asymptotic behavior of the solutions of Hamilton–Jacobi equations", *C. R. Acad. Sci. Paris Sér. I Math.* **330**:11 (2000), 963–968. MR Zbl

[Bernard 2002] P. Bernard, "Connecting orbits of time dependent Lagrangian systems", *Ann. Inst. Fourier* (*Grenoble*) **52**:5 (2002), 1533–1568. MR Zbl

[Bernard and Roquejoffre 2004] P. Bernard and J.-M. Roquejoffre, "Convergence to time-periodic solutions in time-periodic Hamilton–Jacobi equations on the circle", *Comm. Partial Differential Equations* **29**:3-4 (2004), 457–469. MR Zbl

[Contreras et al. 1997] G. Contreras, J. Delgado, and R. Iturriaga, "Lagrangian flows: the dynamics of globally minimizing orbits, II", *Bol. Soc. Brasil. Mat.* (*N.S.*) **28**:2 (1997), 155–196. MR Zbl

[Contreras et al. 2013] G. Contreras, R. Iturriaga, and H. Sánchez-Morgado, "Weak solutions of the Hamilton–Jacobi equation for time periodic Lagrangians", preprint, 2013. arXiv

[Crandall and Lions 1983] M. G. Crandall and P.-L. Lions, "Viscosity solutions of Hamilton–Jacobi equations", *Trans. Amer. Math. Soc.* **277**:1 (1983), 1–42. MR Zbl

[Fathi 1998] A. Fathi, "Sur la convergence du semi-groupe de Lax–Oleinik", *C. R. Acad. Sci. Paris Sér. I Math.* **327**:3 (1998), 267–270. MR Zbl

[Fathi 2005] A. Fathi, "Weak KAM theorem in Lagrangian dynamics", preprint, 2005, Available at http://www.math.unipd.it/~angela/Fathi_7.pdf.

[Fathi and Mather 2000] A. Fathi and J. N. Mather, "Failure of convergence of the Lax–Oleinik semi-group in the time-periodic case", *Bull. Soc. Math. France* **128**:3 (2000), 473–483. MR Zbl

[Ishii 2006] H. Ishii, "Asymptotic solutions for large time of Hamilton–Jacobi equations", pp. 213–227 in *International Congress of Mathematicians, III* (Madrid, 2006), edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Lions 1982] P.-L. Lions, *Generalized solutions of Hamilton–Jacobi equations*, Research Notes in Math. **69**, Pitman, Boston, 1982. MR Zbl

[Mañé 1997] R. Mañé, "Lagrangian flows: the dynamics of globally minimizing orbits", *Bol. Soc. Brasil. Mat.* (*N.S.*) **28**:2 (1997), 141–153. MR Zbl

[Mather 1991] J. N. Mather, "Action minimizing invariant measures for positive definite Lagrangian systems", *Math. Z.* **207**:2 (1991), 169–207. MR Zbl

[Mather 1993] J. N. Mather, "Variational construction of connecting orbits", *Ann. Inst. Fourier* (*Grenoble*) **43**:5 (1993), 1349–1386. MR Zbl

[Roquejoffre 2001] J.-M. Roquejoffre, "Convergence to steady states or periodic solutions in a class of Hamilton–Jacobi equations", *J. Math. Pures Appl.* (9) **80**:1 (2001), 85–104. MR Zbl

[Wang and Yan 2012] K. Wang and J. Yan, "A new kind of Lax–Oleinik type operator with parameters for time-periodic positive definite Lagrangian systems", *Comm. Math. Phys.* **309**:3 (2012), 663–691. MR Zbl

Received November 15, 2015. Revised June 5, 2017.

KAIZHI WANG
SCHOOL OF MATHEMATICAL SCIENCES
SHANGHAI JIAO TONG UNIVERSITY
SHANGHAI
CHINA
kzwang@sjtu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 298, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.298.233

GLOBAL WELL-POSEDNESS FOR THE 2D FRACTIONAL BOUSSINESQ EQUATIONS IN THE SUBCRITICAL CASE

DAOGUO ZHOU, ZILAI LI, HAIFENG SHANG,
JIAHONG WU, BAOQUAN YUAN AND JIEFENG ZHAO

We study the global regularity of solutions to the 2D Boussinesq equations with fractional dissipation, given by $(-\Delta)^{\alpha/2}u$ in the velocity equation and by $(-\Delta)^{\beta/2}\theta$ in the temperature equation. We establish the global regularity for $\frac{2}{3} < \alpha < 1$, $\alpha+\beta > 1$ and $\alpha > \frac{1}{1+\beta}$. This result is for the subcritical regime $\alpha+\beta > 1$ and the point here is to obtain the global regularity for the largest possible range of α.

1. Introduction

This paper examines the global (in time) well-posedness problem on the 2D Boussinesq equations with fractional dissipation. The Boussinesq equations concerned here model large scale atmospheric and oceanic flows that are responsible for cold fronts and the jet stream (see the books by Gill [1982], Majda [2003], Pedlosky [1979]). In addition, the Boussinesq equations also play an important role in the study of Rayleigh–Benard convection [Constantin and Doering 1999]. The standard 2D Boussinesq equations with Laplacian dissipation can be written

$$\begin{cases} u_t + u\cdot\nabla u + \nabla p = \nu\,\Delta u + \theta e_2, \\ \theta_t + u\cdot\nabla\theta = \kappa\,\Delta\theta, \\ \nabla\cdot u = 0, \end{cases} \tag{1-1}$$

where u denotes the 2D velocity field, p the pressure, θ the temperature in the context of thermal convection and the density in the modeling of geophysical fluids, ν the viscosity, κ the thermal diffusivity, and $e_2 = (0, 1)$ is the unit vector in the vertical direction.

The 2D Boussinesq equations have recently attracted considerable attention in the community of mathematical fluid mechanics due to their mathematical significance. Mathematically the 2D Boussinesq equations serve as a lower-dimensional model of the 3D hydrodynamics equations. In fact, the 2D Boussinesq equations retain

MSC2010: primary 35B65, 35Q35; secondary 76B03, 76A10.
Keywords: Boussinesq equations, fractional dissipation, global regularity.

some key features of the 3D Euler and Navier–Stokes equations such as the vortex stretching mechanism. The inviscid 2D Boussinesq equations are identical to the Euler equations for the 3D axisymmetric swirling flows (away from the symmetry axis) (see, e.g., [Majda and Bertozzi 2001]).

Our attention will be focused on the 2D Boussinesq equations with fractional dissipation

$$\begin{cases} u_t + u\cdot\nabla u + \Lambda^\alpha u + \nabla p = \theta e_2, \\ \theta_t + u\cdot\nabla\theta + \Lambda^\beta\theta = 0, \\ \nabla\cdot u = 0, \\ u(x,0) = u_0(x), \quad \theta(x,0) = \theta_0(x), \end{cases} \tag{1-2}$$

where $\Lambda = (-\Delta)^{1/2}$ and the general fractional Laplacian operator Λ^α can be defined via the Fourier transform

$$\widehat{\Lambda^\alpha f}(\xi) = |\xi|^\alpha \widehat{f}(\xi).$$

This generalization allows us to study a family of equations simultaneously and may be physically relevant. In fact, there are geophysical circumstances in which the Boussinesq equations with fractional Laplacian may arise. Flows in the middle atmosphere traveling upward undergo changes due to the changes of atmospheric properties, although the incompressibility and Boussinesq approximations are applicable. The effect of kinematic and thermal diffusion is attenuated by the thinning of atmosphere. This anomalous attenuation can be modeled by using the space fractional Laplacian (see [Gill 1982; Caputo 1967]).

One of the fundamental problems concerning the Boussinesq system is whether or not its solutions remain smooth for all time or they blow up in a finite time. This problem could be extremely difficult. A standard approach to the global regularity problem is to first obtain the local existence and regularity and then extend the local solution to a global one by establishing global a priori bounds for the solution. Due to the divergence-free condition $\nabla\cdot u = 0$, any solution (u,θ) with sufficiently smooth data admits a global L^2-bound for u and a global L^q-bound for θ ($q\in[1,\infty]$). However, when the dissipation or the thermal diffusion is not sufficient, it can be extremely difficult to obtain global bounds for suitable derivatives of u or θ. When the Boussinesq equations are inviscid (no velocity dissipation or thermal diffusion), the equations of $\omega = \nabla\times u$ and $\nabla^\perp\theta$,

$$\begin{cases} \partial_t\omega + (u\cdot\nabla)\omega = \partial_{x_1}\theta, \\ \partial_t\nabla^\perp\theta + (u\cdot\nabla)\nabla^\perp\theta = (\nabla^\perp\theta\cdot\nabla)u, \end{cases}$$

resemble the 3D Euler vorticity equation

$$\partial_t\omega^E + (u^E\cdot\nabla)\omega^E = (\omega^E\cdot\nabla)u^E,$$

where $\nabla^\perp = (-\partial_{x_2}, \partial_{x_1})$, and u^E and ω^E denote the 3D Euler velocity and the corresponding vorticity, respectively.

When Δu and $\Delta\theta$ are present, the global regularity can then be established following a similar proof as that for the 2D Navier–Stokes equations. The issue that arises naturally is how much dissipation is really needed for the global regularity. This problem has attracted considerable interest recently and important progress has been made (see, e.g., [Adhikari et al. 2010; 2011; 2014; Cao and Wu 2013; Constantin and Vicol 2012; Danchin and Paicu 2011; Hmidi et al. 2010; 2011; Hou and Li 2005; KC et al. 2014; Lai et al. 2011; Larios et al. 2013; Li et al. 2016; Li and Titi 2016; Miao and Xue 2011; Ohkitani 2001; Stefanov and Wu 2018; Wu and Xu 2014; Wu et al. 2016; 2015; Yang et al. 2014; Ye 2017; Ye and Xu 2016; Zhao 2010; Zhou 2018; Zhou and Li 2017]). Various approaches and techniques have been developed to obtain the global regularity for (1-2) with smaller and smaller $\alpha \in (0, 2)$ and $\beta \in (0, 2)$.

As pointed out in [Jiu et al. 2014], it is useful to classify α and β into three categories: the subcritical case when $\alpha+\beta > 1$, the critical case when $\alpha+\beta = 1$ and the supercritical case when $\alpha+\beta < 1$. This classification gives us a sense of the level of difficulty for different parameter ranges. The global regularity problem for the supercritical regime $\alpha+\beta < 1$ appears to be out of reach at this moment. Current results for this regime address the eventual regularity of weak solutions [Yang et al. 2014; Wu et al. 2016]. There are exciting developments for the critical regime. Two special critical cases, $\alpha = 1,\ \beta = 0$ and $\beta = 1, \alpha = 0$, were studied and resolved in [Hmidi et al. 2010; 2011]. More general critical cases with $\alpha+\beta = 1$ and $\alpha \in (0, 1)$ were dealt with by Jiu, Miao, Wu and Zhang [Jiu et al. 2014], who established the global regularity for (1-2) with $\alpha+\beta = 1$ and $1 > \alpha > \alpha_0 \equiv \frac{23-\sqrt{145}}{12} \approx 0.9132$. Stefanov and Wu improved the result of Jiu, Miao, Wu and Zhang by further enlarging the range of α with $\alpha+\beta = 1$ and $1 > \alpha > \frac{\sqrt{1777}-23}{24} \approx 0.7981$ [2018]. A very recent work of Wu, Xu, Xue and Ye assesses the global regularity for $\alpha+\beta = 1$ and $\alpha \in (0.7692, 1)$ [Wu et al. 2016].

This paper focuses on the subcritical regime $\alpha+\beta > 1$. The global regularity problem, even in this regime, can be difficult, and there are ranges of subcritical regime for which the global regularity of (1-2) remains unknown. To give an accurate account of current results, we further divide the subcritical regime into two cases: $\alpha \ge \beta$ and $\alpha < \beta$. We refer to the first case as velocity dissipation dominated and the second case as thermal diffusion dominated. For the velocity dominated case, Miao and Xue [2011] was able to establish the global regularity of (1-2) with

$$\alpha \in \Big(\frac{6-\sqrt{6}}{4},\ 1\Big), \quad \beta \in \Big(1-\alpha,\ \min\Big\{\frac{(7+2\sqrt{6})\alpha}{5} - 2,\ \frac{\alpha(1-\alpha)}{\sqrt{6}-2\alpha},\ 2-2\alpha\Big\}\Big).$$

Note that $\frac{6-\sqrt{6}}{4} \approx 0.8876$. Ye [2017] was able to enlarge the range to

$$0.7351 < \alpha < 1, \quad \beta \in \Big(1-\alpha,\ \min\Big\{3-3\alpha,\ \frac{\alpha}{2},\ \frac{3\alpha^2+4\alpha-4}{8(1-\alpha)}\Big\}\Big).$$

For the thermal diffusion dominated case, Constantin and Vicol obtained as a consequence of their nonlinear maximum principle for fractional Laplacian operators the global regularity of (1-2) with $\beta > \frac{2}{2+\alpha}$. In addition, Yang, Jiu and Wu [Yang et al. 2014] obtained the global regularity for a larger range of β, and Ye and Xu [2016] made further improvements on the range of β.

This paper focuses on the velocity dissipation dominated case, $\alpha \geq \beta$. Our primary goal has been to obtain the global regularity for the smallest possible $\alpha \in (0, 1)$ with $\alpha + \beta > 1$ and $\alpha > \beta > 0$. Our main result is stated in Theorem 1.2. A slightly weaker result with a smaller range of α is stated in Theorem 1.1. The main reason for keeping Theorem 1.1 is that Theorem 1.2 is built upon Theorem 1.1 and its proof.

Theorem 1.1. *Let $s > 2$. Assume that $u_0 \in H^s(\mathbb{R}^2)$ and $\nabla \cdot u_0 = 0$, and $\theta_0 \in H^s(\mathbb{R}^2)$. Consider the fractional Boussinesq equations (1-2) with α and β satisfying*

$$0 < \alpha, \beta < 1, \qquad \alpha > \frac{2}{\beta+2}, \tag{1-3}$$

then (1-2) has a unique global (in time) solution (u, θ) satisfying

$$(u, \theta) \in C([0, T]; H^s(\mathbb{R}^2)).$$

Theorem 1.2. *Let $s > 2$. Assume that $u_0 \in H^s(\mathbb{R}^2)$ and $\nabla \cdot u_0 = 0$, and $\theta_0 \in H^s(\mathbb{R}^2)$. Consider the fractional Boussinesq equations (1-2) with α and β satisfying*

$$\frac{2}{3} < \alpha < 1, \quad 0 < \beta < 1, \quad \alpha > \frac{1}{\beta+1}, \tag{1-4}$$

then (1-2) has a unique global (in time) solution (u, θ) satisfying

$$(u, \theta) \in C([0, T]; H^s(\mathbb{R}^2)).$$

The proof of Theorem 1.1 relies on the equation for a combined quantity and the nonlinear maximum principle for fractional Laplacian operators developed by Córdoba and Córdoba [2004] and by Constantin and Vicol [2012]. Due to the presence of the "vortex stretching" term $\partial_{x_1}\theta$, energy estimates on the vorticity equation

$$\partial_t \omega + (u \cdot \nabla)\omega + \Lambda^\alpha \omega = \partial_{x_1}\theta$$

with $\alpha \in (0, 1)$ would not yield any global bound on ω. A well-known practice is to eliminate $\partial_{x_1}\theta$ by considering the combined quantity

$$G = \omega - \mathcal{R}_\alpha \theta \quad \text{with} \quad \mathcal{R}_\alpha = \partial_1 \Lambda^{-\alpha},$$

which satisfies

$$G_t + u \cdot \nabla G + \Lambda^\alpha G = [\mathcal{R}_\alpha, u \cdot \nabla]\theta + \Lambda^{\beta-\alpha}\partial_1 \theta,$$

where $[\mathcal{R}_\alpha, u \cdot \nabla]\theta$ denotes the standard commutator. Combining this equation with that of $\nabla\theta$, applying the nonlinear maximum principle for fractional Laplacian operators and invoking commutator estimates, one derives differential inequalities for $\|G(t)\|_{L^\infty}$ and $\|\nabla\theta(t)\|_{L^\infty}$, which yields Theorem 1.1. Theorem 1.2 involves improved arguments. Its proof makes use of the global L^2 bound for G whenever $\alpha > \frac{2}{3}$ and $\alpha+\beta > 1$, and the pointwise lower bound

$$f(x)\cdot\Lambda^\alpha f(x) \geq \frac{1}{2}\Lambda^\alpha |f(x)|^2 + \frac{|f(x)|^{2+p\alpha/d}}{c\|f\|_{L^p}^{p\alpha/d}}.$$

This lower bound is in terms of the L^p-norms of the functions instead of the L^p-norm of the antiderivative of f, and thus has a higher power than the corresponding lower bound in terms of the L^p-norm of the antiderivative.

The rest of this paper is divided into two sections. Section 2 proves Theorem 1.1 while Section 3 proves Theorem 1.2. Two appendices are also attached. The first one provides the frequency localization operators and Besov spaces, and related facts. Appendix B supplies the proofs for some of the facts used in Sections 2 and 3.

2. Proof of Theorem 1.1

This section proves Theorem 1.1. To do so, we make several preparations. The first is a pointwise inequality for fractional Laplacian operators in [Constantin and Vicol 2012; Córdoba and Córdoba 2004].

Lemma 2.1. *Let $\alpha \in (0,2)$ and $q \in [1,\infty]$. There exists $C = C(d,\alpha,q)$ such that, for any function $f = f(x)$ with $x \in \mathbb{R}^d$ that is sufficiently smooth and decays at infinity,*

$$\nabla f(x)\cdot\Lambda^\alpha\nabla f(x) \geq \frac{1}{2}\Lambda^\alpha|\nabla f(x)|^2 + \frac{|\nabla f(x)|^{2+q\alpha/(d+q)}}{C\,\|f\|_{L^q}^{q\alpha/d+q}}, \quad x \in \mathbb{R}^d.$$

The next lemma states an interpolation inequality between Besov spaces (see, e.g., [Bahouri et al. 2011; Miao et al. 2012; Hajaiej et al. 2011]). The definition of Besov spaces is provided in Appendix A.

Lemma 2.2. *Let $s_1 < s_2$ be real numbers and let $\gamma \in (0,1)$. Let $p \in [1,\infty]$. Then, there exists a constant $C = C(s_1,s_2,\gamma)$ such that*

$$\|f\|_{\dot B^{\gamma s_1+(1-\gamma)s_2}_{p,1}} \leq C\,\|f\|^{\gamma}_{\dot B^{s_1}_{p,\infty}}\,\|f\|^{1-\gamma}_{\dot B^{s_2}_{p,\infty}}.$$

In particular, for any $\sigma \in (0,1)$ and $p \in [1,\infty]$,

$$\|\Lambda^\sigma f\|_{L^p} \leq \|f\|_{\dot B^{\sigma}_{p,1}} \leq C\,\|f\|^{1-\sigma}_{B^0_{p,\infty}}\,\|f\|^{\sigma}_{B^1_{p,\infty}} \leq C\,\|f\|^{1-\sigma}_{L^p}\,\|\nabla f\|^{\sigma}_{L^p}.$$

We will also need the commutator estimates stated in the following lemma. This lemma is taken from [Li et al. 2016, Lemma 2.2].

Lemma 2.3. *Let $j \geq 0$ be an integer. Let $\alpha \in (0, 2)$. Assume $q \in [2, \infty]$ and $q_1, q_2 \in [2, \infty]$ satisfy $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$. Assume $\nabla \cdot u = 0$. Then*

$$\begin{aligned}
\|\Delta_j[\mathcal{R}_\alpha, u \cdot \nabla]\theta\|_{L^q} \leq{} & C\, 2^{(1-\alpha)j}\, \|\nabla u\|_{L^{q_1}}\, \|\Delta_j \theta\|_{L^{q_2}} \\
& + C\, \|\nabla u\|_{L^{q_1}} \sum_{k \leq j-1} 2^{k-j} 2^{(1-\alpha)k}\, \|\Delta_k \theta\|_{L^{q_2}} \\
& + C\, \|\nabla u\|_{L^{q_1}} \sum_{k \geq j-1} 2^{(2-\alpha)(j-k)} 2^{(1-\alpha)k} \|\Delta_k \theta\|_{L^{q_2}} \\
& + C\, \|\nabla u\|_{L^{q_1}} \sum_{k \geq j-1} 2^{j-k}\, 2^{(1-\alpha)k} \|\Delta_k \theta\|_{L^{q_2}},
\end{aligned} \tag{2-1}$$

where C's are constants. In addition, (2-1) *still holds if $\mathcal{R}_\alpha$ is replaced by $\Lambda^{1-\alpha}$. A special consequence of* (2-1) *is the bound*

$$\|[\mathcal{R}_\alpha, u \cdot \nabla]\theta\|_{L^q} \leq C\, \|\nabla u\|_{L^{q_1}}\, \|\theta\|_{B^{1-\alpha}_{q_2,1}}. \tag{2-2}$$

Similarly,

$$\|[\Lambda^{1-\alpha}, u \cdot \nabla]\theta\|_{L^q} \leq C\, \|\nabla u\|_{L^{q_1}}\, \|\theta\|_{B^{1-\alpha}_{q_2,1}}.$$

Alternatively, the commutator can also be bounded as follows. A proof is provided in Appendix B.

Lemma 2.4. *Let $\alpha \in (0, 1)$. Then*,

$$\|[\partial_1 \Lambda^{-\alpha}, u \cdot \nabla]\theta\|_{B^0_{\infty,1}} \leq C(\|\omega\|_2 + \|\omega\|_\infty)\|\theta\|_{B^{1-\alpha+\epsilon}_{\infty,1}} + C\|u\|_2\|\theta\|_2.$$

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. . It suffices to establish a global a priori bound on $\|(u, \theta)\|_{H^s}$. As we know, if one of the global bounds, for any $t > 0$,

$$\int_0^t \|\nabla \omega(\tau)\|_{L^\infty}\, d\tau < \infty \quad \text{or} \quad \int_0^t \|\nabla \theta(\tau)\|_{L^\infty}\, d\tau < \infty \tag{2-3}$$

holds, then $\|(u, \theta)(t)\|_{H^s}$ is globally bounded. The rest of the proof verifies the bounds in (2-3).

The following global bounds follow easily from (1-2):

$$\|\theta(t)\|_{L^q} \leq \|\theta_0\|_{L^q} \quad \text{for any } q \in [1, \infty],$$

$$\|u(t)\|_{L^2}^2 + \int_0^t \|\Lambda^{\frac{\alpha}{2}} u(\tau)\|_{L^2}^2\, d\tau \leq (\|u_0\|_{L^2} + t\, \|\theta_0\|_{L^2})^2.$$

However, direct energy estimates on (1-2) or on the equation of the vorticity $\omega = \nabla \times u$,

$$\begin{cases} \omega_t + u \cdot \nabla \omega + \Lambda^\alpha \omega = \partial_1 \theta, \\ \theta_t + u \cdot \nabla \theta + \Lambda^\beta \theta = 0, \end{cases}$$

would not yield the desired global bound in (2-3), due to the vortex stretching term $\partial_1\theta$. As in [Hmidi et al. 2010; 2011; Miao and Xue 2011; Jiu et al. 2014], the idea is to eliminate $\partial_1\theta$ and work with the combined quantity

$$G = \omega - \mathcal{R}_\alpha \theta \quad \text{with} \quad \mathcal{R}_\alpha = \partial_1 \Lambda^{-\alpha}, \tag{2-4}$$

which satisfies

$$G_t + u\cdot\nabla G + \Lambda^\alpha G = [\mathcal{R}_\alpha, u\cdot\nabla]\theta + \Lambda^{\beta-\alpha}\partial_1\theta, \tag{2-5}$$

where we have used the standard commutator notation

$$[\mathcal{R}_\alpha, u\cdot\nabla]\theta = \mathcal{R}_\alpha(u\cdot\nabla\theta) - u\cdot\nabla\mathcal{R}_\alpha\theta.$$

Following the idea of [Constantin and Vicol 2012], we obtain the differential inequality for $\|G(t)\|_{L^\infty}$,

$$\frac{d}{dt}\|G\|_{L^\infty} + C\frac{\|G\|_{L^\infty}^{1+\alpha/2}}{(\|u\|_{L^2}+\|\Lambda^{-\alpha}\theta\|_{L^2})^{\alpha/2}} \le \|[\mathcal{R}_\alpha, u\cdot\nabla]\theta\|_{L^\infty} + \|\Lambda^{\beta-\alpha}\partial_1\theta\|_{L^\infty} \tag{2-6}$$

and for $\|\nabla\theta\|_{L^\infty}$,

$$\frac{d}{dt}\|\nabla\theta\|_{L^\infty} + C\,\frac{\|\nabla\theta\|_{L^\infty}^{1+\beta}}{\|\theta\|_{L^\infty}^{\beta}} \le \|\nabla u\|_{L^\infty}\|\nabla\theta\|_{L^\infty}. \tag{2-7}$$

We briefly explain the derivation of (2-6). Without loss of generality, we assume G is smooth and decays to zero at infinity. Multiplying (2-5) by G and applying Lemma 2.1 with $q = 2$, we have

$$\begin{aligned}\partial_t|G|^2 + u\cdot\nabla|G|^2 + \Lambda^\alpha|G|^2 + C\frac{|G|^{1+\alpha/2}}{(\|u\|_{L^2}+\|\Lambda^{-\alpha}\theta\|_{L^2})^{\alpha/2}}&\\ \le 2(\|[\mathcal{R}_\alpha, u\cdot\nabla]\theta\|_{L^\infty} + \|\Lambda^{\beta-\alpha}\partial_1\theta\|_{L^\infty})|G|.&\end{aligned} \tag{2-8}$$

For each $t > 0$, there exists $\bar{x} = \bar{x}(t) \in \mathbb{R}^2$ such that

$$G(\bar{x}(t), t) = \|G(t)\|_{L^\infty} = \max_{x\in\mathbb{R}^2}|G(x,t)|.$$

As explained in [Córdoba and Córdoba 2004] and [Constantin et al. 2015, Appendix B],

$$(\partial_t|G|)(\bar{x}(t), t) = \frac{d}{dt}G(\bar{x}(t), t) = \frac{d}{dt}\|G(t)\|_{L^\infty}.$$

In addition, we recall the facts that $(u\cdot\nabla)|G|(\bar{x}(t), t) = 0$ and $(\Lambda^\alpha|G|^2)(\bar{x}(t), t) \ge 0$. Therefore, setting $x = \bar{x}(t)$ in (2-8) and invoking the aforementioned facts yields (2-6). The inequality (2-7) is obtained in a similar fashion.

The terms in (2-6) can be further bounded as follows:

$$\|u(t)\|_{L^2} \le \|u_0\|_{L^2} + t\,\|\theta_0\|_{L^2}, \qquad \|\Lambda^{-\alpha}\theta\|_{L^2} \le C\,\|\theta\|_{L^{2/(1+\alpha)}} \le C\,\|\theta_0\|_{L^{2/(1+\alpha)}}.$$

By (2-2) of Lemma 2.3 and Lemma 2.2,

$$\begin{aligned}\|[\mathcal{R}_\alpha, u\cdot\nabla]\theta\|_{L^\infty} &\le C\,\|\nabla u\|_{L^\infty}\,\|\theta\|_{B^{1-\alpha}_{\infty,1}}\\ &\le C\,\|\nabla u\|_{L^\infty}\,\|\theta\|^{\alpha}_{B^0_{\infty,\infty}}\,\|\theta\|^{1-\alpha}_{B^1_{\infty,\infty}}\\ &\le C\,\|\nabla u\|_{L^\infty}\,\|\theta\|^{\alpha}_{L^\infty}\,\|\nabla\theta\|^{1-\alpha}_{L^\infty}\end{aligned}$$

and

$$\|\Lambda^{\beta-\alpha}\partial_1\theta\|_{L^\infty} \le \|\Lambda^{\beta-\alpha}\partial_1\theta\|_{\dot B^0_{\infty,1}} \le C\,\|\theta\|^{\alpha-\beta}_{L^\infty}\|\nabla\theta\|^{1+\beta-\alpha}_{L^\infty}.$$

Inserting the bounds above in (2-6) yields

$$\begin{aligned}\frac{d}{dt}\|G\|_{L^\infty} + C_1(t)\,\|G\|^{1+\alpha/2}_{L^\infty} &\le C\,\|\nabla u\|_{L^\infty}\,\|\nabla\theta\|^{1-\alpha}_{L^\infty} + C\,\|\nabla\theta\|^{1+\beta-\alpha}_{L^\infty},\\ \frac{d}{dt}\|\nabla\theta\|_{L^\infty} + C_2\,\|\nabla\theta\|^{1+\beta}_{L^\infty} &\le \|\nabla u\|_{L^\infty}\|\nabla\theta\|_{L^\infty},\end{aligned}$$

where

$$C_1(t) = \frac{1}{(\|u_0\|_{L^2} + t\,\|\theta_0\|_{L^2} + \|\theta_0\|_{L^{2/(1+\alpha)}})^{\alpha/2}}.$$

Furthermore, according to Constantin and Vicol [2012],

$$\|\nabla u(t)\|_{L^\infty} \le C\,(1+\|\omega(t)\|_{L^\infty}) + C\,\|\omega(t)\|_{L^\infty}) \log\Bigl(1+\int_0^t (1+\|u(\tau)\|_{L^2}+\|\omega(\tau)\|_{L^\infty}+\|\nabla\theta(\tau)\|_{L^2})^{\gamma(\alpha,\beta)}\,d\tau\Bigr), \tag{2-9}$$

where $\gamma(\alpha,\beta) > 0$ is a constant depending on α and β. Due to

$$\|\omega\|_{L^\infty} \le \|G\|_{L^\infty} + \|\mathcal{R}_\alpha\theta\|_{L^\infty} \le \|G\|_{L^\infty} + C\,\|\theta\|^{\alpha}_{L^\infty}\,\|\nabla\theta\|^{1-\alpha}_{L^\infty}, \tag{2-10}$$

we obtain

$$\begin{aligned}\frac{d}{dt}\|G\|_{L^\infty} + C_1\|G\|^{1+\alpha/2}_{L^\infty} \le{}& C_2\|G\|_{L^\infty}\,\|\nabla\theta\|^{1-\alpha}_{L^\infty}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty})\\ &+ C_3\,\|\nabla\theta\|^{2-2\alpha}_{L^\infty}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) + C_4\|\nabla\theta\|^{1+\beta-\alpha}_{L^\infty},\end{aligned} \tag{2-11}$$

$$\begin{aligned}\frac{d}{dt}\|\nabla\theta\|_{L^\infty} + C_5\,\|\nabla\theta\|^{1+\beta}_{L^\infty} \le{}& C_6\,\|G\|_{L^\infty}\,\|\nabla\theta\|_{L^\infty}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty})\\ &+ C_7\|\nabla\theta\|^{2-\alpha}_{L^\infty}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}),\end{aligned} \tag{2-12}$$

where, for notational convenience, we have written

$$L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) = 1+\log\Bigl(1+\int_0^t (1+\|G\|_{L^\infty}+\|\nabla\theta\|_{L^\infty})^{\gamma(\alpha,\beta)}ds\Bigr). \tag{2-13}$$

We combine (2-11) and (2-12) to prove the global bound (2-3). The argument is as follows. For each $t \ge 0$, we consider two cases:

$$\tfrac{1}{2}C_5\|\nabla\theta\|^{\beta}_{L^\infty} > C_6\,\|G\|_{L^\infty}\;L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) \tag{2-14}$$

and

$$\tfrac{1}{2}C_5\|\nabla\theta\|_{L^\infty}^{\beta} \le C_6\,\|G\|_{L^\infty}\ L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}). \tag{2-15}$$

We start with the first case when (2-14) holds. We split this case into two cases, either

$$C_7\,\|\nabla\theta\|_{L^\infty}^{2-\alpha}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) \le \tfrac{1}{2}C_5\,\|\nabla\theta\|_{L^\infty}^{1+\beta} \tag{2-16}$$

or

$$C_7\,\|\nabla\theta\|_{L^\infty}^{2-\alpha}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) > \tfrac{1}{2}C_5\,\|\nabla\theta\|_{L^\infty}^{1+\beta}. \tag{2-17}$$

When (2-16) is valid, then (2-12) becomes

$$\frac{d}{dt}\|\nabla\theta\|_{L^\infty} + \big(\tfrac{1}{2}C_5\,\|\nabla\theta\|_{L^\infty}^{1+\beta} - C_6\,\|G\|_{L^\infty}\,\|\nabla\theta\|_{L^\infty}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty})\big) < 0,$$

which, due to (2-14), implies that $\|\nabla\theta\|_{L^\infty} < \infty$. Then (2-11) implies $\|G\|_{L^\infty} < \infty$. When (2-17) is valid,

$$C_7\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) > \tfrac{1}{2}C_5\,\|\nabla\theta\|_{L^\infty}^{1+\beta-(2-\alpha)}.$$

Since $1+\beta-(2-\alpha) = \alpha+\beta-1 > 0$, we have

$$\|\nabla\theta\|_{L^\infty}^{1+\beta-(2-\alpha)} \le 2C_5^{-1}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}). \tag{2-18}$$

Due to (2-14), L only grows logarithmically in $\|\nabla\theta\|_{L^\infty}$ and thus (2-18) implies that $\|\nabla\theta\|_{L^\infty} < \infty$. Then (2-11) implies $\|G\|_{L^\infty} < \infty$. We now turn to the second case when (2-15) holds. We also split this case into two cases: either

$$L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) \le C\,\|G\|_{L^\infty}^{\epsilon} \tag{2-19}$$

or

$$L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) > C\,\|G\|_{L^\infty}^{\epsilon}, \tag{2-20}$$

where $\epsilon > 0$ is small such that (2-22) below holds. When (2-19) holds, (2-11) becomes

$$\begin{aligned}\frac{d}{dt}\|G\|_{L^\infty} + C_1\|G\|_{L^\infty}^{1+\alpha/2} \\ \le \widetilde{C}_2\,\|G\|_{L^\infty}^{1+(1-\alpha/\beta)+\tilde\epsilon} + \widetilde{C}_3\,\|G\|_{L^\infty}^{(2-2\alpha/\beta)+\tilde\epsilon} + \widetilde{C}_4\|G\|_{L^\infty}^{(1+\beta-\alpha/\beta)+\tilde\epsilon},\end{aligned} \tag{2-21}$$

where $\widetilde{C}_2$, $\widetilde{C}_3$ and $\widetilde{C}_4$ are constants, and

$$\tilde\epsilon = \epsilon\ \max\Big\{1+\frac{1-\alpha}{\beta}, 1+\frac{2-2\alpha}{\beta}, \frac{1+\beta-\alpha}{\beta}\Big\}.$$

Due to (1-3) or $\alpha > \frac{2}{2+\beta}$, we can choose $\epsilon > 0$ small such that

$$1+\frac{\alpha}{2} > \max\Big\{1+\frac{1-\alpha}{\beta}, \frac{2-2\alpha}{\beta}, \frac{1+\beta-\alpha}{\beta}\Big\} + \tilde\epsilon. \tag{2-22}$$

Then (2-21) implies $\|G\|_{L^\infty} < \infty$ and (2-19) implies $\|\nabla\theta\|_{L^\infty} < \infty$. When (2-20) holds, (2-15) and the logarithmic growth of L in $\|G\|_{L^\infty}$ implies $\|G\|_{L^\infty} < \infty$. Therefore, for each case, the global bounds in (2-3) hold. This argument here can also be understood as a continuation argument. One starts with initial data that falls into one of the cases. Obviously, the corresponding solution can be continued as long as the solution remains in the same case. If, at a certain time, the solution evolves into the opposite case, the solution can also be continued. That is, the solution can be continued forever. The proof of Theorem 1.1 is complete. □

3. Proof of Theorem 1.2

This section is devoted to the proof of Theorem 1.2. We first prove a proposition stating the global L^2-bound for G. This result was obtained by Ye [2017], but we provide a slightly simpler and more transparent proof.

Proposition 3.1. *Consider the equation of G in (2-5). Assume that α and β satisfy*

$$\tfrac{2}{3} < \alpha < 1, \quad 0 < \beta < 1, \quad \alpha+\beta > 1.$$

Then we have the following global bounds, for any $t > 0$,

$$\|G(t)\|_{L^2} < \infty, \qquad \int_0^t \|\Lambda^{\alpha/2} G(\tau)\|_{L^2}^2 \, d\tau < \infty,$$

$$\sup_{j\ge -1} \int_0^t 2^{2\beta j} \|\Delta_j \theta(\tau)\|^2 \, d\tau < \infty, \quad \textit{especially}, \quad \int_0^t \|\Lambda^\sigma \theta(\tau)\|_{L^2}^2 \, d\tau < \infty,$$

where $0 < \sigma < \beta$.

In order to prove Proposition 3.1, we state a lemma and its corollary first.

Lemma 3.2. *Assume $\beta > 0$. Assume θ solves*

$$\theta_t + u\cdot\nabla\theta + \Lambda^\beta \theta = 0, \quad \theta(x,0) = \theta_0(x).$$

Then,

$$\sup_{j\ge -1} \int_0^t 2^{2\beta j} \|\Delta_j \theta(\tau)\|^2 \, d\tau \le C \, \|\theta_0\|_{B^{\beta/2}_{2,\infty}}^2 + \widetilde{C} \int_0^t \|\omega(\tau)\|_{L^2}^2 \, d\tau, \tag{3-1}$$

where ω denotes the vorticity, and C, $\widetilde{C}$ are constants depending on the initial data.

A special consequence of Lemma 3.2 is the following corollary.

Corollary 3.3. *Assume that α and β satisfy*

$$0 < \alpha, \beta < 1, \quad \alpha+\beta > 1.$$

Then

$$\sup_{j\ge -1} \int_0^t 2^{2\beta j} \|\Delta_j \theta(\tau)\|^2 \, d\tau \le C(t, \|(u_0,\theta_0)\|_{H^1}) + C \int_0^t \|G(\tau)\|_{L^2}^2 \, d\tau. \tag{3-2}$$

In particular, *for any* $0 < \sigma < \beta$,

$$\int_0^t \|\Lambda^\sigma \theta(\tau)\|_{L^2}^2 \, d\tau \le C(t, \|(u_0, \theta_0)\|_{H^1}) + C \int_0^t \|G(\tau)\|_{L^2}^2 \, d\tau. \tag{3-3}$$

We provide the proof of Lemma 3.2 and Corollary 3.3.

Proof of Lemma 3.2 and Corollary 3.3. Applying the Fourier localization operator Δ_j with $j \in \mathbb{Z}$ and $j \ge -1$ to the equation of θ and then dotting the resulting equation with $\Delta_j \theta$ yields

$$\frac{1}{2}\frac{d}{dt}\|\Delta_j\theta\|_{L^2}^2 + 2^{2\beta j}\|\Delta_j\theta\|_{L^2}^2 = -\int \Delta_j\theta\,[\Delta_j, u\cdot\nabla\theta]\,dx \le \|\Delta_j\theta\|_{L^2}\,\|[\Delta_j, u\cdot\nabla\theta]\|_{L^2}.$$

Applying a standard commutator estimate (see, e.g, [Hmidi et al. 2011, p. 443])

$$\|[\Delta_j, u\cdot\nabla]\theta\|_{L^2} \le C\,\|\theta\|_{B^0_{\infty,\infty}}\,\|\nabla u\|_{L^2},$$

we obtain

$$\frac{d}{dt}\|\Delta_j\theta\|_{L^2} + 2^{\beta j}\,\|\Delta_j\theta\|_{L^2} \le C\,\|\theta_0\|_{L^\infty}\,\|\omega\|_{L^2}.$$

Integrating in time yields

$$\|\Delta_j\theta(t)\|_{L^2} \le C\,e^{-2^{\beta j}t}\,\|\Delta_j\theta_0\|_{L^2} + C\int_0^t e^{-2^{\beta j}(t-\tau)}\|\omega(\tau)\|_{L^2}\,d\tau.$$

Taking the L^2-norm in time and applying Young's inequality for convolution, we have

$$\left[\int_0^t \|\Delta_j\theta(\tau)\|_{L^2}^2\,d\tau\right]^{\frac{1}{2}} \le C\,2^{-\frac{1}{2}\beta j}\,\|\Delta_j\theta_0\|_{L^2} + C\,2^{-\beta j}\left[\int_0^t \|\omega(\tau)\|_{L^2}^2\,d\tau\right]^{\frac{1}{2}}.$$

Multiplying each side by $2^{\beta j}$ and then squaring each side yields

$$\int_0^t 2^{2\beta j}\|\Delta_j\theta(\tau)\|_{L^2}^2\,d\tau \le C\,2^{\beta j}\,\|\Delta_j\theta_0\|_{L^2}^2 + C_0\int_0^t \|\omega(\tau)\|_{L^2}^2\,d\tau.$$

Taking the supremum with respect to j yields (3-1). To show (3-2), we note that

$$\|\omega\|_{L^2} \le \|G\|_{L^2} + \|\Lambda^{1-\alpha}\theta\|_{L^2}. \tag{3-4}$$

For any $\sigma < \beta$, we choose a large integer j_0 such that

$$\sum_{j\ge j_0+1} 2^{2(\sigma-\beta)j} < \frac{1}{4\widetilde{C}}.$$

Then

$$\|\Lambda^\sigma\theta\|_{L^2}^2 = \sum_{j\le j_0} 2^{2\sigma j}\|\Delta_j\theta\|_{L^2}^2 + \sum_{j\ge j_0+1} 2^{2(\sigma-\beta)j}\, 2^{2\beta j}\|\Delta_j\theta\|_{L^2}^2 \le C(j_0, \|\theta_0\|_{L^2}) + \frac{1}{4\widetilde{C}}\sup_{j\ge -1} 2^{2\beta j}\|\Delta_j\theta\|_{L^2}^2. \tag{3-5}$$

Inserting (3-4) and (3-5) in (3-1) yields (3-2), and (3-3) follows from (3-5). This completes the proof of Lemma 3.2 and Corollary 3.3. □

We also need the following lemma (see [Stefanov and Wu 2018]).

Lemma 3.4. *Let* $1 > \alpha > \frac{1}{2}$, $1 < p_2 < \infty$, $1 < p_1 < \infty$, *and* $1 < p_3 \le \infty$, *so that* $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. *For every* s_1 $(0 \le s_1 < 1-\alpha)$ *and* s_2 $(s_2 > 1-\alpha-s_1)$, *there exists a* $C = C(p_1, p_2, p_3, s_1, s_2)$, *such that*

$$\left|\int_{\mathbb{R}^d} F[\mathcal{R}_\alpha, u_G\cdot\nabla]\theta\, dx\right| \le C\|\Lambda^{s_1}\theta\|_{L^{p_1}}\|F\|_{W^{s_2,p_2}}\|G\|_{L^{p_3}}. \tag{3-6}$$

Similarly, for every s_1 $(0 \le s_1 < 1-\alpha)$ *and* s_2 $(s_2 > 2-2\alpha-s_1)$, *we have*

$$\left|\int_{\mathbb{R}^d} F[\mathcal{R}_\alpha, u_\theta\cdot\nabla]\psi\, dx\right| \le C\|\Lambda^{s_1}\theta\|_{L^{p_1}}\|F\|_{W^{s_2,p_2}}\|\psi\|_{L^{p_3}}. \tag{3-7}$$

Here u_G *denotes the velocity associated with* G, *namely* $u_G = \nabla^\perp(-\Delta)^{-1}G$, *and* $u_\theta = \nabla^\perp(-\Delta)^{-1}\partial_1\Lambda^{-\alpha}\theta$. *The definition of* G *implies that* $u = u_G + u_\theta$.

We now prove Proposition 3.1.

Proof of Proposition 3.1. This proof is obtained by modifying that for the global L^2 bound of G in Stefanov and Wu [2018]. Dotting (2-5) with G and integrating by parts yields

$$\frac{1}{2}\frac{d}{dt}\|G\|_{L^2}^2 + \|\Lambda^{\alpha/2}G\|_{L^2}^2 = J_1 + J_2, \tag{3-8}$$

where

$$J_1 = \int G\,\Lambda^{\beta-\alpha}\partial_1\theta\, dx, \qquad J_2 = \int G\,[\mathcal{R}_\alpha, u\cdot\nabla]\theta\, dx.$$

By Hölder's inequality and Corollary 3.3 with $\sigma = \beta + 1 - \frac{3}{2}\alpha$ $\left(\sigma < \beta \text{ since } \alpha > \frac{2}{3}\right)$,

$$|J_1| \le \|\Lambda^{\alpha/2}G\|_{L^2}\,\|\Lambda^{\beta+1-3\alpha/2}\theta\|_{L^2} \le \tfrac{1}{4}\|\Lambda^{\alpha/2}G\|_{L^2}^2 + \|\Lambda^{\beta+1-3\alpha/2}\theta\|_{L^2}^2.$$

As in [Jiu et al. 2014] and [Stefanov and Wu 2018], we write

$$u = \nabla^\perp\Delta^{-1}\omega = \nabla^\perp\Delta^{-1}G + \nabla^\perp\Delta^{-1}\mathcal{R}_\alpha\theta \equiv u_G + u_\theta.$$

J_2 is then split into two parts accordingly. The term with u_G part is estimated as in [Stefanov and Wu 2018]. For $\alpha > \frac{2}{3}$, we choose $1-\alpha < s < \frac{\alpha}{2}$ and apply Lemma 3.4,

$$\left|\int G\,[\mathcal{R}_\alpha, u_G\cdot\nabla]\theta\,dx\right| \le C\,\|\theta_0\|_{L^\infty}\,\|G\|_{L^2}\,\|G\|_{H^{\alpha/2}} \le \tfrac{1}{4}\|\Lambda^{\alpha/2}G\|_{L^2}^2 + C\,\|G\|_{L^2}^2.$$

To bound the term associated with u_θ, we apply Lemma 3.4 with $s_1 = \beta+1-\frac{3}{2}\alpha$ and $s_2 = \frac{1}{2}(1-\beta)$. Since $\alpha > \frac{2}{3}$ and $\alpha+\beta > 1$, we have $s_1 < \beta$ and $s_2 < \frac{\alpha}{2}$, and $s_1+s_2 > 2-2\alpha$. Therefore

$$\begin{aligned}\left|\int G\,[\mathcal{R}_\alpha, u_\theta\cdot\nabla]\theta\,dx\right| &\le C\,\|\Lambda^{s_1}\theta\|_{L^2}\,\|\theta\|_{L^\infty}\,\|G\|_{H^{s_2}}\\ &\le C\,\|\theta_0\|_{L^\infty}\,\|\Lambda^{\beta+1-3\alpha/2}\theta\|_{L^2}\,\|G\|_{H^{\alpha/2}}\\ &\le \tfrac{1}{4}\,\|\Lambda^{\alpha/2}G\|_{L^2}^2 + C\,\|\Lambda^{\beta+1-3\alpha/2}\theta\|_{L^2}^2.\end{aligned}$$

Inserting the bounds above in (3-8), and applying Corollary 3.3 with $\sigma = \beta+1-\frac{3}{2}\alpha$ and Gronwall's inequality yields the desired global bound. □

In order to prove Theorem 1.2, we also need the following lower bound for the fractional Laplacian operator. The proof of this lemma follows the lines of Constantin and Vicol [2012] and will be provided in Appendix B.

Lemma 3.5. *Let $\alpha \in (0,2)$. For any smooth function f that decays sufficiently fast at infinity, suppose that $\bar{x} \in \mathbb{R}^2$ is a point at which $|f(x)|$ attains its maximum. Then,*

$$f\,\Lambda^\alpha f \ge C\,\frac{|f(\bar{x})|^{2+\alpha}}{\|f\|_{L^2}^\alpha}$$

for a constant $C = C(\alpha)$.

Proof of Theorem 1.2. Making use of Lemma 3.5, we obtain, as in the derivation of (2-6),

$$\frac{d}{dt}\|G\|_{L^\infty} + C\frac{\|G\|_{L^\infty}^{1+\alpha}}{\|G\|_{L^2}^\alpha} \le C\,\|\nabla u\|_{L^\infty}\,\|\nabla\theta\|_{L^\infty}^{1-\alpha} + C\,\|\nabla\theta\|_{L^\infty}^{\beta+1-\alpha},$$

$$\frac{d}{dt}\|\nabla\theta\|_{L^\infty} + C\frac{\|\nabla\theta\|_{L^\infty}^{1+\beta}}{\|\theta\|_{L^\infty}^\beta} \le \|\nabla u\|_{L^\infty}\|\nabla\theta\|_{L^\infty}.$$

We further use (2-9) and (2-10) to obtain

$$\begin{aligned}\frac{d}{dt}\|G\|_{L^\infty} + C_1(t)\,\|G\|_{L^\infty}^{1+\alpha} &\le C_2\,\|G\|_{L^\infty}\,\|\nabla\theta\|_{L^\infty}^{1-\alpha}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty})\\ &\quad + C_3\,\|\nabla\theta\|_{L^\infty}^{2(1-\alpha)}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) + C_4\,\|\nabla\theta\|_{L^\infty}^{1+\beta-\alpha},\end{aligned}$$

$$\frac{d}{dt}\|\nabla\theta\|_{L^\infty} + C_5\,\|\nabla\theta\|_{L^\infty}^{1+\beta} \le C_6\,\|G\|_{L^\infty}\,\|\nabla\theta\|_{L^\infty}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}) \\ +C_7\,\|\nabla\theta\|_{L^\infty}^{2-\alpha}\,L(\|G\|_{L^\infty}, \|\nabla\theta\|_{L^\infty}),$$

where $C_1(t) = C\,(\|G(t)\|_{L^2})^{-1}$, $C_5 = C\,(\|\theta\|_{L^\infty})^{-1}$, and L is as defined in (2-13). We can then argue in a similar fashion as in the proof of Theorem 1.1 that the global bounds $\|G\|_{L^\infty} < \infty$ and $\|\nabla\theta\|_{L^\infty} < \infty$ hold if α and β satisfy (1-4). In fact, if $\frac{2}{3} < \alpha < 1$ and $\alpha > \frac{1}{1+\beta}$, then

$$\alpha > \frac{1-\alpha}{\beta}, \qquad 1+\alpha > \frac{2-2\alpha}{\beta}, \qquad 1+\alpha > \frac{1+\beta-\alpha}{\beta}$$

and the argument in the proof of Theorem 1.1 works here. This completes the proof of Theorem 1.2. □

Appendix A. Frequency localization and Besov spaces

This appendix provides the definition of the Littlewood–Paley decomposition and the definition of Besov spaces. Some related facts used in the previous sections are also included. The material presented in this appendix can be found in several books and many papers (see, e.g., [Bahouri et al. 2011; Bergh and Löfström 1976; Miao et al. 2012; Runst and Sickel 1996; Triebel 1992]).

We start with several notational conventions. $\mathcal{S}$ denotes the usual Schwarz class and $\mathcal{S}'$ its dual, the space of tempered distributions. To introduce the Littlewood–Paley decomposition, we write for each $j \in \mathbb{Z}$,

$$A_j = \{\xi \in \mathbb{R}^d : \; 2^{j-1} \le |\xi| < 2^{j+1}\}.$$

The Littlewood–Paley decomposition asserts the existence of a sequence of functions $\{\Phi_j\}_{j\in\mathbb{Z}} \in \mathcal{S}$ such that

$$\operatorname{supp}\widehat{\Phi}_j \subset A_j, \qquad \widehat{\Phi}_j(\xi) = \widehat{\Phi}_0(2^{-j}\xi) \quad \text{or} \quad \Phi_j(x) = 2^{jd}\Phi_0(2^j x),$$

and

$$\sum_{j=-\infty}^{\infty} \widehat{\Phi}_j(\xi) = \begin{cases} 1 & \text{if } \xi \in \mathbb{R}^d \setminus \{0\}, \\ 0 & \text{if } \xi = 0. \end{cases}$$

Therefore, for a general function $\psi \in \mathcal{S}$, we have

$$\sum_{j=-\infty}^{\infty} \widehat{\Phi}_j(\xi)\widehat{\psi}(\xi) = \widehat{\psi}(\xi) \quad \text{for } \xi \in \mathbb{R}^d \setminus \{0\}.$$

We now choose $\Psi \in \mathcal{S}$ such that

$$\widehat{\Psi}(\xi) = 1 - \sum_{j=0}^{\infty} \widehat{\Phi}_j(\xi), \quad \xi \in \mathbb{R}^d.$$

Then, for any $\psi \in \mathcal{S}$,

$$\Psi * \psi + \sum_{j=0}^{\infty} \Phi_j * \psi = \psi$$

and hence

$$\Psi * f + \sum_{j=0}^{\infty} \Phi_j * f = f \tag{A-1}$$

in $\mathcal{S}'$ for any $f \in \mathcal{S}'$. To define the inhomogeneous Besov space, we set

$$\Delta_j f = \begin{cases} 0 & \text{if } j \le -2, \\ \Psi * f & \text{if } j = -1, \\ \Phi_j * f & \text{if } j = 0, 1, 2, \cdots . \end{cases} \tag{A-2}$$

Besides the Fourier localization operators Δ_j, the partial sum S_j is also a useful notation. For an integer j,

$$S_j \equiv \sum_{k=-1}^{j-1} \Delta_k.$$

For any $f \in \mathcal{S}'$, the Fourier transform of $S_j f$ is supported on the ball of radius 2^j. It is clear from (A-1) that $S_j \to \mathrm{Id}$ as $j \to \infty$ in the distributional sense. In addition, the notation $\widetilde{\Delta}_k$, defined by

$$\widetilde{\Delta}_k = \Delta_{k-1} + \Delta_k + \Delta_{k+1},$$

is also useful and has been used in the previous sections.

Definition A.1. The inhomogeneous Besov space $B^s_{p,q}$ with $s \in \mathbb{R}$ and $p, q \in [1, \infty]$ consists of $f \in \mathcal{S}'$ satisfying

$$\|f\|_{B^s_{p,q}} \equiv \|2^{js} \|\Delta_j f\|_{L^p}\|_{l^q} < \infty,$$

where $\Delta_j f$ is as defined in (A-2).

Many frequently used function spaces are special cases of Besov spaces. The following proposition lists some useful equivalence and embedding relations.

Proposition A.2. *For any $s \in \mathbb{R}$,*

$$H^s \sim B^s_{2,2}.$$

For any $s \in \mathbb{R}$ and $1 < q < \infty$,

$$B^s_{q,\min\{q,2\}} \hookrightarrow W^s_q \hookrightarrow B^s_{q,\max\{q,2\}}.$$

For any noninteger $s > 0$, the Hölder space C^s is equivalent to $B^s_{\infty,\infty}$.

Bernstein's inequalities are useful tools in dealing with Fourier localized functions. These inequalities trade integrability for derivatives. The following proposition provides Bernstein type inequalities for fractional derivatives. The upper bounds also hold when the fractional operators are replaced by partial derivatives.

Proposition A.3. *Let $\alpha \geq 0$. Let $1 \leq p \leq q \leq \infty$.*

(1) *If f satisfies*
$$\operatorname{supp} \widehat{f} \subset \{\xi \in \mathbb{R}^d : |\xi| \leq K 2^j\},$$
for some integer j and a constant $K > 0$, then
$$\|(-\Delta)^\alpha f\|_{L^q(\mathbb{R}^d)} \leq C_1\, 2^{2\alpha j + jd(1/p-1/q)} \|f\|_{L^p(\mathbb{R}^d)}.$$

(2) *If f satisfies*
$$\operatorname{supp} \widehat{f} \subset \{\xi \in \mathbb{R}^d : K_1 2^j \leq |\xi| \leq K_2 2^j\}$$
for some integer j and constants $0 < K_1 \leq K_2$, then
$$C_1\, 2^{2\alpha j} \|f\|_{L^q(\mathbb{R}^d)} \leq \|(-\Delta)^\alpha f\|_{L^q(\mathbb{R}^d)} \leq C_2\, 2^{2\alpha j + jd(1/p-1/q)} \|f\|_{L^p(\mathbb{R}^d)},$$
where C_1 and C_2 are constants depending on α, p and q only.

Appendix B. Proofs of facts used in the previous sections

This appendix provides the proofs of several facts used in Sections 2 and 3.

We first provide several pointwise inequalities involving fractional Laplacian operators. These lower bounds here are in terms of the L^p-norms of the functions instead of the L^p-norms of the antiderivatives. Therefore, these lower bounds have higher powers than the corresponding lower bounds in terms of the antiderivatives. The proofs of these lower bounds follow the ideas of Constantin and Vicol [2012].

Lemma B.1. *Let $p \in [1, \infty)$. Assume $f \geq 0$, $f \in L^p(\mathbb{R}^d)$ and $f \in C^1(\mathbb{R}^d)$. Suppose that f attains its maximum value at the point $\bar{x}$. Then,*

$$\Lambda^\alpha f(\bar{x}) \geq \frac{f(\bar{x})^{1+\alpha p/d}}{c\|f\|_{L^p}^{\alpha p/d}} \tag{B-1}$$

for some constant $c = c(d, \alpha, p)$.

Proof. Let χ be a radially nondecreasing smooth cut-off function, which vanishes on $|x| \leq 1$ and is identically 1 on $|x| \geq 2$, and $|\nabla \chi| \leq 4$. Let $R > 0$ be a number to be specified later. We estimate

$$\begin{aligned}\Lambda^\alpha f(\bar{x}) &= c_{d,\alpha}\int_{\mathbb{R}^d}\frac{f(\bar{x})-f(\bar{x}-y)}{|y|^{d+\alpha}}\,dy\\ &\geq c_{d,\alpha} f(\bar{x})\int_{\mathbb{R}^d}\frac{\chi(y/R)}{|y|^{d+\alpha}}dy - c_{d,\alpha}\left|\int_{\mathbb{R}^d} f(\bar{x}-y)\frac{\chi(y/R)}{|y|^{d+\alpha}}dy\right|\\ &\geq c_{d,\alpha} f(\bar{x})\int_{|y|\geq 2R}\frac{1}{|y|^{d+\alpha}}dy - c_{d,\alpha}\|f\|_{L^p}\left(\int_{\mathbb{R}^d}\left|\frac{\chi(y/R)}{|y|^{d+\alpha}}\right|^{p'}dy\right)^{1/p'}\\ &\geq c_1\frac{f(\bar{x})}{R^\alpha} - c_2\frac{\|f\|_{L^p}}{R^{\alpha+d/p}},\end{aligned}$$

where $c_1 = c_1(d,\alpha)$, and $c_2 = c_2(d,\alpha,\delta)$ are positive constants, which may be computed explicitly. Letting $R^{d/p} = 2c_2\|f\|_{L^p}/(c_1 f(\bar{x}))$ concludes the proof. □

Lemma B.2. *Let $\alpha\in(0,2)$ and let $p\in[1,\infty)$. Assume $f\in L^p(\mathbb{R}^d)$ and $f\in C^1(\mathbb{R}^d)$. Then we have the pointwise bound*

(B-2) $$f(x)\cdot\Lambda^\alpha f(x) \geq \tfrac{1}{2}\Lambda^\alpha|f(x)|^2 + \frac{|f(x)|^{2+p\alpha/d}}{c\|f\|_{L^p}^{p\alpha/d}}$$

for some positive constant $c = c(d,\alpha,p)$.

Proof. Recall the pointwise identity (see [Constantin and Vicol 2012])

(B-3) $$f(x)\cdot\Lambda^\alpha f(x) = \tfrac{1}{2}\Lambda^\alpha(|f|^2)(x) + \tfrac{1}{2}D,$$

where

(B-4) $$D = c_{d,\alpha} PV\int_{\mathbb{R}^d}\frac{|f(x)-f(x+y)|^2}{|y|^{d+\alpha}}\,dy.$$

For χ defined as in the previous proof,

$$\begin{aligned}D &\geq c_{d,\alpha}\int_{\mathbb{R}^d}\frac{|f(x)-f(x+y)|^2}{|y|^{d+\alpha}}\chi(y/R)dy\\ &\geq c_{d,\alpha}|f(x)|^2\int_{\mathbb{R}^d}\frac{\chi(y/R)}{|y|^{d+\alpha}}dy - 2c_{d,\alpha}|f(x)|\left|\int_{\mathbb{R}^d} f(x+y)\frac{\chi(y/R)}{|y|^{d+\alpha}}\,dy\right|\\ &\geq c_{d,\alpha}|f(x)|^2\int_{|y|\geq R}\frac{1}{|y|^{d+\alpha}}dy - 2\,c_{d,\alpha}|f(x)|\,\|f\|_{L^p}\left(\int_{\mathbb{R}^d}\left|\frac{\chi(y/R)}{|y|^{d+\alpha}}\right|^{p'}dy\right)^{1/p'}\\ &\geq c_1\frac{|f(x)|^2}{R^\alpha} - c_2\frac{|f(x)|\,\|f\|_{L^p}}{R^{\alpha+d/p}},\end{aligned}$$

for some positive constants c_1 and c_2 which depend only on d, α, and p. Letting

$$R^{d/p} = \frac{c_2 \|f\|_{L^p}}{2c_1 |f(x)|}$$

concludes the proof of this lemma. □

A special consequence is the following lower bound.

Corollary B.3. *Let $\alpha \in (0, 2)$. Assume f is smooth and decays sufficiently fast at infinity. Assume that $\bar{x} \in \mathbb{R}^d$ is a maximum point at which $|f(x)|$ attains its maximum. Then,*

$$f(\bar{x}) \cdot \Lambda^\alpha f(\bar{x}) \geq \frac{|f(\bar{x})|^{2+p\alpha/d}}{c\|f\|_{L^p}^{p\alpha/d}}$$

where $c = c(d, \alpha, p)$.

Next we provide the proof of Lemma 2.4.

Proof. Write $\mathcal{R}_\alpha = \partial_{x_1} \Lambda^{-\alpha}$, then $\Delta_k \mathcal{R}_\alpha = 2^{(1-\alpha)k} h_k$, where $h_k(x) = 2^{dk} h_0(2^k x)$ and $h_0(x) \in C_0^\infty(\mathbb{R}^d)$. By the notion of paraproducts,

$$\Delta_k[\mathcal{R}_\alpha, u\cdot\nabla]\theta = \sum_{|j-k|\leq 2} \Delta_k[\mathcal{R}_\alpha, S_{j-1}u\cdot\nabla]\Delta_j\theta + \sum_{|j-k|\leq 2} \Delta_k[\mathcal{R}_\alpha, \Delta_j u\cdot\nabla]S_{j-1}\theta + \sum_{j\geq k-4} [\mathcal{R}_\alpha, \Delta_j u\cdot\nabla]\tilde{\Delta}_j\theta := J_1 + J_2 + J_3. \tag{B-5}$$

We estimate the L^∞-norm of the terms on the right.

$$\begin{aligned}
\|J_1\|_{L^\infty} &\leq C2^{(1-\alpha)k} \||x|2^{kd} h_0(2^k x)\|_{L^1} \|\nabla S_{k-1} u\|_{L^\infty} \|\Delta_k \nabla\theta\|_{L^\infty} \\
&\leq C2^{-\alpha k} \|\nabla\Delta_k\theta\|_{L^\infty} \left(\|\nabla\Delta_{-1}\|_{L^\infty} + \sum_{j=0}^{k-2} \|\Delta_j \nabla u\|_{L^\infty} \right) \\
&\leq C2^{-\alpha k} \|\nabla\Delta_k\theta\|_\infty (\|u\|_{L^2} + k\|\omega\|_{L^\infty}).
\end{aligned}$$

For J_2 and J_3, we have

$$\begin{aligned}
\|J_2\|_{L^\infty} &\leq C\|\Delta_k \mathcal{R}_\alpha(\Delta_k u \cdot \nabla S_{k-1}\theta) - \Delta_k(\Delta_k u \cdot \nabla \mathcal{R}_\alpha(S_{k-1}\theta - \Delta_{-1}))\|_{L^\infty} \\
&\qquad + \|\Delta_k(\Delta_k u \cdot \nabla \mathcal{R}_\alpha \Delta_{-1}\theta)\|_{L^\infty} \\
&\leq C2^{(1-\alpha)k} \||x|2^{dk} h_0(2^k x)\|_{L^1} \|\nabla\Delta_k u\|_{L^\infty} \|\nabla S_{k-1}\theta\|_{L^\infty} \\
&\qquad + C\|\Delta_k u\|_{L^\infty} \|\nabla \mathcal{R}_\alpha \Delta_{-1}\theta\|_{L^\infty} \\
&\leq C2^{-k\alpha} \|\Delta_k\omega\|_{L^\infty} \left(\sum_{j=-1}^{k-2} \|\nabla\Delta_j\theta\|_{L^\infty} \right) + C\|\theta\|_{L^2} \|\Delta_k u\|_{L^\infty}.
\end{aligned}$$

$$\begin{aligned}\|J_3\|_{L^\infty} \le& \sum_{j\le 1} \|\nabla\cdot\Delta_k\mathcal{R}_\alpha(\Delta_j u\tilde{\Delta}_j\theta)\|_{L^\infty} + \|\nabla\cdot\Delta_k(\Delta_j u\mathcal{R}_\alpha\tilde{\Delta}_j\theta)\|_{L^\infty}\\ &+ \sum_{j\ge\max(2,k-4)} \|\Delta_k\mathcal{R}_\alpha(\Delta_j u\cdot\nabla\tilde{\Delta}_j\theta)\|_{L^\infty} + \|\Delta_k\nabla\cdot(\Delta_j u\cdot\mathcal{R}_\alpha\tilde{\Delta}_j\theta)\|_{L^\infty}\\ \le& C\|u\|_{L^2}\|\theta\|_{L^2} + \sum_{j\ge\max(2,k-1)} 2^{(1-\alpha)k}\|\Delta_j u\|_{L^\infty}\|\nabla\tilde{\Delta}_j\theta\|_{L^\infty}\\ &+ 2^k\|\Delta_j u\|_{L^\infty}\|\mathcal{R}_\alpha\tilde{\Delta}_j\theta\|_{L^\infty}.\end{aligned}$$

Therefore,

$$\|[\mathcal{R}_\alpha, u\cdot\nabla]\theta\|_{B^0_{\infty,1}} \le \sum_{k\ge -1}\|J_1\|_{L^\infty} + \sum_{k\ge -1}\|J_2\|_{L^\infty} + \sum_{k\ge -1}\|J_3\|_{L^\infty} := I_1 + I_2 + I_3$$

and

$$I_1 \le C(\|\omega\|_{L^2}+\|\omega\|_{L^\infty})\sum_{k\ge -1} 2^{(1-\alpha)+\varepsilon k}\|\Delta_k\theta\|_{L^\infty} \le C(\|\omega\|_{L^2}+\|\omega\|_{L^\infty})\|\theta\|_{B^{1-\alpha+\varepsilon}_{\infty,1}},$$

$$\begin{aligned}I_2 \le& C\sum_{k\ge -1}\|\Delta_k\omega\|_{L^\infty}\sum_{j=-1}^{k-2} 2^{\alpha(j-k)}2^{-\alpha j}\|\nabla\Delta_j\theta\|_{L^\infty} + C\|\theta\|_{L^2}\sum_{k\ge 0}2^{-k}\|\Delta_k\omega\|_{L^\infty} + C\\ \le& C\|\omega\|_{L^\infty}\|\theta\|_{B^{1-\alpha+\epsilon}_{\infty,1}} + C\|\omega\|_{L^\infty},\end{aligned}$$

$$\begin{aligned}I_3 \le& C\|u||_{L^2}\|\theta\|_{L^2} + C\sum_{k\ge -1}\sum_{j\ge\max(2,k-1)} 2^{(1-\alpha)(k-j)}\|\Delta_j\nabla u\|_{L^\infty}2^{-\alpha j}\|\nabla\tilde{\Delta}_j\theta\|_{L^\infty}\\ &+ C\sum_{k\ge -1}\sum_{j\ge\max(2,k-1)} 2^{k-j}\|\nabla\Delta_j u\|_{L^\infty}2^{(1-\alpha)j}\|\tilde{\Delta}_j\theta\|_{L^\infty}\\ \le& C\|u\|_{L^2}\|\theta\|_{L^2} + C\|\omega\|_{L^\infty}\|\theta\|_{B^{1-\alpha}_{\infty,1}}.\end{aligned}$$

Combining these estimates, we have

$$\|[\mathcal{R}_\alpha, u\cdot\nabla]\theta\|_{B^0_{\infty,1}} \le C(\|\omega\|_{L^2} + \|\omega\|_{L^\infty})\|\theta\|_{B^{1-\alpha+\varepsilon}_{\infty,1}} + C\|u\|_{L^2}\|\theta\|_{L^2}$$

for any $\varepsilon > 0$. □

Acknowledgements

D. Zhou was supported by the National Natural Science Foundation of China (NNSFC) (No. 11401176) and the Doctor Fund of Henan Polytechnic University(HPU) (No. B2012-110). Z. Li was supported by the NNSFC (No. 11601128), the Doctor Fund of HPU (No. B2016-57), and the Key Research Projects of University in Henan Province (No. 16A110015). H. Shang was supported by NNSFC (No. 11201124), the Key Research Projects of University in Henan Province (No.

2015GGJS-070) and the Outstanding Youth Foundation of HPU (No. J2014-03). J. Wu was supported by NSF grant DMS 1614246, and by the AT&T Foundation at Oklahoma State University. B. Yuan was supported by NNSFC (No. 11471103). J. Zhao was supported by the Doctoral Fund of HPU (No. B2016-61).

References

[Adhikari et al. 2010] D. Adhikari, C. Cao, and J. Wu, "The 2D Boussinesq equations with vertical viscosity and vertical diffusivity", *J. Differential Equations* **249**:5 (2010), 1078–1088. MR Zbl

[Adhikari et al. 2011] D. Adhikari, C. Cao, and J. Wu, "Global regularity results for the 2D Boussinesq equations with vertical dissipation", *J. Differential Equations* **251**:6 (2011), 1637–1655. MR Zbl

[Adhikari et al. 2014] D. Adhikari, C. Cao, J. Wu, and X. Xu, "Small global solutions to the damped two-dimensional Boussinesq equations", *J. Differential Equations* **256**:11 (2014), 3594–3613. MR Zbl

[Bahouri et al. 2011] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Math. Wissenschaften **343**, Springer, 2011. MR Zbl

[Bergh and Löfström 1976] J. Bergh and J. Löfström, *Interpolation spaces: an introduction*, Grundlehren der Math. Wissenschaften **223**, Springer, 1976. MR Zbl

[Cao and Wu 2013] C. Cao and J. Wu, "Global regularity for the two-dimensional anisotropic Boussinesq equations with vertical dissipation", *Arch. Ration. Mech. Anal.* **208**:3 (2013), 985–1004. MR Zbl

[Caputo 1967] M. Caputo, "Linear models of dissipation whose Q is almost frequency independent, II", *Geophys. J. Roy. Astronomical Soc.* **13**:5 (1967), 529–539.

[Constantin and Doering 1999] P. Constantin and C. R. Doering, "Infinite Prandtl number convection", *J. Statist. Phys.* **94**:1-2 (1999), 159–172. MR Zbl

[Constantin and Vicol 2012] P. Constantin and V. Vicol, "Nonlinear maximum principles for dissipative linear nonlocal operators and applications", *Geom. Funct. Anal.* **22**:5 (2012), 1289–1321. MR Zbl

[Constantin et al. 2015] P. Constantin, A. Tarfulea, and V. Vicol, "Long time dynamics of forced critical SQG", *Comm. Math. Phys.* **335**:1 (2015), 93–141. MR Zbl

[Córdoba and Córdoba 2004] A. Córdoba and D. Córdoba, "A maximum principle applied to quasi-geostrophic equations", *Comm. Math. Phys.* **249**:3 (2004), 511–528. MR Zbl

[Danchin and Paicu 2011] R. Danchin and M. Paicu, "Global existence results for the anisotropic Boussinesq system in dimension two", *Math. Models Methods Appl. Sci.* **21**:3 (2011), 421–457. MR Zbl

[Gill 1982] A. Gill, *Atmosphere-ocean dynamics*, International Geophysics Series **30**, Academic Press, New York, 1982.

[Hajaiej et al. 2011] H. Hajaiej, L. Molinet, T. Ozawa, and B. Wang, "Necessary and sufficient conditions for the fractional Gagliardo–Nirenberg inequalities and applications to Navier–Stokes and generalized boson equations", pp. 159–175 in *Harmonic analysis and nonlinear partial differential equations* (Kyoto, 2010), edited by T. Ozawa and M. Sugimoto, RIMS Kôkyûroku Bessatsu **B26**, Res. Inst. Math. Sci., Kyoto, 2011. MR Zbl

[Hmidi et al. 2010] T. Hmidi, S. Keraani, and F. Rousset, "Global well-posedness for a Boussinesq–Navier–Stokes system with critical dissipation", *J. Differential Equations* **249**:9 (2010), 2147–2174. MR Zbl

[Hmidi et al. 2011] T. Hmidi, S. Keraani, and F. Rousset, "Global well-posedness for Euler–Boussinesq system with critical dissipation", *Comm. Partial Differential Equations* **36**:3 (2011), 420–445. MR Zbl

[Hou and Li 2005] T. Y. Hou and C. Li, "Global well-posedness of the viscous Boussinesq equations", *Discrete Contin. Dyn. Syst.* **12**:1 (2005), 1–12. MR Zbl

[Jiu et al. 2014] Q. Jiu, C. Miao, J. Wu, and Z. Zhang, "The two-dimensional incompressible Boussinesq equations with general critical dissipation", *SIAM J. Math. Anal.* **46**:5 (2014), 3426–3454. MR Zbl

[KC et al. 2014] D. KC, D. Regmi, L. Tao, and J. Wu, "Generalized 2D Euler–Boussinesq equations with a singular velocity", *J. Differential Equations* **257**:1 (2014), 82–108. MR Zbl

[Lai et al. 2011] M.-J. Lai, R. Pan, and K. Zhao, "Initial boundary value problem for two-dimensional viscous Boussinesq equations", *Arch. Ration. Mech. Anal.* **199**:3 (2011), 739–760. MR Zbl

[Larios et al. 2013] A. Larios, E. Lunasin, and E. S. Titi, "Global well-posedness for the 2D Boussinesq system with anisotropic viscosity and without heat diffusion", *J. Differential Equations* **255**:9 (2013), 2636–2654. MR Zbl

[Li and Titi 2016] J. Li and E. S. Titi, "Global well-posedness of the 2D Boussinesq equations with vertical dissipation", *Arch. Ration. Mech. Anal.* **220**:3 (2016), 983–1001. MR Zbl

[Li et al. 2016] J. Li, H. Shang, J. Wu, X. Xu, and Z. Ye, "Regularity criteria for the 2D Boussinesq equations with supercritical dissipation", *Commun. Math. Sci.* **14**:7 (2016), 1999–2022. MR Zbl

[Majda 2003] A. Majda, *Introduction to PDEs and waves for the atmosphere and ocean*, Courant Lecture Notes in Math. **9**, Courant Inst. Math. Sci., New York, 2003. MR Zbl

[Majda and Bertozzi 2001] A. J. Majda and A. L. Bertozzi, *Vorticity and incompressible flow*, Cambridge Texts in Applied Math. **27**, Cambridge Univ. Press, 2001. MR Zbl

[Miao and Xue 2011] C. Miao and L. Xue, "On the global well-posedness of a class of Boussinesq–Navier–Stokes systems", *Nonlinear Differential Equations Appl.* **18**:6 (2011), 707–735. MR Zbl

[Miao et al. 2012] C. X. Miao, W. J. Hong, and Z. F. Zhang, *Littlewood–Paley* 理论及其在流体动力学方程中的应用 , Monographs on Modern Pure Math. **142**, Sci. Press, Beijing, 2012.

[Ohkitani 2001] K. Ohkitani, "Comparison between the Boussinesq and coupled Euler equations in two dimensions", pp. 127–145 in *Tosio Kato's method and principle for evolution equations in mathematical physics* (Sapporo, 2001), edited by S. T. Kuroda et al., Sūrikaisekikenkyūsho Kōkyūroku **1234**, Res. Inst. Math. Sci., Kyoto, 2001. MR

[Pedlosky 1979] J. Pedlosky, *Geophysical fluid dynamics*, Springer, 1979. Zbl

[Runst and Sickel 1996] T. Runst and W. Sickel, *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, de Gruyter Series in Nonlinear Anal. and Appl. **3**, de Gruyter, Berlin, 1996. MR Zbl

[Stefanov and Wu 2018] A. Stefanov and J. Wu, "A global regularity result for the 2D Boussinesq equations with critical dissipation", *Journal d'Analyse Mathématique* (online publication December 2018).

[Triebel 1992] H. Triebel, *Theory of function spaces, II*, Monographs in Math. **84**, Birkhäuser, Basel, 1992. MR Zbl

[Wu and Xu 2014] J. Wu and X. Xu, "Well-posedness and inviscid limits of the Boussinesq equations with fractional Laplacian dissipation", *Nonlinearity* **27**:9 (2014), 2215–2232. MR Zbl

[Wu et al. 2015] J. Wu, X. Xu, and Z. Ye, "Global smooth solutions to the n-dimensional damped models of incompressible fluid mechanics with small initial datum", *J. Nonlinear Sci.* **25**:1 (2015), 157–192. MR Zbl

[Wu et al. 2016] J. Wu, X. Xu, L. Xue, and Z. Ye, "Regularity results for the 2D Boussinesq equations with critical or supercritical dissipation", *Commun. Math. Sci.* **14**:7 (2016), 1963–1997. MR Zbl

[Yang et al. 2014] W. Yang, Q. Jiu, and J. Wu, "Global well-posedness for a class of 2D Boussinesq systems with fractional dissipation", *J. Differential Equations* **257**:11 (2014), 4188–4213. MR Zbl

[Ye 2017] Z. Ye, "Global smooth solution to the 2D Boussinesq equations with fractional dissipation", *Math. Methods Appl. Sci.* **40**:12 (2017), 4595–4612. MR Zbl

[Ye and Xu 2016] Z. Ye and X. Xu, "Global well-posedness of the 2D Boussinesq equations with fractional Laplacian dissipation", *J. Differential Equations* **260**:8 (2016), 6716–6744. MR Zbl

[Zhao 2010] K. Zhao, "2D inviscid heat conductive Boussinesq equations on a bounded domain", *Michigan Math. J.* **59**:2 (2010), 329–352. MR Zbl

[Zhou 2018] D. Zhou, "Global regularity of the two-dimensional Boussinesq equations without diffusivity in bounded domains", *Nonlinear Anal. Real World Appl.* **43** (2018), 144–154. MR Zbl

[Zhou and Li 2017] D. Zhou and Z. Li, "Global well-posedness for the 2D Boussinesq equations with zero viscosity", *J. Math. Anal. Appl.* **447**:2 (2017), 1072–1079. MR Zbl

Received March 7, 2017. Revised January 12, 2018.

DAOGUO ZHOU
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCE
HENAN POLYTECHNIC UNIVERSITY
HENAN
CHINA
daoguozhou@hpu.edu.cn

ZILAI LI
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCE
HENAN POLYTECHNIC UNIVERSITY
HENAN
CHINA
lizl@hpu.edu.cn

HAIFENG SHANG
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCE
HENAN POLYTECHNIC UNIVERSITY
HENAN
CHINA
hfshang@hpu.edu.cn

JIAHONG WU
DEPARTMENT OF MATHEMATICS
OKLAHOMA STATE UNIVERSITY
STILLWATER, OK
UNITED STATES
jiahong.wu@okstate.edu

BAOQUAN YUAN
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCE
HENAN POLYTECHNIC UNIVERSITY
HENAN
CHINA
bqyuan@hpu.edu.cn

JIEFENG ZHAO
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCE
HENAN POLYTECHNIC UNIVERSITY
HENAN
CHINA
zhaojiefeng003@hpu.edu.cn

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 298 No. 1 January 2019