Pacific Journal of Mathematics

Volume 273 No. 1 January 2015

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

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

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

The subscription price for 2015 is US $420/year for the electronic version, and $570/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.1

MAXIMAL ESTIMATES FOR SCHRÖDINGER EQUATIONS WITH INVERSE-SQUARE POTENTIAL

CHANGXING MIAO, JUNYONG ZHANG AND JIQIANG ZHENG

We consider maximal estimates for the solution to an initial value problem of the linear Schrödinger equation with a singular potential. We show a result about the pointwise convergence of solutions to this special variable coefficient Schrödinger equation with initial data $u_0 \in H^s(\mathbb{R}^n)$ for $s > \frac{1}{2}$ or radial initial data $u_0 \in H^s(\mathbb{R}^n)$ for $s \geqslant \frac{1}{4}$ and that the solution does not converge when $s < \frac{1}{4}$.

1. Introduction and statement of main result

We study the maximal estimates for the solution to an initial value problem of the linear Schrödinger equation with an inverse square potential. More precisely, we consider the Schrödinger equation

$$\begin{cases} i\partial_t u - \Delta u + \dfrac{a}{|x|^2}u = 0 & (t,x) \in \mathbb{R}\times\mathbb{R}^n \setminus \{0\},\ a > -(n-2)^2/4, \\ u(x,0) = u_0(x). \end{cases} \tag{1-1}$$

The scale-covariance elliptic operator $P_a := -\Delta + a/|x|^2$ appearing in (1-1) plays a key role in many problems of physics and geometry. The heat and wave flows for the elliptic operator P_a have been studied in the theory of combustion (see [Vazquez and Zuazua 2000]) and in wave propagation on conic manifolds (see [Cheeger and Taylor 1982]). The Schrödinger equation (1-1) arises in the study of quantum mechanics [Kalf et al. 1975]. There has been a lot of interest in developing Strichartz estimates both for the Schrödinger and wave equations with the inverse square potential; we refer the reader to [Burq et al. 2003; 2004; Planchon et al. 2003a; 2003b; Miao et al. 2013]. However, as far as we know, there are few results about the maximal estimates associated with the operator P_a, which arises in the study of pointwise convergence problem for the Schrödinger and wave equations with the inverse square potential. In this paper, we aim to address some maximal estimates in the special settings associated with the operator P_a. As a

MSC2010: 35B65, 35Q55, 47J35.
Keywords: inverse square potential, maximal estimate, spherical harmonics.

direct consequence, we obtain the pointwise convergence result for $u_0 \in H^s(\mathbb{R}^n)$ with $s > \frac{1}{2}$.

In the case of the free Schrödinger equation without potential, i.e., $a = 0$, there is a large amount of literature on developing maximal estimates for its solution, which can be formally written as

$$u(t,x) = e^{it\Delta}u_0(x) = \int_{\mathbb{R}^n} e^{2\pi i(x\cdot\xi - t|\xi|^2)}\hat{u}_0(\xi)\,\mathrm{d}\xi.$$

When $n = 1$, Carleson [1980] proved that the convergence result holds in the sense that $\lim_{t\to 0} u(t) = u_0$ for a.e. x when $u_0 \in H^s(\mathbb{R})$ with $s \geqslant \frac{1}{4}$. Dahlberg and Kenig [1982] showed that the result is sharp in the sense that the solution does not converge when $s < \frac{1}{4}$. When $n \geqslant 2$, Sjölin [1987] and Vega [1988] independently proved convergence results when $u_0 \in H^s(\mathbb{R}^n)$ with $s > \frac{1}{2}$. It follows from the construction of [Dahlberg and Kenig 1982; Vega 1988] that the solution does not converge when $s < \frac{1}{4}$. When $n = 2$, Bourgain [1995] showed that there is a certain $s < \frac{1}{2}$ such that the convergence result holds, and this result was improved by Moyua, Vargas and Vega [Moyua et al. 1996]. Having shown the bilinear restriction estimates for paraboloids, Tao and Vargas [2000] and Tao [2003] showed convergence for $s > \frac{15}{32}$ and $s > \frac{2}{5}$ respectively. This was improved further to $s > \frac{3}{8}$ in [Lee 2006; Shao 2010]. Very recently, Bourgain [2013] made some progress in high dimension $n \geqslant 2$ to show that the convergence result holds for $s > \frac{1}{2} - \frac{1}{4n}$ when $n \geqslant 1$ and that the convergence result needs $s \geqslant (n-2)/(2n)$ when $n \geqslant 5$.

In the situation when $a \neq 0$, (1-1) can be viewed as a special Schrödinger equation with variable singular coefficients. The potential prevents us from using the Fourier transform to give the expression of the solution. With the motivation of regarding the potential term as a perturbation on angular direction in [Burq et al. 2003; Planchon et al. 2003b; Miao et al. 2013], we express the solution by using the Hankel transform of radial functions and spherical harmonics. Instead of the Fourier transform, we utilize the Hankel transform and modify the argument of [Vega 1988] to show that the pointwise convergence result holds when the initial data $u_0 \in H^s(\mathbb{R}^n)$ for $s > \frac{1}{2}$, or when radial initial data $u_0 \in H^s(\mathbb{R}^n)$ for $s \geqslant \frac{1}{4}$, and that the solution does not converge when $s < \frac{1}{4}$.

Let u be the solution to (1-1); we define the maximal function by

$$u^*(x) = \sup_{|t|>0} |u(x,t)|. \tag{1-2}$$

Our main theorems are the following.

Theorem 1.1. *Let $\beta > 1$, $n \geqslant 2$ and $s > \frac{1}{2}$. Then*

$$\int_{\mathbb{R}^n} |u^*(x)|^2 \frac{\mathrm{d}x}{(1+|x|)^\beta} \leqslant C\|u_0\|^2_{H^s(\mathbb{R}^n)}. \tag{1-3}$$

As a direct consequence of Theorem 1.1, we have:

Corollary 1.1. *Let $u_0 \in H^s(\mathbb{R}^n)$ with $s > \frac{1}{2}$ and $n \geqslant 2$. Then*

$$\lim_{t \to 0} u(t,x) = u_0(x), \quad a.e.\ x \in \mathbb{R}^n. \tag{1-4}$$

Theorem 1.2. *Let B^n be the open unit ball in $\mathbb{R}^n$. Assume that there exists a constant C independent of u_0 such that*

$$\int_{B^n} |u^*(x)|^2 \,\mathrm{d}x \leqslant C\|u_0\|^2_{H^s(\mathbb{R}^n)} \quad \text{for all } u_0(x) \in H^s(\mathbb{R}^n). \tag{1-5}$$

Then $s \geqslant \frac{1}{4}$.

With this in mind, Theorem 1.1 is far from being sharp. Assuming that the initial data possesses additional angular regularity, we have:

Theorem 1.3. *Let B^n be the open unit ball in $\mathbb{R}^n$ and $\epsilon > 0$. There exists a constant C independent of u_0 such that*

$$\int_{B^n} |u^*(x)|^2 \,\mathrm{d}x \leqslant C\|u_0\|^2_{H_r^{\frac{1}{4}} H_\theta^{\frac{n-1}{2}+\epsilon}}, \tag{1-6}$$

where for $s, s' \geqslant 0$,

$$H_r^s H_\theta^{s'} = \Big\{ g : \|g\|_{H_r^s H_\theta^{s'}} := \big\| (1-\Delta_\theta)^{\frac{s'}{2}} ((1-\Delta)^{\frac{s}{2}} g) \big\|_{L^2_{r^{n-1}\mathrm{d}r}(\mathbb{R}^+; L^2_\theta(\mathbb{S}^{n-1}))} < \infty \Big\}.$$

Here Δ_θ denotes the Laplace–Beltrami operator on $\mathbb{S}^{n-1}$.

Remark 1.1. i) This result implies that the pointwise convergence of solutions to (1-1) holds for radial initial data $u_0 \in H^s(\mathbb{R}^n)$ with $s \geqslant \frac{1}{4}$.

ii) This result is an analogue of [Cho et al. 2006, Theorem 1.1]. We remark that the parameter ϵ there should be corrected to $\epsilon > \frac{1}{2}$ rather than $\epsilon > 0$. Thus, we generalize and improve the result of Cho et al. by making use of a finer result proved in [Gigante and Soria 2008].

Now we introduce some notation. We use $A \lesssim B$ to denote the statement that $A \leqslant CB$ for some large constant C which may vary from line to line and depend on various parameters; and similarly use $A \ll B$ to denote the statement $A \leqslant C^{-1}B$. We employ $A \sim B$ to denote the statement that $A \lesssim B \lesssim A$. If the constant C depends on a special parameter other than the above, we shall denote it explicitly by subscripts. We briefly write $A+\epsilon$ as $A+$ or $A-\epsilon$ as $A-$ for $0 < \epsilon \ll 1$. Throughout this paper, pairs of conjugate indices are written as p, p', where $1/p + 1/p' = 1$ with $1 \leqslant p \leqslant \infty$.

This paper is organized as follows. In Section 2, we mainly revisit the properties of the Bessel functions and the Hankel transform associated with $-\Delta + a/|x|^2$. Section 3 is devoted to the proofs of the theorems.

2. Preliminaries

We list some results about the Hankel transform and the Bessel functions and then show a characterization of the Sobolev norm in the Hankel transform version.

We begin by recalling the expansion formula with respect to spherical harmonics. For details, we refer to [Stein and Weiss 1971]. For the sake of convenience, let

$$\xi = \rho\omega \quad \text{and} \quad x = r\theta \quad \text{with } \omega, \theta \in \mathbb{S}^{n-1}. \tag{2-1}$$

For any $g \in L^2(\mathbb{R}^n)$, the expansion formula with respect to the spherical harmonics yields

$$g(x) = \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} a_{k,\ell}(r) Y_{k,\ell}(\theta),$$

where

$$\{Y_{k,1}, \dots, Y_{k,d(k)}\}$$

is the orthogonal basis of the space of spherical harmonics of degree k on $\mathbb{S}^{n-1}$, called $\mathcal{H}^k$, having dimension

$$d(k) = \frac{2k+n-2}{k} C_{n+k-3}^{k-1} \simeq \langle k \rangle^{n-2}.$$

We remark that for $n = 2$, the dimension of $\mathcal{H}^k$ is independent of k. Obviously, we have the orthogonal decomposition

$$L^2(\mathbb{S}^{n-1}) = \bigoplus_{k=0}^{\infty} \mathcal{H}^k.$$

By orthogonality, it gives

$$\|g(x)\|_{L^2_\theta(\mathbb{S}^{n-1})} = \|a_{k,\ell}(r)\|_{\ell^2_{k,\ell}}. \tag{2-2}$$

From $-\Delta_\theta Y_{k,\ell}(\theta) = k(k+n-2) Y_{k,\ell}(\theta)$, the fractional power of $1 - \Delta_\theta$ can be written explicitly [Machihara et al. 2005] as

$$(1-\Delta_\theta)^{\frac{s}{2}} g(x) = \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} (1 + k(k+n-2))^{\frac{s}{2}} a_{k,\ell}(r) Y_{k,\ell}(\theta). \tag{2-3}$$

We will need the Fourier transform of $a_{k,\ell}(r) Y_{k,\ell}(\theta)$. Theorem 3.10 of [Stein and Weiss 1971] asserts the Hankel transform formula

$$\hat{g}(\rho\omega) \sim \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} i^k Y_{k,\ell}(\omega) \rho^{-\frac{n-2}{2}} \int_0^\infty J_{k+\frac{n-2}{2}}(2\pi r\rho) a_{k,\ell}(r) r^{\frac{n}{2}} \, \mathrm{d}r. \tag{2-4}$$

Here the Bessel function $J_k(r)$ of order k is defined by the integral

$$J_k(r) = \frac{(\frac{r}{2})^k}{\Gamma(k+\frac{1}{2})\Gamma(\frac{1}{2})} \int_{-1}^{1} e^{isr}(1-s^2)^{\frac{2k-1}{2}}\,ds \quad \text{with } k > -\tfrac{1}{2} \text{ and } r > 0.$$

A simple computation gives the rough estimates

$$|J_k(r)| \leqslant \frac{Cr^k}{2^k\Gamma(k+\frac{1}{2})\Gamma(\frac{1}{2})}\left(1+\frac{1}{k+\frac{1}{2}}\right), \tag{2-5}$$

where C is an absolute constant. This estimate will be mainly used when $r \lesssim 1$. Another well-known asymptotic expansion for the Bessel function is

$$J_k(r) = r^{-\frac{1}{2}}\sqrt{\frac{2}{\pi}}\cos\left(r - \frac{k\pi}{2} - \frac{\pi}{4}\right) + O_k(r^{-\frac{3}{2}}) \quad \text{as } r \to \infty, \tag{2-6}$$

but with a constant depending on k (see [Stein and Weiss 1971]). As pointed out in [Stein 1993], if one seeks a uniform bound for large r and k, the best one can do is $|J_k(r)| \leqslant Cr^{-\frac{1}{3}}$. One will find that this decay doesn't lead to the desired result.

We now recall the properties of Bessel function $J_k(r)$ in [Stein 1993; Stempak 2000].

Lemma 2.1 (asymptotics of the Bessel function). *Assume that $k \in \mathbb{N}$ and $k \gg 1$. Let $J_k(r)$ be the Bessel function of order k defined as above. There exist a large constant C and small constant c independent of k and r satisfying these conditions*:

- *When $r \leqslant \frac{k}{2}$,*

$$|J_k(r)| \leqslant Ce^{-c(k+r)}. \tag{2-7}$$

- *When $\frac{k}{2} \leqslant r \leqslant 2k$,*

$$|J_k(r)| \leqslant Ck^{-\frac{1}{3}}(k^{-\frac{1}{3}}|r-k|+1)^{-\frac{1}{4}}. \tag{2-8}$$

- *When $r \geqslant 2k$,*

$$J_k(r) = r^{-\frac{1}{2}}\sum_{\pm} a_{\pm}(r,k)e^{\pm ir} + E(r,k), \tag{2-9}$$

where $|a_{\pm}(r,k)| \leqslant C$ and $|E(r,k)| \leqslant Cr^{-1}$.

As a consequence of Lemma 2.1, we have:

Lemma 2.2. *Let $R \gg 1$. There exists a constant C independent of k, R such that*

$$\int_R^{2R} |J_k(r)|^2\,\mathrm{d}r \leqslant C. \tag{2-10}$$

Proof. To prove (2-10), we write

$$\int_R^{2R} |J_k(r)|^2 \,\mathrm{d}r = \int_{I_1} |J_k(r)|^2 \,\mathrm{d}r + \int_{I_2} |J_k(r)|^2 \,\mathrm{d}r + \int_{I_3} |J_k(r)|^2 \,\mathrm{d}r$$

where $I_1 = [R, 2R] \cap [0, \frac{k}{2}]$, $I_2 = [R, 2R] \cap \big[\frac{k}{2}, 2k\big]$ and $I_3 = [R, 2R] \cap [2k, \infty]$. By (2-7) and (2-9), we have

$$\int_{I_1} |J_k(r)|^2 \,\mathrm{d}r \leqslant C \int_{I_1} e^{-cr} \,\mathrm{d}r \leqslant Ce^{-cR}, \tag{2-11}$$

and

$$\int_{I_3} |J_k(r)|^2 \,\mathrm{d}r \leqslant C. \tag{2-12}$$

On the other hand, one has by (2-8)

$$\int_{[\frac{k}{2},2k]} |J_k(r)|^2 \,\mathrm{d}r \leqslant C \int_{[\frac{k}{2},2k]} k^{-\frac{2}{3}} (1 + k^{-\frac{1}{3}} |r-k|)^{-\frac{1}{2}} \,\mathrm{d}r \leqslant C.$$

Observing that $[R, 2R] \cap [\frac{k}{2}, 2k] = \varnothing$ unless $R \sim k$, we obtain

$$\int_{I_2} |J_k(r)|^2 \,\mathrm{d}r \leqslant C. \tag{2-13}$$

This together with (2-11) and (2-12) yields (2-10). □

For simplicity, we define

$$\mu(k) = \frac{n-2}{2} + k \quad \text{and} \quad \nu(k) = \sqrt{\mu^2(k) + a} \quad \text{with } a > -(n-2)^2/4. \tag{2-14}$$

We sometime write ν instead of $\nu(k)$. Let f be a Schwartz function defined on $\mathbb{R}^n$. We define the Hankel transform of order ν by

$$(\mathcal{H}_\nu f)(\xi) = \int_0^\infty (r\rho)^{-\frac{n-2}{2}} J_\nu(r\rho) f(r\omega) r^{n-1} \,\mathrm{d}r, \tag{2-15}$$

where $\rho = |\xi|$, $\omega = \xi/|\xi|$ and J_ν is the Bessel function of order ν. In particular, if the function f is radial, then we have

$$(\mathcal{H}_\nu f)(\rho) = \int_0^\infty (r\rho)^{-\frac{n-2}{2}} J_\nu(r\rho) f(r) r^{n-1} \,\mathrm{d}r. \tag{2-16}$$

If $f(x) = \sum\limits_{k=0}^{\infty} \sum\limits_{\ell=1}^{d(k)} a_{k,\ell}(r) Y_{k,\ell}(\theta)$, it follows from (2-4) that

$$\hat{f}(\xi) = \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} 2\pi i^k Y_{k,\ell}(\omega) \big(\mathcal{H}_{\mu(k)} a_{k,\ell}\big)(\rho). \tag{2-17}$$

The following properties of the Hankel transform are obtained in [Burq et al. 2003; Planchon et al. 2003b].

Lemma 2.3. *Let $\mathcal{H}_\nu$ be as above and set*

$$A_{\nu(k)} := -\partial_r^2 - \frac{n-1}{r}\partial_r + \Big[\nu^2(k) - \Big(\frac{n-2}{2}\Big)^2\Big]r^{-2}.$$

(i) $\mathcal{H}_\nu = \mathcal{H}_\nu^{-1}$.

(ii) $\mathcal{H}_\nu$ *is self-adjoint, i.e.*, $\mathcal{H}_\nu = \mathcal{H}_\nu^*$.

(iii) $\mathcal{H}_\nu$ *is an* L^2 *isometry, i.e.*, $\|\mathcal{H}_\nu\phi\|_{L^2_\xi} = \|\phi\|_{L^2_x}$.

(iv) $\mathcal{H}_\nu(A_\nu\phi)(\xi) = |\xi|^2(\mathcal{H}_\nu\phi)(\xi)$, *for* $\phi \in L^2$.

We next recall an almost orthogonality inequality. Denote by P_j and $\tilde{P}_j$ the usual dyadic frequency localization at $|\xi| \sim 2^j$ and the localization with respect to $(-\Delta + a/|x|^2)^{\frac{1}{2}}$. We define the projectors $M_{jj'} = P_j\tilde{P}_{j'}$ and $N_{jj'} = \tilde{P}_j P_{j'}$. More precisely, let f be in the k-th harmonic subspace; then

$$P_j f = \mathcal{H}_{\mu(k)}\beta_j\mathcal{H}_{\mu(k)} f \quad \text{and} \quad \tilde{P}_j f = \mathcal{H}_{\nu(k)}\beta_j\mathcal{H}_{\nu(k)} f,$$

where $\beta_j(\xi) = \beta(2^{-j}|\xi|)$ with $\beta \in C_0^\infty(\mathbb{R}^+)$ supported in $\big[\frac{1}{2}, 2\big]$.

Lemma 2.4 (almost orthogonality inequality [Burq et al. 2003]). *Let $f \in L^2(\mathbb{R}^n)$. There exists a constant C independent of j, j' such that*

$$\|M_{jj'}f\|_{L^2(\mathbb{R}^n)}, \|N_{jj'}f\|_{L^2(\mathbb{R}^n)} \leqslant C2^{-\epsilon|j-j'|}\|f\|_{L^2(\mathbb{R}^n)}, \tag{2-18}$$

where $\epsilon < 1 + \min\big\{\frac{n-2}{2}, \big(\frac{(n-2)^2}{4} + a\big)^{\frac{1}{2}}\big\}$.

As a consequence, we have:

Lemma 2.5. *Let $f \in L^2(\mathbb{R}^n)$ be given by*

$$f(x) = \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} a_{k,\ell}(r)Y_{k,\ell}(\theta).$$

Then for $0 \leqslant s < 1 + \min\big\{\frac{n-2}{2}, \big(\frac{(n-2)^2}{4} + a\big)^{\frac{1}{2}}\big\}$ *and* $s' \geqslant 0$,

$$\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{M\in 2^{\mathbb{Z}}} M^{2s}(1+k)^{2s'}\Big\|b_{k,\ell}(\rho)\chi\Big(\frac{\rho}{M}\Big)\rho^{\frac{n-1}{2}}\Big\|_{L^2_\rho}^2 \sim \|f\|^2_{\dot{H}^s_r H^{s'}_\theta}, \tag{2-19}$$

where $b_{k,\ell}(\rho) = (\mathcal{H}_{\nu(k)}a_{k,\ell})(\rho)$ *and* $\chi \in C_0^\infty(\mathbb{R}^n)$ *such that* $\operatorname{supp}\chi \subset \big[\frac{1}{2}, 1\big]$.

Proof. Note that $-\Delta_\theta Y_{k,\ell} = k(k+n-2)Y_{k,\ell}$. By Lemma 2.3, we have

$$\|f\|^2_{\dot H^0_r H^{s'}_\theta} \sim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}(1+k)^{2s'}\|a_{k,\ell}(r)\|^2_{L^2_{r^{n-1}\,\mathrm{d}r}(\mathbb{R}+)}\|Y_{k,\ell}(\theta)\|^2_{L^2_\theta}$$
$$\sim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}(1+k)^{2s'}\|b_{k,\ell}(\rho)\|^2_{L^2_{\rho^{n-1}\,\mathrm{d}\rho}(\mathbb{R}+)}.$$

By (2-3), it suffices to show (2-19) with $s'=0$. By Lemma 2.3, we have

$$\begin{aligned}\big\|b_{k,\ell}(\rho)\chi\big(\tfrac{\rho}{M}\big)\rho^{\frac{n-1}{2}}\big\|_{L^2_\rho} &= \big\|\chi\big(\tfrac{\rho}{M}\big)\mathcal{H}_\nu[Y_{k,l}(\theta)a_{k,\ell}(r)](\xi)\big\|_{L^2_\xi}\\ &= \big\|\mathcal{H}_\nu\big[\chi\big(\tfrac{\rho}{M}\big)\mathcal{H}_\nu(Y_{k,l}(\theta)a_{k,\ell}(r))(\xi)\big]\big\|_{L^2_x}.\end{aligned}$$

This yields, by letting $j=\log_2 M$,

$$\begin{aligned}\big\|b_{k,\ell}(\rho)\chi\big(\tfrac{\rho}{M}\big)\rho^{\frac{n-1}{2}}\big\|_{L^2_\rho} &= \big\|\big[\mathcal{H}_\nu\chi\big(\tfrac{\rho}{M}\big)\mathcal{H}_\nu\big](Y_{k,l}(\theta)a_{k,\ell}(r))\big\|_{L^2_x}\\ &= \big\|\tilde P_j(Y_{k,l}(\theta)a_{k,\ell}(r))\big\|_{L^2_x}.\end{aligned}$$

Let $g_{k,\ell}(x)=Y_{k,l}(\theta)a_{k,\ell}(r)$ and $\overline{P_{j'}}=P_{j'-1}+P_{j'}+P_{j'+1}$. We have by the triangle inequality and Lemma 2.4

$$\begin{aligned}\text{LHS of (2-19)} &= \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{j\in\mathbb{Z}}2^{2sj}\|\tilde P_j g_{k,\ell}\|^2_{L^2_x}\\ &\lesssim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{j\in\mathbb{Z}}2^{2sj}\Big(\sum_{j'}\|\tilde P_j\overline{P_{j'}}P_{j'}g_{k,\ell}\|_{L^2_x}\Big)^2\\ &\lesssim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{j\in\mathbb{Z}}2^{2sj}\Big(\sum_{j'}2^{-\epsilon|j-j'|}\|P_{j'}g_{k,\ell}\|_{L^2_x}\Big)^2,\end{aligned}$$

where $s<\epsilon<1+\min\big\{\frac{n-2}{2},\big(\frac{(n-2)^2}{4}+a\big)^{\frac12}\big\}$. Let $0<\epsilon_1\ll 1$ be such that $\epsilon_2:=\epsilon-\epsilon_1>s$; then the LHS of (2-19) is bounded above by

$$\begin{aligned}&C\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{j\in\mathbb{Z}}2^{2js}\sum_{j'}2^{-2\epsilon_2|j-j'|}\|P_{j'}g_{k,\ell}\|^2_{L^2(\mathbb{R}^n)}\sum_{j'}2^{-2\epsilon_1|j-j'|}\\ &\qquad\leqslant C\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{j'}2^{2j's}\sum_{j\in\mathbb{Z}}2^{2js}2^{-2\epsilon_2|j|}\|P_{j'}g_{k,\ell}\|^2_{L^2(\mathbb{R}^n)}\\ &\qquad\leqslant C\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{j'}2^{2j's}\|P_{j'}g_{k,\ell}\|^2_{L^2(\mathbb{R}^n)}.\end{aligned}$$

By the definition of $P_{j'}$, Lemma 2.3 and (2-17), we have

$$\begin{aligned}
\text{LHS of (2-19)} &\leqslant C\sum_{j'} 2^{2j's} \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} \Big\| \chi\Big(\frac{\rho}{2^{j'}}\Big)[\mathcal{H}_{\mu(k)} a_{k,\ell}](\rho)\rho^{\frac{n-1}{2}} \Big\|^2_{L^2(\mathbb{R}^+)} \\
&= C\sum_{j'} 2^{2j's} \Big\| \chi\Big(\frac{\rho}{2^{j'}}\Big) \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} 2\pi i^k [\mathcal{H}_{\mu(k)} a_{k,\ell}](\rho) Y_{k,\ell}(\omega) \Big\|^2_{L^2(\mathbb{R}^n)} \\
&= C\sum_{j'} 2^{2j's} \Big\| \chi\Big(\frac{\rho}{2^{j'}}\Big) \hat{f} \Big\|^2_{L^2(\mathbb{R}^n)} \sim \|f\|^2_{\dot{H}^s}.
\end{aligned}$$

We can use a similar argument to prove

$$\text{LHS of (2-19)} \geqslant c\|f\|^2_{\dot{H}^s}.$$

This concludes the proof of Lemma 2.4. □

3. Proof of the main theorems

In this section, we first use the spherical harmonic expansion to write the solution as a linear combination of products of the Hankel transform of radial functions and spherical harmonics. We prove the main theorems by analyzing a property of the Hankel transform. The key ingredients are to use the stationary phase argument and to exploit the asymptotic behavior of the Bessel function.

The expression of the solution. Consider the following Cauchy problem:

$$\begin{cases} i\partial_t u - \Delta u + \dfrac{a}{|x|^2} u = 0, \\ u(x,0) = u_0(x). \end{cases} \tag{3-1}$$

We use the spherical harmonic expansion to write

$$u_0(x) = \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} a^0_{k,\ell}(r) Y_{k,\ell}(\theta). \tag{3-2}$$

Let us consider (3-1) in polar coordinates. Write $v(t,r,\theta) = u(t,r\theta)$ and $g(r,\theta) = u_0(r\theta)$. Then $v(t,r,\theta)$ satisfies

$$\begin{cases} i\partial_t v - \partial_{rr} v - \dfrac{n-1}{r}\partial_r v - \dfrac{1}{r^2}\Delta_\theta v + \dfrac{a}{r^2} v = 0, \\ v(0,r,\theta) = g(r,\theta). \end{cases} \tag{3-3}$$

By (3-2), we have

$$g(r,\theta) = \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} a^0_{k,\ell}(r) Y_{k,\ell}(\theta).$$

Using separation of variables, we can write v as a linear combination of products of radial functions and spherical harmonics:

$$v(t,r,\theta)=\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} v_{k,\ell}(t,r)Y_{k,\ell}(\theta), \tag{3-4}$$

where $v_{k,\ell}$ is given by

$$\begin{cases} i\partial_t v_{k,\ell}-\partial_{rr}v_{k,\ell}-\frac{n-1}{r}\partial_r v_{k,\ell}+\frac{k(k+n-2)+a}{r^2}v_{k,\ell}=0,\\ v_{k,\ell}(0,r)=a^0_{k,\ell}(r)\end{cases}$$

for each $k,\ell\in\mathbb{N}$, $1\leqslant \ell\leqslant d(k)$. Then we can rewrite the above by the definition of $A_{\nu(k)}$ as

$$\begin{cases} i\partial_t v_{k,\ell}+A_{\nu(k)}v_{k,\ell}=0,\\ v_{k,\ell}(0,r)=a^0_{k,\ell}(r).\end{cases} \tag{3-5}$$

Applying the Hankel transform to (3-5), by Lemma 2.3(iv), we have

$$\begin{cases} i\partial_t \tilde{v}_{k,\ell}+\rho^2\tilde{v}_{k,\ell}=0,\\ \tilde{v}_{k,\ell}(0,\rho)=b^0_{k,\ell}(\rho),\end{cases} \tag{3-6}$$

where

$$\tilde{v}_{k,\ell}(t,\rho)=(\mathcal{H}_\nu v_{k,\ell})(t,\rho), \quad b^0_{k,\ell}(\rho)=(\mathcal{H}_\nu a^0_{k,\ell})(\rho). \tag{3-7}$$

Solving this ODE and inverting the Hankel transform, we obtain

$$\begin{aligned} v_{k,\ell}(t,r)&=\int_0^\infty (r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)\tilde{v}_{k,\ell}(t,\rho)\rho^{n-1}\,\mathrm{d}\rho\\ &=\int_0^\infty (r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)e^{it\rho^2}b^0_{k,\ell}(\rho)\rho^{n-1}\,\mathrm{d}\rho.\end{aligned}$$

Therefore we get

$$\begin{aligned} u(x,t)&=v(t,r,\theta)\\ &=\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} Y_{k,\ell}(\theta)\int_0^\infty (r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)e^{it\rho^2}b^0_{k,\ell}(\rho)\rho^{n-1}\,\mathrm{d}\rho\\ &=\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} Y_{k,\ell}(\theta)\mathcal{H}_{\nu(k)}\big[e^{it\rho^2}b^0_{k,\ell}(\rho)\big](r).\end{aligned} \tag{3-8}$$

Proof of Theorem 1.1. By the Sobolev embedding $\dot{H}^{\frac{1}{2}-}(\mathbb{R})\cap\dot{H}^{\frac{1}{2}+}(\mathbb{R})\hookrightarrow L^\infty(\mathbb{R})$, it suffices to show:

Proposition 3.1. *Let $\alpha \geqslant \frac{1}{2}-\frac{\beta}{4}$ and $\beta = 1+$ be such that*

$$2\alpha-1+\frac{\beta}{2} < 1+\min\Big\{\frac{n-2}{2}, \Big(\frac{(n-2)^2}{4}+a\Big)^{\frac{1}{2}}\Big\}.$$

There exists a constant C independent of u_0 such that

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}} |\partial_t^\alpha u(x,t)|^2 \frac{\mathrm{d}t\,\mathrm{d}x}{(1+|x|)^\beta} \leqslant C\|u_0\|^2_{\dot{H}^{2\alpha-1+\frac{\beta}{2}}(\mathbb{R}^n)}. \tag{3-9}$$

Proof. By the Plancherel theorem with respect to time t, we obtain

$$\text{LHS of (3-9)} = \int_{\mathbb{R}^n}\int_{\mathbb{R}} \Big|\tau^\alpha \int_{\mathbb{R}} e^{-it\tau}u(x,t)\,\mathrm{d}t\Big|^2 \frac{\mathrm{d}\tau\,\mathrm{d}x}{(1+|x|)^\beta}.$$

Using (3-8), this is bounded above by

$$\begin{aligned}
&\int_{\mathbb{R}^{n+1}}\Big|\tau^\alpha \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} Y_{k,\ell}(\theta)\int_{\mathbb{R}}\int_0^\infty (r\rho)^{-\frac{n-2}{2}} J_{\nu(k)}(r\rho)e^{it(\rho^2-\tau)}b^0_{k,\ell}(\rho)\rho^{n-1}\,\mathrm{d}\rho\,\mathrm{d}t\Big|^2 \\
&\hspace{28em}\times\frac{\mathrm{d}\tau\,\mathrm{d}x}{(1+|x|)^\beta}\\
&\lesssim \int_{\mathbb{R}^{n+1}}\Big|\tau^\alpha \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} Y_{k,\ell}(\theta)\int_0^\infty (r\rho)^{-\frac{n-2}{2}} J_{\nu(k)}(r\rho)b^0_{k,\ell}(\rho)\rho^{n-1}\delta(\tau-\rho^2)\,\mathrm{d}\rho\Big|^2 \\
&\hspace{28em}\times\frac{\mathrm{d}\tau\,\mathrm{d}x}{(1+|x|)^\beta}\\
&\lesssim \int_{\mathbb{R}^n}\int_0^\infty \Big|\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)} Y_{k,\ell}(\theta)\rho^\alpha (r\sqrt{\rho})^{-\frac{n-2}{2}} J_{\nu(k)}(r\sqrt{\rho})b^0_{k,\ell}(\sqrt{\rho})\rho^{\frac{n-1}{2}}\rho^{-\frac{1}{2}}\Big|^2 \\
&\hspace{28em}\times\frac{\mathrm{d}\rho\,\mathrm{d}x}{(1+|x|)^\beta}.
\end{aligned}$$

By orthogonality, therefore, the LHS of (3-9) is

$$\lesssim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\int_0^\infty\int_0^\infty \big|\rho^{2\alpha+\frac{1}{2}}(r\rho)^{-\frac{n-2}{2}} J_{\nu(k)}(r\rho)b^0_{k,\ell}(\rho)\rho^{n-2}\big|^2 \frac{\mathrm{d}\rho\, r^{n-1}\,\mathrm{d}r}{(1+r)^\beta}.$$

Let χ be a smoothing function equals 1 in $\big[1,\frac{3}{2}\big]$ and vanishes outside $\big[\frac{1}{2},2\big]$. For our purpose, we make a dyadic decomposition to obtain

LHS of (3-9)

$$\begin{aligned}
&\lesssim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{M\in 2^{\mathbb{Z}}}\int_0^\infty\int_0^\infty \Big|\rho^{2\alpha+\frac12}(r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)b^0_{k,\ell}(\rho)\rho^{n-2}\chi\Big(\frac{\rho}{M}\Big)\Big|^2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times\frac{r^{n-1}\,\mathrm{d}r\,\mathrm{d}\rho}{(1+r)^\beta} \\
&\lesssim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{M\in 2^{\mathbb{Z}}} M^{2(n-2+2\alpha+\frac12)+1-n} \\
&\qquad\qquad \times\int_0^\infty\int_0^\infty\Big|(r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)b^0_{k,\ell}(M\rho)\chi(\rho)\Big|^2\frac{r^{n-1}\,\mathrm{d}r\,\mathrm{d}\rho}{(1+\frac{r}{M})^\beta} \\
&\lesssim \sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}\sum_{M\in 2^{\mathbb{Z}}}\sum_{R\in 2^{\mathbb{Z}}} M^{n-2+4\alpha}R^{n-1} \\
&\qquad\qquad \times\int_R^{2R}\int_0^\infty\Big|(r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)b^0_{k,\ell}(M\rho)\chi(\rho)\Big|^2\frac{\mathrm{d}r\,\mathrm{d}\rho}{(1+\frac{r}{M})^\beta}.
\end{aligned}$$

Define

$$G_{k,\ell}(R,M)=\int_R^{2R}\int_0^\infty\big|(r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)b^0_{k,\ell}(M\rho)\chi(\rho)\big|^2\frac{\mathrm{d}r\,\mathrm{d}\rho}{(1+\frac{r}{M})^\beta}.$$

Proposition 3.2. (1) *If $R\lesssim 1$, then*

$$G_{k,\ell}(R,M)\lesssim R^{2\nu(k)-n+3}M^{-n}\min\Big\{1,\Big(\frac{M}{R}\Big)^\beta\Big\}\Big\|b^0_{k,\ell}(\rho)\chi\Big(\frac{\rho}{M}\Big)\rho^{\frac{n-1}{2}}\Big\|^2_{L^2}.$$

(2) *If $R\gg 1$, then*

$$G_{k,\ell}(R,M)\lesssim \min\Big\{1,\Big(\frac{M}{R}\Big)^\beta\Big\}R^{-(n-2)}M^{-n}\Big\|b^0_{k,\ell}(\rho)\chi\Big(\frac{\rho}{M}\Big)\rho^{\frac{n-1}{2}}\Big\|^2_{L^2}.$$

Proof. (1) Since $\rho\sim 1$, we have $r\rho\lesssim 1$. By the property (2-5) of the Bessel function, we obtain

$$\begin{aligned}
G_{k,\ell}(R,M)&\lesssim\int_R^{2R}\int_0^\infty\Big|\frac{(r\rho)^{\nu(k)}(r\rho)^{-\frac{n-2}{2}}}{2^{\nu(k)}\Gamma(\nu(k)+\frac12)\Gamma(\frac12)}b^0_{k,\ell}(M\rho)\chi(\rho)\Big|^2\mathrm{d}\rho\frac{\mathrm{d}r}{(1+\frac{r}{M})^\beta}\\
&\lesssim R^{2\nu(k)-n+3}M^{-n}\min\Big\{1,\Big(\frac{M}{R}\Big)^\beta\Big\}\|b^0_{k,\ell}(\rho)\chi(\tfrac{\rho}{M})\rho^{\frac{n-1}{2}}\|^2_{L^2}.
\end{aligned}$$

(2) Since $\rho\sim 1$, we have $r\rho\gg 1$. We estimate

$$\begin{aligned}
&G_{k,\ell}(R,M)\\
&\qquad\lesssim R^{-(n-2)}\int_0^\infty\big|b^0_{k,\ell}(M\rho)\chi(\rho)\big|^2\int_R^{2R}\big|J_{\nu(k)}(r\rho)\big|^2\frac{\mathrm{d}r}{(1+\frac{r}{M})^\beta}\,\mathrm{d}\rho.
\end{aligned}\tag{3-10}$$

Subcase (i): $R\lesssim M$. Noting that $\rho\sim 1$, we obtain by Lemma 2.2

$$\int_R^{2R} |J_{\nu(k)}(r\rho)|^2 \frac{\mathrm{d}r}{(1+\frac{r}{M})^\beta} \lesssim \int_R^{2R} |J_{\nu(k)}(r\rho)|^2 \,\mathrm{d}r \lesssim 1. \tag{3-11}$$

Subcase (ii): $R \gg M$. Noticing that $\rho \sim 1$ again, we obtain by Lemma 2.2

$$\int_R^{2R} |J_{\nu(k)}(r\rho)|^2 \frac{\mathrm{d}r}{(1+\frac{r}{M})^\beta} \lesssim \Big(\frac{M}{R}\Big)^\beta \int_R^{2R} |J_{\nu(k)}(r\rho)|^2 \,\mathrm{d}r \lesssim \Big(\frac{M}{R}\Big)^\beta. \tag{3-12}$$

Putting (3-11) and (3-12) into (3-10), we have

$$\begin{aligned} G_{k,\ell}(R,M) &\lesssim \min\Big\{1, \Big(\frac{M}{R}\Big)^\beta\Big\} R^{-(n-2)} \int_0^\infty \Big|b^0_{k,\ell}(M\rho)\chi(\rho)\Big|^2 \mathrm{d}\rho \\ &\lesssim \min\Big\{1, \Big(\frac{M}{R}\Big)^\beta\Big\} R^{-(n-2)} M^{-n} \Big\| b^0_{k,\ell}(\rho)\chi\Big(\frac{\rho}{M}\Big)\rho^{\frac{n-1}{2}}\Big\|^2_{L^2}. \end{aligned}$$

Thus we have proved Proposition 3.2. $\square$

Now we return to proving Proposition 3.1. By Proposition 3.2, we show

$$\begin{aligned} &\int_{\mathbb{R}^n}\int_{\mathbb{R}} |\partial_t^\alpha u(x,t)|^2 \frac{\mathrm{d}t\,\mathrm{d}x}{(1+|x|)^\beta} \\ &\lesssim \sum_{k=0}^\infty \sum_{\ell=1}^{d(k)} \sum_{M\in 2^{\mathbb{Z}}} \sum_{\{R\in 2^{\mathbb{Z}}:R\lesssim 1\}} \Big(M^{4\alpha-2} R^{2(\nu(k)+1)} \min\Big\{1, \Big(\frac{M}{R}\Big)^\beta\Big\} \\ &\qquad\qquad \times \| b^0_{k,\ell}(\rho)\chi(\frac{\rho}{M})\rho^{\frac{n-1}{2}}\|^2_{L^2}\Big) \\ &\quad + \sum_{k=0}^\infty \sum_{\ell=1}^{d(k)} \sum_{M\in 2^{\mathbb{Z}}} \sum_{\{R\in 2^{\mathbb{Z}}:R\gg 1\}} M^{4\alpha-2+\beta} R^{1-\beta} \| b^0_{k,\ell}(\rho)\chi(\frac{\rho}{M})\rho^{\frac{n-1}{2}}\|^2_{L^2}. \end{aligned}$$

From $\beta = 1+$, one has

$$\begin{aligned} &\sum_{M\in 2^{\mathbb{Z}}} \sum_{\{R\in 2^{\mathbb{Z}}:R\lesssim 1\}} M^{4\alpha-2} R^{2(\nu(k)+1)} \min\Big\{1, \Big(\frac{M}{R}\Big)^\beta\Big\} \| b^0_{k,\ell}(\rho)\chi(\frac{\rho}{M})\rho^{\frac{n-1}{2}}\|^2_{L^2} \\ &\qquad\qquad \lesssim \sum_{M\in 2^{\mathbb{Z}}} M^{4\alpha-2+\beta} \| b^0_{k,\ell}(\rho)\chi(\frac{\rho}{M})\rho^{\frac{n-1}{2}}\|^2_{L^2}. \end{aligned}$$

Since $\alpha \geqslant \frac{1}{2} - \frac{\beta}{4}$, we have by Lemma 2.5

$$\begin{aligned} \int_{\mathbb{R}^n}\int_{\mathbb{R}} |\partial_t^\alpha u(x,t)|^2 \frac{\mathrm{d}t\,\mathrm{d}x}{(1+|x|)^\beta} &\lesssim \sum_{k=0}^\infty \sum_{\ell=1}^{d(k)} \sum_{M\in 2^{\mathbb{Z}}} M^{4\alpha-2+\beta} \Big\| b^0_{k,\ell}(\rho)\chi\Big(\frac{\rho}{M}\Big)\rho^{\frac{n-1}{2}}\Big\|^2_{L^2} \\ &\leqslant C\|u_0\|^2_{\dot H^{2\alpha-1+\frac{\beta}{2}}(\mathbb{R}^n)}. \end{aligned}$$

$\square$

Finally, we apply Proposition 3.1 with $\alpha = \frac{1}{2}+$ and $\alpha = \frac{1}{2}-$ to prove Theorem 1.1.

Proof of Theorem 1.2. We now construct an example to show Theorem 1.2. The main idea is the stationary phase argument. By (3-8), we recall

$$u(x,t)=\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}Y_{k,\ell}(\theta)\int_0^\infty (r\rho)^{-\frac{n-2}{2}}J_{\nu(k)}(r\rho)e^{it\rho^2}b^0_{k,\ell}(\rho)\rho^{n-1}\,\mathrm{d}\rho, \tag{3-13}$$

where

$$b^0_{k,\ell}(\rho)=(\mathcal{H}_\nu a^0_{k,\ell})(\rho),\quad u_0(x)=u_0(r\theta)=\sum_{k=0}^{\infty}\sum_{\ell=1}^{d(k)}a^0_{k,\ell}(r)Y_{k,\ell}(\theta).$$

In particular we choose $u_0(x)$ to be a radial function such that $(\mathcal{H}_{\nu(0)}u_0)(\xi)=\chi_N(|\xi|)$, where χ_N is a smooth positive function supported in J_N (to be chosen later) and $N\gg 1$. Then

$$u(x,t)=\int_0^\infty (r\rho)^{-\frac{n-2}{2}}J_{\nu(0)}(r\rho)e^{it\rho^2}\chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho. \tag{3-14}$$

Recalling the asymptotic expansion of the Bessel function,

$$J_\nu(r)=r^{-\frac12}\sqrt{\frac{2}{\pi}}\cos\big(r-\frac{\nu\pi}{2}-\frac{\pi}{4}\big)+O_\nu(r^{-\frac32})\quad\text{as } r\to\infty,$$

with a constant depending on ν (see [Stein and Weiss 1971]), we can write

$$\begin{aligned}u(x,t)=C_\nu\int_0^\infty (r\rho)^{-\frac{n-1}{2}}\big(e^{i(r\rho-\frac{\nu\pi}{2}-\frac{\pi}{4})}-e^{-i(r\rho-\frac{\nu\pi}{2}-\frac{\pi}{4})}\big)e^{it\rho^2}\chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho\\ +C_\nu\int_0^\infty (r\rho)^{-\frac{n-2}{2}}O_\nu\big((r\rho)^{-\frac32}\big)e^{it\rho^2}\chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho.\end{aligned}$$

Let us define

$$I_1(r)=C_\nu e^{i(\frac{\nu\pi}{2}+\frac{\pi}{4})}\int_0^\infty (r\rho)^{-\frac{n-1}{2}}e^{i(-r\rho+t\rho^2)}\chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho, \tag{3-15}$$

$$I_2(r)=C_\nu e^{-i(\frac{\nu\pi}{2}+\frac{\pi}{4})}\int_0^\infty (r\rho)^{-\frac{n-1}{2}}e^{i(r\rho+t\rho^2)}\chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho, \tag{3-16}$$

$$I_3(r)=C_\nu\int_0^\infty (r\rho)^{-\frac{n-2}{2}}O_\nu\big((r\rho)^{-\frac32}\big)e^{it\rho^2}\chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho. \tag{3-17}$$

Let $\phi_r(\rho)=t\rho^2-r\rho$. The fundamental idea is to choose sets J_N and $E\subset B^n$, in which $t(r)$ can be chosen, so that $\partial_\rho\phi_r(\rho)=2t(r)\rho-r$ almost vanishes for all $\rho\in J_N$ and $r\in\{|x|:x\in E\}$. To this end, we choose

$$E=\{x:\tfrac{1}{100}\leqslant|x|\leqslant\tfrac18\}\quad\text{and}\quad J_N=[N,N+2N^{\frac12}].$$

Choose $t(r) = r/(2(N+\sqrt{N}))$; then $\partial_\rho \phi_r(N + N^{\frac{1}{2}}) = 0$. Then

$$I_1(r) = C_\nu e^{i(\frac{\nu\pi}{2}+\frac{\pi}{4})} e^{i\phi_r(N+\sqrt{N})} \int_0^\infty (r\rho)^{-\frac{n-1}{2}} \exp \frac{ir[\rho-(N+\sqrt{N})]^2}{2(N+\sqrt{N})} \chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho.$$

Observe that

$$|I_1(r)| \geqslant c_\nu \int_0^\infty (r\rho)^{-\frac{n-1}{2}} \cos \frac{r[\rho-(N+\sqrt{N})]^2}{2(N+\sqrt{N})} \chi_N(\rho)\rho^{n-1}\,\mathrm{d}\rho. \tag{3-18}$$

Moreover, there exists a small constant $c > 0$ such that

$$\cos \frac{r[\rho-(N+\sqrt{N})]^2}{2(N+\sqrt{N})} \geqslant c,$$

since $\dfrac{r[\rho-(N+\sqrt{N})]^2}{2(N+\sqrt{N})} \leqslant \dfrac{\pi}{4}$ for all $r \in \left[\frac{1}{100}, \frac{1}{8}\right]$, $N \gg 1$ and $\rho \in J_N$. Therefore,

$$|I_1(r)| \geqslant c_\nu r^{-\frac{n-1}{2}} \int_0^\infty \chi_N(\rho)\rho^{\frac{n-1}{2}}\,\mathrm{d}\rho \geqslant c_\nu r^{-\frac{n-1}{2}} N^{\frac{n}{2}}. \tag{3-19}$$

On the other hand, let $\varphi_r(\rho) = t\rho^2 + r\rho$, $t = t(r)$ as before; then $\partial_\rho\varphi_r(\rho) = 2t(r)\rho + r \geqslant \frac{1}{200}$ when $\rho \in J_N$ and $r \in \left[\frac{1}{100}, \frac{1}{8}\right]$. Integrating by parts, we obtain

$$|I_2(r)| \leqslant C_\nu r^{-\frac{n}{2}} N^{\frac{n-2}{2}}. \tag{3-20}$$

Obviously, we have

$$|I_3(r)| \leqslant C_\nu r^{-\frac{n}{2}} N^{\frac{n-2}{2}}. \tag{3-21}$$

Combining (3-19)–(3-21), we get for $N \gg 1$ and $r \in \left[\frac{1}{100}, \frac{1}{8}\right]$

$$u^*(x) \geqslant cN^{\frac{n}{2}}. \tag{3-22}$$

On the other hand, let $j_0 = \log_2 N$; we obtain by the definitions of P_j and $\tilde{P}_j$

$$\|u_0(x)\|_{H^s}^2 = \sum_j 2^{2js}\|P_j u_0\|_{L^2}^2 = \sum_j 2^{2js}\|P_j \tilde{P}_{j_0} u_0\|_{L^2}^2.$$

By Lemma 2.4, we choose $s < \epsilon < 1 + \min\left\{\frac{n-2}{2}, \left(\frac{(n-2)^2}{4} + a\right)^{\frac{1}{2}}\right\}$ to obtain

$$\begin{aligned} \|u_0(x)\|_{H^s}^2 &\leqslant C \sum_j 2^{2js-2\epsilon|j-j_0|}\|u_0\|_{L^2}^2 \\ &= CN^{2s}\sum_j 2^{2js-2\epsilon|j|}\|\chi_N\|_{L^2}^2 = N^{2s+n-\frac{1}{2}}. \end{aligned} \tag{3-23}$$

Thus, by (1-5) and (3-22), we must have $s \geqslant \frac{1}{4}$.

Proof of Theorem 1.3. Even though there is a loss of the angular regularity in Theorem 1.3, the result implies the sharp result for the radial initial data. The key ingredient here is the following lemma proved in [Gigante and Soria 2008].

Lemma 3.1. *Let $\tilde{J}_\nu(s) = s^{\frac{1}{2}} J_\nu(s)$ with $s \geqslant 0$, and let*

$$T_\nu g(r) = \int_I \frac{e^{it(r)\rho^2} \tilde{J}_\nu(r\rho)}{\rho^{\frac{1}{4}}} g(\rho)\,\mathrm{d}\rho. \tag{3-24}$$

Then

$$\int_0^1 |T_\nu g(r)|^2 \,\mathrm{d}r \leqslant C \int_I |g(\rho)|^2 \,\mathrm{d}\rho, \tag{3-25}$$

where the constant C is independent of $g \in L^2(I)$, of the interval I, of the measurable function $t(r)$ and of the order $\nu \geqslant 0$.

We also can follow the Carleson approach [1980] to linearize our maximal operator, by making t into a function of r, $t(r)$. By the triangle inequality, we have the estimate

$$\|u^*(x)\|_{L^2(B^n)} \leqslant C \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} \left\| \int_0^\infty (r\rho)^{-\frac{n-2}{2}} J_{\nu(k)}(r\rho) e^{it(r)\rho^2} b^0_{k,\ell}(\rho) \rho^{n-1} \,\mathrm{d}\rho \right\|_{L^2_{r^{n-1}\,\mathrm{d}r}}.$$

Let $g(\rho) = b^0_{k,\ell}(\rho)\rho^{\frac{n-1}{2}+\frac{1}{4}}$; then

$$\|u^*(x)\|_{L^2(B^n)} \lesssim \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} \left\| \int_0^\infty \tilde{J}_{\nu(k)}(r\rho) e^{it(r)\rho^2} \rho^{-\frac{1}{4}} g(\rho)\,\mathrm{d}\rho \right\|_{L^2_r([0,1])}. \tag{3-26}$$

Using Lemma 3.1, we obtain

$$\|u^*(x)\|_{L^2(B^n)} \lesssim C \sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} \|b^0_{k,\ell}(\rho)\rho^{\frac{n-1}{2}+\frac{1}{4}}\|_{L^2_\rho(\mathbb{R}+)}. \tag{3-27}$$

Let $\alpha = (n-1)/2 + \epsilon$ with $\epsilon > 0$, we have by the Cauchy–Schwarz inequality

$$\|u^*(x)\|_{L^2(B^n)} \leqslant C \bigg(\sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} (1+k)^{-2\alpha} \bigg)^{\frac{1}{2}} \bigg(\sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} (1+k)^{2\alpha} \|b^0_{k,\ell}(\rho)\rho^{\frac{n-1}{2}+\frac{1}{4}}\|^2_{L^2_\rho(\mathbb{R}+)} \bigg)^{\frac{1}{2}}.$$

Since $d(k) \simeq \langle k \rangle^{n-2}$, we have by Lemma 2.5

$$\|u^*(x)\|_{L^2(B^n)} \lesssim \left(\sum_{k=0}^{\infty} \sum_{\ell=1}^{d(k)} (1+k)^{2\alpha} \left\| b^0_{k,\ell}(\rho) \rho^{\frac{n-1}{2}+\frac{1}{4}} \right\|^2_{L^2_\rho(\mathbb{R}^+)} \right)^{\frac{1}{2}} \lesssim \|u_0\|_{H_r^{\frac{1}{4}} H_\theta^\alpha}.$$

This completes the proof of Theorem 1.3.

Acknowledgments

The authors thank the referee and the associated editor for their invaluable comments and suggestions which helped improve the paper greatly. This work was supported in part by the NSF of China under grants 11171033, 11231006, and 11371059. The second author was partly supported by the Fundamental Research Foundation of BIT (20111742015) and RFDP (20121101120044), Beijing Natural Science Foundation (1144014), and National Natural Science Foundation of China (11401024). C. Miao was also supported by Beijing Center for Mathematics and Information Interdisciplinary Sciences.

References

[Bourgain 1995] J. Bourgain, "Some new estimates on oscillatory integrals", pp. 83–112 in *Essays on Fourier analysis in honor of Elias M. Stein* (Princeton, NJ, 1991), edited by C. Fefferman et al., Princeton Math. Ser. **42**, Princeton University Press, 1995. MR 96c:42028 Zbl 0840.42007

[Bourgain 2013] J. Bourgain, "On the Schrödinger maximal function in higher dimension", *Proc. Steklov Inst. Math.* **280** (2013), 46–60. Zbl 1291.35253

[Burq et al. 2003] N. Burq, F. Planchon, J. G. Stalker, and A. S. Tahvildar-Zadeh, "Strichartz estimates for the wave and Schrödinger equations with the inverse-square potential", *J. Funct. Anal.* **203**:2 (2003), 519–549. MR 2004m:35025 Zbl 1030.35024

[Burq et al. 2004] N. Burq, F. Planchon, J. G. Stalker, and A. S. Tahvildar-Zadeh, "Strichartz estimates for the wave and Schrödinger equations with potentials of critical decay", *Indiana Univ. Math. J.* **53**:6 (2004), 1665–1680. MR 2005k:35241 Zbl 1084.35014

[Carleson 1980] L. Carleson, "Some analytic problems related to statistical mechanics", pp. 5–45 in *Euclidean harmonic analysis* (College Park, MD, 1979), edited by J. J. Benedetto, Lecture Notes in Math. **779**, Springer, Berlin, 1980. MR 82j:82005 Zbl 0425.60091

[Cheeger and Taylor 1982] J. Cheeger and M. Taylor, "On the diffraction of waves by conical singularities, I", *Comm. Pure Appl. Math.* **35**:3 (1982), 275–331. MR 84h:35091a Zbl 0526.58049

[Cho et al. 2006] Y. Cho, S. Lee, and Y. Shim, "A maximal inequality associated to Schrödinger type equation", *Hokkaido Math. J.* **35**:4 (2006), 767–778. MR 2007m:42016 Zbl 1122.42008

[Dahlberg and Kenig 1982] B. E. J. Dahlberg and C. E. Kenig, "A note on the almost everywhere behavior of solutions to the Schrödinger equation", pp. 205–209 in *Harmonic analysis* (Minneapolis, MN, 1981), edited by F. Ricci and G. Weiss, Lecture Notes in Math. **908**, Springer, Berlin, 1982. MR 83f:35023 Zbl 0519.35022

[Gigante and Soria 2008] G. Gigante and F. Soria, "On the boundedness in $H^{1/4}$ of the maximal square function associated with the Schrödinger equation", *J. Lond. Math. Soc.* (2) **77**:1 (2008), 51–68. MR 2010a:42039 Zbl 1218.42005

[Kalf et al. 1975] H. Kalf, U.-W. Schmincke, J. Walter, and R. Wüst, "On the spectral theory of Schrödinger and Dirac operators with strongly singular potentials", pp. 182–226 in *Spectral theory and differential equations* (Dundee, 1974), edited by W. N. Everitt, Lecture Notes in Math. **448**, Springer, Berlin, 1975. MR 53 #1051 Zbl 0311.47021

[Lee 2006] S. Lee, "On pointwise convergence of the solutions to Schrödinger equations in $\mathbb{R}^2$", *Int. Math. Res. Not.* **2006** (2006), Article ID #32597. MR 2007j:35180 Zbl 1131.35306

[Machihara et al. 2005] S. Machihara, M. Nakamura, K. Nakanishi, and T. Ozawa, "Endpoint Strichartz estimates and global solutions for the nonlinear Dirac equation", *J. Funct. Anal.* **219**:1 (2005), 1–20. MR 2006b:35199 Zbl 1060.35025

[Miao et al. 2013] C. Miao, J. Zhang, and J. Zheng, "Strichartz estimates for wave equation with inverse square potential", *Comm. Contemp. Math.* **15**:6 (2013), Article ID #1350026. MR 3139408 Zbl 1284.35357

[Moyua et al. 1996] A. Moyua, A. Vargas, and L. Vega, "Schrödinger maximal function and restriction properties of the Fourier transform", *Int. Math. Res. Not.* **1996**:16 (1996), 793–815. MR 97k:42042 Zbl 0868.35024

[Planchon et al. 2003a] F. Planchon, J. G. Stalker, and A. S. Tahvildar-Zadeh, "Dispersive estimate for the wave equation with the inverse-square potential", *Discrete Contin. Dyn. Syst.* **9**:6 (2003), 1387–1400. MR 2005h:35212 Zbl 1047.35081

[Planchon et al. 2003b] F. Planchon, J. G. Stalker, and A. S. Tahvildar-Zadeh, "L^p estimates for the wave equation with the inverse-square potential", *Discrete Contin. Dyn. Syst.* **9**:2 (2003), 427–442. MR 2003j:35195 Zbl 1031.35092

[Shao 2010] S. Shao, "On localization of the Schrödinger maximal operator", preprint, 2010. arXiv 1006.2787v1

[Sjölin 1987] P. Sjölin, "Regularity of solutions to the Schrödinger equation", *Duke Math. J.* **55**:3 (1987), 699–715. MR 88j:35026 Zbl 0631.42010

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR 95c:42002 Zbl 0821.42001

[Stein and Weiss 1971] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series **32**, Princeton University Press, 1971. MR 46 #4102 Zbl 0232.42007

[Stempak 2000] K. Stempak, "A weighted uniform L^p-estimate of Bessel functions: a note on a paper of Guo", *Proc. Amer. Math. Soc.* **128**:10 (2000), 2943–2945. MR 2000m:33004 Zbl 0951.33004

[Tao 2003] T. Tao, "A sharp bilinear restriction estimate for paraboloids", *Geom. Funct. Anal.* **13**:6 (2003), 1359–1384. MR 2004m:47111 Zbl 1068.42011

[Tao and Vargas 2000] T. Tao and A. Vargas, "A bilinear approach to cone multipliers, II: Applications", *Geom. Funct. Anal.* **10**:1 (2000), 216–258. MR 2002e:42013 Zbl 0949.42013

[Vazquez and Zuazua 2000] J. L. Vazquez and E. Zuazua, "The Hardy inequality and the asymptotic behaviour of the heat equation with an inverse-square potential", *J. Funct. Anal.* **173**:1 (2000), 103–153. MR 2001j:35122 Zbl 0953.35053

[Vega 1988] L. Vega, "Schrödinger equations: pointwise convergence to the initial data", *Proc. Amer. Math. Soc.* **102**:4 (1988), 874–878. MR 89d:35046 Zbl 0654.42014

Received May 23, 2013. Revised January 1, 2014.

CHANGXING MIAO
INSTITUTE OF APPLIED PHYSICS AND COMPUTATIONAL MATHEMATICS
P. O. BOX 8009
BEIJING, 100088
CHINA
miao_changxing@iapcm.ac.cn

JUNYONG ZHANG
DEPARTMENT OF MATHEMATICS
BEIJING INSTITUTE OF TECHNOLOGY
BEIJING, 100081
CHINA
zhang_junyong@bit.edu.cn

JIQIANG ZHENG
THE GRADUATE SCHOOL OF CHINA ACADEMY OF ENGINEERING PHYSICS
P. O. BOX 2101
BEIJING, 100088
CHINA
zhengjiqiang@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.21

VASSILIEV INVARIANTS OF VIRTUAL LEGENDRIAN KNOTS

PATRICIA CAHN AND ASA LEVI

We introduce a theory of virtual Legendrian knots. A virtual Legendrian knot is a cooriented wavefront on an oriented surface up to Legendrian isotopy of its lift to the unit cotangent bundle and stabilization and destabilization of the surface away from the wavefront. We show that the groups of Vassiliev invariants of virtual Legendrian knots and of virtual framed knots are isomorphic. In particular, Vassiliev invariants cannot be used to distinguish virtual Legendrian knots that are isotopic as virtual framed knots and have equal virtual Maslov numbers.

We work in the smooth category. All maps and manifolds are C^∞. All surfaces are oriented unless explicitly stated otherwise. All curves are immersed.

1. Introduction

The first goal of this paper is to introduce a theory of virtual Legendrian knots. Briefly, a virtual Legendrian knot is a Legendrian knot in the spherical cotangent bundle ST^*F of a surface F, up to Legendrian isotopy and stabilization and destabilization of the surface F. In this paper, virtual framed knots are framed knots in ST^*F up to framed isotopy and stabilization and destabilization of the surface F. The second goal of the paper is to study the group of order $\leq n$ Vassiliev invariants of virtual Legendrian knots. We show that this group is naturally isomorphic to the group of order $\leq n$ Vassiliev invariants of virtual framed knots. As a corollary, order $\leq n$ Vassiliev invariants do not distinguish virtual Legendrian knots that are isotopic as framed virtual knots and have the same virtual Maslov number. Similar theorems were proved by Goryunov [1997], Hill [1997], Fuchs and Tabachnikov [1997], and Chernov [2003] for Legendrian knots in various contact manifolds.

We give three equivalent formulations of virtual Legendrian knot theory. The first, which we describe in the introduction, was motivated by the formulation of virtual knot theory given by Carter, Kamada and Saito [2002], and was suggested to us by V. Chernov (private communication, 2011).

MSC2010: 57M27.
Keywords: Virtual knot, Legendrian knot, Vassiliev invariant.

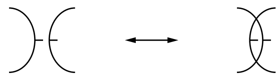

Figure 1. The dangerous tangency move.

First we review the concept of a wavefront as discussed by Arnold [1989]. Suppose that the surface F is made of an isotropic, homogeneous medium, and that light rays are emitted from a point of F. The set of points that these light rays reach at a fixed time t is called a wavefront. As t increases, semicubical cusps may appear, so the wavefront is not necessarily immersed. We say a (cooriented) *wavefront* on an oriented surface F is a cooriented curve on F which is immersed except at a finite number of semicubical cusps. The coorientation represents the direction of propagation of the wavefront. A wavefront is *generic* if it has a finite number of self-intersection points, all of which are transverse double points.

The spherical cotangent bundle ST^*F of a surface F is equipped with the natural contact structure. Arnold [1989; 2004] observed that, due to the Huygens principle, the propagation of a wavefront on F lifts to a Legendrian isotopy in ST^*F. That is, lifting the wavefront to ST^*F according to the direction of its coorienting vector produces a curve in ST^*F which is everywhere tangent to the distribution of contact planes, and during the propagation of the wavefront, the lift of this curve undergoes an isotopy while remaining tangent to the contact planes.

In particular the Huygens principle implies that the *dangerous tangency move* in Figure 1 cannot appear during the propagation of a single front ω, because if two branches of ω become tangent during the propagation in such a way that their coorientations match, they must be tangent for all t.

Hence the Legendrian liftings of two generic wavefronts are Legendrian isotopic if and only if the wavefronts are related by a sequence of the moves in Figure 2 up to certain choices of coorientation, in addition to ambient isotopy. To obtain all the valid choices of coorientation, one should consider the moves in Figure 2 with all possible choices of coorientations on the branches, *except* in the case of the second move, because dangerous tangencies are prohibited; see [Arnold 1994].

A virtual Legendrian knot is a Legendrian knot in the spherical cotangent bundle of a surface, and corresponds to a wavefront on that surface. The definition of virtual Legendrian knot theory suggested by Chernov is as follows: two virtual Legendrian knots are equivalent if their corresponding wavefronts are related by wavefront moves, and stabilization and destabilization of the surface. To stabilize the surface F, we remove two disks from F that are disjoint from the wavefront, and glue the two boundary components together so that the surface remains orientable. To destabilize the surface F, we choose an essential simple closed curve disjoint from the wavefront, cut along it, and glue disks to the two resulting boundary components. Physically, this corresponds to the notion that the medium through

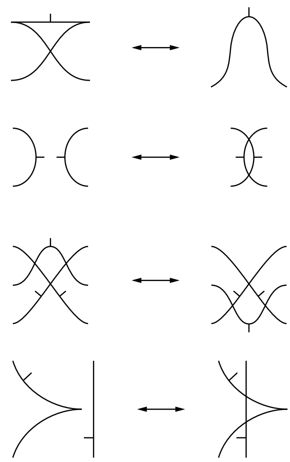

Figure 2. Moves for wavefronts on an oriented surface, with one possible choice of coorientation on each branch.

which the wavefront propagates can change its topology outside a neighborhood of the wavefront. We denote a virtual Legendrian knot by a pair (F, K) where F is a compact oriented surface and K is a wavefront on F, and we let $[(F, K)]_l$ denote its virtual Legendrian isotopy class.

We also give a purely diagrammatic description of virtual Legendrian knot theory in Section 2. Namely, one can view a virtual Legendrian knot as a wavefront in the plane with virtual crossings up to certain moves, in the spirit of Kauffman's original theory of virtual knots [1999].

Questions similar to ours have been studied by many people over the last 20 years. In the rest of this section let $\mathcal{K}$ be $\mathbb{R}$ or $\mathbb{C}$ and $\mathcal{A}$ be any abelian group. Fuchs and Tabachnikov [1997] showed that the vector spaces of $\mathcal{K}$-valued, order $\leq n$ Vassiliev invariants of Legendrian knots with fixed Maslov number in $\mathbb{R}^3$ endowed with the standard contact structure and those of framed knots in $\mathbb{R}^3$ are isomorphic. Around the same time, Goryunov [1997] showed that the vector spaces of $\mathcal{K}$-valued, order $\leq n$ Vassiliev invariants of oriented framed knots in the solid torus $ST^*\mathbb{R}^2$ and those of oriented plane curves without direct self-tangencies are isomorphic.

Generalizing the work of Goryunov, Hill [1997] proved the same result for all planar fronts. Finally, Chernov [2003] was able to show that the groups of $\mathcal{A}$-valued, order $\leq n$ Vassiliev invariants of framed knots in ST^*F, where F is any surface, and those of Legendrian knots in ST^*F are isomorphic.

More precisely, suppose $\mathcal{L}$ is a connected component of the space of Legendrian curves in a contact manifold (M, C) and let $\mathcal{F}$ be the connected component of the space of framed curves in (M, C) that contains $\mathcal{L}$. Chernov proved that the groups of $\mathcal{A}$-valued, order $\leq n$ Vassiliev invariants of framed knots in $\mathcal{F}$ and those of

Legendrian knots in $\mathcal{L}$ are isomorphic for a large class of contact manifolds (M, C). In particular, the theorem holds for $M = ST^*F$ with its natural contact structure. We develop techniques to show that a similar result holds in the virtual category.

Theorem 1.1. *Let $\mathcal{F}$ be a connected component of the space of virtual framed curves and $\mathcal{L} \subset \mathcal{F}$ be a connected component of the space of virtual Legendrian curves contained in $\mathcal{F}$. Let $\mathcal{A}$ be an abelian group and $\mathcal{V}_n^{\mathcal{F}}$ (respectively $\mathcal{V}_n^{\mathcal{L}}$) be the group of $\mathcal{A}$-valued Vassiliev invariants of framed (respectively Legendrian) knots on $\mathcal{F}$ (respectively $\mathcal{L}$) of order $\leq n$. Then the restriction map $\phi : \mathcal{V}_n^F \to \mathcal{V}_n^{\mathcal{L}}$ is an isomorphism.*

It follows that Vassiliev invariants cannot distinguish two virtual Legendrian knots that are homotopic as Legendrian curves and isotopic as framed virtual knots.

Theorem 1.2. *Let $x \in \mathcal{V}_n^{\mathcal{L}}$, let (F, K), (F', L) be representatives of the virtual Legendrian homotopy class $\mathcal{L}$, and let (F, K) be virtually framed isotopic to (F', L'). Then $x\big([(F, K)]_l\big) = x\big([(F', L)]_l\big)$.*

In Section 8 we discuss virtual version of the Maslov number. We prove that, as in the classical case, virtual Legendrian homotopy classes are completely characterized by their Maslov number and their underlying virtual homotopy class.

Theorem 1.3. *Two virtual Legendrian knots are virtual Legendrian homotopic if and only if they have the same virtual Maslov number and are homotopic as virtual knots.*

Part of the motivation for studying virtual knot theory stems from the fact that virtually isotopic classical knots must be isotopic as classical knots. Goussarov, Polyak, and Viro [Goussarov et al. 2000], and the proof can also be found in [Kauffman 1999]. In other words, virtual knot theory extends classical knot theory.

Together with Chernov, we conjecture:

Conjecture 1.4. With $F = S^2$ or $F = \mathbb{R}^2$, two Legendrian knots in ST^*F that are isotopic as virtual Legendrian knots must be Legendrian isotopic in ST^*F.

Kuperberg [2003] later showed that two knots in $F \times I$ that are isotopic as virtual knots must be isotopic as knots in $F \times I$, possibly after a homeomorphism of $F \times I$, provided that F is the surface of smallest genus realizing an element of the virtual isotopy class. We also hope that a similar result holds for virtual Legendrian knots.

Conjecture 1.5. Let K_1 and K_2 be two Legendrian knots in ST^*F that are isotopic as virtual Legendrian knots, and suppose that F is the surface of smallest genus realizing knots in the virtual Legendrian isotopy class of K_1 and K_2. Then, possibly after a contactomorphism of ST^*F, K_1 and K_2 are Legendrian isotopic in ST^*F.

Chernov's definition of virtual Legendrian knot theory can also be generalized to higher dimensions. That is, two Legendrian manifolds L_1 in ST^*M_1 and L_2 in

ST^*M_2 (not necessarily spheres) are virtually Legendrian isotopic if one can be obtained from the other by a sequence of Legendrian isotopies and modifications of contact ST^*M induced by surgery on M in the part of M which does not contain the projection of the Legendrian knot. It's also possible to formulate the first conjecture in higher dimensions, namely two Legendrian knots in $ST^*\mathbb{R}^m$ or ST^*S^m are virtual Legendrian isotopic if and only if they are Legendrian isotopic.

2. Virtual Legendrian knot diagrams

A *virtual Legendrian knot diagram* is a generic wavefront in $\mathbb{R}^2$ with two types of crossings. In addition to ordinary crossings, there are virtual crossings, which are marked with a small circle and obey slightly different Reidemeister moves. The Reidemeister moves involving only ordinary crossings are shown in Figure 2, and these moves are a subset of the possible moves for virtual Legendrian knot diagrams. Again, we can obtain other wavefront moves from these moves by independently reversing the choice of coorientation on any branch, except in the case of the second Reidemeister move, where the dangerous tangency move is forbidden.

The remaining moves for virtual Legendrian front diagrams involve at least one virtual crossing, and are pictured in Figure 3. Other moves can be obtained from these moves by independently reversing the coorientation on any branch. Note that we allow virtual dangerous tangencies. Also note that the move obtained from the first move in Figure 2 by replacing the ordinary crossing with a virtual crossing is *not* allowed.

In Section 7 we will verify that this diagrammatic definition of virtual Legendrian knot theory is equivalent to Chernov's definition given in the Introduction.

If we allow the dangerous self-tangency move in Figure 1, the equivalence relation generated is that of *virtual Legendrian homotopy*, rather than isotopy. We

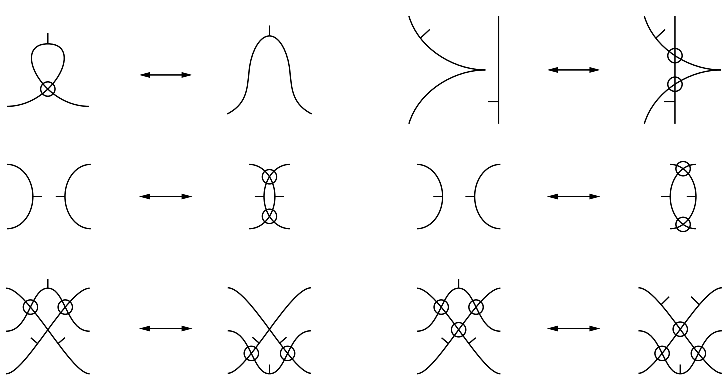

Figure 3. Moves for virtual wavefront diagrams in the plane, with one possible choice of coorientation on each branch.

sometimes refer to a virtual Legendrian homotopy class as a connected component of the space of virtual Legendrian curves.

3. Flat virtual knots

Virtual knot theory was introduced by Kauffman [1999]. We will briefly review the definition of a related object, called a *flat virtual knot*, or *virtual string*, in order to motivate the definition of a virtual Legendrian knot. Virtual strings were introduced by Turaev [2004].

A virtual string is a counterclockwise oriented copy of S^1, called a *core circle*, with arrows whose endpoints are glued to the core circle. The endpoints of the arrows are required to be distinct. We identify two virtual strings if there is a homeomorphism from one to the other preserving the directions of the arrows.

Every generic oriented curve on an oriented surface gives rise to a virtual string, called its *underlying virtual string*. To construct the underlying virtual string of a curve, label the double points of the curve $a_1, \ldots, a_n$. Traverse the curve in the direction of its orientation and record the cyclic order in which the labels a_i appear. Each label appears twice in the cyclic order. Then mark $2n$ points on a counterclockwise oriented copy of S^1, and label these points $a_1, \ldots, a_n$ in the same cyclic order in which the labels appear on the double points during the traversal of the curve. For each $1 \leq i \leq n$, connect the two points labeled a_i by an arrow. In other words, we connect the preimages of each double point of the map $S^1 \to F$ by an arrow. The direction of the arrows are determined by the following rule: At each intersection point of the curve order the two outgoing branches so that their tangent vectors form a positive frame. The head of the arrow in the virtual string should point to the preimage corresponding to the first branch. The underlying virtual string of a curve on a surface is also known as its Gauss diagram.

Two curves on a surface are homotopic if and only if they are related by a sequence of flat Reidemeister moves, pictured at the top of Figure 4, along with other moves that are obtained from those moves by independently reversing the orientation on

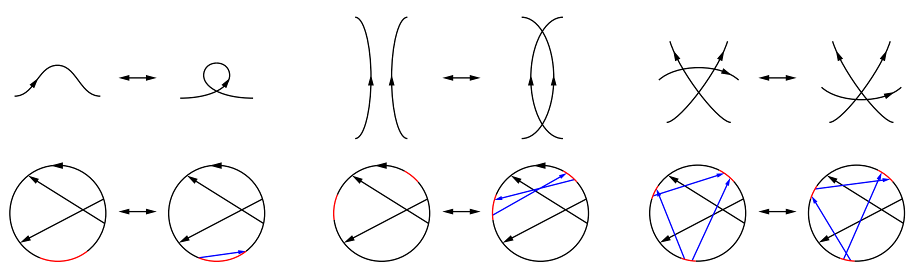

Figure 4. Gauss diagram moves for flat virtual knots and their corresponding moves in the plane.

any branch of the curve in the picture. Each of these moves corresponds to a move on the underlying virtual string (Gauss diagram of the curve). Two virtual strings are virtually homotopic if and only if they are related by a sequence of the Gauss diagram moves in the bottom row of Figure 4, in addition to the Gauss diagram moves obtained from the other versions of the moves in the top row (which differ from the listed moves by the choice of orientation on the branches of the curve.) In Section 4, we will define Gauss diagram moves for virtual Legendrian knots.

4. Gauss diagrams for virtual Legendrian knots

We explain how to associate a Gauss diagram to an oriented virtual Legendrian knot. For planar fronts, our diagrams are similar to the diagrams described by Polyak [1998]. However, unlike Polyak's diagrams, our diagrams are not marked with a basepoint and the signs on our cusps are different from Polyak's.

The Gauss diagram of an oriented and cooriented wavefront K on a surface F with c cusps and n crossings (which we assume are transverse double points) is a counterclockwise oriented copy of S^1, with n arrows glued to S^1 at their endpoints, and c marked points on S^1. Let C be the set of all marked points. Each connected component of $S^1 \setminus C$ is labeled with a coorientation. Furthermore we require that the coorientations on adjacent components of $S^1 \setminus C$ are different, and as a result $|C|$ is even. The endpoints of the arrows are distinct, as are the marked points, and no marked point falls on the endpoint of an arrow. The Gauss diagram of a given front is determined as follows. We view S^1 as the circle parameterizing the curve K, and each pair of preimages of a double point of K is connected by an arrow. At each crossing, we label the outgoing branches of K with a "1" and a "2" so that the ordered pair of their velocity vectors forms a positively oriented frame. The head of the arrow is placed at the preimage corresponding to the branch labeled "1." The marked points of C are the preimages of the cusps of K, and each marked point is equipped with a sign as follows. A cusp of a wavefront is called positive if the outgoing branch of the cusp is in the coorienting half-plane of the cusp. Otherwise, the cusp is called negative (see Figure 5). A virtual Legendrian knot diagram and its Gauss diagram are pictured in Figure 6.

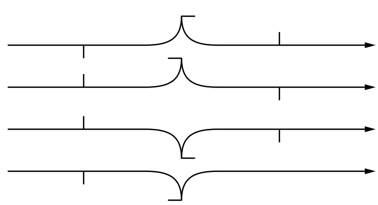

Figure 5. From top to bottom, a positive left cusp, negative left cusp, positive right cusp, and negative right cusp.

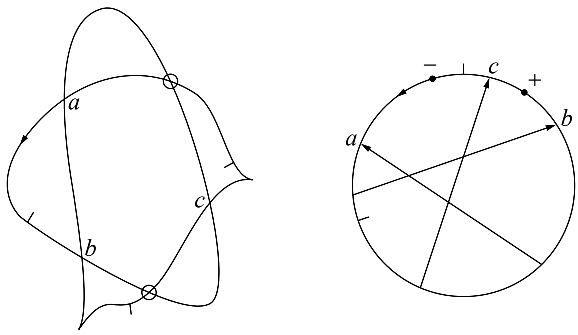

Figure 6. A virtual Legendrian knot diagram and its corresponding Legendrian Gauss diagram.

Each wavefront move in Figure 2 gives rise to a corresponding move on the Gauss diagram of the front in the obvious way, though we do not list these moves explicitly. In Section 7 we will see that the equivalence relation generated by these Gauss diagram moves is simply the equivalence relation of virtual Legendrian knot theory.

5. Virtual framed and virtual Legendrian isotopy in the spherical cotangent bundle

The goal of this section is to define the notions of virtual framed and virtual Legendrian isotopy. The definitions in this section are motivated by the definition of virtual knots given by Carter, Kamada and Saiton [Carter et al. 2002] whose generalization to this context was suggested to us by Chernov. We also introduce flat projection of a virtual blackboard framed knot.

Throughout this section, $\overline{K}$ will typically denote a knot in ST^*F and K will typically denote its projection to the surface.

5A. ***The natural contact structure on the spherical cotangent bundle.*** Let F be an oriented surface and let ST^*F be its spherical cotangent bundle. That is, a point $\omega_p \in ST^*_pF$ is a linear functional on T_pF defined up to multiplication by a positive scalar. Hence ω_p is determined by a choice of 1-dimensional subspace l_ω of T_pF such that $l_\omega = \ker \omega_p$, and a choice of positive half-space, which is a choice of connected component of $T_pF \setminus l_\omega$ on which ω_p is positive. Put $\pi : ST^*F \to F$ to be the usual projection. The contact plane at ω_p is $\pi_*^{-1}(l_\omega)$, which is a 2-dimensional subspace of $T_{\omega_p}ST^*F$.

5B. ***Virtual isotopy in the spherical cotangent bundle of a surface.*** Typically the virtual isotopy class of a knot is the isotopy class of a knot in a thickened surface $F \times I$ up to stabilization and destabilization of F. In this paper we replace $F \times I$ with

ST^*F. We give a careful definition of virtual isotopy in ST^*F in this section, and this definition is based on the formulation of virtual isotopy due in [Carter et al. 2002].

The *surface diagram* of a knot $\overline{K} : S^1 \to ST^*F$ is a triple (F, K, l), where K is the projection of $\overline{K}$ to F, and l is a cooriented line field along K that describes how to lift K to $\overline{K}$. Namely, the point $K(t)$ lifts to the functional in $ST^*_{K(t)}F$ with kernel spanned by $l(t)$, and which is positive on the half-space of $T_{K(t)}F \setminus l(t)$ given by the coorientation of $l(t)$.

Now put

$$(F_1, K_1, l_1) \sim (F_2, K_2, l_2)$$

if there exists a compact oriented surface F_3 and orientation preserving embeddings $\phi_1 : F_1 \to F_3$ and $\phi_2 : F_2 \to F_3$ such that the lifts of the surface diagrams $(F_3, \phi_1(K_1), \phi_{1*}(l_1))$ and $(F_3, \phi_2(K_2), \phi_{2*}(l_2))$ are isotopic in ST^*F_3. Here ϕ_{i*} is the usual differential from $TF_i \to TF_3$, and $\phi_{i*}(l_i)$ is again a cooriented line field. We abuse notation and also let ϕ_{i*} denote the natural map from ST^*F_i to ST^*F_3. At first glance it appears that this map goes in the incorrect direction, but ϕ_i is an embedding so its differential, and hence the induced map on cotangent spaces, are isomorphisms. So, we abuse notation and let $\phi_{i*} : ST^*F_i \to ST^*F_3$. Note that the lift to ST^*F_3 of $(F_3, \phi_i(K_i), \phi_{i*}(l_i))$ is simply $\phi_{i*}(\overline{K}_i)$.

Two virtual knots (F, K, l) and (F', K', l') are *virtually isotopic* if there is a sequence of knot diagrams (F_i, K_i, l_i), $1 \le i \le m$, such that

$$(F, K, l) = (F_1, K_1, l_1) \sim (F_2, K_2, l_2) \sim \cdots \sim (F_m, K_m, l_m) = (F', K', l').$$

5C. ***Virtual framed isotopy in the spherical cotangent bundle.*** Next we define virtual isotopy for framed knots in ST^*F where F is an oriented surface. A virtual framed knot is a knot $\overline{K} : S^1 \to ST^*F$ equipped with a transverse vector field ν considered up to the equivalence relation we define below. We denote this virtual framed knot $(F, \overline{K}^\nu)$. Because we will not need to work with the projection of a virtual framed knot, we will sometimes write (F, K^ν) rather than $(F, \overline{K}^\nu)$; the meaning of the notation will be clear from context.

Let $\phi : F \to F'$ be an orientation preserving embedding, and as above, let $\phi_* : ST^*F \to ST^*F'$ be the induced map on the spherical cotangent bundles. Let $\phi_{**} : TST^*F \to TST^*F'$ be the differential of ϕ_*.

Put $(F_1, \overline{K}_1^{\nu_1}) \sim_f (F_2, \overline{K}_2^{\nu_2})$ if there exists a compact oriented surface F_3 and orientation preserving embeddings $\phi_1 : F_1 \to F_3$ and $\phi_2 : F_2 \to F_3$ such that $(F_3, \phi_{1*}(\overline{K}_1)^{\phi_{1**}(\nu_1)})$ is framed isotopic to $(F_3, \phi_{2*}(\overline{K}_2)^{\phi_{2**}(\nu_2)})$ in ST^*F_3.

We say $(F_1, \overline{K}_1^{\nu_1})$ and $(F_m, \overline{K}_m^{\nu_m})$ are *virtually framed isotopic* if there exists a sequence of virtual framed knots $\overline{K}_i^{\nu_i}$, $1 < i < m$, satisfying

$$(F_1, \overline{K}_1^{\nu_1}) \sim_f (F_2, \overline{K}_2^{\nu_2}) \sim_f \cdots \sim_f (F_m, \overline{K}_m^{\nu_m}).$$

The blackboard framing. We will sometimes consider knots with a certain framing, which we call the blackboard framing. Fix an orientation of ST^*F and of F. A virtual topological knot (F, K, l) is in general position if its velocity vector, $\overline{K}'(t)$, never points in the direction of the S^1 fiber. The orientations of ST^*F and F determine an orientation of the S^1 fibers of ST^*F. Let $\partial\theta$ be the vector field corresponding to the positive orientation of the S^1 fibers. Then for a virtual topological knot in general position the vector field $\partial\theta$ is always transverse to $\overline{K}$ and thus is a framing vector field. We call this framing the *blackboard framing*. We can associate this framing to any virtual topological knot in general position to obtain a virtual framed knot with the blackboard framing, $\overline{K}^{\partial\theta}$.

Proposition 5.1. *Let $(F, \overline{K}^\nu)$ be a virtual framed knot. Then there is a blackboard framed virtual knot $(F, \bar{L}^{\partial\theta})$ in the framed virtual isotopy class $[(F, \overline{K}^\nu)]_f$.*

Proof. It is enough to show that the nonvirtual framed isotopy class of $\overline{K}^\nu$ contains a blackboard framed knot $\bar{L}^{\partial\theta}$. We may assume, possibly by first performing a small perturbation, that the velocity vector of $\overline{K}$ never points in the direction of the S^1 fiber, where $\overline{K}$ denotes the unframed knot in ST^*F obtained from $\overline{K}^\nu$ by forgetting its framing. Now consider the blackboard framed knot $\overline{K}^{\partial\theta}$ which coincides with $\overline{K}^\nu$ as an unframed knot. Let $j = m(\overline{K}^\nu, \overline{K}^{\partial\theta})$ be the relative number of twists of the framings of the two knots (see Section 10A). Next consider the surface diagram (F, K, l) consisting of the projection K of $\overline{K}$ to F and the cooriented line field l which describes how to lift K to $\overline{K}$. We can replace a portion of this surface diagram, where the line field is locally constant, with a small kink (see Figure 7). Adding the top or bottom kink in Figure 7 to (F, K, l) will yield new framed knots $\overline{K}_1$ and $\overline{K}_2$ respectively, which, once isotoped to coincide with $\overline{K}$ as embeddings,

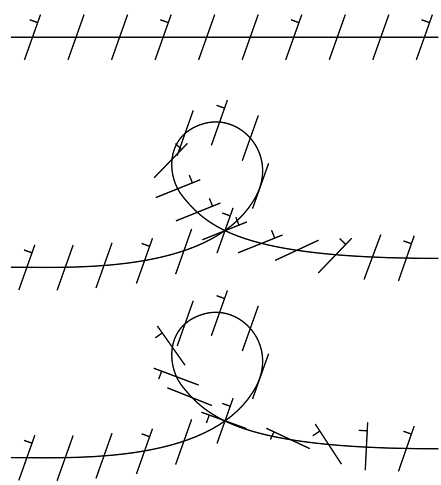

Figure 7. Adding a twist to the framing of a blackboard framed knot.

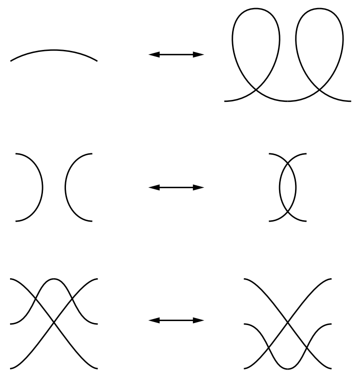

Figure 8. Moves for flat virtual blackboard framed knot projections.

satisfy $m(\overline{K}_i, \overline{K}) = \pm 1$, $i = 1, 2$. The sign depends on the chosen orientation of the fiber. We add $|j|$ copies of one of the kinks in Figure 7 so that the lift $\bar{L}$ of the resulting surface diagram is isotopic to $\overline{K}$, and the blackboard framed knot $\bar{L}^{\partial\theta}$ is framed isotopic to $\overline{K}^{\nu}$. □

The flat projection of a virtual blackboard framed knot. To prove Proposition 10.3 we will use an invariant that is defined on the flat diagram of a blackboard framed virtual framed knot. The flat projection of a virtual blackboard framed knot, $\overline{K}^{\partial\theta}$, is the immersed curve $\pi(\overline{K})$.

If two blackboard framed knots in the spherical cotangent bundle of the same surface are framed homotopic, then their flat projections are related by a sequence of moves in Figure 8.

Thus if two blackboard framed virtual knots $\overline{K}_1^{\partial\theta}$ and $\overline{K}_2^{\partial\theta}$ are virtually framed homotopic then there exists a sequence of moves in Figure 8, in addition to stabilization and destabilization of the surface that takes the flat projection of $\overline{K}_1^{\partial\theta}$ to that of $\overline{K}_2^{\partial\theta}$.

5D. ***Virtual Legendrian isotopy.*** Let $\overline{K}$ be a Legendrian knot in ST^*F in general position. Then $\overline{K}$ projects to a cooriented wavefront K on F. Furthermore, two Legendrian knots $\overline{K}_1$ and $\overline{K}_2$ are Legendrian isotopic in ST^*F if and only if their wavefronts K_1 and K_2 on F are related by a sequence of the moves in Figure 2, excluding the dangerous self-tangency move.

Put $(F_1, K_1) \sim_l (F_2, K_2)$ if there exists a compact oriented surface F_3 and orientation preserving embeddings $\phi_1 : F_1 \to F_3$ and $\phi_2 : F_2 \to F_3$ such that $\phi_1(K_1)$ and $\phi_2(K_2)$ are related by a sequence of moves for wavefronts on F_3, or equivalently, if $\overline{\phi_1(K_1)}$ and $\overline{\phi_1(K_1)}$ are Legendrian isotopic in ST^*F_3.

We say (F, K) and (F', K') are *virtually Legendrian isotopic* if there exists a sequence of pairs satisfying

$$(F, K) = (F_1, K_1) \sim_l (F_2, K_2) \sim_l \cdots \sim_l (F_m, K_m) = (F', K').$$

Remark 5.2. If in this definition we allowed dangerous tangency moves as well then we would have the definition of a virtual Legendrian homotopy.

Remark 5.3. A virtual Legendrian knot in a cooriented contact structure has a natural framing, at each point given by the unit vector in the normal bundle to the velocity vector of the knot on which the cooriented one form evaluates to one. Given a virtual Legendrian knot $\overline{K}$, let $\overline{K}^{st}$ denote the virtual framed knot given by $\overline{K}$ with this framing.

6. Flat framed virtual knot diagrams

In this section we give reformulate the theory of flat virtual framed knots in terms of planar diagrams.

Given a virtual knot with a blackboard framing, $\overline{K}^{\partial\theta}$ we associate to it a planar flat virtual framed knot diagram. This is obtained first by forgetting the framing and cooriented line field, leaving a generic immersed curve K on the surface F, denoted (F, K). Then we construct the virtual string associated to this curve as described in Section 3. This virtual string describes a flat virtual knot diagram in the plane in the usual way, which is unique up to any combination of virtual moves (a sequence of virtual moves is sometimes called the detour move).

If two virtual framed knots are virtual framed homotopic then their planar flat virtual diagrams are related by a sequence of moves in Figure 9.

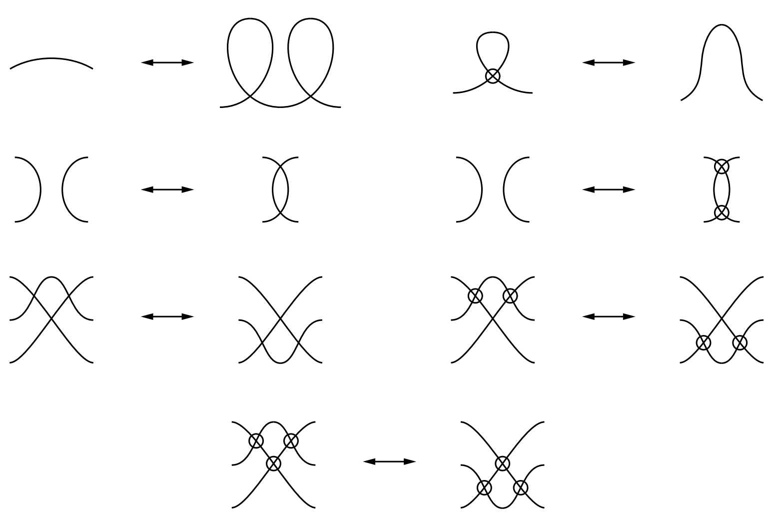

Figure 9. Moves for flat virtual framed knot diagrams.

7. Equivalent definitions of virtual Legendrian isotopy

Let $\mathcal{LGD}$ be the set of Legendrian Gauss diagrams and let $\sim_{lgd}$ be the equivalence relation generated by Gauss diagram moves. Recall that the set of Gauss diagram moves is precisely the set of moves on Gauss diagrams obtained by translating each wavefront move to the Gauss diagram. Let $LGD = \mathcal{LGD}/\sim_{lgd}$.

Let $\mathcal{VLD}$ be the set of virtual Legendrian knot diagrams. Let $\sim_{vld}$ be the equivalence relation generated by virtual Legendrian knot diagram moves. Put $VLD = \mathcal{VLD}/\sim_{vld}$.

The theory of Legendrian Gauss diagrams up to Gauss diagram moves is equivalent to the theory of virtual Legendrian knot diagrams up to virtual wavefront moves.

Theorem 7.1. *The map $g : \mathcal{VLD} \to \mathcal{LGD}$ given by associating a* (*unique*) *Legendrian Gauss diagram to a virtual Legendrian knot diagram induces a bijection $g_\sim : VLD \to LGD$.*

Proof. First we verify that $g_\sim$ is well-defined. Indeed, the Legendrian Gauss diagrams of two equivalent virtual Legendrian knot diagrams differ by Gauss diagram moves, as moves involving virtual crossings do not affect the Gauss diagram.

The map $g_\sim$ is clearly surjective, as any Legendrian Gauss diagram gives rise to a virtual Legendrian knot diagram. One simply draws all cusps and crossings present in the Legendrian Gauss diagram in the plane, and connects such cusps and crossings by arcs, creating virtual crossings where these arcs cross.

It remains to check that $g_\sim$ is injective. Suppose we have two virtual Legendrian knot diagrams D_1 and D_2 with the same Gauss diagram. We will show that D_1 and D_2 are related by virtual Legendrian knot diagram moves. We first change D_2 by a regular isotopy so that a small neighborhood of each of its cusps and double points coincides with a small neighborhood of each corresponding cusp and double point of D_1. The resulting diagrams differ only in how the cusps and double points are connected by arcs. To move an arc a_2 of D_2 so that it coincides with the corresponding arc a_1 of D_1, we move a_2 by a fixed endpoint homotopy, such that any crossings created during that homotopy are virtual. This move is known as the *detour move*, and is simply a sequence of the moves pictured in Figure 3. □

Now let $\mathcal{SFD}$ be the set of all front diagrams on orientable surfaces, i.e., all pairs (F, K) where F is an oriented surface and K is a cooriented wavefront on F. Let $\sim_{sfd}$ be the equivalence relation generated by the relation $\sim_l$ defined in Section 5D, and let $SFD = \mathcal{SFD}/\sim_{sfd}$. In other words, SFD consists of pairs (F, K) of oriented surfaces F with cooriented wavefront diagrams K on F considered up to wavefront moves and stabilization and destabilization of the surface F.

Next we show that the theory of Legendrian Gauss diagrams up to Gauss diagram moves is equivalent to the theory of fronts on surfaces up to wavefront moves.

Theorem 7.2. *The map $h : \mathcal{SFD} \to \mathcal{LGD}$ that assigns a Legendrian Gauss diagram to a wavefront on an oriented surface induces a bijection $h_\sim : SFD \to LGD$.*

Proof. First we check that $h_\sim$ is well-defined. That is, suppose we have two pairs (F, K) and (F', K'), such that for some sequence of pairs $\{(F_i, K_i)\}_{i=1}^n$, we have

$$(F, K) = (F_1, K_1) \sim_l (F_2, K_2) \sim_l \ldots (F_n, K_n) = (F', K').$$

We need to check that (F, K) and (F', K') have equivalent Legendrian Gauss diagrams, but since stabilization and destabilization do not affect the Legendrian Gauss diagram of a wavefront on a surface, this is clear.

Next we verify that $h_\sim$ is surjective. To do this, we construct a wavefront diagram on a surface given a Legendrian Gauss diagram. First we build the disk-band surface realizing the underlying flat virtual knot of the wavefront. For a detailed explanation of this procedure, see [Turaev 2004]. Then we insert positive and negative cusps according to the markings on the Legendrian Gauss diagram.

Finally we verify $h_\sim$ is injective. To do this we must show that if two pairs (F, K) and (F', K') have equivalent Gauss diagrams, then (F, K) and (F', K') are equivalent. Again, the argument is completely analogous to the case of virtual strings, which is carefully described in [Turaev 2004]. □

8. Virtual versions of the classical invariants

In this section we define the Maslov number for virtual Legendrian knots. We do not discuss virtual analogues of the Bennequin number in this work. The virtual Maslov number $\mu(K_v)$ where K_v is a planar virtual front diagram, is defined to be the number of positive cusps minus the number of negative cusps; see Figure 5. Clearly, if K_v corresponds to the front (F, K) on a surface then $\mu(K_v)$ is equal to the (nonvirtual) Maslov number of the front K on F.

A *positive* (*respectively negative*) *stabilization* of the virtual Legendrian knot K is obtained by inserting a pair of positive (respectively negative) cusps at any point along K. Let K^{n_1,n_2} denote the virtual Legendrian knot obtained from K by applying n_1 positive stabilizations and n_2 negative stabilizations.

Proposition 8.1. *The virtual Legendrian knot K is virtual Legendrian homotopic to $K^{n,n}$ for any positive integer n.*

Proof. One can add a pair of positive cusps and a pair of negative cusps using a Legendrian homotopy; see Figure 10. This sequence of moves is due to Fuchs and Tabachnikov [1997]. This Legendrian homotopy is also a virtual Legendrian homotopy. □

Figure 10. Adding a pair of negative cusps and a pair of positive cusps via a (virtual) Legendrian homotopy.

Theorem 8.2. *Two virtual Legendrian knots are virtual Legendrian homotopic if and only if they have the same virtual Maslov number and are homotopic as virtual knots.*

Proof. One can verify that the virtual Maslov number is an invariant of virtual Legendrian homotopy by checking its invariance under all virtual wavefront moves, as well as the dangerous tangency move. Hence two virtual Legendrian homotopic virtual Legendrian knots have the same virtual Maslov number and (clearly) are homotopic as virtual knots.

Now suppose that K and L are Legendrian knots with the same virtual Maslov number that are homotopic as virtual knots. We will show below that the assumption that K and L are homotopic as virtual knots implies that for any two sufficiently large positive integers n_1 and n_2, there exist integers n_3 and n_4 such that K^{n_1,n_2} is virtual Legendrian homotopic to L^{n_3,n_4}. In particular, this will be true for some $p = n_1 = n_2$ large enough. Then, since K and $K^{p,p}$ are virtual Legendrian homotopic by Proposition 8.1, and for suitable n_3 and n_4, $K^{p,p}$ and L^{n_3,n_4} are virtual Legendrian homotopic, we have $\mu(L) = \mu(K) = \mu(K^{p,p}) = \mu(L^{n_3,n_4})$. Then $\mu(L) = \mu(L^{n_3,n_4})$ implies $n_3 = n_4$. Put $q = n_3 = n_4$. Again by Proposition 8.1, L and $L^{q,q}$ are virtual Legendrian homotopic. Hence K and L are virtual Legendrian homotopic.

It remains to show that if K and L are virtually homotopic then for sufficiently large positive integers n_1 and n_2, there exist integers n_3 and n_4 such that K^{n_1,n_2} is virtual Legendrian homotopic to L^{n_3,n_4}. We do this in the next lemma. □

Lemma 8.3. *Let (F, K) and (F', L) be two virtually homotopic Legendrian knots. Then for sufficiently large positive integers n_1 and n_2 there exist integers n_3 and n_4 such that K^{n_1,n_2} is virtual Legendrian homotopic to L^{n_3,n_4}.*

Proof. We let $\bar{K}$ be a Legendrian knot in ST^*F, let $\bar{L}$ be a Legendrian knot in ST^*F', and we have a sequence of pairs

$$(F, K) = (F_1, K_1), (F_2, K_2), \dots, (F_n, K_n) = (F', L)$$

of curves $\overline{K}_i$ in ST^*F_i such that the cooriented line fields (F_i, K_i, l_i) on F_i and $(F_{i+1}, K_{i+1}, l_{i+1})$ on F_{i+i} lifting to $\overline{K}_i$ and $\overline{K}_{i+1}$ respectively can be realized as homotopic cooriented line fields on a third surface G_i (meaning their lifts are homotopic in ST^*G_i).

On each surface G_i, this homotopy looks locally (within a Darboux chart) like a sequence of Reidemeister moves and crossing changes. We show how to imitate this virtual homotopy by a virtual Legendrian homotopy by replacing topological Reidemeister moves with Legendrian Reidemeister moves, and by replacing topological crossing changes with Legendrian crossing changes.

The argument is now local, so we simply consider the case where $\overline{K}$ and $\bar{L}$ are homotopic as Legendrian knots in ST^*F for a fixed surface F. Furthermore we assume for now that the homotopy between $\overline{K}$ and $\bar{L}$ is contained in a single Darboux chart, so that $\overline{K}$ and $\bar{L}$ are Legendrian knots in the standard contact $\mathbb{R}^3$. We consider the topological knot projections of $\overline{K}$ and $\bar{L}$ to the xz-plane. Because $\overline{K}$ and $\bar{L}$ are homotopic, their topological projections are related by a sequence of Reidemeister moves of Types 1–3 and crossing changes; see Figure 11.

Fuchs and Tabachnikov [1997] proved that if K and L are topologically isotopic Legendrian knots in the standard contact $\mathbb{R}^3$, then for sufficiently large n_1 and n_2, there exist n_3 and n_4 such that K^{n_1,n_2} is Legendrian isotopic to L^{n_3,n_4}. We prove the same statement, replacing isotopy with homotopy, and imitate their proof. Let κ and λ be the front projections of K and L. First we know κ and λ are related by a topological isotopy K_t with projection κ_t, such that $K_0 = K$, $K_1 = L$, and for some $0 = t_0 < t_1 < t_2 < \cdots < t_n = 1 \in [0, 1]$, $\kappa_i = \kappa_{t_i}$ and $\kappa_{i+1} = \kappa_{t_{i+1}}$ are related by a single topological Reidemeister move, topological crossing change (see Figure 11), or a passage through a vertical tangency (an ambient isotopy during which a strand without a vertical tangency passes through a vertical tangency, and after which two new vertical tangencies appear locally).

Next convert each topological knot diagram κ_i into a front diagram by replacing vertical tangencies with cusps, and correcting "wrong crossings" as in [Fuchs and Tabachnikov 1997]; see Figure 12. These operations clearly do not affect the topological type of the knot.

Fuchs and Tabachnikov then show how to replace topological Reidemeister moves and passages through vertical tangencies with Legendrian versions of these

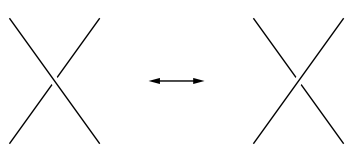

Figure 11. A crossing change.

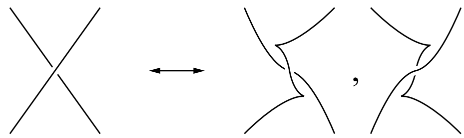

Figure 12. Converting a "wrong crossing" in a topological knot diagram to a front projection of a Legendrian knot in standard contact $\mathbb{R}^3$.

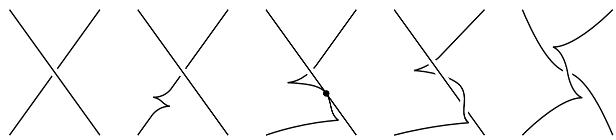

Figure 13. After adding cusp pairs, we can replace a topological crossing change with a Legendrian crossing change.

moves, possibly after adding extra positive or negative cusp pairs. It remains to show that we can do the same for topological crossing changes. This is shown in Figure 13. □

9. Vassiliev invariants

In this section we define finite order Vassiliev invariants for virtual Legendrian, framed and topological knots. For the rest of the paper let $\mathcal{A}$ be an abelian group.

A self-intersection point of an immersed curve is called transverse if the two velocity vectors to the curve at that point are linearly independent. An n-singular virtual Legendrian (resp. framed, topological) knot is a virtual Legendrian (resp. framed, topological) immersed curve with n transverse self-intersection points.

In an oriented 3-manifold a transverse self-intersection point of an immersed curve can be resolved in two ways. We call the resolution positive if the velocity vector to one strand, the velocity vector to the second strand and the vector from the second to the first strand form a positive 3-frame. Otherwise the self-intersection is called negative. Given an immersed curve with $(n+1)$ transverse self-intersection points there are 2^{n+1} possible ways to resolve these self-intersection points. The *sign of a resolution* is the product of the signs of the resolutions of all of the individual double points.

An $\mathcal{A}$-valued virtual Legendrian (resp. framed, topological) knot invariant is a function from the set of virtual Legendrian (resp. framed, topological) isotopy classes to $\mathcal{A}$.

A Vassiliev invariant of virtual Legendrian (resp. framed, topological) of order $\leq n$ is an $\mathcal{A}$-valued virtual Legendrian (resp. framed, topological) knot invariant that vanishes on the signed sum of the 2^{n+1} resolutions of any $(n+1)$-singular virtual Legendrian (resp. framed, topological) knot.

10. Construction of the isomorphism

In what follows, recall that $[\,\cdot\,]_f$ denotes a framed virtual isotopy class, $[\,\cdot\,]_l$ denotes a Legendrian virtual isotopy class, and $[\,\cdot\,]$ denotes a topological virtual isotopy class.

When studying Vassiliev invariants of virtual Legendrian knots, we consider invariants of virtual Legendrian knots from a single connected component of the space of virtual Legendrian curves (or equivalently, knots within a fixed virtual homotopy class and with a fixed virtual Maslov number). Similarly, when studying Vassiliev invariants of virtual framed knots or virtual Legendrian knots, we consider invariants of knots from a single connected component of the space of virtual framed curves.

In the following sections we will show that Vassiliev invariants cannot be used to distinguish virtually homotopic virtual Legendrian knots with the same virtual Maslov number that are isotopic as framed virtual knots. First we prove the following seemingly weaker statement:

Theorem 10.1. *Let $x \in \mathcal{V}_n^{\mathcal{L}}$. Suppose that (F, K) and (F', L) are two virtually homotopic virtual Legendrian knots with the same virtual Maslov number, such that (F, K) is virtual framed isotopic to (F', L) with their natural framings. Then $x([(F, K)]_l) = x([(F', L)]_l)$.*

Using this theorem we are then able to prove the following stronger statement:

Theorem 10.2. *Let $\mathcal{F}$ be a connected component of the space of virtual framed curves and $\mathcal{L} \subset \mathcal{F}$ be a connected component of the space of virtual Legendrian curves contained in $\mathcal{F}$. Let $\mathcal{A}$ be an abelian group and $\mathcal{V}_n^{\mathcal{F}}$ be the group of $\mathcal{A}$ valued Vassiliev invariants on $\mathcal{F}$ of order $\leq n$. Define $\mathcal{V}_n^{\mathcal{L}}$ likewise. The restriction map $\phi : \mathcal{V}_n^{\mathcal{F}} \to \mathcal{V}_n^{\mathcal{L}}$ is an isomorphism.*

It is clear that Theorem 10.2 implies Theorem 10.1. We will now show that Theorem 10.1 implies Theorem 10.2. We will prove Theorem 10.1 in a later section.

Now we outline the construction of an inverse, ψ, to the map ϕ defined above. If the framed isotopy class of (F, K^v) contains a Legendrian representative (F_l, K_l) then we define $\psi(x)([(F, K^v)]_f) = x([(F_l, K_l)]_l)$. By Theorem 10.1, this is well-defined. If every virtual framed isotopy class were realizable by a virtual Legendrian knot, then the existence of ψ would follow immediately from Theorem 10.1. The Bennequin inequality tells us that not every (nonvirtual) framed isotopy class is realizable by a Legendrian knot in, for example, $\mathbb{R}^3$ with the standard contact structure. However, no Bennequin inequality is known for virtual Legendrian knots.

Therefore, it could be that every virtual framed isotopy class is realizable by a virtual Legendrian knot; this question is currently open.

10A. ***The relative number of twists of two framed knots.*** Given two virtual framed knots $\overline{K}_1^{\nu_1}$ and $\overline{K}_2^{\nu_2}$ that coincide as pointwise embeddings in the same spherical cotangent bundle ST^*F, we can measure the relative number of twists of their framings as follows. Let $\nu_1^{\perp}$ be a vector in TST^*F orthogonal to both $\overline{K}'_1(t)$ and $\nu_1(t)$ such that the triple $\{\overline{K}'_1(t), \nu_1(t), \nu_1^{\perp}\}$ is a positive frame in TST^*F. The frame consisting of $\nu_1^{\perp}$ and the vector $\nu_1(t)$ gives a trivialization of the normal bundle of $\overline{K}$. Define $m(\overline{K}_1^{\nu_1}, \overline{K}_2^{\nu_2})$ to be the total number of rotations of ν_2 with respect to this trivialization.

In Proposition 10.3 we use m to characterize when two framed virtual knots that coincide as unframed knots are homotopic as virtual framed curves.

For the proof of Proposition 10.3 we need the following definition. Given a virtual framed knot $\overline{K}^{\nu}$ consider one of its corresponding flat virtual framed knot diagrams in $\mathbb{R}^2$ (see Section 6). Let r be the rotation number of the flat framed knot diagram in the plane and let v be the number of virtual crossings. Put $\rho(\overline{K}^{\nu}) = r + v \mod 2$. One can check that this quantity does not change under any moves in Figure 9, and thus is well-defined across all possible virtual framed knot diagrams for a given virtual framed knot. Furthermore this also shows that it is invariant under virtual framed homotopy.

Proposition 10.3. *Let $\overline{K}_1^{\nu_1}$ and $\overline{K}_2^{\nu_2}$ be virtual framed knots* (*resp. singular virtual framed knots with n transverse double points*) *that coincide pointwise as embeddings* (*resp. immersions*) *in ST^*F. Then $\overline{K}_1^{\nu_1}$ and $\overline{K}_2^{\nu_2}$ are virtual framed homotopic if and only if $m(\overline{K}_1^{\nu_1}, \overline{K}_2^{\nu_2})$ is even.*

Proof. If $m(\overline{K}_1^{\nu_1}, \overline{K}_2^{\nu_2})$ is even, then $\overline{K}_1^{\nu_1}$ and $\overline{K}_2^{\nu_2}$ are homotopic as framed knots in ST^*F because one can pass through a small kink to change m by two.

Now suppose $m(\overline{K}_1^{\nu_1}, \overline{K}_2^{\nu_2})$ is odd. Since m is odd we must have $\rho(\overline{K}_1^{\nu_1}) \neq \rho(\overline{K}_2^{\nu_2})$. Thus $\overline{K}_1^{\nu_1}$ cannot be virtual framed homotopic to $\overline{K}_2^{\nu_2}$. □

Suppose that $\overline{K}_1^{\nu_1}$ and $\overline{K}_2^{\nu_2}$ coincide as smooth embeddings, and $m(\overline{K}_1^{\nu_1}, \overline{K}_2^{\nu_2}) = i$. Then we write $\overline{K}_2^{\nu_2} = (\overline{K}_1^{\nu_1})^i$.

We want to prove that Theorem 10.1 implies that the inverse ψ described above exists. Chernov [2003] showed that an analogue of Theorem 10.1 for ordinary Legendrian and framed knots in most contact manifolds implies that an analogue Theorem 10.2 holds, i.e., the inverse ψ exists. The proof in [Chernov 2003] that ψ exists is mostly local. One can check that the same proof will work in the virtual category provided the following two propositions hold:

Proposition 10.4. *Let $\mathcal{F}$ be a connected component in the space of virtual framed curves* (*resp. singular virtual framed curves*), *and $\mathcal{L} \subset \mathcal{F}$ be a connected component* *in the space of virtual Legendrian curves (resp. singular virtual Legendrian curves). Let* $\overline{K}^\nu \in \mathcal{F}$ *be a virtual framed knot (resp. singular knot). Then there exists* $i \in \mathbb{Z}$ *and a virtual Legendrian knot (resp. singular knot)* $\bar{L} \in \mathcal{L}$ *such that* $\bar{L} \in [(\overline{K}^\nu)^{2i}]_f$*. Furthermore if there exists a virtual Legendrian knot (resp. singular knot)* $\bar{L} \in \mathcal{L}$ *such that* $[\bar{L}]_f = [\overline{K}^\nu]_f$ *then there exists a virtual Legendrian knot (resp. singular knot)* $\bar{L}' \in \mathcal{L}$ *such that* $[\bar{L}']_f = [(\overline{K}^\nu)^{-2}]_f$*.*

Proof. Chernov [2003] shows that for some $i \in \mathbb{Z}$, there exists a Legendrian knot $\bar{L}$ in the ordinary (nonvirtual) framed isotopy class of $(\overline{K}^\nu)^{2i}$ that is also contained in the given connected component of the space of Legendrian curves. This $\bar{L}$ suffices here also because of Proposition 10.3. We will construct this knot $\bar{L}$ when the given knot is nonsingular, but the singular case is similar and is done in [Chernov 2003] for nonvirtual knots. To construct $\bar{L}$, one first forgets the framing of the framed knot $\overline{K}^\nu$ to obtain a knot K. Then C^0-approximate the knot K by a Legendrian knot, and add sufficiently many positive or negative cusp pairs, so that the resulting Legendrian knot $\overline{K}_l$ is in the given virtual Legendrian homotopy class $\mathcal{L}$. This Legendrian knot $\overline{K}_l$ is framed isotopic to a framed knot $\overline{K}^{\nu j}$ such that $\overline{K}^{\nu j}$ and $\overline{K}^\nu$ coincide as unframed knots and $m(\overline{K}^{\nu j}, \overline{K}^\nu) = j$. But because $\mathcal{L} \subset \mathcal{F}$, by Proposition 10.3, $j = 2i$.

To show that if the framed isotopy class of $\overline{K}^\nu$ in ST^*F is realizable by a Legendrian knot, then the framed isotopy class of $(\overline{K}^\nu)^{-2}$ in ST^*F is realizable by a Legendrian knot $\bar{L}'$, simply perform the Legendrian homotopy in Figure 10 to a small arc of the Legendrian knot L in the virtual framed isotopy class of $\overline{K}^\nu$. The resulting Legendrian knot is in the virtual framed isotopy class of $(\overline{K}^\nu)^{-2}$. □

Proposition 10.5. *Let* $\mathcal{F}$ *be a connected component of the space of virtual framed curves, let* $\overline{K}^\nu \in \mathcal{F}$ *and let* $K_u = (F, K, l)$ *be an unframed virtual knot obtained by forgetting the framing on* $\overline{K}^\nu$*. Let* $[K_u]$ *be the class of virtual topological knots that contains* K_u *and* $\overline{K}_1^{\nu_1} = (F_1, \overline{K}_1^{\nu_1}) \in \mathcal{F}$ *be a virtual framed knot with* $(F_1, K_1, l_1) \in [K_u]$*. Then* $\overline{K}_1^{\nu_1}$ *and* $(\overline{K}^\nu)^{2i}$ *are virtual framed isotopic for some* $i \in \mathbb{Z}$*.*

Proof. This follows immediately from Proposition 10.3. □

We know how to define ψ on virtual framed isotopy classes containing a virtual Legendrian knot. Let $\overline{K}^\nu \in \mathcal{F}$ and i be the largest integer such that $[(\overline{K}^\nu)^{2i}]_f$ contains a virtual Legendrian knot in $\mathcal{L}$. (If no such i exists, then every $[(\overline{K}^\nu)^{2i}]_f$ contains a virtual Legendrian knot, so there is no problem defining ψ on all of $\mathcal{F}$.) In this situation we know how to define ψ on the framed isotopy classes $[K^{2j}]_f$ for $j \leq i$. The following definition, analogous to the definition in [Chernov 2003], extends ψ to the virtual framed isotopy classes $[K^{2j}]_f$ for $j > i$.

Definition 10.6. Fix $\overline{K}^\nu \in \mathcal{F}$ and let j be the maximal integer such that $[(\overline{K}^\nu)^{2j}]$ contains a virtual Legendrian knot in $\mathcal{L}$. For $l > j$ define

$$\psi(x)((\overline{K}^\nu)^{2l}) = \sum_{i=1}^{n+1}\left((-1)^{i+1}\frac{(n+1)!}{i!\,(n+1-i)!}\psi(x)((\overline{K}^\nu)^{2l-2i})\right)$$

This definition extends x such that it is a Vassiliev invariant of virtual framed knots of order $\leq n$ and also $\phi \circ \psi = id_{\mathcal{V}_n^{\mathcal{L}}}$ and $\psi \circ \phi = id_{\mathcal{V}_n^{\mathcal{F}}}$ as we wanted. The proof in our case is directly analogous to one given in [Chernov 2003].

11. Proof of Theorem 10.1

Fix a connected component $\mathcal{F}$ of the space of virtual framed curves and a connected component $\mathcal{L}$ of the space of virtual Legendrian curves such that $\mathcal{L} \subset \mathcal{F}$.

In this section we will prove the following theorem:

Theorem 11.1. *Let $x \in \mathcal{V}_n^{\mathcal{L}}$. Suppose (F, K) and (F', L) are two virtual Legendrian knots in $\mathcal{L}$, such that (F, K) is virtually framed isotopic to (F', L). Then $x([(F, K)]_l) = x([(F', L)]_l)$.*

Recall from Section 8 that the virtual Legendrian knot $(F, K^{n,m})$ is obtained by adding n positive and m negative cusp pairs to the diagram of (F, K).

A crucial tool in the proof of Theorem 11.1 is the following lemma:

Lemma 11.2. *Let (F, K) and (F', L) be two virtual Legendrian knots that are isotopic as virtual framed knots, and let $x \in \mathcal{V}_n^{\mathcal{L}}$. Suppose there exists $p \in \mathbb{Z}$ such that $(F, K^{p,p})$ and $(F', L^{p,p})$ are virtually Legendrian isotopic. Then $x([(F, K)]_l) = x([(F', L)]_l)$.*

In order to use the previous lemma we must first show that the integer p in Lemma 11.2 exists whenever (F, K) and (F', L) are in the same connected component of the space of virtual Legendrian curves and are isotopic as virtual framed knots.

To do this, we show that for n_1, n_2 large enough, there exists n_3, n_4 such that (F, K^{n_1,n_2}) and (F', L^{n_3,n_4}) are virtually Legendrian isotopic; this holds without the assumptions that K and L are virtually framed isotopic and are homotopic as virtual Legendrian curves. Then we show that we can assume that $n_1+n_2 = n_3+n_4$ (provided K and L are virtually framed isotopic) and $n_1 - n_2 = n_3 - n_4$ (provided K and L are homotopic as virtual Legendrian curves). It will follow that $n_1 = n_2 = n_3 = n_4$.

Lemma 11.3. *Let (F, K) and (F', L) be virtually isotopic Legendrian knots. Then there exist n_1, n_2, n_3 and n_4 such that (F, K^{n_1,n_2}) is virtually Legendrian isotopic to (F', L^{n_3,n_4}).*

Proof. The proof is the same as the proof of Lemma 8.3, except that because we are now using isotopy rather than homotopy, we do not need to consider crossing changes. □

Theorem 11.4. *Let (F, K) and (F', L) be two virtual Legendrian knots in the same virtual framed isotopy class $\mathcal{F}$. Then given large enough $n_1, n_2, n_3, n_4 \in \mathbb{Z}$ so that (F, K^{n_1,n_2}) and (F', L^{n_3,n_4}) are isotopic as virtual Legendrian knots, we have $n_1 + n_2 = n_3 + n_4$.*

Proof. Throughout the proof below we write K^ν instead of $\overline{K}^\nu$ to denote a framed knot in ST^*F to increase readability. $\overline{K}$ will still denote the lift to ST^*F of a wavefront K.

Since (F, K) and (F', L) are virtual framed isotopic we have a sequence of pairs:

$$(F, K^{st}) = (F_1, K_1^{\nu_1}) \sim_f (F_2, K_2^{\nu_2}) \sim_f \cdots \sim_f (F_m, K_m^{\nu_m}) = (F', L^{st}).$$

More precisely we have surfaces G_i and maps $\phi_i : F_i \to G_i$ and $\psi_i : F_{i+1} \to G_i$ such that $\phi_{i*}(K_i^{\nu_i})$ is framed isotopic to $\psi_{i*}(K_{i+1}^{\nu_{i+1}})$ on G_i via the framed isotopy h_t^i for all $1 \le i < n$.

Now, by the argument in Lemma 11.3 (or rather, Lemma 8.3) we can approximate the previous framed isotopy with a Legendrian isotopy after adding sufficiently many positive and negative cusp pairs to K and L. This yields a sequence

$$(F, K^{n_1,n_2}) = (F_1, L_1) \sim_l (F_2, L_2) \sim_l \cdots \sim_l (F_m, L_m) = (F', L^{n_3,n_4})$$

with the same surfaces G_i and maps ϕ_i and ψ_i as above. However, now we have a sequence of Legendrian isotopies $l_t^i : S^1 \to ST^*G_i$ such that the image of l_0^i is equal to $\phi_{i*}(\bar{L}_i)$, and the image of l_1^i is equal to $\psi_{i*}(\bar{L}_{i+1})$. Furthermore for all $t \in [0, 1]$, the image of l_t^i is contained in a small torus around the image of h_t^i.

We can use the fact that both of these images are contained in a small torus at each time t to show that $n_1 + n_2 = n_3 + n_4$.

Given two homotopic (virtual) framed knots $K_1^{\nu_1}$ and $K_2^{\nu_2}$ in ST^*F which lie in a solid torus T, where T is embedded in ST^*F, we can identify T with the standard torus in $\mathbb{R}^3$. Then we define $\mathrm{slkd}(K_1^{\nu_1}, K_2^{\nu_2})$ to be the difference of the self-linking numbers of the images of $K_1^{\nu_1}$ and $K_2^{\nu_2}$ under this identification. Since both K_1 and K_2 are homotopic to the longitude of the torus, $\mathrm{slkd}(K_1^{\nu_1}, K_2^{\nu_2})$ does not depend on the choice of identification. Furthermore, one can check that $\mathrm{slkd}(\overline{K}, \overline{K}^{n_1,n_2}) = n_1 + n_2$, and $\mathrm{slkd}(\bar{L}, \bar{L}^{n_3,n_4}) = n_3 + n_4$. This argument is similar to an argument in [Chernov 2003].

We now know that slkd does not change as t varies on a fixed surface G_i. Thus $\mathrm{slkd}(\phi_{i*}(K_i^{\nu_i}), \phi_{i*}(\bar{L}_i)) = \mathrm{slkd}(\psi_{i*}(K_{i+1}^{\nu_{i+1}}), \psi_{i*}(\bar{L}_{i+1}))$. So to finish the proof we just need to show the following:

$$\operatorname{slkd}(\psi_{i*}(K_{i+1}^{\nu_{i+1}}), \psi_{i*}(\bar{L}_{i+1})) = \operatorname{slkd}(\phi_{i+1*}(K_{i+1}^{\nu_{i+1}}), \phi_{i+1*}(\bar{L}_{i+1}))$$

However, since slkd does not depend on the identification of the torus in G_i or G_{i+1} with the standard torus in $\mathbb{R}^3$ this equality is clear. □

Theorem 11.5. *Let (F, K) and (F', L) be two virtual Legendrian knots in the same virtual isotopy class that are homotopic as virtual Legendrian curves. If $n_1, n_2, n_3, n_4 \in \mathbb{Z}$ are large enough, so that (F, K^{n_1,n_2}) and (F', L^{n_3,n_4}) are isotopic as virtual Legendrian knots, then $n_1 - n_2 = n_3 - n_4$.*

Proof. From Lemma 11.3 we have $n_1, n_2, n_3, n_4 \in \mathbb{Z}$ so that (F, K^{n_1,n_2}) and (F', L^{n_3,n_4}) are isotopic as virtual Legendrian knots. Since (F, K) and (F', L) are homotopic as virtual Legendrian curves we have that $\mu(F, K) = \mu(F', L)$. For the same reason we get that $\mu(F, K^{n_1,n_2}) = \mu(F', L^{n_3,n_4})$.

Finally, since (F, K^{n_1,n_2}) is obtained from (F, K) by adding n_1 upward cusp pairs and n_2 downward cusp pairs we have $\mu(F, K^{n_1,n_2}) - \mu(F, K) = n_1 - n_2$. Similarly we have $\mu(F', L^{n_3,n_4}) - \mu(F', L) = n_3 - n_4$. So from the equalities above we can conclude that $n_1 - n_2 = n_3 - n_4$ □

In order to prove the final theorem we need the following combinatorial lemma:

Lemma 11.6. *For $0 \le i < p$,*

$$\sum_{k=\lceil \frac{i}{z+1} \rceil}^{p} (-1)^{k+1} \binom{p}{k} \binom{k(z+1)}{i} = 0.$$

Proof. We show this by comparing coefficients of a polynomial.

$$\begin{aligned}
x^p \left(\sum_{k=1}^{z+1} (-1)^k \binom{z+1}{k} x^{k-1} \right)^p &= (1 - (1-x)^{z+1})^p \\
&= \sum_{k=0}^{p} (-1)^k \binom{p}{k} (1-x)^{k(z+1)} \\
&= \sum_{k=0}^{p} (-1)^k \binom{p}{k} \left(\sum_{j=0}^{k(z+1)} (-1)^j \binom{k(z+1)}{j} x^j \right) \\
&= \sum_{j=0}^{p(z+1)} \left(\sum_{k=\lceil \frac{i}{z+1} \rceil}^{p} (-1)^{k+j} \binom{p}{k} \binom{k(z+1)}{j} \right) x^j. \quad \square
\end{aligned}$$

Theorem 11.7. *Let $x \in \mathcal{V}_n^{\mathcal{L}}$, and let $(F, K), (F', L) \in \mathcal{L}$. If there exists p such that $(F, K^{p,p})$ and $(F', L^{p,p})$ are virtual Legendrian isotopic, then $x([(F, K)]_l) = x([(F', L)]_l)$.*

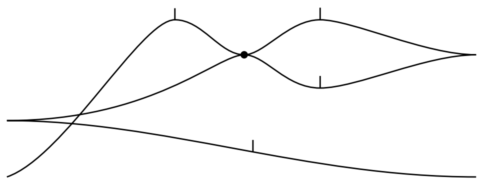

Figure 14. The singular Legendrian arc used to construct K^z.

Proof. Throughout the proof below we drop the $[\,\cdot\,]_l$ notation, with the understanding that x is always an invariant of virtual Legendrian knots. Fix a point q in the image of (F, K) and denote by (F, K^z) the singular virtual Legendrian knot obtained from (F, K) by adding z copies of Figure 14 in a neighborhood of q. For a singular knot (F, K_s) denote by $d(F, K_s)$ the sum of all the signed resolutions of the double points of $\overline{K}_s$, i.e., direct self-tangencies of K_s. The homotopy in Figure 10 implies that $d(F, K^1) = (F, K) - (F, K^{1,1})$. So we have that $x(F, K) = x(F, K) - x(F, K^{1,1})$. By iterating this process we can conclude that $x(F, K^z) = \sum_{j=0}^{z}(-1)^j\binom{z}{j}x(F, K^{j,j})$.

To obtain the result we will begin with $x(F, K)$ and, after adding a linear combination of singular knots on which x is zero, we will change the argument of x to (F', L). First we make some observations that will be used in the computation:

(1) $x(F, K) = x\big(\sum_{k=0}^{p}(-1)^k\binom{p}{k}(F, K^{k(z+1)})\big)$, since for $k > 0$, $(F, K^{k(z+1)})$ has at least $z+1$ double points.

(2) $x(F, K^{m,m}) = x(F', L^{m,m})$ for $m \geq p$ since one can change $(F, K^{p,p})$ to $(F', L^{p,p})$ through a virtual Legendrian isotopy.

(3) $x(F, K^z) = \sum_{j=0}^{z}(-1)^j\binom{z}{j}x(F, K^{j,j})$, as noted above.

Now, calculate:

$$
\begin{aligned}
x(F, K) &= \sum_{k=0}^{p}(-1)^k\binom{p}{k}x(F, K^{k(z+1)}) && \text{by (1) above}\\
&= \sum_{k=0}^{p}(-1)^k\binom{p}{k}\Bigg(\sum_{j=0}^{k(z+1)}(-1)^j\binom{k(z+1)}{j}x(F, K^{j,j})\Bigg) && \text{by (3)}\\
&= \sum_{j=0}^{p(z+1)}\Bigg(\sum_{k=\lceil\frac{i}{z+1}\rceil}^{p}(-1)^{k+j}\binom{p}{k}\binom{k(z+1)}{j}\Bigg)x(F, K^{j,j})\\
&= \sum_{j=p}^{p(z+1)}\Bigg(\sum_{k=\lceil\frac{i}{z+1}\rceil}^{p}(-1)^{k+j}\binom{p}{k}\binom{k(z+1)}{j}\Bigg)x(F, K^{j,j}) && \text{by 11.6}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=p}^{p(z+1)} \Bigg(\sum_{k=\lceil \frac{i}{z+1} \rceil}^{p} (-1)^{k+j} \binom{p}{k} \binom{k(z+1)}{j} \Bigg) x(F', L^{j,j}) && \text{by (2)} \\
&= \sum_{j=0}^{p(z+1)} \Bigg(\sum_{k=\lceil \frac{i}{z+1} \rceil}^{p} (-1)^{k+j} \binom{p}{k} \binom{k(z+1)}{j} \Bigg) x(F', L^{j,j}) \\
&= \sum_{k=0}^{p} (-1)^{k} \binom{p}{k} \Bigg(\sum_{j=0}^{k(z+1)} (-1)^{j} \binom{k(z+1)}{j} x(F', L^{j,j}) \Bigg) \\
&= \sum_{k=0}^{p} (-1)^{k} \binom{p}{k} x(F', K_2^{k(z+1)}) && \text{by (3)} \\
&= x(F', L). && \square
\end{aligned}
$$

Now by Theorems 11.4, 11.5 and 11.7 we can conclude that Vassiliev invariants cannot distinguish virtually framed isotopic virtual Legendrian knots with the same Maslov number.

Theorem 11.8. *Let $x \in \mathcal{V}_n^{\mathcal{L}}$, (F, K), $(F', L) \in \mathcal{L}$ and let (F, K) be virtually framed isotopic to (F', L) then $x([F, K)]_l) = x([(F', L)]_l)$.*

Acknowledgment

We would like to thank our advisor, Vladimir Chernov, for suggesting the problem and for many valuable discussions.

References

[Arnold 1989] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Graduate Texts in Mathematics **60**, Springer, New York, 1989. MR 90c:58046 Zbl 0386.70001

[Arnold 1994] V. I. Arnold, *Topological invariants of plane curves and caustics*, University Lecture Series **5**, Amer. Math. Soc., Providence, RI, 1994. MR 95h:57003 Zbl 0858.57001

[Arnold 2004] V. I. Arnold, *Lectures on partial differential equations*, Springer, Berlin, 2004. Translated from the second Russian edition by Roger Cooke. MR 2004j:35002 Zbl 1076.35002

[Carter et al. 2002] J. S. Carter, S. Kamada, and M. Saito, "Stable equivalence of knots on surfaces and virtual knot cobordisms", *J. Knot Theory Ramifications* **11**:3 (2002), 311–322. MR 2003f:57011 Zbl 1004.57007

[Chernov 2003] V. Tchernov, "Vassiliev invariants of Legendrian, transverse, and framed knots in contact three-manifolds", *Topology* **42**:1 (2003), 1–33. MR 2003f:57029 Zbl 1017.57004

[Fuchs and Tabachnikov 1997] D. Fuchs and S. Tabachnikov, "Invariants of Legendrian and transverse knots in the standard contact space", *Topology* **36**:5 (1997), 1025–1053. MR 99a:57006 Zbl 0904.57006

[Goryunov 1997] V. Goryunov, "Finite order invariants of framed knots in a solid torus and in Arnold's J^+-theory of plane curves", pp. 549–556 in *Geometry and physics* (Aarhus, 1995), edited by J. E. Andersen et al., Lecture Notes in Pure and Appl. Math. **184**, Dekker, New York, 1997. MR 99h:57007 Zbl 0871.57007

[Goussarov et al. 2000] M. Goussarov, M. Polyak, and O. Viro, "Finite-type invariants of classical and virtual knots", *Topology* **39**:5 (2000), 1045–1068. MR 2001i:57017 Zbl 1006.57005

[Hill 1997] J. W. Hill, "Vassiliev-type invariants of planar fronts without dangerous self-tangencies", *C. R. Acad. Sci. Paris Sér. I Math.* **324**:5 (1997), 537–542. MR 98c:57008 Zbl 0881.57003

[Kauffman 1999] L. H. Kauffman, "Virtual knot theory", *European J. Combin.* **20**:7 (1999), 663–690. MR 2000i:57011 Zbl 0938.57006

[Kuperberg 2003] G. Kuperberg, "What is a virtual link?", *Algebr. Geom. Topol.* **3** (2003), 587–591. MR 2004f:57012 Zbl 1031.57010

[Polyak 1998] M. Polyak, "Invariants of curves and fronts via Gauss diagrams", *Topology* **37**:5 (1998), 989–1009. MR 2000a:57080 Zbl 0961.57015

[Turaev 2004] V. Turaev, "Virtual strings", *Ann. Inst. Fourier* (*Grenoble*) **54**:7 (2004), 2455–2525. MR 2006f:57016 Zbl 1066.57022

Received May 28, 2013. Revised September 15, 2013.

PATRICIA CAHN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PENNSYLVANIA
DAVID RITTENHOUSE LAB
209 SOUTH 33RD STREET
PHILADELPHIA, PA 19104-6395
UNITED STATES
pcahn@math.upenn.edu

ASA LEVI
DEPARTMENT OF MATHEMATICS
DARTMOUTH COLLEGE
6188 KEMENY HALL
HANOVER, PA 03755-3551
UNITED STATES
Asa.Levi.gr@dartmouth.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.47

SOME RESULTS ON THE GENERIC VANISHING OF KOSZUL COHOMOLOGY VIA DEFORMATION THEORY

JIE WANG

We study the deformation-obstruction theory of Koszul cohomology groups of g^r_d's on singular nodal curves. We compute the obstruction classes for Koszul cohomology classes on singular curves to deform to a smooth one. In the case where the obstructions are nontrivial, we obtain some partial results for generic vanishing of Koszul cohomology groups.

1. Introduction

In this paper, we apply deformation theory to study the syzygies of general curves in $\mathbb{P}^r$ with fixed genus and degree. Let L be a basepoint-free g^r_d on a smooth curve X. The Koszul cohomology group $K_{p,q}(X, L)$ is the cohomology of the Koszul complex at the (p, q)-spot

$$\longrightarrow \bigwedge^{p+1} H^0(L) \otimes H^0(X, L^{q-1}) \xrightarrow{d_{p+1,q-1}} \bigwedge^p H^0(L) \otimes H^0(X, L^q) \xrightarrow{d_{p,q}} H^0(X, L^{q+1}),$$

where

$$d_{p,q}(v_1 \wedge \cdots \wedge v_p \otimes \sigma) = \sum_i (-1)^i v_1 \wedge \cdots \wedge \hat{v}_i \wedge .. \wedge v_p \otimes v_i \sigma.$$

The Koszul cohomology groups $K_{p,q}(X, L)$ completely determine the shape of a minimal free resolution of the section ring

$$R = R(X, L) = \bigoplus_{k \geq 0} H^0(X, L^k),$$

and therefore carry a lot of information about the extrinsic geometry of X.

We are interested in Green's question:

Problem 1.1. *What is the variational theory of the $K_{p,q}(X, L)$? What do they look like for X a general curve and L a general g^r_d?*

MSC2010: 14H51.

Keywords: Koszul cohomology, general curves, deformation theory, generic vanishing, maximal rank conjecture.

If (X, L) is general in $\mathcal{G}^r_{g,d}$ (in this paper, this means the Brill–Noether number $\rho = g - (r+1)(g-d+r)$ is nonnegative and (X, L) is a general point of the unique component of $\mathcal{G}^r_{g,d}$ which dominates $\mathcal{M}_g$), it is well known that we only have to determine $K_{p,1}(X, L)$, or equivalently $K_{p-1,2}(X, L)$ for $1 \le p \le r-1$ (see Section 2).

Problem 1.1 seems too difficult to answer in full generality. For an arbitrary g^r_d on a general curve X, the simplest cases — determining $K_{1,1}(X, L)$ or $K_{0,2}(X, L)$ — are still unknown. The maximal rank conjecture (MRC) [Eisenbud and Harris 1983] predicts that the multiplication map

$$\mathrm{Sym}^2 H^0(X, L) \xrightarrow{\mu} H^0(X, L^2) \tag{1-1}$$

is either injective or surjective, that is,

$$\min\{k_{1,1}(X, L), k_{0,2}(X, L)\} = 0. \tag{1-2}$$

Geometrically, this means that the number of quadrics in $\mathbb{P}^r$ containing X is as simple as the Hilbert function of $X \subset \mathbb{P}^r$ allows.

There are many partial results about (1-2) using the so-called "méthode d'Horace", originally proposed by Hirschowitz. This amounts to a degeneration argument to a carefully chosen singular curve in projective space and a proof of the statement on such a curve by a delicate inductive argument. We refer to, for instance, [Ballico and Fontanari 2010a; 2010b] for some recent results in this direction.

For higher syzygies, again there are many results (see [Aprodu 2004; Ballico 1996; Ein 1987; Farkas 2009]). One breakthrough result is Voisin's solution [2002; 2005] of the generic Green's conjecture, which solves Problem 1.1 for the case $L = K_X$.

For the vanishing of $K_{p,1}$, there is [Aprodu 2004; 2005], which proved the generic version of the Green–Lazarsfeld gonanity conjecture. This conjecture predicts that, for smooth curve X of gonanity d and L a sufficiently positive line bundle on X,

$$K_{h^0(L)-d,1}(X, L) = 0.$$

Note that Problem 1.1 does not have any assumption on the positivity of L.

It seems that the method of all of the above results amounts to degenerating to special curves, often a carefully chosen singular one, and verifying the statements on these special curves. Given the fact that sometimes such special curves are difficult to find and the inductive arguments can get technical, we would like to take a slightly different point of view. We will consider one-parameter degeneration to the simplest possible singular curve, namely, the union of two smooth curves meeting at a node. Of course, there is no hope of directly verifying the vanishing statements we would like to prove on these curves (see Section 3), but we are able to compute the obstructions for the "extra" Koszul classes of the singular fiber to deform to nearby fibers. If one could prove these "extra" Koszul classes are obstructed, we conclude the general fiber has the vanishing property we need. We feel this point of view has a good chance to generalize.

More precisely, let the property $\mathrm{GV}(p)^r_{g,d}$ mean that, for general $L' = g^r_d$ on general curve C of genus g, we have

$$\min\{k_{p,1}(C, L'), k_{p-1,2}(C, L')\} = 0. \tag{1-3}$$

Problem 1.2. *Does* $\mathrm{GV}(p)^r_{g,d}$ *imply* $\mathrm{GV}(p)^r_{g+1,d+1}$?

If this holds, one could set up an inductive argument. At each step r is fixed and g, d go up by 1, or equivalently, r and h^1 are fixed, and g goes up by 1.

In the case $p = 1$, the maximal rank conjecture predicts the answer should always be affirmative. For higher syzygies, this is not always the case, but one would like to prove some generic vanishing results for some special $\{g, r, d\}$.

We give a simple condition to guarantee $\mathrm{GV}(p)^r_{g,d}$ implies $\mathrm{GV}(p)^r_{g+1,d+1}$ from a deformation-theoretic point of view. We study the deformation theory of Koszul cohomology groups on the simplest kind of singular curve X_0: the union of a general curve C of genus g and an elliptic curve E meeting at a node u. L_0 is carefully chosen (see Section 3) such that:

(a) (X_0, L_0) is smoothable to $L_t = g^r_{d+1}$ on a smooth curve X_t of genus $g + 1$.

(b) $L_0|_C = L'$ and therefore $\min\{k_{p,1}(C, L_0|_C), k_{p-1,2}(C, L_0|_C)\} = 0$.

(c) $L_0|_E = \mathcal{O}_E(v)$ for another general point $v \in E$.

We prove:

Theorem 1.3. *Let* $C \subset \mathbb{P}^r$ *be a general curve*, $|L'|$ *a general* g^r_d *on* C *and* $M_{L'}$ *the kernel bundle defined by the sequence*

$$0 \longrightarrow M_{L'} \longrightarrow H^0(L') \otimes \mathcal{O}_C \xrightarrow{\mathrm{ev}} L' \longrightarrow 0.$$

Then the following hold:

(a) *If* $K_{p,1}(C, L') = 0$, *then* $K_{p,1}(X_t, L_t) = 0$.

(b) *If* $K_{p-1,2}(C, L') = 0$ *and*

$$h^0\big(C, \textstyle\bigwedge^{r-p} M_{L'} \otimes K_C\big) = h^0\big(C, \textstyle\bigwedge^{r-p} M_{L'} \otimes K_C(2u)\big) \tag{1-4}$$

for a general point $u \in C$, *then* $K_{p-1,2}(X_t, L_t) = 0$.

In other words, $\mathrm{GV}(p)^r_{g,d}$ *always implies* $\mathrm{GV}(p)^r_{g+1,d+1}$ *if* (1-4) *holds.*

The upshot is that under such degeneration, we could explicitly compute generators of $K_{p,q}(X_0, L_0)$. Unfortunately (X_0, L_0) does not satisfy (1-3). However we could compute the obstructions for the "extra" Koszul classes to deform to $K_{p,q}(X_t, L_t)$. If every "extra" Koszul class is obstructed, we conclude that (1-3) holds for (X_t, L_t). Condition (1-4) is a sufficient condition for the "extra" Koszul classes to be obstructed.

In the case $p = 1$ (maximal rank conjecture), this sufficient condition turns out to be very geometric.

Theorem 1.4. *Let $C \subset \mathbb{P}^r$ be a general curve embedded by $|L'|$ a general g^r_d and suppose one of the following two conditions holds*:

(a) *μ in* (1-1) *is injective.*

(b) *μ is surjective and there exists a quadric $Q \in \operatorname{Ker}(\mu)$ containing C but not containing the tangential variety $TC := \bigcup_{u\in C} T_u C$.*

Then $(\mathrm{MRC})^r_{g+1,d+1}$ *holds as well.*

To apply Theorem 1.4 to the maximal rank conjecture, one has to verify a hypothesis in (b) which seems geometrically interesting in its own right. Hopefully there will be some other applications.

Starting from the fact that rational normal curves and canonical curves are projectively normal, we verify hypothesis (b) in some special cases and get some partial results.

Corollary 1.5. *Let (X, L) be a general pair in $\mathcal{G}^r_{g,d}$ with $h^1(L) \le 1$. Suppose*

$$d > \tfrac{5}{4}g + \tfrac{9}{4} \quad \text{if } h^1(L) = 0, \text{ or}$$
$$d > \tfrac{5}{4}g + \tfrac{3}{4} \quad \text{if } h^1(L) = 1;$$

then (X, L) is projectively normal.

It is a very well-known result of Green and Lazarsfeld [1986] that any very ample line bundle L on X with

$$\deg(L) \ge 2g_X + 1 - 2h^1(L) - \operatorname{Cliff}(X) \tag{1-5}$$

is projectively normal, and the bound is sharp. Notice that (1-5) implies that $h^1(L) \le 1$.

If X is general,

$$\operatorname{Cliff}(X) = \left\lfloor \frac{g_X - 1}{2} \right\rfloor;$$

thus, the Green–Lazarsfeld theorem predicts projective normality for general curves if d is bigger than roughly $3g/2$. Corollary 1.5 thus says that if L is also general, we could improve the lower bound of d to roughly $5g/4$.

The bounds in Corollary 1.5 are weaker than the bounds in [Ballico and Fontanari 2010b].

We could also fix a small r and let h^1 be arbitrarily large.

Corollary 1.6. *The maximal rank conjecture* (*for quadrics*) *holds if $r \le 4$.*

The reason we can get rid of the restriction on the degree of the line bundle for small r is that we can always verify the hypothesis on TC in Theorem 1.4(b) if $r \leq 4$. Thus $(\mathrm{MRC})^r_{g,d}$ always implies $(\mathrm{MRC})^r_{g+1,d+1}$.

For higher syzygies, we do not expect analogously that $\min\{k_{p,1}, k_{p-1,2}\} = 0$ for $p \geq 2$. We refer the audience to Section 2 for a counterexample. Nevertheless, we do wish to obtain certain vanishing results or effective upper bounds on $k_{p,q}$.

The difficulty in generalizing the inductive argument to higher syzygies is twofold. First there are relatively few known cases to start the induction with. There is essentially a single known starting series of examples for vanishing of syzygies, namely, Voisin's solution to the generic Green conjecture. Besides Voisin's theorem, Farkas [2006] proved that properties $\mathrm{GV}(2)^7_{16,21}$ and $\mathrm{GV}(3)^{10}_{22,30}$ hold. Secondly, for higher syzygies, the sufficient condition for "extra" Koszul classes to be obstructed is not as geometric.

Nevertheless we summarize our results on higher syzygies as follows:

Theorem 1.7. *Let X be a general curves of genus g and L a general g^r_d on X. Then*:

(a) *If $g \geq r+1$, then $K_{p,1}(X, L) = 0$ for $p \geq \lfloor (r+1)/2 \rfloor$.*

(b) *If $h^1(L) = 1$ (which implies that $g \geq r+1$), then*

$$\begin{aligned} K_{p-1,2}(X, L) &= 0 && \text{for } 1 \leq p \leq r - \left\lfloor \frac{g}{2} \right\rfloor, \\ k_{p-1,2}(X, L) &\leq (g - 2r + 2p - 1)\binom{r-1}{p-1} && \text{for } p > r - \left\lfloor \frac{g}{2} \right\rfloor. \end{aligned}$$

Combining Corollaries 1.5 and 1.7(a), we can determine $k_{p,q}(X, L)$ for L a general g^r_d with $r \leq 4$.

Corollary 1.8. *For a general pair (X, L) in $\mathcal{G}^r_{g,d}$ with $r \leq 4$, $g \geq r+1$, we have*

$$\min\{k_{p,1}(X, L), k_{p-1,2}(X, L)\} = 0.$$

The organization of this paper is as follows. In Section 2, we review some basic facts about Koszul cohomology of general curves. In Section 3, we study the Koszul cohomology of the central fiber (X_0, L_0). We explicitly write down the generators of the "extra" Koszul classes in $K_{r-p,0}(X_0, L_0; \omega_{X_0}) \cong K_{p-1,2}(X_0, L_0)^\vee$. Section 4 contains a computation of the obstructions for these classes to deform, and Section 5 gives a sufficient condition for the obstruction classes to be linearly independent, and a proof of Theorem 1.3. In Section 6, we focus on the $p = 1$ case and prove Theorem 1.4 and Corollaries 1.5 and 1.8. Finally, in Section 7, we consider higher syzygies for line bundles with $h^1 = 1$. In some special range of p, we are able to prove some vanishing results as in Theorem 1.7.

2. Koszul cohomology of general curves

We first summarize several special properties of Koszul cohomology groups on general curves over $\mathbb{C}$. We refer to [Aprodu and Nagel 2010] and [Eisenbud 1992] for general facts about Koszul cohomology.

Proposition 2.1. *Suppose X is a general curve and L is a complete g^r_d on X.*

(a) $K_{p,0}(X, L) = 0$ *except when* $p = 0$ *and* $k_{0,0}(X, L) = 1$.

(b) $K_{p,q}(X, L) = 0$ *for* $q \geq 4$.

(c) $K_{p,3}(X, L) = 0$ *except when* $p = r - 1$ *and* $k_{r-1,3}(X, L) = h^1(L)$.

Proof. Statement (a) follows from the definition of Koszul cohomology.

To prove (b) and (c), we use the following facts:

(i) The multiplication map

$$H^0(X, L) \otimes H^0(X, K_X \otimes L^{-1}) \longrightarrow H^0(X, K_X)$$

is injective. This is the Gieseker–Petri theorem.

(ii) $H^0(X, K_X \otimes L^{-2}) = 0$. This is a direct consequence of (i) (see [Arbarello and Cornalba 1981]).

Statement (b) follows from (ii) and the duality theorem of Koszul cohomology (see [Aprodu and Nagel 2010, Section 2.3]):

$$K_{p,q}(X, L) = K_{r-1-p,2-q}(X, L; K_X)^\vee. \tag{2-1}$$

To prove (c), we first apply (2-1) and note that the Koszul differential $d_{r-1-p,-1}$ factors as

$$\begin{tikzcd}
\bigwedge^{r-1-p} H^0(L) \otimes H^0(K_X \otimes L^{-1}) \arrow[r, "d_{r-1-p,-1}"] \arrow[d, "\lrcorner\otimes \mathrm{Id}"'] & \bigwedge^{r-2-p} H^0(L) \otimes H^0(K_X) \\
\bigwedge^{r-2-p} H^0(L) \otimes H^0(L) \otimes H^0(K_X \otimes L^{-1}) \arrow[ur, "\mathrm{Id}\otimes\mu"'] &
\end{tikzcd}$$

By (i), both $\lrcorner \otimes \mathrm{Id}$ and $\mathrm{Id} \otimes \mu$ are injective. □

As a consequence, we have the following corollary:

Corollary 2.2. *Let X be a general curve and L a globally generated g^r_d with $r \geq 1$.*

(a) *L is normally generated if and only if the multiplication map*

$$\mu : S^2 H^0(X, L) \longrightarrow H^0(X, L^2)$$

is surjective.

(b) *If L is normally generated, the homogeneous ideal I_X is generated by quadrics and cubics.*

Proof. The only possible nonzero $K_{0,q}$ for $q \geq 2$ is $K_{0,2}(X, L) = \operatorname{Coker}(\mu)$. If $K_{0,2}(X, L) = 0$, L is normally generated. Since $k_{1,q}$ is the number of minimal generators of I_X of degree $q+1$, (b) follows. □

Moreover, since taking cohomology does not change the Euler characteristic of the complex, we have for any $1 \leq p \leq r-1$ that

$$\begin{aligned} &k_{p,1}(X, L) - k_{p-1,2}(X, L) \\ &\qquad = \sum_{i+j=p+1} (-1)^{j+1} \dim_{\mathbb{C}}\Bigl(\bigl(\textstyle\bigwedge^i V\bigr) \otimes H^0(X, L^j)\Bigr) \\ &\qquad = \binom{r+1}{p}(g-d+r) - \binom{r+1}{p+1}g + \binom{r-1}{p}d + \binom{r}{p+1}(g-1). \end{aligned}$$

Denote this number by $b_p(X, L)$, which depends only on g, r, d, p. Therefore to determine the Koszul cohomology of (X, L), it suffices to determine either row $q = 1$ or $q = 2$.

Remark. Based on the maximal rank conjecture, one might expect that analogously

$$\min\{k_{p,1}(X, L), k_{p-1,2}(X, L)\} = 0 \tag{2-2}$$

for general (X, L). But this is not the case. In fact, F. Schreyer proved in his thesis (see [Green 1984, 4.a.2] for more details) that for any curve X of genus g, there exists a number d_0 such that if $\deg(L) = d \geq d_0$, then

$$K_{p,2}(X, L) \neq 0 \quad \text{if } r-1 \geq p \geq r-g.$$

On the other hand, it follows from a theorem of Green and Lazarsfeld [1984] (see also [Aprodu and Nagel 2010, Corollary 3.39]) that for d large,

$$K_{p,1}(X, L) \neq 0 \quad \text{if } 1 \leq p \leq r - \left\lfloor \frac{g}{2} \right\rfloor - 2.$$

Thus, for $r-g+1 \leq p \leq r - \lfloor g/2 \rfloor - 2$, (2-2) does not hold.

3. Koszul cohomology of the central fiber

Let L' be a g^r_d on a smooth curve C of genus g and $X_0 = C \cup E$ the reducible nodal curve consisting of C and an elliptic curve E meeting at a general point u. Let L_0 be the line bundle on X_0 such that

$$L_0|_C = L'$$

and

$$L_0|_E = \mathcal{O}_E(v),$$

where $v \neq u$. We would like to study the relations between $K_{p,q}(C, L')$ and $K_{p,q}(X_0, L_0)$ in this section.

First, observe that by construction any (global) section of L' on C extends uniquely to a section of L_0 on X_0; thus we have a natural isomorphism

$$\phi : H^0(C, L') \cong H^0(X_0, L_0).$$

Moreover, by the Riemann–Roch theorem, $h^1(C, L') = h^1(X_0, L_0)$, and there is a natural identification

$$H^0(C, K_C \otimes L'^{-1}) \cong H^0(X_0, \omega_{X_0} \otimes L_0^{-1}).$$

A first consequence is:

Proposition 3.1. *If $K_{p,1}(C, L') = 0$, then $K_{p,1}(X_0, L_0) = 0$.*

Proof. Consider the commutative diagram

$$\begin{tikzcd}
\bigwedge^{p+1} H^0(L_0) \arrow[r] \arrow[d, "\cong"] & \bigwedge^{p} H^0(L_0) \otimes H^0(L_0) \arrow[r] \arrow[d, "\cong"] & \bigwedge^{p-1} H^0(L_0) \otimes H^0(L_0^2) \arrow[d] \\
\bigwedge^{p+1} H^0(L') \arrow[r] & \bigwedge^{p} H^0(L') \otimes H^0(L') \arrow[r] & \bigwedge^{p-1} H^0(L') \otimes H^0(L'^2)
\end{tikzcd}$$

where the vertical arrows are restriction maps to C. The hypothesis says that the lower row is exact in the middle. A simple diagram chase gives the conclusion. □

The argument in Proposition 3.1 does not generalize to the case $q = 2$ because $H^0(C, L'^2)$ is not isomorphic to $H^0(X_0, L_0^2)$. Instead we dualize using (2-1),

$$K_{p-1,2}(C, L')^\vee \cong K_{r-p,0}(C, L'; K_C),$$

and compare $K_{r-p,0}(C, L'; K_C)$ with $K_{r-p,0}(X_0, L_0; \omega_{X_0})$.

Here, ω_{X_0} is the dualizing sheaf of X_0. Its restriction to C and E are line bundles $K_C(u)$ and $K_E(u)$ respectively. A global section of the dualizing sheaf consists of (global) one-forms on C and E, viewed as sections of $K_C(u)$ and $K_E(u)$ respectively which vanish at u.

Figure 1 describes the various line bundles in question on X_0 and their restrictions to each component. The S-shaped curve is C and the straight line is E.

Choose a basis $\{\omega_0, \dots, \omega_{g-1}\}$ of $H^0(C, K_C)$ and a basis $\{\omega_g\}$ of $H^0(E, K_E)$. For $0 \le i \le g-1$, we will think of ω_i as a section in $H^0(K_C(u))$ which vanishes on u, and then extend it over E by the zero section. Such a section belongs to $H^0(\omega_{X_0})$, and we still denote it by ω_i. Similarly, we obtain $\omega_g \in H^0(\omega_{X_0})$ with $\omega_g|_C = 0$.

In this way, we obtain a natural identification

$$\psi : H^0(C, K_C) \oplus H^0(E, K_E) \cong H^0(C, K_C(u)) \oplus H^0(E, K_E(u)) \cong H^0(X_0, \omega_{X_0})$$

and

$$H^0(X_0, \omega_{X_0}) = \operatorname{span}\{\omega_i \mid i = 1, \dots, g\}.$$

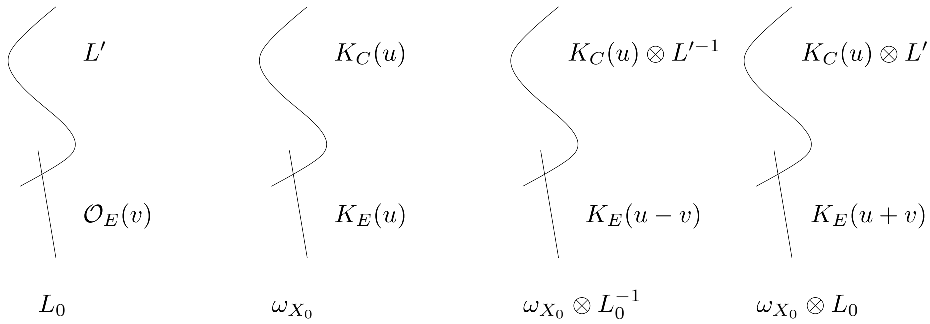

Figure 1. The line bundles on the central fiber.

Notice also that every section in $H^0(X_0, \omega_{X_0})$ vanishes at u.

Now suppose $K_{r-p,0}(C, L'; K_C)=0$. We want to show that $K_{r-p,0}(X_0, L_0; \omega_{X_0})$ can be generated by pure tensors in

$$\bigwedge^{r-p} H^0(X_0, L_0)\otimes H^0(X_0, \omega_{X_0}).$$

To this end, consider the commutative diagram

$$\begin{tikzcd}
\bigwedge^{r-p+1} H^0(L')\otimes H^0(K_C\otimes L'^{-1}) \arrow[r, "\cong"] \arrow[d] & \bigwedge^{r-p+1} H^0(L_0)\otimes H^0(\omega_{X_0}\otimes L_0^{-1}) \arrow[d, "\delta_{-1}"] \\
\bigwedge^{r-p} H^0(L')\otimes H^0(K_C) \arrow[r, "\phi\otimes\psi"] \arrow[d] & \bigwedge^{r-p} H^0(L_0)\otimes H^0(\omega_{X_0}) \arrow[d, "\delta_0"] \\
\bigwedge^{r-p-1} H^0(L')\otimes H^0(K_C\otimes L') \arrow[r] & \bigwedge^{r-p-1} H^0(L_0)\otimes H^0(\omega_{X_0}\otimes L_0).
\end{tikzcd}$$

The top horizontal arrow is an isomorphism since any section in $H^0(X_0, \omega_{X_0}\otimes L_0^{-1})$ restricts to zero on the E component.

Now let $\{\sigma_0, \dots, \sigma_r\}$ be a basis of $H^0(C, L')$. Extend each σ_k uniquely to X_0 to form a basis of $H^0(X_0, L_0)$, still denoting them by σ_k.

We can write any element in $\mathrm{Ker}(\delta_0)$ as

$$\sum_{\substack{k_1,\dots,k_{r-p}\\ j\le g-1}} \alpha_{k_1,\dots,k_{r-p},j}\, \sigma_{k_1}\wedge\cdots\wedge\sigma_{k_{r-p}}\otimes\omega_j + \sum_{k_1,\dots,k_{r-p}} \beta_{k_1,\dots,k_{r-p}}\sigma_{k_1}\wedge\cdots\wedge\sigma_{k_{r-p}}\otimes\omega_g.$$

Since the image under δ_0 of the second term β restricts to 0 on C (since ω_g does), so does the image of the first term α. By our assumption, the left column of the above diagram is exact in the middle and therefore $\alpha\in\mathrm{Im}(\delta_{-1})$.

We conclude that

$$\sum_{k_1,\dots,k_{r-p}} \beta_{k_1,\dots,k_{r-p}} \sigma_{k_1} \wedge \cdots \wedge \sigma_{k_{r-p}} \otimes \omega_g \in \operatorname{Ker}(\delta_0),$$

and this can happen only if

$$\sum_{k_1,\dots,k_{r-p}} \beta_{k_1,\dots,k_{r-p}} \sigma_{k_1} \wedge \cdots \wedge \sigma_{k_{r-p}} \in \bigwedge^{r-p} V,$$

where $V \subset H^0(X_0, L_0)$ is the codimension-one subspace consisting of sections which restrict to zero on E. Also, it is easy to see that a basis of

$$\bigwedge^{r-p} V \otimes \mathbb{C} \cdot \omega_g$$

is linearly independent even modulo $\operatorname{Im}(\delta_{-1})$.

We have proved:

Lemma 3.2. *If* $K_{r-p,0}(C, L'; K_C) = 0$, *we have an isomorphism*

$$K_{r-p,0}(X_0, L_0; \omega_{X_0}) \xrightarrow{\cong} \bigwedge^{r-p} V \otimes \mathbb{C} \cdot \omega_g.$$

4. Infinitesimal calculations

In this section, we carry out the computation of first-order obstructions described in the introduction. We will use the deformation theory of complexes, which was developed in [Green and Lazarsfeld 1987]. The general set-up is as below.

Let S be a smooth variety and $F^\bullet$ a bounded complex of locally free sheaves on S:

$$\cdots \longrightarrow F^{p+1} \xrightarrow{d_{p+1}} F^p \xrightarrow{d_p} F^{p-1} \longrightarrow \cdots.$$

Given a point $t \in S$, denote by $F^\bullet(t)$ the complex of vector spaces at t determined by the fibers of $F^\bullet$; that is,

$$F^\bullet(t) = F^\bullet \otimes \mathbb{C}(t),$$

where $\mathbb{C}(t)$ is the residue field of S at t.

The deformation theory of $H^i(F^\bullet(t))$ as t moves near $0 \in S$ is controlled by the derivative complex, which associates to a tangent vector $v \in T_0S$ a complex

$$\cdots \longrightarrow H^{p+1}(F^\bullet(0)) \xrightarrow{D_v(d_{p+1})} H^p(F^\bullet(0)) \xrightarrow{D_v(d_p)} H^{p-1}(F^\bullet(0)) \longrightarrow \cdots.$$

A (co)homology class $[c] \in H^p(F^\bullet(0))$ deforms to first order along v if and only if $D_v(d_p)([c]) = 0 \in H^{p-1}(F^\bullet(0))$.

To describe the $D_v(d_p)$, recall that a tangent vector $v \in T_0S$ corresponds to an embedding of the dual numbers D into S, so one gets a short exact sequence

$$0 \longrightarrow \mathbb{C}(0) \longrightarrow D \longrightarrow \mathbb{C}(0) \longrightarrow 0.$$

Tensoring the sequence with $F^\bullet$ yields a short exact sequence of complexes, which in turn gives rise to connecting homomorphisms

$$\begin{tikzcd} H^p(F^\bullet \otimes \mathbb{C}(0)) \arrow[r, "D_v(d_p)"] \arrow[d, equal] & H^{p-1}(F^\bullet \otimes \mathbb{C}(0)) \arrow[d, equal] \\ H^p(F^\bullet(0)) & H^{p-1}(F^\bullet(0)) \end{tikzcd}$$

One checks that $D_v(d_p) \circ D_v(d_{p+1}) = 0$.

Now let (X_0, L_0) be the pair constructed in the previous section. We will further assume that both (C, L') and the crossing point u are general. (X_0, L_0) determines a limit linear series in the sense of [Eisenbud and Harris 1986]. By counting Brill–Noether numbers, it is easy to see this limit linear series is deformable to general pairs (X_t, L_t). Let $\mathcal{L} \to \mathcal{X} \to \Delta$ be the total space of a one-parameter family of general pairs $(X_t, L_t) \in \mathcal{G}^r_{g+1,d+1}$ degenerating to (X_0, L_0). We will apply the deformation theory described above to the Koszul complex computing $K_{r-p,0}(X_t, L_t; \omega_{X_t})$:

$$\textstyle\bigwedge^{r-p+1} H^0(L_t) \otimes H^0(\omega_{X_t} \otimes L_t^{-1}) \hookrightarrow \bigwedge^{r-p} H^0(L_t) \otimes H^0(\omega_{X_t}) \\ \xrightarrow{\delta_t} \textstyle\bigwedge^{r-p-1} H^0(L_t) \otimes H^0(\omega_{X_t} \otimes L_t).$$

By the Gieseker–Petri theorem, the left arrow is injective for all t (even at time zero), so $k_{r-p,0}(X_t, L_t; \omega_{X_t})$ can only increase at $t = 0$ if $\mathrm{Ker}(\delta_t)$ does. We would like to compute the derivative of δ_t at $t = 0$:

$$K_{r-p,0}(X_0, L_0; \omega_{X_0}) \xrightarrow{D(\delta_t)|_{t=0}} K_{r-p-1,1}(X_0, L_0; \omega_{X_0}). \tag{4-1}$$

To illustrate the idea, let us first take a look at the simpler case when $p = r - 1$. (On the other hand, the main case we are interested in is the case $p = 1$.) The general case is just notationally more complicated. In this special case, the Koszul differential δ_t becomes the multiplication map μ_t

$$\textstyle\bigwedge^2 H^0(L_t) \otimes H^0(\omega_{X_t} \otimes L_t^{-1}) \hookrightarrow H^0(L_t) \otimes H^0(\omega_{X_t}) \xrightarrow{\mu_t} H^0(\omega_{X_t} \otimes L_t),$$

and the derivative map is

$$K_{1,0}(X_0, L_0; \omega_{X_0}) \xrightarrow{D(\mu_t)|_{t=0}} K_{0,1}(X_0, L_0; \omega_{X_0}).$$

For simplicity, denote $\omega_g \in H^0(X_0, \omega_{X_0})$ by ω. If $K_{1,0}(C, L', K_C) = 0$, then by Lemma 3.2 we have

$$K_{1,0}(X_0, L_0; \omega_{X_0}) \xrightarrow{\cong} V \otimes \mathbb{C} \cdot \omega \subset V \otimes H^0(X_0, \omega_{X_0}).$$

Remark. Even if $K_{1,0}(C, L', K_C) \neq 0$, we nevertheless have the containment $V \otimes \mathbb{C} \cdot \omega \subset K_{1,0}(X_0, L_0; \omega_{X_0})$.

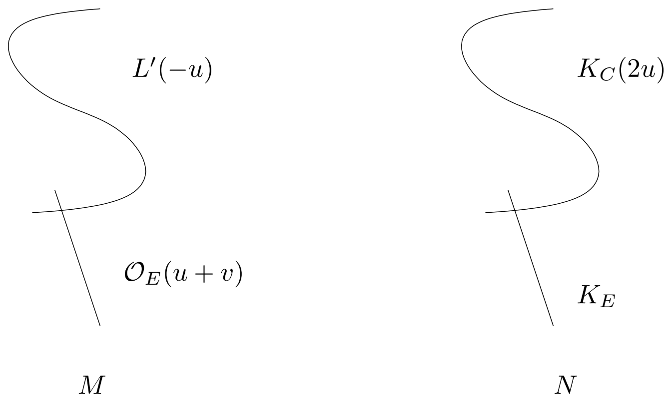

Figure 2. The twisted line bundles on the central fiber.

So, let $\sigma \in V$. By the description of the derivative complex at the beginning of this section, to compute

$$D(\mu_t)|_{t=0}(\sigma \otimes \omega),$$

we have to lift $\sigma \otimes \omega$ to first order in t, apply the Koszul differential μ_t to the lifting, then restrict the outcome divided by t to X_0.

So, let $\tilde{\sigma}$, $\tilde{\omega}$ be sections of $\mathcal{L}$ and $\omega_{\mathcal{X}/\Delta}$ extending σ and ω, respectively.

Since $\tilde{\sigma}$ vanishes on E and $\tilde{\omega}$ vanishes on C, we can write

$$\tilde{\sigma} = \tilde{\sigma}' s_E \tag{4-2}$$

and

$$\tilde{\omega} = \tilde{\omega}' s_C, \tag{4-3}$$

where s_E and s_C are sections of $\mathcal{O}_{\mathcal{X}}(E)$ and $\mathcal{O}_{\mathcal{X}}(C)$ vanishing precisely on E and C, respectively, and $\tilde{\sigma}'$ and $\tilde{\omega}'$ are global sections of

$$M := \mathcal{L}(-E)|_{X_0} \cong \mathcal{L}(C)|_{X_0} \quad \text{and} \quad N := \omega_{\mathcal{X}/\Delta}(-C)|_{X_0} \cong \omega_{\mathcal{X}/\Delta}(E)|_{X_0},$$

respectively. Notice that tensoring $\mathcal{L}$ by $\mathcal{O}_{\mathcal{X}}(-E)$ will increase the degree by 1 on the E component and decrease the degree by 1 on the C component. The line bundles M and N are described in Figure 2. Notice that $M \otimes N \cong \omega_{X_0} \otimes L_0$.

By the construction of the derivative complex,

$$D(\mu_t)|_{t=0}(\sigma \otimes \omega) = \frac{\tilde{\sigma} \cdot \tilde{\omega}}{t}\bigg|_{X_0} = \frac{(\tilde{\sigma}' s_E) \cdot (\tilde{\omega}' s_C)}{t}\bigg|_{X_0} = (\tilde{\sigma}'\tilde{\omega}')|_{X_0} \bmod \operatorname{Im} \mu_0. \tag{4-4}$$

The general case is just notationally more complicated.

Let

$$\sigma_{i_1} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes \omega \in \bigwedge^{r-p} V \otimes H^0(E, K_E).$$

We will compute its image under $D(\delta_t)|_{t=0}$. Similarly to the simpler case, we have to lift $\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega$ to first order, apply the Koszul differential δ_t to the lifting, then restrict the outcome divided by t to X_0.

To this end, let $\tilde{\sigma}_{i_k}$, $\tilde{\omega}$ be sections of $\mathcal{L}$ and $\omega_{\mathcal{X}/\Delta}$ extending σ_{i_k} and ω, respectively. Since $\tilde{\sigma}_{i_k}$ vanishes on E and $\tilde{\omega}$ vanishes on C, we can write

$$\tilde{\sigma}_{i_k}=\tilde{\sigma}'_{i_k}s_E \tag{4-5}$$

and

$$\tilde{\omega}=\tilde{\omega}'s_C, \tag{4-6}$$

as before.

We compute

$$\begin{aligned} D(\delta_t)|_{t=0}&(\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega)\\ &=\frac{\delta(\tilde{\sigma}_{i_1}\wedge\cdots\wedge\tilde{\sigma}_{i_{r-p}}\otimes\tilde{\omega})}{t}\Big|_{X_0}\\ &=\sum_{k=1}^{r-p}(-1)^k\frac{\tilde{\sigma}_{i_1}\wedge\cdots\wedge\hat{\tilde{\sigma}}_{i_k}\wedge\cdots\wedge\tilde{\sigma}_{i_{r-p}}\otimes(\tilde{\sigma}'_{i_k}s_E)(\tilde{\omega}'s_C)}{t}\Big|_{X_0}\\ &=\sum_{k=1}^{r-p}(-1)^k\sigma_{i_1}\wedge\cdots\wedge\hat{\sigma}_{i_k}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes(\tilde{\sigma}'_{i_k}\tilde{\omega}')|_{X_0}\ \operatorname{mod}\ \operatorname{Im}\delta_0. \end{aligned} \tag{4-7}$$

5. The study of obstruction classes

As explained in the introduction, our goal is to show that the rank of the obstruction map

$$K_{r-p,0}(X_0,L_0;\omega_{X_0})\xrightarrow{D(\delta_t)|_{t=0}}K_{r-p-1,1}(X_0,L_0;\omega_{X_0})$$

is as big as it could be, as this would imply $K_{r-p,0}(X_t,L_t;K_{X_t})$ is as small as it could be for $t\neq 0$.

Again let us analyze the simpler case $p=r-1$ first. By [Arbarello and Sernesi 1978], the multiplication map

$$H^0(X_t,L_t)\otimes H^0(X_t,\omega_{X_t})\xrightarrow{\mu_t}H^0(X_t,\omega_{X_t}\otimes L_t)$$

is already surjective for the general fiber, which implies $K_{1,0}(X_t,L_t;K_{X_t})$ is of the expected dimension. So, we are not proving anything new here, but it is helpful to redo this case via infinitesimal methods, because this method has the potential to generalize.

First notice that

$$H^0(X_0,L_0)\otimes H^0(X_0,\omega_{X_0})\xrightarrow{\mu_0}H^0(X_0,\omega_{X_0}\otimes L_0)$$

is not surjective. The problem is that any section in $H^0(X_0, \omega_{X_0})$ vanishes at u, but there is a section in $H^0(X_0, \omega_{X_0} \otimes L_0)$ not vanishing at u. Moreover, μ_0 is exactly of corank one. This is because on the E component, μ_0 becomes

$$H^0(\mathcal{O}_E(v)) \otimes H^0(\mathcal{O}_E(u)) \longrightarrow H^0(\mathcal{O}_E(u+v)),$$

which is of corank 1 (see Figure 1). Since by [Arbarello and Sernesi 1978] (or by the induction hypothesis if one wants an independent proof), the map

$$H^0(C, L') \otimes H^0(K_C) \longrightarrow H^0(L' \otimes K_C)$$

is surjective, we see that if a section $\tau \in H^0(X_0, \omega_{X_0} \otimes L_0)$ vanishes at u, then it is in the image of μ_0. Therefore $k_{1,0}(X_t, L_t; K_{X_t})$ jumps up by one at $t = 0$.

Now by the computation of the obstruction class in (4-4),

$$(\tilde{\sigma}'\tilde{\omega}')|_{X_0}$$

is in the image of

$$H^0(X_0, M) \otimes H^0(X_0, N) \longrightarrow H^0(X_0, \omega_{X_0} \otimes L_0).$$

Since there are always sections in $H^0(X_0, M)$ and $H^0(X_0, N)$ not vanishing at u, we can easily choose $\tilde{\sigma}'$ and $\tilde{\omega}'$ such that $(\tilde{\sigma}'\tilde{\omega}')|_{X_0}$ does not vanish at u. (Notice that any (global) sections of M and N will extend to nearby fibers.) Therefore there is at least a one-dimensional subspace of $K_{1,0}(X_0, L_0; K_{X_0})$ that does not deform to a nearby fiber, namely $(\tilde{\sigma}'s_E)(\tilde{\omega}'s_C)$. This means $K_{1,0}(X_t, L_t; K_{X_t})$ is of the expected dimension for $t \neq 0$. This proves the simpler case.

The case for general p is much more delicate. There are two possible ways to show the obstruction classes in (4-7) are not in the image of δ_0.

The easier way is to mimic the simpler case is to show $(\tilde{\sigma}'_{i_k}\tilde{\omega}')|_{X_0}$ does not lie in the image of

$$H^0(X_0, L_0) \otimes H^0(X_0, \omega_{X_0}) \xrightarrow{\mu_0} H^0(X_0, \omega_{X_0} \otimes L_0).$$

(As we have seen before, μ_0 is of corank one.) This will be the case if $(\tilde{\sigma}'_{i_k}\tilde{\omega}')|_{X_0}$ does not vanish at u. Then the obstruction class in (4-7) has no chance to be in $\mathrm{Im}(\delta_0)$.

To make this idea more precise, choose a basis $\{\sigma_1, \dots, \sigma_r\}$ of V adapted to u, that is, $\sigma_k|_C$ vanishes to order exactly k along u (therefore $\sigma_k|_E = 0$ for $k \geq 1$). Using the same notation as (4-2) and (4-3), we have that

$$(\tilde{\sigma}'_1\tilde{\omega}')|_{X_0}$$

is not in the image of μ_0 because $\sigma_1|_C$ vanishes to order exactly 1 at u, and any extension $\tilde{\sigma}_1 = \tilde{\sigma}'_1 \cdot s_E$ we choose would have $\tilde{\sigma}'_1$ nonvanishing at u (because $s_E|_C$ vanishes to order 1 at u, $\tilde{\sigma}'_1$ does not vanish). Similarly the extension $\tilde{\omega}'$ does not vanish at u. (Although the choice of extensions is not unique, different choices give the same obstruction class modulo $\mathrm{Im}(\delta_0)$.)

However, for $k \geq 2$, because $\sigma_k|_C$ vanishes to order at least 2 at u, we could choose a suitable extension $\tilde{\sigma}_k$ (modulo $\mathrm{Im}(\delta_0)$ this does not depend on the choice of extension) such that $\tilde{\sigma}_k = \tilde{\sigma}_k'' s_E^2$, and therefore

$$(\tilde{\sigma}_k'\tilde{\omega}')|_E = (\tilde{\sigma}_k'' s_E \tilde{\omega}')|_E = 0.$$

Thus for any $1 = i_1 < i_2 < i_3 < \cdots < i_{r-p} \leq r$,

$$D(\delta_t)|_{t=0}(\sigma_1 \wedge \sigma_{i_2} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes \omega)$$
$$= -\sigma_{i_2} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes (\tilde{\sigma}_1'\tilde{\omega}')|_{X_0} + \sum_{k=2}^{r-p} (-1)^k \sigma_{i_1} \wedge \cdots \wedge \hat{\sigma}_{i_k} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes (\tilde{\sigma}_{i_k}'\tilde{\omega}')|_{X_0}.$$

By looking at its restriction to E, we see immediately that the set

$$\{D(\delta_t)|_{t=0}(\sigma_{\mathbf{1}} \wedge \sigma_{i_2} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes \omega) \mid 2 \leq i_2 < i_3 < \cdots < i_{r-p} \leq r\}$$

is linearly independent in $K_{r-p-1,1}(X_0, L_0; \omega_{X_0})$.

Thus at this point the rank of $D(\delta_t)|_{t=0}$ is at least

$$\binom{r-1}{p},$$

and therefore

$$k_{r-p,0}(X_t, L_t; \omega_{X_t}) \leq \binom{r}{p} - \binom{r-1}{p} = \binom{r-1}{p-1} \tag{5-1}$$

for $t \neq 0$.

The second way to show obstructions are nontrivial is more delicate. As we have already seen, for $2 \leq i_1 < \cdots < i_{r-p} \leq r$, restricting to E does not give any information about $D(\delta_t)|_{t=0}(\sigma_{i_1} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes \omega)$, since they all restrict to zero on E. We will have to study the restriction of $D(\delta_t)|_{t=0}(\sigma_{i_1} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes \omega)$ to C.

Equation (4-7) restricted to C becomes

$$D(\delta_t)|_{t=0}(\sigma_{i_1} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes \omega)|_C$$
$$= \sum_{k=1}^{r-p} (-1)^k \sigma_{i_1} \wedge \cdots \wedge \hat{\sigma}_{i_k} \wedge \cdots \wedge \sigma_{i_{r-p}} \otimes (\tilde{\sigma}_{i_k}'\tilde{\omega}')|_C \mod \mathrm{Im}\,\delta_0$$

Here $\tilde{\sigma}_{i_k}'|_C \in H^0(C, L'(-u))$ and is equal to σ_{i_k} for $i_k \geq 1$, if we abuse notation by thinking of σ_{i_k} as sections of $L'(-u)$ instead of L'. (Thus σ_1 is a section of $L'(-u)$ which does not vanish at u and σ_2 vanishes to order 1 at u, etc.) On the other hand, $\tilde{\omega}'|_C \in H^0(K_C(2u))$ and does not vanish at u; denote it by ω'. With the notation above, the obstruction class becomes

$$\text{(5-2)}\quad \sum_{k=1}^{r-p}(-1)^k\sigma_{i_1}\wedge\cdots\wedge\hat{\sigma}_{i_k}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes(\sigma_{i_k}\omega')=\delta_0(\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega').$$

Remark. Here we are still using δ_0 to denote the restriction to C of the original Koszul differential δ_0 on X_0. Equation (5-2) does not mean

$$D(\delta_t)|_{t=0}(\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega)|_C\in\operatorname{Im}(\delta_0),$$

since $\omega'\notin H^0(K_C)$.

Now the nontriviality of obstruction classes on X_0 boils down to a question on (C', L'):

Theorem 5.1. *Let C be a general curve of genus g, L' a g^r_d on C such that $K_{r-p,0}(C, L'; K_C)=0$ and $\{\sigma_0,\dots,\sigma_r\}$ is a basis of $H^0(C, L')$ adapted to a general point $u\in C$, and let $\omega'\in H^0(C, K_C(2u))\setminus H^0(C, K_C)$. Consider the obstruction classes*

$$\text{(5-3)}\quad \{\delta_0(\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega')\mid 2\le i_1<\cdots<i_{r-p}\le r\}\subset K_{r-p-1,1}(C, L'; K_C).$$

(a) *If these classes are linearly independent in $K_{r-p-1,1}(C, L', K_C)$, then*

$$K_{r-p,0}(X_t, L_t;\omega_{X_t})\cong K_{p-1,2}(X_t, L_t)^\vee=0.$$

(b) *On the other hand, if these classes span $K_{r-p-1,1}(C, L'; K_C)$, then*

$$k_{r-p,0}(X_t, L_t;\omega_{X_t})\le\binom{r-1}{p-1}-k_{r-p-1,1}(C, L'; K_C)=-b_{p+1}(X_t, L_t),$$

which implies $K_{p,1}(X_t, L_t)=0$.

Proof. The hypothesis in case (a) implies that $D(\delta_t)|_{t=0}$ in (4-1) is either injective, which means no elements in $K_{r-p,0}(X_0, L_0;\omega_{X_0})$ will extend to nearby. In case (b), the rank of $D(\delta_t)|_{t=0}$ is

$$k_{r-p-1,1}(C, L'; K_C)+\binom{r-1}{p},$$

which implies that

$$k_{r-p,0}(X_t, L_t;\omega_{X_t})\le\binom{r-1}{p-1}-k_{r-p-1,1}(C, L'; K_C)=-b_{p+1}(X_t, L_t).$$

Therefore only a subspace of $K_{r-p,0}(X_0, L_0;\omega_{X_0})$ of correct dimension will extend to nearby fibers to first order. □

Now we give a sufficient condition for the obstruction classes in (5-3) to be linearly independent. Consider the diagram of complexes

$$\begin{tikzcd}
\bigwedge^{r-p+1} H^0(L')\otimes H^0(K_C\otimes L'^{-1}) \arrow[r, equal] \arrow[d, hook] & \bigwedge^{r-p+1} H^0(L')\otimes H^0(K_C\otimes L'^{-1}) \arrow[d, hook] \\
\bigwedge^{r-p} H^0(L')\otimes H^0(K_C) \arrow[r, hook, "\alpha"] \arrow[d, "\delta_0"] & \bigwedge^{r-p} H^0(L')\otimes H^0(K_C(2u)) \arrow[d, "\delta"] \\
\bigwedge^{r-p-1} H^0(L')\otimes H^0(K_C\otimes L') \arrow[r] & \bigwedge^{r-p-1} H^0(L')\otimes H^0(K_C\otimes L'(2u))
\end{tikzcd}$$

Lemma 5.2. *Under the same assumptions as* Theorem 5.1, *if the right column of the above diagram is exact in the middle* (*the left column is exact by assumption*), *then the obstruction classes in* (5-3) *are linearly independent modulo* $\mathrm{Im}(\delta_0)$. *As a consequence of* Theorem 5.1 *then*, $K_{r-p,0}(X_t, L_t; \omega_{X_t}) = 0$.

Proof. The assumption implies that

$$\mathrm{Ker}(\delta) = \mathrm{Ker}(\delta_0) \cong \bigwedge^{r-p+1} H^0(L')\otimes H^0(K_C\otimes L'^{-1}).$$

If a linear combination of the obstruction classes $\delta_0(\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega')$ is equal to $\delta_0(c)$ for some $c\in\bigwedge^{r-p} H^0(L')\otimes H^0(K_C)$, then the same linear combination of the $\{\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega'\}$ minus $\alpha(c)$ is in $\mathrm{Ker}(\delta)=\mathrm{Ker}(\delta_0)$. This contradicts the fact that

$$\{\sigma_{i_1}\wedge\cdots\wedge\sigma_{i_{r-p}}\otimes\omega'\}$$

is linearly independent in $\bigwedge^{r-p} H^0(L')\otimes H^0(K_C(2u))$ modulo the image of α. □

To end this section, we give a proof of Theorem 1.3.

Proof of Theorem 1.3. There are two cases:

(a) $K_{p,1}(C, L') = 0$: By Proposition 3.1, $K_{p,1}(X_0, L_0) = 0$, and $\mathrm{GV}(p)^r_{g+1,d+1}$ follows from upper-semicontinuity of Koszul cohomology.

(b) $K_{p-1,2}(C, L')\cong K_{r-p,0}(C, L, ; K_C)^\vee = 0$: Starting from the defining sequence for the kernel bundle $M_{L'}$,

$$0\longrightarrow M_{L'}\longrightarrow H^0(L')\otimes\mathcal{O}_C\longrightarrow L'\longrightarrow 0,$$

taking the $(r-p)$-th wedge, twisting by K_C (respectively $K_C(2u)$) and then taking global sections, we get

$$\begin{aligned}0\longrightarrow \bigwedge^{r-p} M_{L'}\otimes K_C\longrightarrow \bigwedge^{r-p} H^0(L')\otimes H^0(K_C)&\\ \xrightarrow{\delta_0} \bigwedge^{r-p-1} H^0(L')\otimes H^0(K_C\otimes L')\longrightarrow 0,&\end{aligned}$$

and therefore

$$\operatorname{Ker}(\delta_0) = H^0\big(C, \textstyle\bigwedge^{r-p} M_{L'} \otimes K_C\big). \tag{5-4}$$

Similarly,

$$\operatorname{Ker}(\delta) = H^0\big(C, \textstyle\bigwedge^{r-p} M_{L'} \otimes K_C(2u)\big). \tag{5-5}$$

If

$$h^0\big(C, \textstyle\bigwedge^{r-p} M_{L'} \otimes K_C\big) = h^0\big(C, \textstyle\bigwedge^{r-p} M_{L'} \otimes K_C(2u)\big),$$

we conclude that

$$\operatorname{Ker}(\delta_0) \cong \operatorname{Ker}(\delta),$$

which implies $K_{r-p,0}(X_t, L_t; \omega_{X_t}) \cong K_{p,2}(X_t, L_t)^\vee = 0$ by Lemma 5.2. □

6. Some applications to the maximal rank conjecture

In the case $p = 1$, we can reduce condition (1-4) in Theorem 1.3 to a statement about the tangential variety TC of C; namely, the existence of a quadric containing C but not containing TC. The condition on the tangential variety is quite interesting in its own right. Theorem 1.4 follows immediately from Theorem 1.3 and Lemma 6.1.

Lemma 6.1. *For a general $L' = g^r_d$ on a general curve C of genus g such that $K_{0,2}(C, L') = 0$ (i.e., μ in (1-1) is surjective), if there exists a quadric $Q \subset \mathbb{P}^r$ containing $\phi_{|L'|}(C)$ but not containing its tangential surface $TC := \bigcup_{u \in C} T_u C \subset \mathbb{P}^r$, then*

$$H^0\big(C, \textstyle\bigwedge^{r-1} M_{L'} \otimes K_C\big) = H^0\big(C, \textstyle\bigwedge^{r-1} M_{L'} \otimes K_C(2u)\big).$$

Proof. Notice that

$$\textstyle\bigwedge^r M_{L'} \cong L'^{-1},$$

and therefore

$$\textstyle\bigwedge^{r-1} M_{L'}^\vee \cong M_{L'} \otimes L'.$$

By the Riemann–Roch theorem, it suffices to show that

$$h^0(M_{L'} \otimes L'(-2u)) = h^0(M_{L'} \otimes L') - 2r.$$

The $\geq$ part is automatically true, and only the $\leq$ part needs to be proved.

We have a diagram with exact columns

$$\begin{tikzcd}
0 \arrow[d] & 0 \arrow[d] \\
H^0(M_{L'} \otimes L'(-2u)) \arrow[r, hook] \arrow[d] & H^0(M_{L'} \otimes L') \arrow[d] \\
H^0(L') \otimes H^0(L'(-2u)) \arrow[r, hook] \arrow[d, "\mu'"] & H^0(L') \otimes H^0(L') \arrow[d, "\mu"] \\
H^0(L'^2(-2u)) \arrow[r, hook] & H^0(L'^2) \arrow[d] \\
 & 0
\end{tikzcd}$$

We need to show

$$\dim_{\mathbb{C}} \operatorname{Ker}(\mu') \leq \dim_{\mathbb{C}} \operatorname{Ker}(\mu) - 2r.$$

Let $H_u := H^0(L') \otimes H^0(L'(-2u))$, and $\overline{H}_u$ its image in

$$\frac{H^0(L') \otimes H^0(L')}{\bigwedge^2 H^0(L')} \cong S^2 H^0(L').$$

$\overline{H}_u$ is the space of quadrics which contain the tangent line of C at u.

We have

$$\operatorname{Ker}(\mu') = \operatorname{Ker}(\mu) \cap H_u.$$

By hypothesis, $\overline{\operatorname{Ker}(\mu)} \not\subseteq \overline{H}_u$ for general u (since $Q \notin \overline{H}_u$), and it follows that

$$\begin{aligned}
\dim_{\mathbb{C}}(\overline{\operatorname{Ker}(\mu')}) = \dim_{\mathbb{C}}(\overline{\operatorname{Ker}(\mu) \cap H_u}) &\leq \dim_{\mathbb{C}}(\overline{\operatorname{Ker}(\mu)} \cap \overline{H}_u) \\
&\leq \dim_{\mathbb{C}}(\overline{\operatorname{Ker}(\mu)}) - 1 =: m - 1.
\end{aligned}$$

Thus

$$\begin{aligned}
\dim_{\mathbb{C}}(\operatorname{Ker}(\mu')) &\leq m - 1 + \dim_{\mathbb{C}}\big(\textstyle\bigwedge^2 H^0(L') \cap H_u\big) \\
&= m - 1 + \dim_{\mathbb{C}}\big(\textstyle\bigwedge^2 H^0(L'(-2u))\big) \\
&= m - 1 + \binom{r-1}{2} = m + \binom{r+1}{2} - 2r \\
&= \dim_{\mathbb{C}}(\operatorname{Ker}(\mu)) - 2r. \qquad \square
\end{aligned}$$

Let us go the proof of Corollary 1.5. The numerical assumption in Corollary 1.5 turns out to be a technical assumption needed to verify the assumption about TC in Theorem 1.4(b). This is equivalent to the numerical assumption in Lemma 6.2. By the Appendix, if L' is a general nonspecial g^r_{2r-3} on a general curve C of genus $r-3$, the number of quadrics containing TC is at most

$$\binom{r-4}{2}.$$

Lemma 6.2. *Let $C \subset \mathbb{P}^r$ be a general curve of genus g embedded by L' a general g^r_d with $h^1(L') \le 1$. Suppose*

$$\binom{r+2}{2} - (2d-g+1) > \binom{r-4}{2}$$

(i.e., the number of independent quadrics containing C is at least $\binom{r-4}{2}$); then there exists a quadric Q on $\mathbb{P}^r$ containing C but not containing TC.

Proof. Degenerate (C, L') to (C_0, L'_0), where C_0 is a nodal curve with two smooth components Y and Z meeting at a general point u. Depending on the value of $h^1(L')$, there are two cases:

(a) $h^1(L') = 0$: In this case, $L' = g^r_{g+r}$ for $g \ge 0$. If $0 \le g \le r-3$, take $g_Y = 0$, $g_Z = g$, $L'_0|_Y = \mathcal{O}_{\mathbb{P}^1}(r)$ and $L'_0|_Z = g^0_g$. (One could easily show such a (C_0, L'_0) can deform to (C, L').) Since there are only

$$\binom{r-2}{2}$$

quadrics containing the tangential variety of the rational normal curve in $\mathbb{P}^r$ (see the Appendix) and in this range of g, the number of quadrics containing C is at least

$$\binom{r+2}{2} - (2d-g+1) = \binom{r+2}{2} - (g+2r+1) > \binom{r-2}{2},$$

we conclude that there exists a quadric containing the nearby fiber C but not containing TC.

If $g > r-3$, we take $g_Y = r-3$, $g_Z = g-r+3$, $L'_0|_Y = g^r_{2r-3}$ (a general one) and $L'_0|_Z = g^0_{g-r+3}$. By Proposition A.1 in the Appendix, the number of quadrics containing TC for nearby C is at most $\binom{r-4}{2}$. By the numerical hypothesis, we get our conclusion.

(2) $h^1(L') = 1$: The argument is similar to the above, except that we need to deal with $L' = g^r_{g+r-1}$ for $g \ge r+1$. Again, if $r+1 \le g \le 2r-2$, we take $g_Y = 0$, $Z = C$, $L'_0|_Y = \mathcal{O}_{\mathbb{P}^1}(r)$ and $L'_0|_Z = L'(-ru) = g^0_{g-1}$.

If $g > 2r-2$, take $g_Y = r-3$, $g_Z = g-r+3$, $L'_0|_Y = g^r_{2r-3}$, $L'_0|_Z = g^0_{g-r+2}$. Here $L'_0|_Z$ comes from a general g^r_{g+2} on Z twisted by $\mathcal{O}_Z(-ru)$. The rest of the argument is exactly the same as in case (a). $\square$

Proof of Corollary 1.5. First notice that, by Corollary 2.2, to show projective normality of a general pair, it suffices to show (1-1) is surjective. We will fix h^1 and r and induct on g.

For the $h^1 = 0$ case, we start with the fact that a rational normal curve is projectively normal (i.e., $(\mathrm{MRC})^r_{0,r}$ holds). For the $h^1 = 1$ case, we use the fact that a general canonical curve is projectively normal (i.e., $(\mathrm{MRC})^r_{r+1,2r}$ holds). Now assuming $(\mathrm{MRC})^r_{g,d}$ holds, by Lemma 6.2, as long as

$$\binom{r+2}{2} - (2d - g + 1) > \binom{r-4}{2}, \tag{6-1}$$

Theorem 1.4(b) is satisfied, which implies $(\mathrm{MRC})^r_{g+1,d+1}$ (which is equivalent to projective normality). Plugging $d = g + r - h^1$ into (6-1), we immediately get the bound on d as in the statement of the theorem. □

Proof of Corollary 1.8. The $r = 1, 2$ cases is trivial. The arguments for $r = 3, 4$ are completely similar, so we will only prove the case $r = 4$. Again, we induct on g. First suppose we have proved (MRC) for the base cases $g = 5h^1$, $L = g^4_{4h^1+4}$. Then notice that for $r = 4$, $\binom{r-4}{2} = 0$, and therefore there is no quadric containing the tangential variety in Theorem 1.4(b). Thus $(\mathrm{MRC})^4_{g,d}$ implies $(\mathrm{MRC})^4_{g+1,d+1}$. It remains to prove (MRC) for the base cases. When $h^1 \le 1$, $(\mathrm{MRC})^4_{5h^1,4h^1+4}$ is clear. If $h^1 \ge 2$, we need to show μ in (1-1) is injective. For $h^1 = 2$, $(\mathrm{MRC})^4_{10,12}$ is well known and is proved in [Farkas and Popa 2005]. If $h^1 \ge 3$, we could degenerate again to $C_0 = Y \cup Z$ with $g_Y = 10$, $g_X = 5h^1 - 10$, $L_0|_Y = g^4_{12}$, $L_0|_Z = g^0_{4h^1-8} = g^4_{4h^1-4}(-4u)$. Again it is easy to check such (C_0, L_0) is smoothable in $\mathcal{G}^4_{5h^1,4h^1+4}$ (see [Wang 2013, Corollary 6.1] for details). The injectivity of μ in this case follows from the same argument as in Proposition 3.1. □

It was also proved in [Farkas 2009] that for any integer $s \ge 1$, $(\mathrm{MRC})^{2s}_{s(2s+1),2s(s+1)}$ holds. In this case, $\rho = 0$ and $h^1 = s$. Thus by Theorem 1.4 (a), we have the following corollary:

Corollary 6.3. $(\mathrm{MRC})^{2s}_{s(2s+1)+k,2s(s+1)+k}$ *holds for all* $s \ge 1$, $k \ge 0$; *that is*, (MRC) *holds if* $r = 2h^1$.

7. Higher syzygies

As we mentioned in the introduction, the difficulty in generalizing the inductive argument to higher syzygies is due to the lack of known cases to start the induction with and the lack of an analog of Theorem 1.4 for higher syzygies. Nevertheless, we collect some vanishing results we could obtain in this section.

Proposition 7.1. *For* L *a general* g^r_d *on a general curve* X *with* $g \ge r + 1$, $K_{p,1}(X, L) = 0$ *for* $p \ge \lfloor (r+1)/2 \rfloor$.

Proof. Start with the case $g_X = r+1$, $L' = K_X$. Thanks to Voisin's solution to the generic Green conjecture, $K_{p,1}(X, K_X) = 0$ for $p \geq \lfloor (r+1)/2 \rfloor$. If $g > r+1$, we degenerate to $X_0 = Y \cup Z$ with $g_Y = r+1$, $g_Z = g-r-1$, $L_0|_Y = K_Y$, $L_0|_Z = g^0_{d-2r} = g^r_{d-r}(-ru)$. The statement then follows from the same argument as in Proposition 3.1. $\square$

Remark. Using the same degeneration as in Proposition 7.1, we also have

$$k_{p,1}(X, L) \leq k_{p.1}(Y, K_Y) = \left[\binom{r-1}{p} - \binom{r-1}{p-1}\right] r + \binom{r+1}{p} - \binom{r+1}{p+1}$$

for $1 \leq p < \lfloor (r+1)/2 \rfloor$. We will improve this bound using an infinitesimal argument.

Even though we do not have an analog of Theorem 1.4, when g is not too big compared to p, the inductive argument still go through:

Lemma 7.2. *Let C be a general curve of genus g and L' a g^r_d with $h^1(L') = 1$. If $p \leq r - \lfloor (g+1)/2 \rfloor$, then the sequence*

$$\begin{aligned} \textstyle\bigwedge^{r-p+1} H^0(L') \otimes H^0(K_C \otimes L'^{-1}) &\hookrightarrow \textstyle\bigwedge^{r-p} H^0(L') \otimes H^0(K_C(2u)) \\ &\xrightarrow{\delta} \textstyle\bigwedge^{r-p-1} H^0(L') \otimes H^0(K_C \otimes L'(2u)) \end{aligned}$$

is exact in the middle.

Proof. Let $t \in H^0(K_C \otimes L'^{-1})$ be a generator. Multiplication by t gives an embedding

$$H^0(L') \xrightarrow{\cdot t} H^0(K_C).$$

Denote its image by W. Let C' be the image of C under the map given by $|K_C(2u)|$. C' is of arithmetic genus $g+1$ and has a cusp. We can identify $H^0(C, K_C(2u))$ with $H^0(C', \omega_{C'})$, where $\omega_{C'}$ is the dualizing sheaf of C'. With the above notation, we can identify the cohomology group in question with $K_{r-p,1}(C', \omega_{C'}; W) \subset K_{r-p,1}(C', \omega_{C'})$. Since curves in $K3$ surfaces satisfy the Green conjecture (see [Voisin 2002; 2005]), by degenerating C' to a cuspidal curve in $K3$ surface, we have

$$K_{r-p,1}(C', \omega_{C'}) = 0$$

for $r - p \geq \lfloor (g+1)/2 \rfloor$. $\square$

Again we start our induction with a general curve of genus $r+1$, and $L' = K_C$. For $p < \lfloor (r+1)/2 \rfloor$, we have

$$K_{p-1,2}(C, K_C)^\vee \cong K_{r-p,1}(C, K_C) = 0.$$

Now we apply the construction in Section 3. By Lemma 5.2 and Lemma 7.2, we get:

Proposition 7.3. *For L a general g^r_d on a general curve X with $h^1(L) = 1$:*

(a) $k_{p-1,2}(X, L) = k_{r-p,0}(X, L; K_X) = 0$ *if* $p \leq r - \lfloor g/2 \rfloor$.

(b) $k_{p-1,2}(X, L) \leq (g - 2r + 2p - 1)\binom{r-1}{p-1}$ *if* $p > r - \lfloor g/2 \rfloor$.

Proof. Again we start our induction on g with a general curve of genus $r + 1$ and $L' = K_C$. We always have

$$K_{p-1,2}(C, K_C)^\vee \cong K_{r-p,1}(C, K_C) = 0$$

for $r - p \geq \lfloor (r + 1)/2 \rfloor$.

Now we apply the construction in Section 3. If $\lfloor g/2 \rfloor \leq r - p$, Lemmas 5.2 and 7.2 apply, and we get (a).

When $\lfloor g/2 \rfloor$ gets past $r - p$ (or equivalently $g > 2r - 2p + 1$), we nevertheless have estimate (5-1) for each attached elliptic tail. Thus the bound in (b) follows. □

Combining the results of Propositions 7.1 and 7.3, we get Theorem 1.7.

Remark. For line bundles with $h^1 = 1$, the assumption $p \leq r - \lfloor g/2 \rfloor$ is equivalent to the condition that $d \geq 2g - 2 + p - \lfloor (g - 1)/2 \rfloor$. Thus Proposition 7.3 is the generic version of the generalized Green–Lazarsfeld conjecture [1986] for special linear series. However this generic version is known to follow from the generic Green conjecture (see [Aprodu and Farkas 2011, Proposition 4.30]). It seems to the author that the bound in (b) is new.

Appendix

We prove the following statement, which is needed in the proof of Corollary 1.5.

Proposition A.1. *For a general curve C of genus $r - 3$ embedded in $\mathbb{P}^r$ by a general g^r_{2r-3}, the number of quadrics containing TC is at most*

$$\binom{r-4}{2}.$$

Consider the rational normal curve C of degree d in $\mathbb{P}^d$. It is well known that there are

$$\binom{d}{2}$$

independent quadrics containing C. Denote them by $\Delta_{a,b}$ for $0 \leq a < b \leq d - 1$, where $\Delta_{a,b}$ is the 2×2 minor corresponding to columns a and b of the matrix

$$\begin{pmatrix} x_0 & x_1 & x_2 & \cdots & x_{d-2} & x_{d-1} \\ x_1 & x_2 & x_3 & \cdots & x_{d-1} & x_d \end{pmatrix},$$

with the usual convention that $\Delta_{a,b} = -\Delta_{b,a}$.

It is proved in [Eisenbud 1992] that there are

$$\binom{d-2}{2}$$

quadrics containing TC. They are

$$\Gamma_{a,b} = \Delta_{a+2,b} - 2\Delta_{a+1,b+1} + \Delta_{a,b+2}$$

for $0 \le a, b \le d-3$.

Now, consider the projection C' of C to $\mathbb{P}^r$ ($d = 2r-3$, $r \ge 3$) given by

$$t \longrightarrow [1, t^2, t^4, \dots, t^{2r-6}, t^{2r-5}, t^{2r-4}, t^{2r-3}].$$

C' has arithmetic genus $r-3$ and has a unique singular point at $t = 0$ locally isomorphic to $\operatorname{Spec}(\mathbb{C}[t^2, t^{2r-5}])$.

Lemma A.2. *The complete linear system $|\mathcal{O}_{C'}(1)|$ has projective dimension r; that is, $C' \subset \mathbb{P}^r$ is linearly normal. As a consequence, C' is smoothable in $\mathbb{P}^r$.*

Proof. Let $L_l = \operatorname{span}\{P_1, P_3, \dots, P_{2k-1}\} \subset \mathbb{P}^{2r-3}$, where

$$P_i = [0, 0, \dots, 1^{(i\text{-th})}, \dots, 0], \quad i = 0, 1, \dots, r,$$

and denote by $C_k \subset \mathbb{P}^{2r-3-k}$ the projection of C with center L_k. The curve C_k has a unique singular point locally isomorphic to $\operatorname{Spec}(\mathbb{C}[t^2, t^{2k+1}])$. Note that $C' = C_{r-3} \subset \mathbb{P}^r$. We use induction to show that the complete linear system $\mathcal{O}_{\mathbb{P}^{2r-3-k}}(1)|_{C_k}$ has projective dimension $2r-3-k$. The natural projection map $Pr_k : C_k \to C_{k+1}$ induces an inclusion $H^0(\mathcal{O}_{C_{k+1}}(1)) \subset H^0(\mathcal{O}_{C_k}(1))$. By the inductive hypothesis, $h^0((\mathcal{O}_{C_k}(1)) = h^0(\mathcal{O}_{\mathbb{P}^{2r-3-k}}(1)) = 2r-2-k$. Since we obtain C_{k+1} from C_k by projection from a point, $h^0((\mathcal{O}_{C_{k+1}}(1)) \ge h^0((\mathcal{O}_{C_k}(1)) - 1$. Since C_{k+1} has arithmetic genus one higher than C_k, $H^0(\mathcal{O}_{C_{k+1}}(1)) \not\subseteq H^0(\mathcal{O}_{C_k}(1))$. Thus $h^0((\mathcal{O}_{C_{k+1}}(1)) = 2r-3-k$. For the last statement, note that the curve C' has only a plane curve singularity, and thus is smoothable (as an abstract curve). Moreover, since $h^0(\mathcal{O}_{C'}(1)) = r+1$, $\mathcal{O}_{C'}(1)$ is a complete nonspecial g^r_{2r-3}. For any one-parameter smoothing (C_t, L_t) of the pair $(C', \mathcal{O}_{C'}(1))$, since h^0 of the central fiber does not jump up, all $r+1$ global sections of $\mathcal{O}_{C'}(1)$ deform to L_t. □

Proof of Proposition A.1. We could explicitly compute the quadrics containing TC': they are just quadrics in $\mathbb{P}^{2r-3}$ containing TC with singular locus containing the center of projection $L_{r-3} = \operatorname{span}\{P_1, P_3, \dots, P_{2r-7}\}$. Now if we think of each quadric $\Gamma_{a,b}$ as a $(2r-2) \times (2r-2)$ symmetric matrix, we are just looking for matrices $Q \in S_\Gamma := \operatorname{span}\{\Gamma_{a,b} \mid 0 \le a < b \le 2r-6\}$ such that $L_{r-3} \subset \operatorname{Ker} Q$. (We think of Q as a linear operator on $\mathbb{C}^{2r-2}$ and L_{r-3} as a subspace of $\mathbb{C}^{2r-2}$.)

Notice that each $\Gamma_{a,b}$, as a matrix, can have nonzero entries at the (i, j)-spot $(0 \le i, j \le 2r-3)$ only if

$$i + j = a + b + 3.$$

Stated differently, each $\Gamma_{a,b}$, as a matrix, is supported on one of the diagonals.

For each $4 \le k \le 4r-10$, there are $(\lfloor k/2 \rfloor - 1)$ of the $\Gamma_{a,b}$ contributing to nonzero entries on the line

$$i + j = k, \quad \text{for } 4 \le k \le 2r-3,$$

and $(2r - 4 - \lfloor (k+1)/2 \rfloor)$ of the $\Gamma_{a,b}$ if $2r - 3 < k \le 4r - 10$.

Write

$$S_\Gamma = \bigoplus_{k=4}^{4r-10} S_k,$$

where $S_k = \operatorname{span}\{\Gamma_{a,b} \mid 0 \le a < b \le 2r-6,\ a+b = k-3\}$. It is obvious that if $Q \in S_\Gamma$ vanishes on L_{r-3}, then its S_k component also vanishes on L_{r-3}. Thus it suffices to count how many quadrics in each S_k vanish on L_{r-3}.

Let's just consider the case $4 \le k \le 2r-3$; the other case is similar.

When k is odd, vanishing on L_{r-3} imposes $(k-1)/2$ independent conditions on S_k, more than the dimension of S_k. Thus no quadric in S_k vanishes on L_{r-3}.

When k is even, vanishing on L_{r-3} only imposes $\lceil k/4 \rceil$ independent conditions. We conclude that for $4 \le k \le 2r-3$, there are

$$\sum_{\substack{8 \le k \le 2r-4 \\ k \text{ even}}} \left(\frac{k}{2} - 1 - \left\lceil \frac{k}{4} \right\rceil \right) = \left\lfloor \frac{r^2 - 8r + 16}{4} \right\rfloor$$

quadrics containing TC' (if $r \le 5$ there are none!).

Similarly, for $2r - 3 < k \le 4r - 10$, we count that there are

$$\left\lfloor \frac{r^2 - 8r + 16}{4} \right\rfloor - \left\lfloor \frac{r-4}{2} \right\rfloor$$

quadrics containing TC'.

So we get a total of

$$\left\lfloor \frac{r^2 - 8r + 16}{4} \right\rfloor + \left\lfloor \frac{r^2 - 8r + 16}{4} \right\rfloor - \left\lfloor \frac{r-4}{2} \right\rfloor = \binom{r-4}{2}$$

quadrics containing TC'. By specializing to C', we conclude our proof. $\square$

Acknowledgements

This work is a continuation of my thesis project. I would like to thank my thesis adviser Herb Clemens for suggesting the problem and method, and his constant support for this work. I would also like to thank Aaron Bertram, Gavril Farkas and Joe Harris for generously sharing their ideas on this problem. Last but not least, I thank the referee for the helpful comments and suggestions to improve the paper.

References

[Aprodu 2004] M. Aprodu, "Green–Lazarsfeld gonality conjecture for a generic curve of odd genus", *Int. Math. Res. Not.* **2004**:63 (2004), 3409–3416. MR 2005k:14012 Zbl 1072.14036

[Aprodu 2005] M. Aprodu, "Remarks on syzygies of d-gonal curves", *Math. Res. Lett.* **12**:2-3 (2005), 387–400. MR 2006d:14028 Zbl 1084.14032

[Aprodu and Farkas 2011] M. Aprodu and G. Farkas, "Green's conjecture for curves on arbitrary $K3$ surfaces", *Compos. Math.* **147**:3 (2011), 839–851. MR 2012e:14006 Zbl 1221.14039

[Aprodu and Nagel 2010] M. Aprodu and J. Nagel, *Koszul cohomology and algebraic geometry*, University Lecture Series **52**, American Mathematical Society, Providence, RI, 2010. MR 2011f:14051 Zbl 1189.14001

[Arbarello and Cornalba 1981] E. Arbarello and M. Cornalba, "Su una congettura di Petri", *Comment. Math. Helv.* **56**:1 (1981), 1–38. MR 82k:14029 Zbl 0505.14002

[Arbarello and Sernesi 1978] E. Arbarello and E. Sernesi, "Petri's approach to the study of the ideal associated to a special divisor", *Invent. Math.* **49**:2 (1978), 99–119. MR 80c:14020 Zbl 0399.14019

[Ballico 1996] E. Ballico, "On the minimal free resolution of general embeddings of curves", *Pacific J. Math.* **172**:2 (1996), 315–319. MR 97d:14042 Zbl 0848.14007

[Ballico and Fontanari 2010a] E. Ballico and C. Fontanari, "Normally generated line bundles on general curves", *J. Pure Appl. Algebra* **214**:6 (2010), 837–840. MR 2011b:14013 Zbl 1184.14039

[Ballico and Fontanari 2010b] E. Ballico and C. Fontanari, "Normally generated line bundles on general curves, II", *J. Pure Appl. Algebra* **214**:8 (2010), 1450–1455. MR 2011b:14014 Zbl 1185.14028

[Ein 1987] L. Ein, "A remark on the syzygies of the generic canonical curves", *J. Differential Geom.* **26**:2 (1987), 361–365. MR 89a:14031 Zbl 0632.14024

[Eisenbud 1992] D. Eisenbud, "Green's conjecture: an orientation for algebraists", pp. 51–78 in *Free resolutions in commutative algebra and algebraic geometry* (Sundance, UT, 1990), edited by D. Eisenbud and C. Huneke, Res. Notes Math. **2**, Jones and Bartlett, Boston, 1992. MR 93e:13020 Zbl 0792.14015

[Eisenbud and Harris 1983] D. Eisenbud and J. Harris, "Divisors on general curves and cuspidal rational curves", *Invent. Math.* **74**:3 (1983), 371–418. MR 85h:14019 Zbl 0527.14022

[Eisenbud and Harris 1986] D. Eisenbud and J. Harris, "Limit linear series: basic theory", *Invent. Math.* **85**:2 (1986), 337–371. MR 87k:14024 Zbl 0598.14003

[Farkas 2006] G. Farkas, "Syzygies of curves and the effective cone of $\overline{\mathcal{M}}_g$", *Duke Math. J.* **135**:1 (2006), 53–98. MR 2008a:14037 Zbl 1107.14019

[Farkas 2009] G. Farkas, "Koszul divisors on moduli spaces of curves", *Amer. J. Math.* **131**:3 (2009), 819–867. MR 2010f:14030 Zbl 1176.14006

[Farkas and Popa 2005] G. Farkas and M. Popa, "Effective divisors on $\overline{\mathcal{M}}_g$, curves on $K3$ surfaces, and the slope conjecture", *J. Algebraic Geom.* **14**:2 (2005), 241–267. MR 2006a:14043 Zbl 1081.14038

[Green 1984] M. L. Green, "Koszul cohomology and the geometry of projective varieties", *J. Differential Geom.* **19**:1 (1984), 125–171. MR 85e:14022 Zbl 0559.14008

[Green and Lazarsfeld 1984] M. L. Green and R. Lazarsfeld, "The nonvanishing of certain Koszul cohomology groups", *J. Differential Geom.* **19**:1 (1984), 168–170. Appendix to [Green 1984]. MR 85e:14022 Zbl 0559.14008

[Green and Lazarsfeld 1986] M. L. Green and R. Lazarsfeld, "On the projective normality of complete linear series on an algebraic curve", *Invent. Math.* **83**:1 (1986), 73–90. MR 87g:14022 Zbl 0594.14010

[Green and Lazarsfeld 1987] M. L. Green and R. Lazarsfeld, "Deformation theory, generic vanishing theorems, and some conjectures of Enriques, Catanese and Beauville", *Invent. Math.* **90**:2 (1987), 389–407. MR 89b:32025 Zbl 0659.14007

[Voisin 2002] C. Voisin, "Green's generic syzygy conjecture for curves of even genus lying on a $K3$ surface", *J. Eur. Math. Soc.* (*JEMS*) **4**:4 (2002), 363–404. MR 2003i:14040 Zbl 1080.14525

[Voisin 2005] C. Voisin, "Green's canonical syzygy conjecture for generic curves of odd genus", *Compos. Math.* **141**:5 (2005), 1163–1190. MR 2006c:14053 Zbl 1083.14038

[Wang 2013] J. Wang, "On the projective normality of line bundles of extremal degree", *Math. Ann.* **355**:3 (2013), 1007–1024. MR 3020151 Zbl 1269.14003

Received June 6, 2013. Revised April 14, 2014.

JIE WANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GEORGIA
ATHENS, GA 30602
UNITED STATES
jiewang@math.uga.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.75

CONFORMAL METRICS WITH CONSTANT CURVATURE ONE AND FINITELY MANY CONICAL SINGULARITIES ON COMPACT RIEMANN SURFACES

QING CHEN, WEI WANG, YINGYI WU AND BIN XU

A conformal metric g with constant curvature one and finitely many conical singularities on a compact Riemann surface Σ can be thought of as the pullback of the standard metric on the 2-sphere by a multivalued locally univalent meromorphic function f on $\Sigma\backslash\{\text{singularities}\}$, called the developing map of the metric g. When the developing map f of such a metric g on the compact Riemann surface Σ has reducible monodromy, we show that, up to some Möbius transformation on f, the logarithmic differential $d(\log f)$ of f turns out to be an abelian differential of the third kind on Σ, which satisfies some properties and is called a character 1-form of g. Conversely given such an abelian differential ω of the third kind satisfying the above properties, we prove that there exists a unique 1-parameter family of conformal metrics on Σ such that all these metrics have constant curvature one, the same conical singularities, and have ω as one of their character 1-forms. This provides new examples of conformal metrics on compact Riemann surfaces of constant curvature one and with singularities. Moreover we prove that the developing map is a rational function for a conformal metric g with constant curvature one and finitely many conical singularities with angles in $2\pi\mathbb{Z}_{>1}$ on the two-sphere.

1. Introduction

Let Σ be a compact Riemann surface and p a point on Σ. A conformal metric g on Σ has a *conical singularity* at p with *singular angle* $2\pi\alpha > 0$ if in a neighborhood of p, $g = e^{2\varphi}|dz|^2$, where z is a local complex coordinate defined in the neighborhood of p with $z(p) = 0$ and $\varphi - (\alpha - 1)\ln|z|$ is continuous in the neighborhood. Let

Chen is supported in part by the National Natural Science Foundation of China (grant no. 11271343). Wu is supported in part by the National Natural Science Foundation of China (grant no. 11071249) and the President Fund of UCAS. Xu is supported in part by Anhui Provincial Natural Science Foundation (grant no. 1208085MA01) and the Fundamental Research Funds for the Central Universities (grant no. WK0010000020).

MSC2010: primary 32Q30; secondary 34M35.

Keywords: conformal metric of constant curvature one, conical singularity, developing map, character 1-form.

$p_1, \dots, p_n$ be points of Σ and g a conformal metric on Σ with conical singularity at p_j of singular angle $2\pi\alpha_j > 0$ for $j = 1, \dots, n$. Then we say that the metric g represents the divisor $D := \sum_{j=1}^{n}(\alpha_j - 1)P_j$. The Gauss–Bonnet formula says that the integral of the curvature on Σ equals 2π times

$$\chi(\Sigma) + \deg D,$$

where $\chi(\Sigma)$ denotes the Euler number of Σ and $\deg D = \sum_j(\alpha_j - 1)$ the degree of the divisor D. A classical problem is whether there exists a conformal metric on Σ of constant curvature K representing the divisor D. If $K \le 0$, then the unique metric exists if and only if the left-hand side $\chi(\Sigma) + \deg D \le 0$; see [McOwen 1988; Troyanov 1991].

If $\chi(\Sigma) + \deg D > 0$, or equivalently $K \equiv 1$ if we multiply the original metric by some constant, the problem turns to be quite subtle and is still open now, except that there are some partial results. Troyanov [1989] considered the case of two points on the sphere and proved that the necessary and sufficient condition in this case is $\alpha_1 = \alpha_2$. A more general result also due to him [1991, Theorem 4] says that there exists a metric of constant positive curvature if

$$0 < \chi(\Sigma) + \deg D < \min\{2, 2\min\alpha_j\}. \tag{1}$$

Luo and Tian [1992] proved that the above condition is also necessary and the metric is unique, provided that Σ is the 2-sphere and all angles lie in $(0, 2\pi)$. In case that Σ is a sphere and the divisor D is supported at three points, [Umehara and Yamada 2000; Eremenko 2004; Furuta and Hattori 1998; Fujimori et al. 2011] give a necessary and sufficient condition for the existence of the metric, which is also unique if and only if none of the three angles belongs to $2\pi\mathbb{Z}_{>0}$.

We attack the problem by using the idea of a developing map, due to R. Bryant [1988, pp. 333–4], Umehara and Yamada [2000, p. 76] and Eremenko [2004, p. 3350]. Let g be a conformal metric of constant curvature one on Σ representing the divisor D. Let $\Sigma^* = \Sigma \backslash \{p_1, \dots, p_n\}$. Every point p in Σ^* has a neighborhood U_p isometric (so conformal) to an open set $\mathfrak{U}_p$ of the Riemann sphere $\overline{\mathbb{C}}$ endowed with the standard metric g_{st}. Denoting by $\mathfrak{f}_p : U_p \to \mathfrak{U}_p$ this isometry (conformal map), Umehara and Yamada [2000] and Eremenko [2004] claimed that $\mathfrak{f}_p$ can be extended to the whole of Σ^* by analytic continuation such that the extension gives a multivalued locally univalent meromorphic function f on Σ^*, whose monodromy belongs to the group $\mathrm{PSU}(2)$ of orientation-preserving isometries of $\overline{\mathbb{C}}$ (see Lemma 2.1). Hence the metric g can be thought of as the pullback

$$g = \frac{4|f'(z)|^2|dz|^2}{(1 + |f(z)|^2)^2}$$

under f of g_{st}. Moreover, prompted by [Umehara and Yamada 2000, (2.10)] and [Eremenko 2004, (2)], we show in Lemma 3.1 that the Schwarzian $\{f, z\}$ of f in a neighborhood U_j of p_j with complex coordinate z with $z(p_j) = 0$ has the form

$$\{f, z\} := \frac{f'''(z)}{f'(z)} - \frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2 = \frac{c_j}{z^2} + \frac{d_j}{z} + \psi_j(z),$$

where $c_j = (1-\alpha_j^2)/2$, the d_j are constants and the ψ_j are holomorphic functions in U_j, dependent on the complex coordinate z. Since the value of c_j is independent of the choice of the complex coordinate z, we say that f is *compatible with the divisor* $D = \sum_j (\alpha_j - 1) P_j$. We now arrive at:

Definition 1.1 [Umehara and Yamada 2000, p. 76]. Let g be a conformal metric on Σ of constant curvature one representing the divisor D. We call a multivalued locally univalent meromorphic function f on Σ^* a *developing map* of the metric g if $g = f^* g_{\mathrm{st}}$.

On the other hand, if there exists a multivalued meromorphic function f on Σ^* which is compatible with the divisor D and has monodromy in $\mathrm{PSU}(2)$, then there exists a conformal metric $g = f^* g_{\mathrm{st}}$ with constant curvature one and representing D (see Lemma 3.2). Therefore we can sum up the above into a necessary and sufficient condition (see Theorem 3.4) for the existence problem of conical conformal metrics of constant curvature one on Σ.

In this manuscript, we mainly focus on a special class of conical conformal metrics of constant curvature one, called *reducible metrics*, which we can classify by using abelian differentials of the third kind.

Definition 1.2 [Umehara and Yamada 2000, p. 76]. We call a conformal metric g on Σ of constant curvature one and with finitely many conical singularities an *irreducible metric* if the monodromy group of a developing map of the metric g cannot be diagonalized, that is, the monodromy group has no fixed point on the Riemann sphere $\overline{\mathbb{C}}$ (see Lemma 4.1). We call g *reducible* if the monodromy group has at least one fixed point on $\overline{\mathbb{C}}$. We call a reducible metric (*non*)*trivial* if the monodromy of a developing map of the metric is (non)trivial. Lemma 2.2 tells us that these definitions do not depend on the choice of a developing map.

A trivial reducible metric is a pullback of g_{st} under some rational function on Σ (Lemma 4.2). Each subgroup of $\mathrm{PSU}(2)$ having at least one fixed point on $\overline{\mathbb{C}}$ is abelian, and, up to conjugacy, can be thought of as a subgroup of the standard maximal torus

$$\mathrm{U}(1) = \left\{\mathrm{diag}\left(e^{\sqrt{-1}\theta}, e^{-\sqrt{-1}\theta}\right) \mid \theta \in \mathbb{R}\right\}/\{\pm I_2\}$$

of PSU(2) (see Lemma 4.1). Each fractional transformation in U(1) is multiplication by $e^{2\sqrt{-1}\theta}$. Therefore, for a nontrivial reducible metric g on Σ, by Lemma 2.2 there exist exactly two developing maps, say f and $1/f$, of the metric g, whose monodromies belong to U(1).

Definition 1.3. Let g be a nontrivial reducible metric on a compact Riemann surface Σ. We call a developing map f of g *multiplicative* if the monodromy of f belongs to U(1). Such an f is unique up to taking the reciprocal, and the logarithmic differential

$$\omega := d(\log f) = \frac{df}{f}$$

of the multiplicative developing map f is a meromorphic 1-form on Σ^*. Actually, ω can be extended to be an abelian differential of the third kind on Σ (see Lemma 4.3), which we call the *character 1-form* of the reducible metric g. Hence the character 1-form of a nontrivial reducible metric is unique up to sign.

Let g be a trivial reducible metric on Σ. By Lemma 4.2, there exists a rational function $f : \Sigma \to \overline{\mathbb{C}}$ such that $g = f^* g_{\mathrm{st}}$. By Lemma 2.2, each developing map of the metric g is a rational function, and multiplicative. We call a *character 1-form* of the metric g the logarithmic differential of a developing map of g. The character 1-forms of the trivial reducible metric g are automatically abelian differentials of the third kind on Σ.

To set up the notation for stating the properties of character 1-forms, we need say something more about the standard metric g_{st} on the Riemann sphere $\overline{\mathbb{C}}$, which is a trivial reducible metric. The set of all developing maps of g_{st} can be identified with the group PSU(2). Up to taking the reciprocal, any two developing maps of g_{st}, fixing 0 and ∞, respectively, differ by a multiple complex constant with modulus 1. Up to sign, the logarithmic differentials of all the developing maps, leaving the set $\{0, \infty\}$ invariant, coincide with the abelian differential $\Theta := d(\log w) = dw/w$, which has two simple poles of 0 and ∞. The residues of Θ at 0 and ∞ equal 1 and -1, respectively. The algebraic dual $X := w\partial/\partial w$ of Θ is a meromorphic vector field with two simple zeroes of 0 and ∞. The index of X equals 1 at both 0 and ∞. $\Phi(w) = 4|w|^2/(1+|w|^2)$ is a smooth Morse function on $\overline{\mathbb{C}}$, whose complex gradient field $\Phi^{\cdot,w}\partial/\partial w$ equals X. Moreover, Φ has only two critical points, which are the minimal point 0 and the maximal point ∞. Consider a multiplicative developing map f of a reducible metric g on Σ. Up to sign, the character 1-form $\omega = df/f = d(\log f)$ equals the pullback $f^*\Theta$ of Θ by f. Denote by $Y := (f(z)/f'(z))(\partial/\partial z)$ the algebraic dual vector field of ω, which is a meromorphic vector field on Σ. Then Y equals the complex gradient field $\Psi^{\cdot,z}\partial/\partial z$ of the smooth function $\Psi(z) = 4|f(z)|^2/(1+|f(z)|^2)$ on Σ^*, which can be continuously extended to Σ (see Lemma 4.3).

Using the above notations, we state more precisely the properties of the character 1-form of a reducible metric.

Theorem 1.4. *Let g be a reducible metric representing the divisor*

$$D=\sum_{j=1}^{n}(\alpha_j-1)P_j \quad \text{with } 1\neq\alpha_j>0.$$

Let f be a developing map of g, and let f be multiplicative if g is nontrivial such that the character 1-form $\omega=df/f$ of g equals $f^*\Theta$. Let Y be the algebraic dual vector field of ω, and $\Psi(z)=4|f(z)|^2/(1+|f(z)|^2)$. Then the following statements hold:

(1) *The set of zeroes of the meromorphic vector field Y coincides with the extremal point set of the function Ψ. Each zero of Y is simple, and Y vanishes at each point p_j where $\alpha_j>0$ is a noninteger. The set of poles of Y coincides with the saddle point set of Ψ. Each pole of Y is some conical singularity p_j of the reducible metric g, where α_j is an integer greater than 1 and the order of the pole p_j of Y equals α_j-1.*

(2) *Let $p_1,\dots,p_J$ be the saddle points of Ψ, let $p_{J+1},\dots,p_n$ be the singular extremal points of Ψ, and let $e_1,\dots,e_S$ be the smooth extremal points of Ψ on Σ^*. Then the canonical divisor of the character 1-form ω has the form*

$$(\omega)=\sum_{j=1}^{J}(\alpha_j-1)P_j-\sum_{k=J+1}^{n}P_k-\sum_{\ell=1}^{S}E_\ell.$$

In particular, each pole of ω is simple; that is, ω is an abelian differential of the third kind. The residue of ω at the pole e_ℓ equals 1 or -1, where e_ℓ is a minimal or maximal point of Ψ; the residue of ω at the pole p_k equals α_j or $-\alpha_j$, where p_k is a minimal or maximal point of Ψ. Moreover the real part of ω is exact on $\Sigma':=\Sigma\backslash\{p_{J+1},\dots,p_n,e_1,\dots,e_S\}$:

$$2\Re\omega=d(\log|f|^2).$$

(3) *The developing map f extends over $\Sigma'\cup\Sigma^*$ and has the expression*

$$f(z)=C\exp\left(\int^z\omega\right)$$

for some nonzero complex constant C. In particular, the local monodromy of f around each p_j $(1\le j\le J)$ is trivial, and the limit $\lim_{p\to p_j}f(p)$ exists and belongs to $\mathbb{C}\backslash\{0\}$. If we continue analytically a function element $\mathfrak{f}$ of f along a simple and sufficiently small loop winding around p_k $(J+1\le k\le n)$ counterclockwise, then we obtain $\mathfrak{f}\exp(2\pi\sqrt{-1}\alpha_k)$. The limit $\lim_{p\to p_k}|f(p)|$ exists, and equals 0 or $+\infty$, provided p_k is a minimal or maximal point of $\Psi(z)$. This is also the case for e_ℓ.

Using abelian differentials of the third kind with the above properties, we can construct new examples of conformal metrics with constant curvature one and with finitely many conical singularities.

Theorem 1.5. *Let ω be an abelian differential of the third kind having poles on a compact Riemann surface Σ, whose residues are all nonzero real numbers and whose real part is exact outside the set of poles of ω. Then there exists a unique 1-parameter family $\{g_\lambda : \lambda \in (0, +\infty)\}$ of reducible metrics on Σ such that ω is one of the character 1-forms of each metric $g_\lambda := f_\lambda^\star g_{\mathrm{st}}$, where*

$$f_\lambda(z) = \lambda \cdot \exp\left(\int^z \omega\right)$$

is a multivalued locally univalent meromorphic function on $\Sigma\backslash\{\text{poles of } \omega\}$ with monodromy in $\mathrm{U}(1)$. Suppose that the canonical divisor of ω has the form

$$(\omega) = \sum_{j=1}^{J} (\alpha_j - 1) P_j - \sum_{k=J+1}^{N} Q_k,$$

where the α_j are integers > 1. Then the divisor D represented by g_λ has the form

$$D = \sum_{j=1}^{J} (\alpha_j - 1) P_j + \sum_{k=J+1}^{N} (|\operatorname{Res}_{Q_k}(\omega)| - 1) Q_k.$$

Moreover, the g_λ are trivial reducible metrics if and only if the integral of ω on each loop in $\Sigma\backslash\{\text{poles of } \omega\}$ is $2\pi\sqrt{-1}$ times an integer. In particular, each residue of ω is an integer.

Remark 1.6. Each reducible metric does not satisfy Troyanov's condition (1) (see Corollary 4.5). Therefore we obtain a class of new examples of conformal metrics of constant curvature one with finitely many singularities, since there exist plenty of abelian differentials of the third kind satisfying the condition in Theorem 1.5 (see [Springer 1957, Corollary 8-3]).

Troyanov's condition (1) for the corresponding irreducible metrics depends only on the values of angles. Example 4.7 shows that the existence of reducible metrics does not depend only on angles but also on the position of singularities.

Remark 1.7. Umehara and Yamada [2000] called a nontrivial reducible metric $\mathscr{H}^1$-reducible, and a trivial reducible metric $\mathscr{H}^3$-reducible.

Remark 1.8. Besides the previous reference, our motivation for defining character 1-form comes from [Chen and Wu 2011; Chen et al. 2013], where the authors use character 1-form to completely classify HCMU metrics of nonconstant curvature on compact Riemann surfaces. We will say more about this in the ending of Section 4.

Theorem 1.9. *A conformal metric of constant curvature one, representing an effective $\mathbb{Z}$-divisor D on the 2-sphere is a trivial reducible metric. That is, it is the pullback under some rational function f on the 2-sphere of the standard metric g_{st} on the Riemann sphere $\overline{\mathbb{C}}$.*

Remark 1.10. The case where D is an effective $\mathbb{Z}$-divisor supported at two or three points was proved in [Troyanov 1989; Furuta and Hattori 1998; Umehara and Yamada 2000; Eremenko 2004]. However Theorem 1.9 does not hold in general for compact Riemann surfaces of nonzero genus (see Example 4.6).

Since the rational function f in Theorem 1.9 has ramified divisor D, the theorem reduces the problem of characterizing conformal metrics of constant curvature one and representing an effective $\mathbb{Z}$-divisor D on the two-sphere to the following classical one: Which kind of divisors D can be a ramification divisor of some rational function, and how many equivalent classes of rational functions have the prescribed ramification divisor? Here we say that two rational functions are *equivalent* if one of them is given by the postcomposition of the other with a Möbius transformation.

When the points in the support of the ramified divisor lie in a general position, L. Goldberg [1991] solved a special case of the latter problem, where each ramified order equals one; I. Scherbak [2002] gave a complete answer for the general case. [Eremenko and Gabrielov 2002; Eremenko et al. 2006] proved that there exists a real rational function in each equivalence class if each point in the support of D is real.

Remark 1.11. Since we only consider conformal metrics with finite area in this manuscript, the singularities of "zero angle" would not show up (see [Bryant 1988, Proposition 4]).

We explain the organization of this paper. In Section 2, we shall first make a detailed exposition on developing map in Lemmas 2.1 and 2.2. We compute, in Lemma 3.1 of Section 3, the Schwarzian of a developing map of a conformal metric with constant curvature one representing a divisor D. In Lemma 3.2 we show that the converse of Lemma 3.1 also holds. Then Theorem 1.9 follows from these two lemmas. In Section 4 we prove Theorems 1.4 and 1.5 as applications of Lemma 3.2. These two theorems are applied to give some examples of irreducible and reducible metrics in Corollary 4.5 and Examples 4.6 and 4.7. In Section 5, we give an alternative proof of a theorem of Troyanov [1989], as an application of Theorems 1.9 and 1.4. Moreover we discuss the nonuniqueness of reducible metrics representing a given divisor D. Last, we propose some questions about both irreducible and reducible metrics.

2. Existence of developing maps and their monodromy

Lemma 2.1. *Let g be a conformal metric on a compact Riemann surface Σ of constant curvature one, representing the divisor $D = \sum_{j=1}^{n}(\alpha_j - 1)P_j$ with $\alpha_j > 0$. Then there exists a multivalued locally univalent holomorphic map f from $\Sigma^* := \Sigma \backslash \{p_1, \dots, p_n\}$ to the Riemann sphere $\overline{\mathbb{C}}$ such that the monodromy of f belongs to* PSU(2) *and*

$$g = f^* g_{\mathrm{st}},$$

where $g_{\mathrm{st}} = 4|dw|^2/(1+|w|^2)^2$ is the standard metric over $\overline{\mathbb{C}}$.

Proof. Denote by $d(\,\cdot\,,\cdot\,)$ the distance on Σ induced by the metric g. Choose an arbitrary point p in Σ^* and fix it. Take a positive number $r = r_p$ sufficiently small such that $d(p, \{p_1, \dots, p_n\}) > r$ and there exists a geodesic polar coordinate chart in the open metric ball $B(p, r) = \{q \in \Sigma \mid d(p, q) < r\} \subset \Sigma^*$. Choose a positively oriented orthonormal basis $\{e^1, e^2\} = \{e^1_p, e^2_p\}$ of the tangent space $T_p\Sigma$. Choose an arbitrary point $\mathfrak{p} \in \overline{\mathbb{C}}$ and fix it. Since the Gauss curvature of $(B(p, r), g)$ is constant and equals one, by a theorem of Riemann [Petersen 2006, p. 136], there exists an open metric ball $\mathfrak{B}(\mathfrak{p}, r)$ in the Riemann sphere $(\overline{\mathbb{C}}, g_{\mathrm{st}})$ and an orientation-preserving isometry $\mathfrak{f}_p$ from $(B(p, r), g)$ onto $(\mathfrak{B}(\mathfrak{p}, r), g_{\mathrm{st}})$. Let $\mathfrak{e}^1_{\mathfrak{p}} := \mathfrak{f}_*(e^1)$ and $\mathfrak{e}^2_{\mathfrak{p}} := \mathfrak{f}_*(e^2)$, which also form a positively oriented orthonormal basis of $T_{\mathfrak{p}}\overline{\mathbb{C}}$. Then $\mathfrak{f}_p$ is a conformal map from $B(p, r)$ to $\mathfrak{B}(\mathfrak{p}, r)$.

Take an arbitrary point q in Σ^* and a curve $L : [0, 1] \to \Sigma^*$ joining p to q. Then there exists some

$$0 < \delta < \min\bigl(d\bigl(\gamma([0, 1]), \{p_1, \dots, p_n\}\bigr), r_p\bigr)$$

such that there exists a geodesic polar coordinate chart in the open metric ball $B(a, \delta)$ for each point a on the curve L. If we properly divide the interval $0 \le t \le 1$ into n subintervals for sufficiently large n, $0 = \gamma_0 < \gamma_1 < \cdots < \gamma_n = 1$, then the curve L splits into n subarcs $L_1, L_2, \dots, L_n$ with L_g $(g = 1, \dots, n)$ joining $L(\gamma_{g-1}) =: c_{g-1}$ to $L(\gamma_g) =: c_g$. Moreover, if we denote by $B_0, B_1, \dots, B_n$ the open metric balls with centers $a = c_0, c_1, \dots, c_n$ and with radius δ, then the closed arcs L_g lie completely in B_{g-1} for $g = 1, \dots, n$. Let f_0 be the restriction to B_0 of the conformal map $\mathfrak{f}_p : B(p, r_p) \to \mathfrak{B}(\mathfrak{p}, r_p)$. Then f_0 is an isometry (conformal map) from B_0 onto $\mathfrak{B}_0 := \mathfrak{B}(\mathfrak{p}, \delta)$. Choose a positively oriented orthonormal basis $\{e^1_{c_1}, e^2_{c_2}\}$ of $T_{c_1}\Sigma$. Since $c_1 \in B_0$, we let $\mathfrak{c}_1 := f_0(c_1) \in \mathfrak{B}_0$, $\mathfrak{e}^1_{c_1} := (f_0)_*(e^1_{c_1})$ and $\mathfrak{e}^2_{c_1} := (f_0)_*(e^2_{c_1})$. Then there exists a unique isometry $f_1 : B_1 \to \mathfrak{B}_1 := \mathfrak{B}(\mathfrak{c}_1, \delta)$ such that $f_1(c_1) = \mathfrak{c}_1$ and f_1 maps $\{e^1_{c_1}, e^2_{c_1}\}$ to $\{\mathfrak{e}^1_{\mathfrak{c}_1}, \mathfrak{e}^2_{\mathfrak{c}_1}\}$. Then $f_1 = f_0$ on $B_0 \cap B_1$. Since $c_2 \in L_2 \subset B_1$, f_1 is an analytic continuation of f_0 from point c_0 to c_2 along the arc $L_0 \cup L_1$. In this way, we obtain $f_0, \dots, f_n$, which are recursively defined on $B_0, \dots, B_n$ and give an analytic continuation of $\mathfrak{f}_p$ from p to q along the curve L. Using the same argument as [Siegel 1969, pp. 13–15], we can show that this analytic continuation is independent of the choice of division points on L. Moreover, if L^* is another curve in Σ^* joining a to b that is homotopic to L, then the result of doing analytic continuation of $\mathfrak{f}_p$ along L^* is the same as along L. Summing up, we obtain a multivalued locally isometric (univalent conformal) map f from (Σ^*, g) to $(\overline{\mathbb{C}}, g_{\mathrm{st}})$.

At last, we prove that all the monodromy of f belongs to PSU(2). Suppose $L : [0, 1] \to \Sigma^*$ is a closed curve with $L(0) = L(1) = p$. We use the notation in the previous paragraph. Recall that f_0 maps $p = c_0$ to $\mathfrak{p}$, $(f_0)_*$ maps $\{e^1, e^2\}$ to $\{\mathfrak{e}^1_{\mathfrak{p}}, \mathfrak{e}^2_{\mathfrak{p}}\}$, f_n maps $p = c_n$ to $\mathfrak{c}_n$ and $(f_n)_*$ maps $\{e^1, e^2\}$ to $\{\mathfrak{e}^1_{\mathfrak{c}_n}, \mathfrak{e}^2_{\mathfrak{c}_n}\}$. Then there exists a unique isometry $\mathfrak{L} \in \mathrm{PSU}(2)$ of $(\overline{\mathbb{C}}, g_{\mathrm{st}})$ such that $\mathfrak{L}(\mathfrak{p}) = \mathfrak{c}_n$ and $\mathfrak{L}_*$ maps $\{\mathfrak{e}^1_{\mathfrak{p}}, \mathfrak{e}^2_{\mathfrak{p}}\}$ to $\{\mathfrak{e}^1_{\mathfrak{c}_n}, \mathfrak{e}^2_{\mathfrak{c}_n}\}$. Therefore $f_n = \mathfrak{L} \circ f_0$. □

Lemma 2.2. *Any two developing maps f_1, f_2 of the metric g are related by a fractional linear transformation $\mathfrak{L} \in$ PSU(2), i.e., $f_2 = \mathfrak{L} \circ f_1$. In particular, any two developing maps of g have mutually conjugate monodromy in* PSU(2). *Then we call this conjugate class the monodromy of the metric g. The space of developing maps of the metric g has a one-to-one correspondence with the quotient group of* PSU(2) *by the monodromy group of a developing map of g.*

Proof. Take a point $p \in \Sigma^*$ and a positively oriented orthonormal basis $\{e^1, e^2\}$ of $T_p\Sigma^*$. Let $\mathfrak{f}_j$ be a function element of f_j near p for $j = 1, 2$. Denote $\mathfrak{p}_j := \mathfrak{f}_j(p)$ and $\mathfrak{e}^k_{\mathfrak{p}_j} := (\mathfrak{f}_j)_*(e^k)$ for $j, k = 1, 2$. Then there exists a unique $\mathfrak{L} \in \mathrm{PSU}(2)$ such that $\mathfrak{L}(\mathfrak{p}_1) = \mathfrak{p}_2$, and $\mathfrak{L}_*$ maps $\{\mathfrak{e}^1_{\mathfrak{p}_1}, \mathfrak{e}^2_{\mathfrak{p}_1}\}$ to $\{\mathfrak{e}^1_{\mathfrak{p}_2}, \mathfrak{e}^2_{\mathfrak{p}_2}\}$. Then we obtain the equality $\mathfrak{f}_2 = \mathfrak{L} \circ \mathfrak{f}_1$ near p, which implies $f_2 = \mathfrak{L} \circ f_1$. It follows from direct computation that the monodromy of f_1 and f_2 are mutually conjugate.

Given a developing map f and a fractional linear transformation $\mathfrak{L} \in \mathrm{PSU}(2)$, we can see $\mathfrak{L} \circ f = f$ if and only if there exists a point $p \in \Sigma^*$ and a functional element $\mathfrak{f}_p$ near p such that $\mathfrak{L} \circ \mathfrak{f}_p$ is another function element of f near p. That is, $\mathfrak{L}$ belongs to the image of the monodromy representation of $\pi_1(\Sigma^*, p)$ with respect to $\mathfrak{f}$. Therefore PSU(2) acts in this way transitively on the set of all developing maps with isotropy group isomorphic to the monodromy group. □

Remark 2.3. The developing maps also exist for flat or hyperbolic conformal metrics with finitely many conical or cusp singularities, and analogues of Lemmas 2.2 and 3.2 hold.

3. The Schwarzian of a developing map

Lemma 3.1. *Let g be a conformal metric of constant curvature one on a compact Riemann surface Σ, and suppose g represents a divisor $D = \sum_{j=1}^{n} (\alpha_j - 1) P_j$, where $\alpha_j > 0$ for all j. Suppose that $f : \Sigma^* = \Sigma \backslash \{p_1, \ldots, p_n\} \to \overline{\mathbb{C}}$ is a developing* *map of g. Then the Schwarzian $\{f, z\}$ of f equals*

$$\{f, z\} = \frac{1-\alpha_j^2}{2z^2} + \frac{d_j}{z} + \psi_j(z)$$

in a neighborhood U_j of p_j with complex coordinate z and $z(p_j) = 0$, where the d_j are constants and the ψ_j are holomorphic functions in U_j, depending on the complex coordinate z.

Proof. If we rewrite the metric $g = 4|f'(z)|^2|dz|^2/(1+|f(z)|^2)^2$ as $g = e^{2u}|dz|^2$, then we find $u = \log|f'(z)| + \log 2 - \log(1+|f|^2)$. The lemma on page 300 of [Troyanov 1989] tells us that

$$\eta(z) = 2\left(\frac{\partial^2 u}{\partial z^2} - \left(\frac{\partial u}{\partial z}\right)^2\right)dz^2$$

defines a *projective connection* compatible with the divisor D. The interested reader could find in the same reference the definition of the projective connection, which we will not use in this paper. The compatibility of the projective connection η with the divisor D on page 300 of [Troyanov 1989] means that

$$\eta(z) = \left(\frac{1-\alpha_j^2}{2z^2} + \frac{d_j}{z} + \phi_j(z)\right)dz^2, \quad \phi_j \text{ holomorphic},$$

where z is the complex coordinate near p_j. Since the developing map f is a projective multivalued function on Σ^*, its Schwarzian $\{f, z\}$ with respect to the complex coordinate z near p_j is a single-valued function of z. At last, we find

$$\begin{aligned}
2\left(\frac{\partial^2 u}{\partial z^2} - \left(\frac{\partial u}{\partial z}\right)^2\right) &= 2\frac{\partial}{\partial z}\left(\frac{f''(z)}{2f'(z)} - \frac{f'(z)\bar{f}}{1+|f|^2}\right) - 2\left(\frac{f''(z)}{2f'(z)} - \frac{f'(z)\bar{f}}{1+|f|^2}\right)^2 \\
&= \left(\frac{f'''(z)}{f'(z)} - \left(\frac{f''(z)}{f'(z)}\right)^2 - \frac{2f''(z)\bar{f}(z)}{1+|f|^2} + 2\left(\frac{f'(z)\bar{f}}{1+|f|^2}\right)^2\right) \\
&\quad - \left(\frac{1}{2}\left(\frac{f''(z)}{f'(z)}\right)^2 - \frac{2f''(z)\bar{f}(z)}{1+|f|^2} + 2\left(\frac{f'(z)\bar{f}}{1+|f|^2}\right)^2\right) \\
&= \frac{f'''(z)}{f'(z)} - \frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2 = \{f, z\}. \qquad \square
\end{aligned}$$

A multivalued locally univalent meromorphic function h on Σ^* is said to be *projective* if any two function elements $\mathfrak{h}_1, \mathfrak{h}_2$ of h near a point $p \in \Sigma^*$ are related by a fractional linear transformation $L \in \mathrm{PGL}(2, \mathbb{C})$, i.e., $\mathfrak{h}_1 = L \circ \mathfrak{h}_2$.

Lemma 3.2. *Let $f : \Sigma^* \to \overline{\mathbb{C}}$ be a projective multivalued locally univalent meromorphic function, and suppose the monodromy of f belongs to a maximal compact subgroup of* PGL(2, $\mathbb{C}$). *If f is compatible with the divisor $D = \sum_j (\alpha_j - 1)P_j$, then* *there exists a neighborhood U_j of p_j with complex coordinate z and $L_j \in \mathrm{PGL}(2, \mathbb{C})$ such that $z(p_j) = 0$ and $g_j = L_j \circ f$ has the form $g_j(z) = z^{\alpha_j}$, where $0 < \alpha_j \neq 1$ and $c_j = (1-\alpha_j^2)/2$. Moreover there exists $\mathfrak{L} \in \mathrm{PGL}(2, \mathbb{C})$ such that the pullback $(\mathfrak{L} \circ f)^* g_{\mathrm{st}}$ of the standard metric g_{st} by $\mathfrak{L} \circ f$ is a conformal metric of constant curvature one which represents the divisor $D = \sum_j (\alpha_j - 1) P_j$. In particular, if the monodromy of f belongs to* PSU(2), *then the fractional linear transformation $\mathfrak{L}$ turns out to be the identity map.*

Proof. Recall the well-known fact that every maximal compact group of $\mathrm{PGL}(2, \mathbb{C})$ is conjugate to the subgroup PSU(2). There exists a fractional linear transformation $\mathfrak{L}$ such that the monodromy of $\mathfrak{L} \circ f$ belongs to PSU(2). Hence we may assume that this is the case for f without loss of generality.

We first show the first statement of Lemma 3.2: that there exists a neighborhood U_j of p_j with complex coordinate z and some $L_j \in \mathrm{PGL}(2, \mathbb{C})$ such that $g_j = L_j \circ f$ has the form z^{α_j}. Since f is compatible with D, we could choose a neighborhood U_j of p_j and a complex coordinate x on U_j such that $x(p_j) = 0$ and

$$\{f, x\} = \frac{c_j}{x^2} + \frac{d_j}{x} + \phi_j(x),$$

where $\phi_j(x)$ is holomorphic in U_j and $c_j := (1-\alpha_j^2)/2$. By [Yoshida 1987, Proposition, p. 39], in the neighborhood U_j there are two linearly independent solutions u_0 and u_1 of the equation

$$\frac{d^2u}{dx^2} + \frac{1}{2}\left(\frac{c_j}{x^2} + \frac{d_j}{x} + \phi_j(x)\right)u = 0,$$

with single-valued coefficient such that $f(x) = u_1(x)/u_0(x)$. Actually we have $u_0 = (df/dx)^{-1/2}$ and $u_1 = f(x)u_0$. Moreover, if f changes projectively, i.e.,

$$f \mapsto \frac{af+b}{cf+d} \quad \text{with } ad-bc=1,$$

then u_0 and u_1 change linearly, i.e.,

$$\begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \mapsto \begin{pmatrix} d & c \\ b & a \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix},$$

and vice versa.

Define an operator $L_j := x^2(d^2/dx^2) + q_j(x)$ with $q_j(x) = (c_j + d_j x + x^2\phi_j(x))/2$. Then both u_0 and u_1 are solutions of the equation $L_j u = 0$. Since the monodromy of f belongs to PSU(2), the cyclic group generated by the local monodromy of the equation $L_j u = 0$ around $x = 0$ is contained in a maximal compact subgroup of $\mathrm{PGL}(2, \mathbb{C})$ conjugate to PSU(2). Note that the equation $L_j u = 0$ has a regular singularity at 0. We could apply the Frobenius method (see [Yoshida 1987, §2.5]) to solve it. Note that the indicial equation

$$f(s) = s(s-1) + \frac{1-\alpha_j^2}{4} = 0 \tag{2}$$

of the differential equation $L_j u = 0$ at $x = 0$ has roots $s_0 = (1-\alpha_j)/2$ and $s_1 = (1+\alpha_j)/2$, and $s_1 - s_0 = \alpha_j > 0$. Let $\sum_{k=0}^{\infty} b_k x^k$ be the power series expansion of $q_j(x)$, where $b_0 = c_j/2$. Let s be a parameter. Then $u(s,x) = x^s \sum_{k=0}^{\infty} c_k(s) x^k$, with $c_0(s) \equiv 1$, is an solution of $L_j u = 0$ if and only if the equation

$$f(s+n)c_n + R_n = 0 \tag{$\sharp_n$}$$

holds for all $n = 0, 1, 2, \ldots$, where

$$R_0 = 0, \quad \text{and, for } n > 0, \quad R_n = R_n(c_1, \ldots, c_{n-1}, s) = \sum_{i=0}^{n-1} c_i b_{n-i}.$$

Note that the equation ($\sharp_0$) is exactly the indicial equation (2). Since $f(s_1 + n) \neq 0$ for all $n \geq 1$, we find that $u(s_1, x)$ is a solution of the equation.

Case 1. Suppose that $s_1 - s_0 = \alpha_j$ is not an integer. Then, by the same reasoning, $u(s_0, x)$ is another solution, which is linearly independent of $u(s_1, x)$. Summing up, we have

$$u(s_0, x) = x^{s_0}(1+\psi_0) \quad \text{and} \quad u(s_1, x) = x^{s_1}(1+\psi_1),$$

where both ψ_0 and ψ_1 are holomorphic functions vanishing at 0. Here we take a smaller neighborhood of 0 than U_j to assure the convergence of the power series defining ψ_k if necessary. Since both $u_0(x)$ and $u_1(x)$ are linear combinations of $u(s_0, x)$ and $u(s_1, x)$, $f(x) = u_1(x)/u_0(x)$ equals some fractional linear transform of $u(s_1, x)/u(s_0, x)$. For simplicity of notation, we may assume $f(x) = u(s_1, x)/u(s_0, x)$ equals x^{α_j} times a holomorphic function $\varphi_j(x)$ with $\varphi_j(0) = 1$. Therefore we could choose another complex coordinate $z = z(x)$ of U_j under which $f = f(z) = z^{\alpha_j}$.

Case 2. Suppose that $m := s_1 - s_0 = \alpha_j$ is an integer ≥ 2.

Subcase 2.1. If $R_m = 0$, we can solve the equation ($\sharp_n$) for $s = s_0$ for all $n \geq 1$ by choosing c_m arbitrarily and obtain another solution $u(s_0, x)$ linearly independent of $u(s_1, x)$. An argument similar to that in Case 1 completes the proof.

Subcase 2.2. Suppose $R_m \neq 0$. Define

$$u^* = x^{s_0} \sum_{k=0}^{\infty} c_k(s_0) x^k,$$

where $c_0 = 1$, the c_j $(1 \leq j < m)$ are determined by ($\sharp_j$), while c_m is arbitrarily fixed, and the c_j $(j > m)$ are determined also by ($\sharp_j$). Then the linear combination of u^* and $(\partial/\partial s)u(s, x)|_{s=s_1}$

$$U_0(x) := f'(s_1)u^* - R_m \frac{\partial}{\partial s} u(s, x)|_{s=s_1}$$

is a solution. It should be mentioned that since f and R_n are holomorphic with respect to s, this is also the case for the c_n. Then we correct a typo in [Yoshida 1987, p. 23] and find the two linearly independent solutions, given by

$$\begin{pmatrix} U_0(x) \\ u(s_1, x) \end{pmatrix} = \begin{pmatrix} x^{s_0} & x^{s_1} \log x \\ 0 & x^{s_1} \end{pmatrix} \begin{pmatrix} f'(s_1) \sum_{k=0}^{\infty} c_k(s_0) x^k - R_m x^m \sum_{k=0}^{\infty} c'_k(s_1) x^k \\ \sum_{k=0}^{\infty} c_k(s_1) x^k \end{pmatrix}.$$

Then the local monodromy of the equation $L_j u = 0$ at $x = 0$ is the conjugacy class in $\mathrm{PGL}(2, \mathbb{C})$ of the matrix

$$M = \begin{pmatrix} 1 & 2\pi\sqrt{-1} \\ 0 & 1 \end{pmatrix}.$$

However, the cyclic group generated by M is a free abelian group, which has no limit point under the usual topology of $\mathrm{PGL}(2, \mathbb{C})$, contradicting the fact that the monodromy of the equation $L_j u = 0$ is contained in a compact group of $\mathrm{PGL}(2, \mathbb{C})$. That is, we rule out Subcase 2.2.

Summing up, we prove the statement where α_j is an integer ≥ 2. Moreover we can also see that in this case the local monodromy at p_j is trivial; that is, p_j is an apparent singularity of the equation $L_j u = 0$ and the multivalued function f.

Since f is locally univalent on Σ^* and has monodromy belonging to $\mathrm{PSU}(2)$, $f^* g_{\mathrm{st}}$ is a well-defined smooth Riemannian metric on Σ^* with constant curvature one. The first statement proved just now implies that this metric has conical singularities at p_j with angles $2\pi\alpha_j$. □

Remark 3.3. Lemma 3.2 has some overlap with [Bryant 1988, Proposition 4], in the sense that both say the same thing near each singularity.

We sum up the above two lemmas:

Theorem 3.4. *There exists a conformal metric of constant curvature one representing a divisor D on a compact Riemann surface Σ if and only if there is a projective multivalued meromorphic function on $\Sigma^* = \Sigma \setminus \mathrm{Supp}\, D$ compatible with the divisor D and having monodromy in* $\mathrm{PSU}(2)$.

Proof of Theorem 1.4. Let f be a developing map of the metric g representing an effective $\mathbb{Z}$-divisor $\sum_j n_j P_j$ on the sphere. By Lemma 3.1, f has regular singularity of weight $(1-n_j^2)/2$ at p_j. By Lemma 2.1, the monodromy of f belongs to $\mathrm{PSU}(2)$. By Lemma 3.2, there exists $\mathfrak{L}_j \in \mathrm{PGL}(2, \mathbb{C})$ and a complex coordinate z near p such that $\mathfrak{L}_j \circ f$ has the form $f(z) = z^{n_j+1}$ near p_j, which implies the local monodromy of f at p_j is trivial. Since the sphere is simply connected, the monodromy of f is trivial, that is, f is a single-valued meromorphic function outside $\{p_j\}$. Moreover, f can be extended meromorphically onto the whole sphere; that is, f is a rational function on the sphere. $\square$

At a point $p \in \Sigma$ near which $f = f(z)$ is univalent holomorphic, we find that the Schwarzian $\{f, z\}$ is holomorphic, and vice versa (see [Yoshida 1987, Remark, p. 44]). Actually we can prove a more general result.

Lemma 3.5. *Let U be an open disk containing 0 in the complex plane $\mathbb{C}$ with coordinate w and f a projective multivalued meromorphic function on $U\backslash\{0\}$ with regular singularity of weight zero at 0. That is, $\{f, w\}$ equals d/w plus a holomorphic function $\phi(w)$, where both the constant d and $\phi(w)$ depend on the coordinate w. Assume that the subgroup of* $\mathrm{PGL}(2, \mathbb{C})$ *generated by the local monodromy of f at 0 is precompact in* $\mathrm{PGL}(2, \mathbb{C})$. *Then there exists $\mathfrak{L} \in \mathrm{PGL}(2, \mathbb{C})$ and another complex coordinate z of U such that $\mathfrak{L} \circ f(z) = z$ and $z(0) = 0$.*

Proof. Use the same argument as Case 2 of the proof of Lemma 3.2. Also note that the indicial equation here has two roots, 0 and 1. $\square$

This lemma has the following geometric consequence.

Proposition 3.6. *The conic singularity with angle 2π of a conformal metric with constant curvature one is actually a smooth point of the metric.*

4. Proof of Theorems 1.5 and 1.9

Lemma 4.1. (1) *A subgroup G of* $\mathrm{PSU}(2)$ *can be diagonalized if and only if G has a fixed point on $\overline{\mathbb{C}}$. Such a group is contained in some maximal torus $\mathbb{T}$ of* $\mathrm{PSU}(2)$. *In particular, G is abelian.*

(2) *There exists an abelian subgroup of* $\mathrm{PSU}(2)$ *which has no fixed point on $\overline{\mathbb{C}}$.*

Proof. (1) Consider the natural unitary representation ρ of $\mathrm{SU}(2)$ on $V \cong \mathbb{C}^2$ endowed with the natural Hermitian inner product $\langle \cdot, \cdot \rangle$. For each subgroup H of $\mathrm{SU}(2)$, ρ restricts to a faithful unitary representation ρ_H of H on V. Let $\widetilde{G} \subset \mathrm{SU}(2)$ be the lifting of $G \subset \mathrm{PSU}(2)$. We say that G can be *diagonalized* if the representation $(\rho_{\widetilde{G}}, V)$ can be decomposed into the direct sum of two one-dimensional subspaces:

$$V = \mathbb{C}e_1 \oplus \mathbb{C}e_2 \quad \text{and} \quad \langle e_k, e_\ell \rangle = \delta_{k\ell}.$$

Then, up to conjugacy, $\widetilde{G}$ can be viewed as a subgroup of the standard maximal torus

$$\mathrm{U}(1) = \left\{ \mathrm{diag}\left(e^{\sqrt{-1}\theta}, e^{-\sqrt{-1}\theta}\right) \mid \theta \in \mathbb{R} \right\}$$

of $\mathrm{SU}(2)$. Hence $G = \widetilde{G}/\{\pm I_2\}$ is abelian.

Looking at the Riemann sphere $\overline{\mathbb{C}}$ as the complex one-dimensional projective space $\mathbb{P}(V)$ with the natural projection $\pi : V\backslash\{0\} \to \mathbb{P}(V)$, we can see that both

$\pi(e_1)$ and $\pi(e_2)$ are two distinct fixed points of the G-action on $\mathbb{P}(V)$ if G can be diagonalized.

Suppose that G has a fixed point $\pi(e_1)$ on $\mathbb{P}(V)$ with $e_1 \in V$ and $\langle e_1, e_1\rangle = 1$. Then e_1 is a common eigenvector of all the elements in $\widetilde{G}$. Since $\rho_{\widetilde{G}}$ is a unitary representation on V, it can be decomposed into $V = \mathbb{C}e_1 \oplus \mathbb{C}e_2$, where e_2 is a unit vector orthogonal to e_1. That is, G can be diagonalized.

(2) The abelian subgroup D_2 of PSU(2) generated by

$$z \mapsto -z \quad \text{and} \quad z \mapsto \frac{1}{z}$$

has no fixed point on $\overline{\mathbb{C}}$. □

Lemma 4.2. *A trivial reducible metric g representing a divisor D on a compact Riemann surface Σ is the pullback f^*g_{st} of g_{st} by some rational function f on Σ.*

Proof. Let $f : \Sigma^* \to \overline{\mathbb{C}}$ be a developing map of the metric g. Since the monodromy of f is trivial, so is the local monodromy of f around each point in the support of D. By Lemma 3.2 the divisor D must be an effective $\mathbb{Z}$-divisor. Using Lemma 3.2 again, we find that the holomorphic map $f : \Sigma^* \to \overline{\mathbb{C}}$ has a holomorphic extension to Σ. □

Lemma 4.3. *Suppose that g is a reducible metric on Σ and f a multiplicative developing map of g. Then the holomorphic 1-form $d(\log f)$ on Σ^* can be extended to be an abelian differential of the third kind on Σ. The function $\Psi = 4|f(z)|^2/(1+|f(z)|^2)$ on Σ^* can be extended continuously to Σ.*

Proof. The proof is contained in the following proof of Theorem 1.4. □

Proof of Theorem 1.4. (1–2) We show that if a point $p \in \Sigma^*$ is a zero of $Y(z) = (f(z)/f'(z))(\partial/\partial z)$, then p is simple. We choose a function element $\mathfrak{f}$ near p. Since the monodromy of f belongs to U(1), $Y = (\mathfrak{f}(z)/\mathfrak{f}'(z))\partial/\partial z$, which is independent of the choice of the function element $\mathfrak{f}$ and the complex coordinate z. Since $\mathfrak{f}$ is a univalent meromorphic function near p, there exist $\mathfrak{L} \in \mathrm{PGL}(2, \mathbb{C})$ and a complex coordinate z near p with $z(p) = 0$ such that $\mathfrak{L} \circ f = z$. Then $\mathfrak{f} = (az+b)/(cz+d)$ with $ad - bc = 1$ near p, and $Y = (az+b)(cz+d)\partial/\partial z$. It is clear that p cannot be a pole of Y. Since $Y = 0$ at $z(p) = 0$, $bd = 0$.

Case 1. Since $ad - bc = 1$ and $bd = 0$, we assume $b = 0$, $d \neq 0$ in this case. Then $ad = 1$ and $\mathfrak{f}(z) = az/(cz+d)$. Then $Y = az(cz+d)\partial/\partial z$ has a simple zero at $z(p) = 0$. Hence $\omega = dz/(az(cz+d))$ has residue 1 at p. Since $f(0) = 0$, p is a minimal point of Ψ and $\Psi(p) = 0$.

Case 2. Similarly, when $d = 0$ and $b \neq 0$, Y also has simple zero at p, ω has residue -1, $\lim_{q\to p} |f(q)| = +\infty$, $\lim_{q\to p} \Psi(q) = 4$, and p is a maximal point of Ψ.

We show that each point $q \in \{p_1, \dots, p_n\}$ must be a simple zero of Y, provided the conical angle of the metric g at q equals $2\pi s > 0$ and s is a noninteger. By Lemmas 3.1 and 3.2, we can choose a function element $\mathfrak{f}$ near q and a complex coordinate z near q such that $\mathfrak{f} = (az^s + b)/(cz^s + d)$ with $ad - bc = 1$. On the other hand, since the monodromy of f belongs to $\mathrm{U}(1)$, so does the local monodromy of $\mathfrak{f}$. Then there exists $\theta \in \mathbb{R}$ such that

$$e^{2\pi\sqrt{-1}\theta}\mathfrak{f} = e^{2\pi\sqrt{-1}\theta}\frac{az^s + b}{cz^s + d} = \frac{ae^{2\pi\sqrt{-1}s}z^s + b}{ce^{2\pi\sqrt{-1}s}z^s + d}.$$

This is equivalent to the following equalities holding:

$$\begin{aligned}
&ace^{2\pi\sqrt{-1}s}\big(1 - e^{2\pi\sqrt{-1}\theta}\big) = 0,\\
&\big(ade^{2\pi\sqrt{-1}s} + bc\big) - e^{2\pi\sqrt{-1}\theta}\big(bce^{2\pi\sqrt{-1}s} + ad\big) = 0,\\
&bd\big(1 - e^{2\pi\sqrt{-1}\theta}\big) = 0.
\end{aligned}$$

Solving the equation, we find that either $c = b = 0$ or $a = d = 0$, that is, $\mathfrak{f}(z)$ equals μz^s ($\mu \neq 0$) or λz^{-s} ($\lambda \neq 0$). Hence $Y = \pm sz(\partial/\partial z)$ has a simple zero at $z(q) = 0$. Since $\mathfrak{f}(0)$ equals 0 or ∞, Ψ is continuous at q, which is a minimal or maximal point of Ψ achieving value 0 or 4 if and only if ω has residue $s > 0$ or $-s < 0$ at p.

Let p be a singular point of the metric g with conical angle 2π times an integer $n > 1$. Then

$$\mathfrak{f}(z) = \frac{az^n + b}{cz^n + d} \quad \text{and} \quad Y = \frac{(az^n + b)(cz^n + d)}{nz^{n-1}}\frac{\partial}{\partial z}.$$

Case A. Assume $bd \neq 0$. Then p is a pole of Y with order $n-1$ and a zero of ω with order $n-1$, and $\lim_{z\to p} f(z) = f(0) = b/d \in \mathbb{C}\setminus\{0\}$. Moreover Ψ is continuous at p, which is a saddle point of Ψ.

Case B. Assume $bd = 0$. Then it is easy to check that p is a simple zero of Y. If $b = 0$ (resp. $b \neq 0$), then $\lim_{q\to p} |f(p)|$ equals 0 (resp. $+\infty$), where ω has residue n (resp. $-n$). Ψ is continuous at p, where it achieves the minimal value 0 or the maximal value 4.

(3) The local monodromy property of f follows from Lemmas 3.1 and 3.2. □

Lemma 4.4. (1) *Let g be a reducible metric on Σ. Then each character 1-form ω of g has at least two poles.*

(2) *Besides the assumption of* (1), *assume that ω has no zero and has only two poles. Then Σ is the Riemann sphere $\overline{\mathbb{C}}$ and g has two singularities with the same angle, say $\alpha > 0$. Moreover, if the two singularities are assumed to be* 0 *and ∞, then $\omega = \alpha(dz/z)$ up to sign.*

Proof. (1) Let f be the multiplicative developing map such that $\omega = d(\log f)$. Since $\Psi = 4|f|^2/(1+|f|^2)$ is a nonconstant continuous function on Σ, it must achieve its minimum and maximum. Either a minimal point or a maximal one of Ψ is a pole of ω by Theorem 1.4.

(2) $\Sigma = \overline{\mathbb{C}}$ follows from $\deg(\omega) = -2$. By the residue theorem, the two residues of ω have different signs at the two poles, say 0 and ∞. It follows from Theorem 1.4 that g has exactly two singularities 0 and ∞ with the same angle, say α. Assume that ω has residues α and $-\alpha$ at 0 and ∞, respectively. Then $\omega = \alpha(dz/z)$. □

Proof of Theorem 1.5. We divide the proof into the following two cases:

Case 1. Assume that the integral of ω at some loop in $\Sigma' := \Sigma\backslash\{\text{poles of } \omega\}$ does not belong to the set $2\pi\sqrt{-1}\mathbb{Z}$. Since $\Re\omega$ is exact on Σ', solving the equation

$$\omega = d(\log f)$$

on Σ', up to a complex multiple with modulus one, we obtain a 1-parameter family of multivalued locally univalent meromorphic functions

$$f_\lambda(z) = \lambda \cdot \exp\left(\int^z \omega\right), \quad \lambda \in (0, +\infty).$$

Moreover, f_λ has nontrivial monodromy belonging to $\mathrm{U}(1)$ and $f_\lambda^\star g_{\mathrm{st}}$ is a nontrivial reducible metric with character 1-form ω. Conversely, if g is a reducible metric such that ω is one of its character 1-forms, then there exists a developing map $\tilde{f}$ of g such that $\omega = d(\log \tilde{f})$. Since $\tilde{f}$ has nontrivial monodromy in $\mathrm{U}(1)$, g is nontrivial. Solving the equation $\omega = d\tilde{f}/\tilde{f}$, up to a complex multiple with modulus one, we find that $\tilde{f}$ equals f_λ for some $\lambda > 0$. Therefore such a reducible metric g is unique. By the argument in the proof of Theorem 1.4, we find the divisor D represented by g equals

$$\sum_{j=1}^{J}(\alpha_j - 1)P_j + \sum_{k=J+1}^{N}(|\mathrm{Res}_{Q_k}(\omega)| - 1)Q_k.$$

Case 2. Assume that the monodromy given by ω is trivial; that is, the integral of ω at each loop in $\Sigma' := \Sigma\backslash\{\text{poles of } \omega\}$ belongs to the set $2\pi\sqrt{-1}\mathbb{Z}$. The pullback $f^* g_{\mathrm{st}}$ with $f(z) = \exp\left(\int^z \omega\right)$ is a trivial reducible metric such that f is one of its developing map and a rational function on Σ and ω is one of its character 1-forms. Conversely, if g is a reducible metric with ω one of its character 1-forms, then $g = \tilde{f}^* g_{\mathrm{st}}$ with $\tilde{f}(z) = \lambda \exp\left(\int^z \omega\right)$ for some $\lambda > 0$. Moreover, $\tilde{f}$ is a rational function uniquely determined by ω and λ. Therefore, such a reducible metric g lies in the 1-parameter family $\{f_\lambda^* g_{\mathrm{st}} : \lambda \in (0, +\infty)\}$. By a similar argument to the proof of Theorem 1.4, we can show that the effective divisor represented by g equals

$$\sum_{j=1}^{J}(\alpha_j - 1)P_j + \sum_{k=J+1}^{N}(|\operatorname{Res}_{Q_k}(\omega)| - 1)Q_k. \qquad \square$$

Corollary 4.5. *Under the notation of* Theorem 1.5, *we have*

$$\chi(\Sigma) + \deg D \geq \min(2, 2\min_j \alpha_j).$$

In particular, *the divisor D does not satisfy Troyanov's condition* (1). *In other words*, *if D satisfies condition* (1), *then each conformal metric which has constant curvature one and represents D is irreducible.*

Proof. The character 1-form ω has at least two poles since the continuous function $\Psi = 4|f|^2/(1+|f|^2)$ in Theorem 1.4 has at least a minimal point and a maximal one. Then, using the equality

$$-\chi(\Sigma) = \deg(\omega) = \sum_{j=1}^{J}(\alpha_j - 1) - (N - J),$$

we have

$$\begin{aligned}\chi(\Sigma) + \deg D &= \chi(\Sigma) + \sum_{j=1}^{J}(\alpha_j - 1) + \sum_{k=J+1}^{N}(|\operatorname{Res}_{Q_k}(\omega)| - 1)\\ &= \sum_{k=J+1}^{N}|\operatorname{Res}_{Q_k}(\omega)| \geq \min(2, 2\min_j \alpha_j). \qquad \square\end{aligned}$$

Example 4.6. Consider a conformal metric g on the two-sphere with constant curvature one and finitely many conical singularities $p_1, \ldots, p_n$. Let the angle at p_j be $2\pi\alpha_j$. If $n \geq 3$ and each α_j is a noninteger, then g is irreducible. Otherwise, by Theorem 1.4, the character 1-form of g would have at least three poles and have no zeroes, a contradiction. In particular, we consider the conformal metric g of constant curvature one with three angles π, π, π at 0, 1 and ∞ on the two-sphere. Then the developing maps of g coincide with those projective solutions with monodromy in PSU(2) of the Gauss hypergeometric equation

$$z(1-z)\frac{d^2u}{dz^2} + (c - (a+b+1)z)\frac{du}{dz} - abu = 0,$$

where

$$|1-c| = |c-a-b| = |a-b| = \tfrac{1}{2}$$

(see [Yoshida 1987, Section 5.3]). Then the monodromy group of a developing map of g is conjugate to D_2 in Lemma 4.1 and thus abelian, but g is irreducible.

It follows from [Troyanov 1991, Theorem 4] that on a torus there exists a conformal metric g with constant curvature one and a conical singularity p with angle $2\pi\alpha$, where $1<\alpha<3$. Then Corollary 4.5 tell us that g is not reducible. The existence of such a irreducible metric having one angle 4π implies that Theorem 1.9 does not hold on a torus.

Example 4.7. We at first consider an elementary example of reducible metrics on the Riemann sphere $\overline{\mathbb{C}}$. Let a, b be two positive numbers. Consider the 1-form

$$\omega=\left(\frac{a}{z}+\frac{b}{z-1}\right)dz,$$

which has residue $-a-b$ at ∞. Note that $a/(a+b)$ is the zero of ω. Hence ω satisfies the condition in Theorem 1.5 and thus gives a reducible metric on the two-sphere with angles $2\pi a, 2\pi b, 2\pi(a+b)$ and 4π at $0, 1, \infty$ and $a/(a+b)$, respectively.

On the other hand, suppose that g is a reducible metric on the two-sphere having angles $2\pi a, 2\pi b, 2\pi(a+b)$ and 4π at $0, 1, \infty$ and $\lambda\in\mathbb{C}\backslash\{0,1\}$, respectively, where $a, b, a+b$ are not integers. Then $\lambda=a/(a+b)$. Actually, letting ω be a character 1-form of g, we can see from Theorem 1.4 that $0, 1, \infty$ are simple poles of ω, and λ is the zero of ω. By the residue theorem, we may assume that

$$\begin{aligned}\mathrm{Res}_0(\omega)&=a,\\ \mathrm{Res}_1(\omega)&=b,\\ \mathrm{Res}_\infty(\omega)&=-a-b,\end{aligned}$$

which implies that

$$\omega=\left(\frac{a}{z}+\frac{b}{z-1}\right)dz \quad\text{and}\quad \lambda=\frac{a}{a+b}.$$

We then give an explicit construction of a general reducible metric g on the Riemann sphere $\overline{\mathbb{C}}$. Suppose that $z_1,\ldots,z_k$, $k\geq 2$, are all the poles of ω lying in $\mathbb{C}$. Then

$$\omega=\left(\sum_{j=1}^{k}\frac{a_j}{z-z_j}\right)dz,$$

where $a_j\in\mathbb{R}\backslash\{0\}$ is the residue of ω at z_j for each j. Hence the zero set of ω is determined by the poles of ω and their residues. In other words, the set $\{(z_j,a_j): j=1,2,\ldots,k\}$ determines the positions of the singularities of g corresponding to the saddle points of the function Ψ and the values of their conical angles. Solving the ordinary differential equation $d\log f=\omega$, we find that there exists some $\lambda > 0$ such that the multivalued function

$$f_\lambda(z) = \lambda \cdot \prod_{j=1}^{k} (z - z_j)^{a_j}$$

is a developing map of the metric g. This implies that the existence of reducible metrics does not only depend on angles, but also on the position of singularities.

Remark 4.8. The monodromy of a developing map is irreducible for a hyperbolic conformal metric with finitely many conical or cusp singularities. Moreover it is also the case for a flat conformal metric with finitely many conical singularities unless the metric is a smooth one on a torus. We leave the proof to interested readers.

We conclude this section by saying something more about the relationship between reducible metrics and HCMU metrics. As a generalization with singularities on compact Riemann surfaces of Calabi's extremal Kähler metric on compact complex manifolds [Calabi 1982; 1985], X. Chen [2000] first introduced the concepts of the HCMU metric and the extremal metric with singularities, and proved some fundamental results. In particular, a conformal metric $\tilde{g}$ on a compact Riemann surface with singularities is called *HCMU* if and only if it has finite area and finite Calabi energy, and the complex gradient $K^{,z}\partial/\partial z$ of the Gauss curvature $K = K_{\tilde{g}}$ is a holomorphic vector field outside the singularities. Wang and Zhu [2000] and Lin and Zhu [2002] obtained some interesting results and generalized some results of X. Chen. Recently [Chen and Wu 2011; Chen et al. 2013] completely classified the nonconstant curvature HCMU metrics with conical or cusp singularities, by using the *character 1-form*

$$\tilde{\omega} = \frac{dz}{K_{\tilde{g}}^{,z}}$$

of an HCMU metric $\tilde{g}$. The properties that $\tilde{\omega}$ is an abelian differential of the third kind with real residues and its real part is exact outside the set of simple poles play a crucial rule in the classification. The following observation puts both HCMU metrics and metrics of constant curvature in the same philosophical frame.

Observation 4.9. Given a nonconstant curvature HCMU metric $\tilde{g}$ with singularities on a compact Riemann surface Σ, there exists a multivalued locally univalent meromorphic function $\tilde{f}$ on Σ^* having monodromy in the abelian group

$$\{\exp(\sqrt{-1}\theta) \mid \theta \in \mathbb{R}\},$$

and a football HCMU metric g_{fb} (see [Chen 2000; Chen et al. 2005]) over $\overline{\mathbb{C}}$ such that $\tilde{g} = \tilde{f}^* g_{fb}$. Moreover the character 1-form of $\tilde{g}$ coincides with the logarithmic differential $d(\log f)$ of f, up to a constant.

This is not used in this paper and its proof will be left elsewhere.

5. Discussion

As an application of Theorems 1.9 and 1.4, we shall show that if g is a conformal metric on the sphere $\overline{\mathbb{C}}$ of constant curvature one, representing the divisor $D = (\alpha - 1)P + (\beta - 1)Q$, where $\alpha, \beta > 0$, then $\alpha = \beta$.

Proof. Case 1. We assume that at least one of α and β is not an integer. Suppose that this is the case for α. Since the punctured sphere $\overline{\mathbb{C}}\backslash\{p, q\}$ has fundamental group isomorphic to $\mathbb{Z}$, the metric g is a reducible metric. Let f be one of its developing maps. By Lemmas 3.1 and 3.2, the local monodromy of f at p is nontrivial. Hence g is nontrivial. We may assume that f is multiplicative so that $\omega = df/f$ is the character 1-form of g. Theorem 1.4 tells us that p is a simple pole of ω with residue $\pm\alpha$, and q is either a simple pole or a zero point of ω. If q is a simple pole too, then the residue equals $\pm\beta$. Since the canonical divisor (ω) has degree -2, by Theorem 1.4, we find that ω has exactly two poles of p, q, which implies $\alpha = \beta$. It is impossible that q is a zero point of ω; otherwise ω must have at least two simple poles in $\overline{\mathbb{C}}\backslash\{p, q\}$, each of which has residue ± 1 by Theorem 1.4. This contradicts the fact that $\alpha \notin \mathbb{Z}$.

Case 2. Suppose that both α and β are integers ≥ 2. By Theorem 1.9, each developing map f of g is a rational function on $\overline{\mathbb{C}}$. Since the ramification divisor R_f of f equals $(\alpha-1)P+(\beta-1)Q$, f has degree $d = (\alpha+\beta)/2$ by the Riemann–Hurwitz theorem. Suppose $\alpha \neq \beta$, say $\alpha > \beta$. Then f has expression $z \mapsto z^{\alpha}$ near p, which implies f has degree $\geq \alpha > d$. Contradiction! □

We observe that the reducible metrics representing a given divisor are not unique in general:

Proposition 5.1. *Suppose that there exists a reducible metric g representing the divisor D on the two-sphere. Denote by $\mathbb{M}(D)$ the space of conformal metrics of constant curvature one representing D, by $\mathbb{A}(D)$ that of reducible metrics representing D.*

(1) *If D is supported at two points p_1 and p_2, $\mathbb{M}(D) = \mathbb{A}(D)$, and g is unique if and only if the two angles are equal and do not belong to $2\pi\mathbb{Z}_{>1}$. If the two angles are equal and belong to $2\pi\mathbb{Z}_{>1}$, then $\mathbb{A}(D)$ is connected and has dimension* 1.

(2) *If D is supported at three points and $\mathbb{A}(D) \neq \varnothing$, then $\mathbb{M}(D) = \mathbb{A}(D)$ is connected. Moreover if g is trivial,* $\dim \mathbb{A}(D) = 3$; *otherwise* $\dim \mathbb{A}(D) = 1$.

(3) *Suppose that D is supported at more than three points and $\mathbb{A}(D) \neq \varnothing$. Then, if g is trivial,* $\dim \mathbb{A}(D) = 3$; *otherwise* $\dim \mathbb{A}(D) \geq 1$.

Proof. The first statement was proved by Troyanov [1989, Theorem I, p. 298]. The second was shown in [Umehara and Yamada 2000, Corollary 2.3].

Suppose that D is supported at more than three points. Following [Umehara and Yamada 2000, (2.5)], we define

$$I_g := \{g_a = (a \star f)^* g \mid a \in \mathrm{PSL}(2, \mathbb{C}), a \cdot \mathrm{Im}\, \rho_f \cdot a^{-1} \subset \mathrm{PSU}(2)\},$$

where $a \star f$ denotes the Möbius transformation of f by a and

$$\rho_f : \pi_1(\Sigma^*) \to \mathrm{PSU}(2)$$

denotes the monodromy representation of the developing map f of the metric g. Each metric g_a in I_g has a developing map $a \star f$, which has the same Schwarzian as f and monodromy conjugate to that of f. Hence I_g is contained in $\mathbb{A}(D)$. Then it follows from [Umehara and Yamada 2000, Lemma B, p. 92] that if g is trivial, $\dim \mathbb{A}(D) \geq 3$; otherwise $\dim \mathbb{A}(D) \geq 1$.

We consider the moduli of trivial reducible metrics representing an effective $\mathbb{Z}$-divisor D, which can be reduced to the space of rational functions with the same ramification divisor D. We say that two rational functions are *equivalent* if one of them is given by the postcomposition of the other with a Möbius transformation. It follows from [Umehara and Yamada 2000, Lemma B, p. 92] that the trivial reducible metrics having developing maps of the same type form a moduli of the three-dimensional hyperbolic space $\mathscr{H}^3$. The beautiful theorem in [Scherbak 2002] says that there is a least upper bound given by the Schubert calculus for the number of equivalent classes of all the rational functions with ramification divisor D, which can be achieved by a generic choice of the support of D. Hence we obtain the corresponding information for the number of connected components of $\mathbb{A}(D)$. □

It is time to propose some questions interesting to us.

Question 5.2. Does there exist a divisor on some compact Riemann surface which is represented by both an irreducible metric and a reducible one? It does not happen for the divisors satisfying the Troyanov condition (1) by Corollary 4.5. It also does not happen on the two-sphere under either of the following two conditions:

(1) The support of D consists of three points or less (see Proposition 5.1).

(2) Each α_j is a noninteger for $j = 1, 2, \ldots, n$, with $n \geq 3$ (see Example 4.6).

Question 5.3. Suppose $\mathbb{A}(D)$ is nonempty for a divisor D on a compact Riemann surface. Study the moduli space $\mathbb{A}(D)$ of reducible metrics representing a divisor D on a compact Riemann surface, such as its dimension and the number of its components. We know the answer on the two-sphere in the case that D has support at two or three points or D is an effective $\mathbb{Z}$-divisor in Proposition 5.1.

Question 5.4. Suppose that there exists an irreducible metric g representing $D = \sum_j (\alpha_j - 1) P_j$. Is g the unique metric of constant curvature one representing D? Luo and Tian [1992] showed this is the case on the two-sphere if each α_j lies in

(0, 1). Moreover, Umehara and Yamada [2000] gave a positive answer if D is a divisor supported at three points on the two-sphere.

Question 5.5. Suppose that there exists an irreducible metric g representing $D = \sum_j (\alpha_j - 1) P_j$. Does there exist an irreducible metric representing any divisor D' sufficiently near D? On the two-sphere, if each α_j lies in (0, 1), then the necessary and sufficient condition is a topologically open one for the existence of a irreducible metric representing D given by Troyanov [1991] and Luo and Tian [1992]. On the two-sphere, if D has support at three points, so is the necessary and sufficient condition given by Umehara and Yamada [2000]. S. K. Donaldson [2012] proved an openness theorem for Kähler Einstein metrics on a Fano manifold with conical singularity along the anticanonical divisor.

Acknowledgments

Xu would like to express his deep gratitude to Professor Xiuxiong Chen for his constant moral support and lots of invaluable conversations. He thanks Professor Masaaki Umehara so much for many valuable comments and discussions through emails. He is also very grateful to Professor Ryoichi Kobayashi, Dr. Zhi Chen and Dr. Jinxing Xu for their stimulating conversations. In particular, the penetrating questions from Dr. Jinxing Xu gave him the impetus to fix a gap in the old version of the manuscript.

Note added in proof

After this paper was accepted, the authors learned that Daniele Bartolucci, Francesca De Marchis and Andrea Malchiodi [Bartolucci et al. 2011] proved a general existence result for the problem of prescribing the Gaussian curvature on surfaces of positive genus with conical singularities in supercritical regimes. Two years later, Bartolucci and Malchiodi [2013] removed the assumption on the genus. As a consequence, their results imply a new general existence theorem for conformal metrics with constant curvature one and finitely many conical singularities which do not satisfy the Troyanov condition (1). The authors observed that these metrics are irreducible. Moreover, Bartolucci, De Marchis and Malchiodi [Bartolucci et al. 2011] proved the existence of multiple solutions on surfaces of genus bigger than one, which implies that the answer to Question 5.4 is negative.

References

[Bartolucci and Malchiodi 2013] D. Bartolucci and A. Malchiodi, "An improved geometric inequality via vanishing moments, with applications to singular Liouville equations", *Comm. Math. Phys.* **322**:2 (2013), 415–452. MR 3077921 Zbl 1276.58005

[Bartolucci et al. 2011] D. Bartolucci, F. De Marchis, and A. Malchiodi, "Supercritical conformal metrics on surfaces with conical singularities", *Int. Math. Res. Not.* **2011**:24 (2011), 5625–5643. MR 2863376 Zbl 1254.30066

[Bryant 1988] R. L. Bryant, "Surfaces of mean curvature one in hyperbolic space", pp. 321–347 in *Théorie des variétés minimales et applications* (Palaiseau, 1983–1984), Astérisque **154–155**, Société Mathématique de France, Paris, 1988. MR 955072 Zbl 0635.53047

[Calabi 1982] E. Calabi, "Extremal Kähler metrics", pp. 259–290 in *Seminar on Differential Geometry*, edited by S.-T. Yau, Ann. of Math. Stud. **102**, Princeton University Press, 1982. MR 83i:53088 Zbl 0487.53057

[Calabi 1985] E. Calabi, "Extremal Kähler metrics, II", pp. 95–114 in *Differential geometry and complex analysis*, edited by I. Chavel and H. M. Farkas, Springer, Berlin, 1985. MR 86h:53067 Zbl 0574.58006

[Chen 2000] X. Chen, "Obstruction to the existence of metric whose curvature has umbilical Hessian in a K-surface", *Comm. Anal. Geom.* **8**:2 (2000), 267–299. MR 2001k:53060 Zbl 0971.53029

[Chen and Wu 2011] Q. Chen and Y. Wu, "Character 1-form and the existence of an HCMU metric", *Math. Ann.* **351**:2 (2011), 327–345. MR 2012g:58020 Zbl 1227.53076

[Chen et al. 2005] Q. Chen, X. Chen, and Y. Wu, "The structure of HCMU metric in a K-surface", *Int. Math. Res. Not.* **2005**:16 (2005), 941–958. MR 2006f:53046 Zbl 1090.53036

[Chen et al. 2013] Q. Chen, Y. Wu, and B. Xu, "HCMU metrics with cusp and conic singularities", preprint, 2013. arXiv 1302.6655

[Donaldson 2012] S. K. Donaldson, "Kähler metrics with cone singularities along a divisor", pp. 49–79 in *Essays in mathematics and its applications*, edited by P. M. Pardalos and T. M. Rassias, Springer, Heidelberg, 2012. MR 2975584

[Eremenko 2004] A. Eremenko, "Metrics of positive curvature with conic singularities on the sphere", *Proc. Amer. Math. Soc.* **132**:11 (2004), 3349–3355. MR 2005h:53054 Zbl 1053.53025

[Eremenko and Gabrielov 2002] A. Eremenko and A. Gabrielov, "Rational functions with real critical points and the B. and M. Shapiro conjecture in real enumerative geometry", *Ann. of Math.* (2) **155**:1 (2002), 105–129. MR 2003c:58028 Zbl 0997.14015

[Eremenko et al. 2006] A. Eremenko, A. Gabrielov, M. Shapiro, and A. Vainshtein, "Rational functions and real Schubert calculus", *Proc. Amer. Math. Soc.* **134**:4 (2006), 949–957. MR 2007d:14103 Zbl 1110.14052

[Fujimori et al. 2011] S. Fujimori, Y. Kawakami, M. Kokubu, W. Rossman, M. Umehara, and K. Yamada, "CMC-1 trinoids in hyperbolic 3-space and metrics of constant curvature one with conical singularities on the 2-sphere", *Proc. Japan Acad. Ser. A Math. Sci.* **87**:8 (2011), 144–149. MR 2012k:53013 Zbl 1242.53070

[Furuta and Hattori 1998] M. Furuta and Y. Hattori, "2-dimensional singular spherical space forms", manuscript, 1998.

[Goldberg 1991] L. R. Goldberg, "Catalan numbers and branched coverings by the Riemann sphere", *Adv. Math.* **85**:2 (1991), 129–144. MR 92b:14014 Zbl 0732.14013

[Lin and Zhu 2002] C. S. Lin and X. Zhu, "Explicit construction of extremal Hermitian metrics with finite conical singularities on S^2", *Comm. Anal. Geom.* **10**:1 (2002), 177–216. MR 2003c:58010 Zbl 1021.58008

[Luo and Tian 1992] F. Luo and G. Tian, "Liouville equation and spherical convex polytopes", *Proc. Amer. Math. Soc.* **116**:4 (1992), 1119–1129. MR 93b:53034 Zbl 0806.53012

[McOwen 1988] R. C. McOwen, "Point singularities and conformal metrics on Riemann surfaces", *Proc. Amer. Math. Soc.* **103**:1 (1988), 222–224. MR 89m:30089 Zbl 0657.30033

[Petersen 2006] P. Petersen, *Riemannian geometry*, 2nd ed., Graduate Texts in Mathematics **171**, Springer, New York, 2006. MR 2007a:53001 Zbl 1220.53002

[Scherbak 2002] I. Scherbak, "Rational functions with prescribed critical points", *Geom. Funct. Anal.* **12**:6 (2002), 1365–1380. MR 2004c:14101 Zbl 1092.14065

[Siegel 1969] C. L. Siegel, *Topics in complex function theory, I: Elliptic functions and uniformization theory*, Interscience Tracts in Pure and Applied Mathematics **25**, Wiley, New York, 1969. MR 41 #1977 Zbl 0184.11201

[Springer 1957] G. Springer, *Introduction to Riemann surfaces*, Addison-Wesley, Reading, MA, 1957. MR 19,1169g Zbl 0078.06602

[Troyanov 1989] M. Troyanov, "Metrics of constant curvature on a sphere with two conical singularities", pp. 296–306 in *Differential geometry* (Peñíscola, 1988), edited by F. J. Carreras et al., Lecture Notes in Math. **1410**, Springer, Berlin, 1989. MR 90m:53057 Zbl 0697.53037

[Troyanov 1991] M. Troyanov, "Prescribing curvature on compact surfaces with conical singularities", *Trans. Amer. Math. Soc.* **324**:2 (1991), 793–821. MR 91h:53059 Zbl 0724.53023

[Umehara and Yamada 2000] M. Umehara and K. Yamada, "Metrics of constant curvature 1 with three conical singularities on the 2-sphere", *Illinois J. Math.* **44**:1 (2000), 72–94. MR 2001f:53072 Zbl 0958.30029

[Wang and Zhu 2000] G. Wang and X. Zhu, "Extremal Hermitian metrics on Riemann surfaces with singularities", *Duke Math. J.* **104**:2 (2000), 181–210. MR 2001i:58017 Zbl 0980.58009

[Yoshida 1987] M. Yoshida, *Fuchsian differential equations: with special emphasis on the Gauss–Schwarz theory*, Aspects of Mathematics **11**, Vieweg, Braunschweig, 1987. MR 90f:32025 Zbl 0618.35001

Received July 20, 2013. Revised December 18, 2013.

QING CHEN
WU WEN-TSUN KEY LABORATORY OF MATH, USTC, CHINESE ACADEMY OF SCIENCES
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI, 230026
CHINA
qchen@ustc.edu.cn

WEI WANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, IRVINE
IRVINE, CA 92697-3875
UNITED STATES
wangw10@math.uci.edu

YINGYI WU
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF CHINESE ACADEMY OF SCIENCES
BEIJING, 100049
CHINA
wuyy@ucas.ac.cn

BIN XU
WU WEN-TSUN KEY LABORATORY OF MATH, USTC, CHINESE ACADEMY OF SCIENCES
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI, 230026
CHINA
bxu@ustc.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.101

$\mathbb{Q}$-BASES OF THE NÉRON–SEVERI GROUPS OF CERTAIN ELLIPTIC SURFACES

MASAMICHI KURODA

P. Stiller computed the rank of the Néron–Severi group (known as the Picard number) for several families of elliptic surfaces. However, he did not give the generators of these groups. In this paper we give $\mathbb{Q}$-bases of these groups explicitly. If these surfaces are rational, we also show that they are $\mathbb{Z}$-bases.

1. Introduction

An elliptic surface is a surface which has a surjective map onto a curve such that the generic fiber is a curve of genus one (see [Kodaira 1963a; 1963b]). The Néron–Severi group is the group of divisors modulo algebraic equivalence (see [Hartshorne 1977, Exercise V.1.7]). This group is known to be a finitely generated abelian group, and its rank is called the Picard number. P. Stiller [1987] computed the Picard numbers of several families of elliptic surfaces by studying the action of certain automorphisms on the cohomology group. However he did not give the generators of these groups. The purpose of this paper is to give explicit $\mathbb{Q}$-bases of the Néron–Severi groups of Stiller's list [1987, Examples 1–5] of elliptic surfaces.

We explain briefly how to construct such $\mathbb{Q}$-bases. Let $\mathcal{E}$ be an elliptic surface. We denote by $\mathrm{NS}(\mathcal{E})$ the Néron–Severi group of $\mathcal{E}$. T. Shioda [1972] proved that $\mathrm{NS}(\mathcal{E})$ is generated by fibral divisors and horizontal divisors. Here we mean by a fibral divisor a sum of irreducible components of fibers, and by a horizontal divisor a sum of images of sections. Let $\mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ ($n \in \mathbb{N}$) be one of the families of elliptic surfaces of [Stiller 1987], and let E_n be the generic fiber of $\mathcal{E}_n$ for each n. The E_n are elliptic curves over the function field $\mathbb{C}(t)$. Computing the Picard number of an elliptic surface is equivalent to determining the Mordell–Weil rank (i.e., the rank of the Mordell–Weil group) of the generic fiber. Stiller [1987, Examples 1–5] proved that for each family $\mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ ($n \in \mathbb{N}$), there exists a finite set Adm_i ($1 \le i \le 5$) of natural numbers such that the Mordell–Weil rank* of $E_n/\mathbb{C}(t)$ is $r = \sum_{\substack{d \mid n \\ d \in \mathrm{Adm}_i}} \varphi(d)$,

MSC2010: 14J27.
Keywords: elliptic surfaces, Néron–Severi groups.

*The referee pointed out that a related or a similar formula is obtained in [Silverman 2000] for arbitrary elliptic surfaces defined over number fields under the Tate conjecture.

where φ is the Euler totient function. In this paper we shall construct r rational points of E_n in an ad hoc manner, and show the linear independence of the associated divisors in $\mathrm{NS}(\mathcal{E}_n)$. If $\mathcal{E}_n$ is rational, then we further show that they form a $\mathbb{Z}$-basis.

This paper is organized as follows. Section 2 is a quick review of some basic results on the Néron–Severi groups of elliptic surfaces. Section 3 is the heart of this paper. We give a number of $\mathbb{Q}$-bases or $\mathbb{Z}$-bases of the Néron–Severi groups of Stiller's list of elliptic surfaces. In Section 4, we give an alternative proof of Stiller's computations of the Picard numbers.

2. The Néron–Severi group of an elliptic surface

In this paper, we mean by an elliptic surface a surjective morphism $f : \mathcal{E} \to C$ onto a curve with a section (say zero section) such that the generic fiber of f is an elliptic curve. Let $f : \mathcal{E} \to \mathbb{P}^1_{\mathbb{C}}$ be a nonsplit minimal elliptic surface, and let $E/\mathbb{C}(t)$ be the generic fiber. There is a natural group isomorphism between the Mordell–Weil group of $E/\mathbb{C}(t)$, denoted by $E(\mathbb{C}(t))$, and the group of sections of $\mathcal{E}$ over $\mathbb{P}^1_{\mathbb{C}}$, denoted by $\mathcal{E}(\mathbb{P}^1_{\mathbb{C}})$ (see [Silverman 1994, Proposition 3.10(c)]):

$$\begin{aligned} E(\mathbb{C}(t)) &\xrightarrow{\sim} \mathcal{E}(\mathbb{P}^1_{\mathbb{C}}), \\ P = (x_P, y_P) &\longmapsto (\sigma_P : t \mapsto (x_P(t), y_P(t), t)). \end{aligned} \tag{1}$$

According to the Mordell–Weil theorem, $E(\mathbb{C}(t))$ is a finitely generated group. In the following, for each $P \in E(\mathbb{C}(t))$, we denote by (P) the image in $\mathcal{E}$ of the section corresponding to P. For simplicity, we denote by ∞ the image of the zero section, that is, the section corresponding to zero element $O \in E(\mathbb{C}(t))$.

The singular fibers are classified by Kodaira [1963a; 1963b]. We shall follow Kodaira's notation. Let $\Sigma(\mathcal{E})$ be the finite set of points t in $\mathbb{P}^1_{\mathbb{C}}$ such that $\mathcal{E}_t := f^{-1}(t)$ is a singular fiber. For each $t \in \mathbb{P}^1_{\mathbb{C}}$, let m_t be the number of irreducible components of the fiber $\mathcal{E}_t$, and we denote by $F_{t,a}$ $(0 \le a \le m_t - 1)$ the irreducible components. If $t \in \mathbb{P}^1_{\mathbb{C}} \setminus \Sigma(\mathcal{E})$, then $\mathcal{E}_t = F_{t,0}$ is a smooth fiber, and we have

$$\{t \in \mathbb{P}^1_{\mathbb{C}} \mid m_t \ge 2\} = \{t \in \Sigma(\mathcal{E}) \mid m_t \ge 2\}.$$

We fix a general fiber $C_0 := \mathcal{E}_{t_0}$, $t_0 \in \mathbb{P}^1_{\mathbb{C}} \setminus \Sigma(\mathcal{E})$, and we take $F_{t,0}$ to be the unique component of $\mathcal{E}_t$ intersecting with ∞.

Let $E(\mathbb{C}(t))_{\mathrm{free}}$ denote the quotient group $E(\mathbb{C}(t))/E(\mathbb{C}(t))_{\mathrm{tor}}$, where $E(\mathbb{C}(t))_{\mathrm{tor}}$ is the torsion subgroup. Let r be the Mordell–Weil rank of $E/\mathbb{C}(t)$, that is, the rank of $E(\mathbb{C}(t))$, and we take r generators $P_1, \dots, P_r$ of $E(\mathbb{C}(t))_{\mathrm{free}}$. We put

$$D_i = (P_i) - \infty \in \mathrm{Div}(\mathcal{E}) \quad (1 \le i \le r),$$

where $\mathrm{Div}(\mathcal{E})$ is the group of divisors on $\mathcal{E}$.

Proposition 2.1 [Shioda 1972, Theorem 1.1]. *The free part of the Néron–Severi group* $\mathrm{NS}(\mathcal{E})$ *of the elliptic surface* $\mathcal{E}$, *denoted by* $\mathrm{NS}(\mathcal{E})_{\mathrm{free}}$, *is generated by the divisors*

$$C_0, \infty, D_1, \dots, D_r, F_{t,a} \quad (t \in \Sigma(\mathcal{E}), 1 \le a \le m_t - 1). \tag{2}$$

In particular, the Picard number ρ *of the elliptic surface* $\mathcal{E}$ *is given by*

$$\rho = r + 2 + \sum_{t \in \Sigma(\mathcal{E})} (m_t - 1).$$

Stiller computed the Mordell–Weil rank r, but did not give D_i's explicitly. We will give r linearly independent points of the Mordell–Weil group in Section 3. Note that these points are not always generators of the group. We end this section by introducing a practical way to show the linear independence of the divisors $C_0, \infty, D_1, \dots, D_r, F_{t,a}$ $(t \in \Sigma(\mathcal{E}), 1 \le a \le m_t - 1)$, or equivalently the intersection matrix M of these divisors has a nonzero determinant.

For each $P_i \in E(\mathbb{C}(t))$, we have $D_i \cdot C_0 = ((P_i) - \infty) \cdot C_0 = 1 - 1 = 0$. Then there exists a fibral divisor $\Phi_i \in \mathrm{Div}(\mathcal{E}) \otimes \mathbb{Q}$ such that

$$(D_i + \Phi_i) \cdot F = 0 \ \text{ for all fibral divisors } F \in \mathrm{Div}(\mathcal{E}).$$

More explicitly the divisor Φ_i is obtained in the following way (see [Silverman 1994, Proposition 8.3]). We set $a_{t0}(P_i) = 0$ for all $t \in \mathbb{P}^1_{\mathbb{C}}$. Further, when $m_t \ge 2$ we consider the following system of linear equations:

$$\sum_{k=1}^{m_t - 1} a_{tk}(P_i) F_{t,k} \cdot F_{t,l} = -D_i \cdot F_{t,l} \quad (1 \le l \le m_t - 1).$$

This is a system of $m_t - 1$ equations in the $m_t - 1$ variables $a_{tk}(P_i)$. Since the intersection matrix $(F_{t,i} \cdot F_{t,j})_{1 \le i,j \le m_t - 1}$ has a nonzero determinant, this system of equations has a unique solution in rational numbers $a_{tk}(P_i) \in \mathbb{Q}$. Then the divisor

$$\Phi_i := \sum_{t \in \mathbb{P}^1_{\mathbb{C}}} \sum_{k=0}^{m_t - 1} a_{tk}(P_i) F_{t,k} = \sum_{t \in \{t_1, \dots, t_s\}} \sum_{k=1}^{m_t - 1} a_{tk}(P_i) F_{t,k} \in \mathrm{Div}(\mathcal{E}) \otimes \mathbb{Q}$$

has the desired property, where we set $\{t_1, \dots, t_s\} = \{t \in \Sigma(\mathcal{E}) \mid m_t \ge 2\}$. Note that since $F_{t_\alpha,k} \cdot F_{t_\beta,l} = 0$ $(\alpha \ne \beta)$, we have

$$0 = (D_i + \Phi_i) \cdot F_{t_\alpha,j} = \left(D_i + \sum_{k=1}^{m_{t_\alpha} - 1} a_{t_\alpha k}(P_i) F_{t_\alpha,k} \right) \cdot F_{t_\alpha,j}.$$

We fix a uniformizer $u_t \in \mathbb{C}(t)$ at t, that is, $\mathrm{ord}_t(u_t) = 1$. Let $f^* : \mathrm{Div}(\mathbb{P}^1_{\mathbb{C}}) \to \mathrm{Div}(\mathcal{E})$ be a homomorphism defined by extending $(t) \mapsto \sum_{j=0}^{m_t - 1} \mathrm{ord}_{F_{t,j}}(u_t \circ f) F_{t,j}$ linearly.

For each two points $t_1, t_2 \in \mathbb{P}^1_{\mathbb{C}} \setminus \Sigma(\mathcal{E})$ with $t_1 \neq t_2$, since C_0 is algebraically equivalent to $f^*(t_i)$ $(i = 1, 2)$, we have

$$C_0^2 = f^*(t_1) \cdot f^*(t_2) = 0.$$

Now it is not hard to show the following lemma.

Lemma 2.2 [Cox and Zucker 1979]. *Let M be the intersection matrix of divisors $C_0, \infty, D_1, \ldots, D_r, F_{t,a}$ $(t \in \Sigma(\mathcal{E}), 1 \leq a \leq m_t - 1)$, and let M_α $(1 \leq \alpha \leq s)$ be the intersection matrix of divisors $F_{t_\alpha,1}, \ldots, F_{t_\alpha,m_{t_\alpha}-1}$. Put*

$$N = \begin{bmatrix} (D_1 + \Phi_1) \cdot D_1 & \cdots & (D_1 + \Phi_1) \cdot D_r \\ \vdots & \ddots & \vdots \\ (D_r + \Phi_r) \cdot D_1 & \cdots & (D_r + \Phi_r) \cdot D_r \end{bmatrix}.$$

Then we have

$$\det M = -\det N \prod_{\alpha=1}^{s} \det M_\alpha.$$

In particular, $\det M \neq 0$ *if and only if* $\det N \neq 0$ *since M_α has nonzero determinant for each α. Note that each M_α gives one of the root lattices A_n, D_n, E_6, E_7 or E_8, and* $\det M_\alpha$ *equals the number of simple components of the singular fiber $\mathcal{E}_{t_\alpha}$.*

3. Stiller's list of elliptic surfaces

In this section, we give explicit $\mathbb{Q}$-bases of the Néron–Severi groups of the elliptic surfaces in Stiller's Examples 1–5. If these surfaces are rational, then we also show that they are $\mathbb{Z}$-bases. Note that these Néron–Severi groups are torsion-free, by [Cox and Zucker 1979] or [Shioda 1990].

We give a proof in detail in the case of Stiller's Example 4. In the other cases we just give results, because the argument is the same.

Throughout this paper, ζ_n will denote a primitive n-th root of unity for a natural number n.

Stiller's Example 4. This example is the minimal elliptic surface whose generic fiber is the elliptic curve defined by the equation

$$Y^2 = 4X^3 - 3u^{4n}X + u^{5n}(u^n - 2) \quad (u \in \mathbb{P}^1_{\mathbb{C}}, n \in \mathbb{N}) \tag{3}$$

over $\mathbb{C}(u)$. We perform the change of variables

$$X = \frac{2^3(3x+1)}{3^6 t^{2n}}, \quad Y = \frac{-2^5\sqrt{2}\, y}{3^7\sqrt{3} t^{3n}}, \quad u = \sqrt[n]{\frac{-4}{27}}\, t^{-1}.$$

Then the defining equation (3) becomes

$$y^2 = x^3 + x^2 + t^n \quad (t \in \mathbb{P}^1_{\mathbb{C}}, n \in \mathbb{N}). \tag{4}$$

Let E_n be the elliptic curve defined by (4) and $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ be the associated elliptic surface. For the later use, we write down the construction of $\mathcal{E}_n$. Put

$$\overline{X}_1 = \bigl\{([X:Y:Z], t) \in \mathbb{P}^2_{\mathbb{C}} \times \mathbb{A}^1_t \mid Y^2Z = X^3 + X^2Z + t^nZ^3\bigr\},$$
$$g_1 : \overline{X}_1 \to \mathbb{A}^1_t; \quad ([X:Y:Z], t) \mapsto t.$$

Let us write $n = 6l + k$ with $1 \le k \le 6$. By putting $\bar{x} = x/t^{2(l+1)}$, $\bar{y} = y/t^{3(l+1)}$, $\bar{t} = 1/t$, we obtain the minimal Weierstrass form

$$\bar{y}^2 = \bar{x}^3 + \bar{t}^{2(l+1)}\bar{x}^2 + \bar{t}^{6-k}.$$

over $t = \infty$. Put

$$\overline{X}_2 = \bigl\{([\overline{X}:\overline{Y}:\overline{Z}], \bar{t}) \in \mathbb{P}^2_{\mathbb{C}} \times \mathbb{A}^1_{\bar{t}} \mid \overline{Y}^2\overline{Z} = \overline{X}^3 + \bar{t}^{2(l+1)}\overline{X}^2\overline{Z} + \bar{t}^{6-k}\overline{Z}^3\bigr\},$$
$$g_2 : \overline{X}_2 \to \mathbb{A}^1_{\bar{t}}; \quad ([\overline{X}:\overline{Y}:\overline{Z}], \bar{t}) \mapsto \bar{t}.$$

By gluing the surfaces $\overline{X}_1$ and $\overline{X}_2$, we obtain a projective surface W together with a surjective morphism $g : W \to \mathbb{P}^1_{\mathbb{C}}$. The surface W has singularities in $g^{-1}(0)$ and $g^{-1}(\infty)$. Taking the minimal resolution of singularities of W, we obtain $\mathcal{E}_n$ with $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$.

Using Tate's algorithm (see [Silverman 1994]), one can show that the surface $\mathcal{E}_n$ has singular fibers of type I_n over 0, type I_1 over $\zeta_n^i \sqrt[n]{-4/27}$ $(0 \le i \le n-1)$ and type II^* (resp. IV^*, I_0^*, IV, II, I_0) over ∞ as $n \equiv 1$ (resp. 2, 3, 4, 5, 0) modulo 6. Stiller computed the Mordell–Weil rank $r = \operatorname{rank}(E_n(\mathbb{C}(t)))$ and hence the Picard number $\rho = \operatorname{rank}(\mathrm{NS}(\mathcal{E}_n))$, which is given in Table 1. The result for r can be summarized in the following way. Put $\mathrm{Adm}_4 = \{2, 3, 4, 5\}$. Then

$$r = \sum_{\substack{d \mid n \\ d \in \mathrm{Adm}_4}} \varphi(d), \tag{5}$$

where φ is the Euler function. We now define $\varphi(d)$ rational points of E_d for each $d \in \mathrm{Adm}_4$.

Definition 3.1. For $d \in \mathrm{Adm}_4$ and $j \in (\mathbb{Z}/d\mathbb{Z})^\times$, we define $\mathbb{C}(t)$-rational points $P_{d,j}$ of E_d as follows.

$$P_{2,1} = (0, -t),$$
$$P_{3,j} = (-\zeta_3^j t, -\zeta_3^j t),$$
$$P_{4,1} = \bigl(\sqrt{2}t + 2t^2, \sqrt{2}t + 3t^2 + 2\sqrt{2}t^3\bigr),$$
$$P_{4,3} = \bigl(-2 + 2(-1)^{1/4}t, -2\sqrt{-1} + 4\sqrt{-1}(-1)^{1/4}t + t^2\bigr),$$
$$P_{5,j} = \bigl(2^{-2/5}\zeta_5^{2j}t^2, 2^{-2/5}\zeta_5^{2j}t^2 + 2^{-3/5}\zeta_5^{3j}t^3\bigr).$$

n	r	ρ
$6l+1,\ l\geq 0$	$r=\begin{cases}4 & \text{if } l\equiv 4 \mod 5,\\ 0 & \text{otherwise.}\end{cases}$	$\rho=\begin{cases}n+13\\ n+9\end{cases}$
$6l+2,\ l\geq 0$	$r=\begin{cases}7 & \text{if } 3l+1\equiv 0 \mod 10,\\ 5 & \text{if } 3l+1\equiv 5 \mod 10,\\ 3 & \text{if } 3l+1\equiv 2,4,6,8 \mod 10,\\ 1 & \text{otherwise.}\end{cases}$	$\rho=\begin{cases}n+14\\ n+12\\ n+10\\ n+8\end{cases}$
$6l+3,\ l\geq 0$	$r=\begin{cases}6 & \text{if } l\equiv 2 \mod 5,\\ 2 & \text{otherwise.}\end{cases}$	$\rho=\begin{cases}n+11\\ n+7\end{cases}$
$6l+4,\ l\geq 0$	$r=\begin{cases}7 & \text{if } 3l+2\equiv 0 \mod 10,\\ 5 & \text{if } 3l+2\equiv 5 \mod 10,\\ 3 & \text{if } 3l+2\equiv 2,4,6,8 \mod 10,\\ 1 & \text{otherwise.}\end{cases}$	$\rho=\begin{cases}n+10\\ n+8\\ n+6\\ n+4\end{cases}$
$6l+5,\ l\geq 0$	$r=\begin{cases}4 & \text{if } l\equiv 0 \mod 5,\\ 0 & \text{otherwise.}\end{cases}$	$\rho=\begin{cases}n+5\\ n+1\end{cases}$
$6l+6,\ l\geq 0$	$r=\begin{cases}9 & \text{if } l+1\equiv 0 \mod 10,\\ 7 & \text{if } l+1\equiv 5 \mod 10,\\ 5 & \text{if } l+1\equiv 2,4,6,8 \mod 10,\\ 3 & \text{otherwise.}\end{cases}$	$\rho=\begin{cases}n+10\\ n+8\\ n+6\\ n+4\end{cases}$

Table 1. The Mordell–Weil rank r and the Picard number ρ.

For any d that divides n, there is the surjective map $\rho : E_n \to E_d$ given by $(x, y, t) \mapsto (x, y, t^{n/d})$, and then the inverse image $\rho^*(P_{d,j})$ defines a $\mathbb{C}(t)$-rational point of E_n. In what follows, we use the same symbol $P_{d,j}$ for $\rho^*(P_{d,j})$ since the context will prevent any confusion.

Theorem 3.2. *Let $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}} (n \in \mathbb{N})$ be the elliptic surfaces associated to the elliptic curves $E_n : y^2 = x^3 + x^2 + t^n (n \in \mathbb{N})$ over $\mathbb{P}^1_{\mathbb{C}}$. Then, for each $n \in \mathbb{N}$, $\mathrm{NS}(\mathcal{E}_n)$ has a $\mathbb{Q}$-basis $C_0, \infty, D_{d,j}, F_{t,a}$ ($d \in \mathrm{Adm}_4$, $d|n$, $j \in (\mathbb{Z}/d\mathbb{Z})^\times$, $t \in \Sigma(\mathcal{E}_n)$, $1 \leq a \leq m_t - 1$), where $D_{d,j} = (P_{d,j}) - \infty$ and the other notations are the same as Section 2. Moreover if $\mathcal{E}_n$ is rational* (*i.e.*, $n \leq 6$), *these divisors form a $\mathbb{Z}$-basis.*

Proof. In the cases of $n = 6l+1$ with $l \not\equiv 4 \mod 5$ or $n = 6l+5$ with $l \not\equiv 0 \mod 5$, we have $r = 0$ by Table 1, so the assertion follows immediately from Proposition 2.1. We consider the other cases. All we have to do is to show the linear independence of the divisors in Theorem 3.2. Let $d(n)$ be the least common multiple of all numbers in $\{d \in \mathrm{Adm}_4 : d|n\}$. Note that the set is nonempty in these cases. Since n is divisible by $d(n)$, there exists a rational map

$$\begin{aligned}\mathcal{E}_n &\to \mathcal{E}_{d(n)},\\ (x, y, t) &\mapsto (x, y, t^{n/d(n)}).\end{aligned}$$

This map induces an injection

$$\begin{aligned}E_{d(n)}(\mathbb{C}(t)) &\hookrightarrow E_n(\mathbb{C}(t)),\\ (x(t), y(t)) &\mapsto (x(t^{n/d(n)}), y(t^{n/d(n)})).\end{aligned}$$

We obtain $\operatorname{rank}(E_{d(n)}(\mathbb{C}(t))) = \operatorname{rank}(E_n(\mathbb{C}(t)))$ by (5), hence linearly independent points of $E_{d(n)}(\mathbb{C}(t))$ remain so in $E_n(\mathbb{C}(t))$. There exists an injection

$$\begin{aligned}E_{d(n)}(\mathbb{C}(t)) &\hookrightarrow E_{60}(\mathbb{C}(t)),\\ (x(t), y(t)) &\mapsto \big(x(t^{60/d(n)}), y(t^{60/d(n)})\big).\end{aligned}$$

In particular, points in $E_{d(n)}(\mathbb{C}(t))$ are linearly independent if and only if their images are independent in $E_{60}(\mathbb{C}(t))$. Therefore it is sufficient to show the assertion in the case of $n = 60$. Recall that the surface $\mathcal{E}_{60}$ has singular fibers of type I_{60} over 0, type I_1 over $\zeta_{60}^i \sqrt[60]{-4/27}$ $(0 \le i \le 59)$ and type I_0 over ∞.

We want to show that the divisors C_0, ∞, $D_{2,1}$, $D_{3,1}$, $D_{3,2}$, $D_{4,1}$, $D_{4,3}$, $D_{5,1}$, $D_{5,2}$, $D_{5,3}$, $D_{5,4}$, $F_{0,1}, \dots, F_{0,59}$ are a $\mathbb{Q}$-basis of $\mathrm{NS}(\mathcal{E}_{60})$. Equivalently the matrix

$$N = \begin{bmatrix} (D_{2,1} + \Phi_{2,1}) \cdot D_{2,1} & \cdots & (D_{2,1} + \Phi_{2,1}) \cdot D_{5,4} \\ \vdots & \ddots & \vdots \\ (D_{5,4} + \Phi_{5,4}) \cdot D_{2,1} & \cdots & (D_{5,4} + \Phi_{5,4}) \cdot D_{5,4} \end{bmatrix}$$

has a nonzero determinant (see Section 2 for the notations).

Firstly we compute the self intersection numbers ∞^2, $(P_{2,1})^2$, $(P_{3,i})^2$, $(P_{4,j})^2$, $(P_{5,k})^2$.

Lemma 3.3. *Let $n = 6l + k$ with $l \ge 0$, $1 \le k \le 6$. Then the canonical divisor $K_{\mathcal{E}_n}$ of the elliptic surface $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ is $K_{\mathcal{E}_n} \cong f^*\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}(l-1)$, and we have $(P)^2 = \infty^2 = -(l+1)$ for each point $P \in E_n(\mathbb{C}(t))$. In particular, if n equals 60, then we have*

$$(P_{2,1})^2 = (P_{3,i})^2 = (P_{4,j})^2 = (P_{5,k})^2 = \infty^2 = -10. \tag{6}$$

Proof. By Kodaira's canonical bundle formula (see [1963a; 1963b]), we get

$$K_{\mathcal{E}_n} \cong f^*\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}(d-2), \quad d = \frac{1}{12} \sum_{t \in \Sigma(\mathcal{E}_n)} \varepsilon(t), \tag{7}$$

where $\varepsilon(t)$ is defined as follows:

type	I_n	II	III	IV	I_n^*	II^*	III^*	IV^*
$\varepsilon(t)$	n	2	3	4	$n+6$	10	9	8

By the reduction type of the singular fibers of $\mathcal{E}_n$, we have $d = l + 1$. Thus

$$K_{\mathcal{E}_n} \cong f^*\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}(d-2) = f^*\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}(l-1). \tag{8}$$

By (1), each point $P \in E_n(\mathbb{C}(t))$ corresponds to a section $\sigma_P : \mathbb{P}^1_{\mathbb{C}} \to \mathcal{E}_n$. The translation-by-$P$ map on E_n can be uniquely extended to a map $\tau_P : \mathcal{E}_n \to \mathcal{E}_n$ by the minimality of $\mathcal{E}_n$ (see [Silverman 1994, Proposition 9.1]). It follows that $\tau_P^* D_1 \cdot \tau_P^* D_2 = D_1 \cdot D_2$ for any two divisors $D_1, D_2 \in \mathrm{Div}(\mathcal{E}_n)$. Hence $(P)^2 = \tau_P^*(P) \cdot \tau_P^*(P) = \infty^2$.

Since ∞ is isomorphic to $\mathbb{P}^1_{\mathbb{C}}$, ∞ is of genus zero. Thus, by the adjunction formula, we get

$$\tfrac{1}{2}(\infty^2 + K_{\mathcal{E}_n} \cdot \infty) + 1 = 0, \text{ that is, } \infty^2 = -(K_{\mathcal{E}_n} \cdot \infty + 2).$$

On the other hand, we can compute

$$\begin{aligned} K_{\mathcal{E}_n} \cdot \infty = (f^*\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}(l-1)) \cdot \infty &= (l-1) f^*(\mathcal{O}_{\mathbb{P}^1_{\mathbb{C}}}(1)) \cdot \infty \\ &= (l-1) C_0 \cdot \infty = l - 1 \end{aligned}$$

by (8). Therefore we get $(P)^2 = \infty^2 = -(l+1)$ for all $P \in E_n(\mathbb{C}(t))$. □

Next we compute the intersection numbers of the divisors ∞, $(P_{2,1})$, $(P_{3,i})$, $(P_{4,j})$, $(P_{5,k})$, $F_{0,a}$ $(1 \le a \le 59)$ in $\mathcal{E}_{60}$. In the affine surface $X_1 : y^2 = x^3 + x^2 + t^{60}$, the divisors $(P_{2,1})$, $(P_{3,i})$, $(P_{4,j})$ and $(P_{5,k})$ are given by

$$\begin{aligned} &(P_{2,1}) = \bigl(x = 0,\ y = -t^{30}\bigr), \\ &(P_{3,i}) = \bigl(x = -\zeta_3^i t^{20},\ y = -\zeta_3^i t^{20}\bigr), \\ &(P_{4,1}) = \bigl(x = \sqrt{2} t^{15} + 2t^{30},\ y = \sqrt{2} t^{15} + 3t^{30} + 2\sqrt{2} t^{45}\bigr), \\ &(P_{4,3}) = \bigl(x = -2 + 2(-1)^{1/4} t^{15},\ y = -2\sqrt{-1} + 4\sqrt{-1}(-1)^{1/4} t^{15} + t^{30}\bigr), \\ &(P_{5,k}) = \bigl(x = 2^{-2/5} \zeta_5^{2k} t^{24},\ y = 2^{-2/5} \zeta_5^{2k} t^{24} + 2^{-3/5} \zeta_5^{3k} t^{36}\bigr). \end{aligned}$$

Thus $(P_{4,3})$ does not pass through the singular point $(0, 0, 0)$ of X_1, however $(P_{2,1})$, $(P_{3,i})$, $(P_{4,1})$ and $(P_{5,k})$ pass through this point. Since this singular point is of type A_{59}, we can resolve it blowing up 30 times. We denote by $(x_{(m)}, y_{(m)}, t_{(m)})$ the coordinates in the neighborhood of the singular point after the m-th blowing-up $(1 \le m \le 29)$, and denote by $(P_{2,1})^{(m)}$, $(P_{3,i})^{(m)}$, $(P_{4,1})^{(m)}$, $(P_{5,k})^{(m)}$ the m-th blowing-up of $(P_{2,1})$, $(P_{3,i})$, $(P_{4,1})$, $(P_{5,k})$, respectively. These curves are given by

$$\begin{aligned} &(P_{2,1})^{(m)} = \bigl(x_{(m)} = 0,\ y_{(m)} = -t_{(m)}^{30-m}\bigr), \\ &(P_{3,i})^{(m)} = \bigl(x_{(m)} = -\zeta_3^i t_{(m)}^{20-m},\ y_{(m)} = \zeta_3^i t_{(m)}^{20-m}\bigr), \\ &(P_{4,1})^{(m)} = \bigl(x_{(m)} = \sqrt{2} t^{15-m} + 2t^{30-m},\ y_{(m)} = \sqrt{2} t^{15-m} + 3t^{30-m} + 2\sqrt{2} t^{45-m}\bigr), \\ &(P_{5,k})^{(m)} = \bigl(x_{(m)} = 2^{-2/5} \zeta_5^{2k} t_{(m)}^{24-m},\ y_{(m)} = 2^{-2/5} \zeta_5^{2k} t_{(m)}^{24-m} + 2^{-3/5} \zeta_5^{3k} t_{(m)}^{36-m}\bigr). \end{aligned}$$

In particular, in $\mathcal{E}_{60}$ the divisors $(P_{3,i})$ (resp. $(P_{4,1})$, $(P_{5,k})$) intersect with either of two $\mathbb{P}^1_{\mathbb{C}}$ which appear by the 20-th (resp. 15, 24-th) blowing-up, and the divisors $(P_{2,1})$ intersect with unique $\mathbb{P}^1_{\mathbb{C}}$ which appears by the 30-th blowing-up. Hence we may assume that $(P_{2,1})$ (resp. $(P_{3,i})$, $(P_{4,1})$, $(P_{4,3})$, $(P_{5,k})$) intersects with $F_{0,30}$ (resp. $F_{0,20}$, $F_{0,15}$, $F_{0,0}$, $F_{0,24}$). In addition, in $X_1 \cap (t \neq 0)$,

$$
\begin{aligned}
&(P_{2,1}) \cap (P_{3,i}) = \varnothing,\\
&(P_{2,1}) \cap (P_{4,1}) = \bigl\{\bigl(0, -\tfrac{1}{2}, t\bigr) \mid t^{15} = -\tfrac{1}{\sqrt{2}}\bigr\},\\
&(P_{2,1}) \cap (P_{4,3}) = \bigl\{(0, -\sqrt{-1}, t) \mid t^{15} = (-1)^{3/4}\bigr\},\\
&(P_{2,1}) \cap (P_{5,k}) = \varnothing,\\
&(P_{3,1}) \cap (P_{3,2}) = \varnothing,\\
&(P_{3,i}) \cap (P_{4,1}) = \bigl\{\bigl(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, t\bigr) \mid t^5 = -\zeta_3^{2i} \tfrac{1}{\sqrt{2}}\bigr\},\\
&(P_{3,i}) \cap (P_{4,3}) = \bigl\{(x(t), y(t), t) \mid \zeta_3^{2i} t^{10} - (1+\sqrt{-1})((-1)^{1/4}\zeta_3^{i} t^5 - 1) = 0\bigr\},\\
&(P_{3,i}) \cap (P_{5,k}) = \varnothing,\\
&(P_{4,1}) \cap (P_{4,3}) = \varnothing,\\
&(P_{4,1}) \cap (P_{5,k}) = \bigl\{(x(t), y(t), t) \mid 2^{7/10}\zeta_5^{3k} t^6 + \zeta_5^{4k} t^3 + 2^{3/10} = 0\bigr\},\\
&(P_{4,3}) \cap (P_{5,k}) = \bigl\{(x(t), y(t), t) \mid \zeta_5^{3k} t^6 \bigl(\zeta_5^{3k} t^6 - 2^{4/5}(-1)^{1/4}\zeta_5^{4k} t^3 + 2^{3/5}\sqrt{-1}\,\bigr)\\
&\qquad\qquad - 2^{1/5}(1+\sqrt{-1})\bigl(2^{1/5}(-1)^{1/4}\zeta_5^{4k} t^3 - \sqrt{-1}\,\bigr) = 0\bigr\},
\end{aligned}
$$

$$(P_{5,k_1}) \cap (P_{5,k_2}) = \varnothing \quad (k_1 \neq k_2).$$

and the local intersection numbers of the divisors $(P_{2,1})$, $(P_{3,i})$, $(P_{4,j})$, $(P_{5,k})$ at these intersection points are all one.

On the other hand, in the ∞-model X_2 (the surface obtained by the variable transformation $\bar{x} = x/t^{20}$, $\bar{y} = y/t^{30}$, $\bar{t} = 1/t$) or its projection $\overline{X}_2$, the divisors $(P_{2,1})$, $(P_{3,i})$, $(P_{4,j})$ and $(P_{5,k})$ are given by

$$
\begin{aligned}
&(P_{2,1}) = (\bar{x} = 0, \bar{y} = -1),\\
&(P_{3,i}) = (\bar{x} = -\zeta_3^{i}, \bar{y} = -\zeta_3^{i} \bar{t}^{10}),\\
&(P_{4,1}) = \bigl\{([\sqrt{2}\bar{t}^{20} + 2\bar{t}^5 : \sqrt{2}\bar{t}^{30} + 3\bar{t}^{15} + 2^{2/3} : \bar{t}^{15}], \bar{t}) \mid \bar{t} \in \mathbb{A}^1\bigr\},\\
&(P_{4,3}) = \bigl(\bar{x} = -2\bar{t}^{20} + 2(-1)^{1/4}\bar{t}^5, \bar{y} = -2\sqrt{-1}\bar{t}^{30} + 4\sqrt{-1}(-1)^{1/4}\bar{t}^{15} + 1\bigr),\\
&(P_{5,k}) = \bigl\{([2^{-2/5}\zeta_5^{2k}\bar{t}^2 : 2^{-2/5}\zeta_5^{2k}\bar{t}^{12} + 2^{-3/5}\zeta_5^{3k} : \bar{t}^6], \bar{t}) \mid \bar{t} \in \mathbb{A}^1\bigr\}.
\end{aligned}
$$

Thus when $\bar{t}$ equals 0 (t equals ∞), the divisors $(P_{4,1})$, $(P_{5,k})$ and ∞ intersect at $([0 : 1 : 0], 0) \in \overline{X}_2$ and the other pairs of $(P_{2,1})$, $(P_{3,i})$, $(P_{4,j})$, $(P_{5,k})$ and ∞ do not intersect. Moreover the local intersection number of $(P_{4,1})$ and ∞ at this point is five, and the numbers of the other pairs of $(P_{4,1})$, $(P_{5,k})$ and ∞ are two. From the above and (6), we obtain

$$(P_{2,1})\cdot(P_{d,l})=\begin{cases}-10 & (d=2),\\ 0 & (d=3,5),\\ 15 & (d=4),\end{cases}$$

$$(P_{3,i})\cdot(P_{d,l})=\begin{cases}-10 & (d=3,\, i=l),\\ 0 & (d=3,\, i\neq l),\\ 5 & (d=4,\, j=1),\\ 10 & (d=4,\, j=3),\\ 0 & (d=5),\end{cases}$$

$$(P_{4,j})\cdot(P_{d,l})=\begin{cases}-10 & (d=4,\, j=l),\\ 0 & (d=4,\, j\neq l),\\ 8 & (d=5,\, j=1),\\ 12 & (d=5,\, j=3),\end{cases}$$

$$(P_{5,k})\cdot(P_{5,l})=\begin{cases}-10 & (k=l),\\ 2 & (k\neq l).\end{cases}$$

Finally we give $\Phi_{d,l}$ for $d\in\mathrm{Adm}_4$, $l\in(\mathbb{Z}/d\mathbb{Z})^\times$ by the method mentioned in Section 2 and compute $(D_{d,l}+\Phi_{d,l})\cdot D_{d',l'}$ where $d,\, d'\in\mathrm{Adm}_4$, $l\in(\mathbb{Z}/d\mathbb{Z})^\times$, $l'\in(\mathbb{Z}/d'\mathbb{Z})^\times$. Recall that $\Phi_{d,l}$ is defined by $\Phi_{d,l}=\sum_{i=1}^{59}a_i(P_{d,l})F_{0,i}$ and

$$(a_1(P_{d,l}),\dots,a_{59}(P_{d,l}))=-(D_{d,l}\cdot F_{0,1},\dots,D_{d,l}\cdot F_{0,59})(F_{0,i}\cdot F_{0,j})^{-1}.$$

For integers $1\le m\le 59$, since $\infty\cdot F_{0,m}=0$, we have, in Kronecker delta notation,

$$D_{2,1}\cdot F_{0,m}=(P_{2,1})\cdot F_{0,m}=\delta_{m,30},$$

$$D_{3,i}\cdot F_{0,m}=(P_{3,i})\cdot F_{0,m}=\delta_{m,20},$$

$$D_{4,1}\cdot F_{0,m}=(P_{4,1})\cdot F_{0,m}=\delta_{m,15},$$

$$D_{4,3}\cdot F_{0,m}=(P_{4,3})\cdot F_{0,m}=0,$$

$$D_{5,k}\cdot F_{0,m}=(P_{5,k})\cdot F_{0,m}=\delta_{m,24}.$$

Since the reduction type of $(\mathcal{E}_{60})_0$ is I_{60}, we have the intersection matrix

$$(F_{0,i}\cdot F_{0,j})_{1\le i,j\le 59}=\begin{bmatrix}-2 & 1 & \cdots & 0\\ 1 & -2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 1\\ 0 & \cdots & 1 & -2\end{bmatrix}.$$

It is an easy exercise to show that the j-th row of the inverse of this matrix is $\frac{-1}{60}[60-j,\, 2\cdot(60-j),\,\dots,\, j\cdot(60-j),\, j\cdot(59-j),\dots j\cdot 2,\, j]$. Therefore we obtain

$a_1(P_{4,3}) = \cdots = a_{59}(P_{4,3}) = 0$ and

$$\begin{aligned}
(a_1(P_{2,1}), \ldots, a_{59}(P_{2,1})) &= \tfrac{1}{60}[30, 2\cdot 30, \ldots, 30\cdot 30, \ldots, 30\cdot 2, 30],\\
(a_1(P_{3,i}), \ldots, a_{59}(P_{3,i})) &= \tfrac{1}{60}[40, 2\cdot 40, \ldots, 20\cdot 40, \ldots, 20\cdot 2, 20],\\
(a_1(P_{4,1}), \ldots, a_{59}(P_{4,1})) &= \tfrac{1}{60}[45, 2\cdot 45, \ldots, 15\cdot 45, \ldots, 15\cdot 2, 15],\\
(a_1(P_{5,k}), \ldots, a_{59}(P_{5,k})) &= \tfrac{1}{60}[36, 2\cdot 36, \ldots, 24\cdot 36, \ldots, 24\cdot 2, 24].
\end{aligned}$$

In particular, $\Phi_{4,3} = 0$ and $\Phi_{3,i}$, $\Phi_{5,k}$ do not depend on i, k, so we put $\Phi_2 = \Phi_{2,1}$, $\Phi_3 = \Phi_{3,i}$, $\Phi_4 = \Phi_{4,1}$, $\Phi_5 = \Phi_{5,k}$. We can compute $(D_{d,l} + \Phi_d)\cdot D_{d',l'}$, where $d, d' \in \mathrm{Adm}_4$, $l \in (\mathbb{Z}/d\mathbb{Z})^\times$, $l' \in (\mathbb{Z}/d'\mathbb{Z})^\times$, as follows:

$$(D_{2,1} + \Phi_2)\cdot D_{d,l} = \begin{cases} -5 & (d = 2),\\ 0 & (d = 3),\\ \frac{15}{2} & (d = 4, l = 1),\\ 5 & (d = 4, l = 3),\\ 0 & (d = 5),\end{cases}$$

$$(D_{3,i} + \Phi_3)\cdot D_{d,l} = \begin{cases} -\frac{20}{3} & (d = 3, i = l),\\ \frac{10}{3} & (d = 3, i \neq l),\\ 0 & (d = 4, 5),\end{cases}$$

$$(D_{4,j} + \Phi_4)\cdot D_{d,l} = \begin{cases} -\frac{75}{4} & (d = 4, j = l = 1),\\ -20 & (d = 4, j = l = 3),\\ -15 & (d = 4, j \neq l),\\ 0 & (d = 5),\end{cases}$$

$$(D_{5,k} + \Phi_5)\cdot D_{d,l} = \begin{cases} -\frac{48}{5} & (d = 5, k = l),\\ \frac{12}{5} & (d = 5, k \neq l).\end{cases}$$

Since $(D_{d,l} + \Phi_d)\cdot F = 0$ for all fibral divisors F, we have $(D_{d,l} + \Phi_d)\cdot D_{d',l'} = (D_{d',l'} + \Phi_{d'})\cdot D_{d,l}$. Thus we obtain

$$N = \begin{bmatrix}
-5 & 0 & 0 & \frac{15}{2} & 5 & 0 & 0 & 0 & 0\\
0 & -\frac{20}{3} & \frac{10}{3} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \frac{10}{3} & -\frac{20}{3} & 0 & 0 & 0 & 0 & 0 & 0\\
\frac{15}{2} & 0 & 0 & -\frac{75}{4} & -15 & 0 & 0 & 0 & 0\\
5 & 0 & 0 & -15 & -20 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -\frac{48}{5} & \frac{12}{5} & \frac{12}{5} & \frac{12}{5}\\
0 & 0 & 0 & 0 & 0 & \frac{12}{5} & -\frac{48}{5} & \frac{12}{5} & \frac{12}{5}\\
0 & 0 & 0 & 0 & 0 & \frac{12}{5} & \frac{12}{5} & -\frac{48}{5} & \frac{12}{5}\\
0 & 0 & 0 & 0 & 0 & \frac{12}{5} & \frac{12}{5} & \frac{12}{5} & -\frac{48}{5}
\end{bmatrix},$$

and $\det N = -2^8 3^5 5^4 \neq 0$. Therefore the divisors C_0, ∞, $D_{2,1}$, $D_{3,i}$, $D_{4,j}$, $D_{5,k}$, $F_{0,1}, \ldots, F_{0,59}$ form a $\mathbb{Q}$-basis of $\mathrm{NS}(\mathcal{E}_{60})$.

Similarly in the cases of $n \leq 6$, we can compute $\det M = \pm 1$, where M is the intersection matrix of the divisors in Theorem 3.2 (see Lemma 2.2). In particular, the divisors form a $\mathbb{Z}$-basis of $\mathrm{NS}(\mathcal{E}_n)$. □

Stiller's Example 1. This example is the minimal elliptic surface whose generic fiber is the elliptic curve defined by

$$Y^2 = 4X^3 - 3u^{3n}X - u^{5n} \quad (u \in \mathbb{P}^1_{\mathbb{C}}, n \in \mathbb{N})$$

over $\mathbb{C}(u)$. By changing the variables suitably, the defining equation becomes

$$y^2 = x^3 + t^n x + t^n \quad (t \in \mathbb{P}^1_{\mathbb{C}}). \tag{9}$$

We denote by E_n the elliptic curve defined by (9) and by $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ the associated elliptic surface. Stiller [1987] proved that the Mordell–Weil rank $r = \mathrm{rank}(E_n(\mathbb{C}(t)))$ is given by

$$r = \sum_{\substack{d \mid n \\ d \in \mathrm{Adm}_1}} \varphi(d), \tag{10}$$

where φ is the Euler function and $\mathrm{Adm}_1 = \{1, 2, 3, 7, 8, 10, 12, 15, 18, 20, 42\}$.

Remark 3.4. For use in Section 4, we note that $d \in \mathrm{Adm}_1$ if and only if each $j \in \{j \in \mathbb{N} : 9d \leq 12j \leq 10d\}$ is not relatively prime to d. Such d's are called *admissible* in [Stiller 1987].

Definition 3.5. For $d \in \mathrm{Adm}_1$ and $j \in (\mathbb{Z}/d\mathbb{Z})^\times$, we define $\mathbb{C}(t)$-rational points $P_{d,j}$ of E_d as follows.

$$\begin{aligned}
P_{1,1} &= \big(-1, \sqrt{-1}\big), \\
P_{2,1} &= \big(\sqrt{-1}\, t, -t\big), \\
P_{3,j} &= \big(-\zeta_3^j t, \sqrt{-1}\zeta_3^{2j} t^2\big), \\
P_{7,j} &= \big(-\zeta_7^{2j} t^2 - \zeta_7^{3j} t^3, \sqrt{-1}(\zeta_7^{3j} t^3 + \zeta_7^{4j} t^4 + \zeta_7^{5j} t^5)\big), \\
P_{8,j} &= \left(\sum_{i=0}^{2} a_i(8, j) t^{i+2}, \sum_{i=0}^{3} b_i(8, j) t^{i+3}\right), \\
P_{10,j} &= \big(2^{2/5}\zeta_{10}^{4j} t^4, -\zeta_{10}^{5j} t^5 - 2^{1/5}\zeta_{10}^{7j} t^7\big), \\
P_{12,j} &= \left(\sum_{i=0}^{2} a_i(12, j) t^{4+i}, \sum_{i=0}^{3} b_0(12, j) t^{6+i}\right),
\end{aligned}$$

$$
\begin{aligned}
P_{15,j} &= \big(-\zeta_{15}^{5j}t^5 - 3^{1/5}\zeta_{15}^{6j}t^6 - 3^{2/5}\zeta_{15}^{7j}t^7, \\
&\qquad \sqrt{-1}(3^{3/5}\zeta_{15}^{8j}t^8 + 3^{4/5}\zeta_{15}^{9j}t^9 + 2\zeta_{15}^{10j}t^{10} + 3^{1/5}\zeta_{15}^{11j}t^{11})\big), \\
P_{18,j} &= \left(\sum_{i=0}^{2} a_i(18,j)(\zeta_{18}^{j}t)^{6+2i}, \sum_{i=0}^{3} b_i(18,j)(\zeta_{18}^{j}t)^{9+2i}\right), \\
P_{20,j} &= \left(\sum_{i=0}^{2} a_i(20,j)(\zeta_{20}^{j}t)^{6+2i}, \sum_{i=0}^{3} b_i(20,j)(\zeta_{20}^{j}t)^{9+2i}\right), \\
P_{42,j} &= \left(\sum_{i=1}^{5} a_i(\zeta_{42}^{j}t)^{12+2i}, \sum_{i=1}^{7} b_i(\zeta_{42}^{j}t)^{19+2i}\right),
\end{aligned}
$$

where the coefficients $a_k(d,j)$, $b_k(d,j)$ are given by Table 2 and the set of complex numbers $(a_1,\dots,a_5,b_1,\dots,b_7)$ are solutions of a system of equations

$$
\begin{aligned}
b_7^2 &= a_5^3, \\
2b_6b_7 &= 3a_4a_5^2 + a_5, \\
2b_5b_7 + b_6^2 &= 3a_3a_5^2 + 3a_4^2a_5 + a_4, \\
2b_4b_7 + 2b_5b_6 &= 3a_2a_5^2 + 6a_3a_4a_5 + a_4^3 + a_3, \\
2b_3b_7 + 2b_4b_6 + b_5^2 &= 3a_1a_5^2 + (6a_2a_4 + 3a_3^2)a_5 + 3a_3a_4^2 + a_2, \\
2b_2b_7 + 2b_3b_6 + 2b_4b_5 &= (6a_1a_4 + 6a_2a_3)a_5 + 3a_2a_4^2 + 3a_3^2a_4 + a_1, \\
2b_1b_7 + 2b_2b_6 + 2b_3b_5 + b_4^2 &= (6a_1a_3 + 3a_2^2)a_5 + 3a_1a_4^2 + 6a_2a_3a_4 + a_3^3, \\
2b_1b_6 + 2b_2b_5 + 2b_3b_4 &= 6a_1a_2a_5 + (6a_1a_3 + 3a_2^2)a_4 + 3a_2a_3^2, \\
2b_1b_5 + 2b_2b_4 + b_3^2 &= 3a_1^2a_5 + 6a_1a_2a_4 + 3a_1a_3^2 + 3a_2^2a_3, \\
2b_1b_4 + 2b_2b_3 &= 3a_1^2a_4 + 6a_1a_2a_3 + a_2^3, \\
2b_1b_3 + b_2^2 &= 3a_1^2a_3 + 3a_1a_2^2, \\
2b_1b_2 &= 3a_1^2a_2, \\
b_1^2 &= a_1^3 + 1.
\end{aligned}
$$

The system is given by comparing the coefficients of $(b_1t^{21} + b_2t^{23} + \cdots + b_7t^{33})^2$ with those of $(a_1t^{14} + a_2t^{16} + \cdots + a_5t^{22})^3 + t^{42}(a_1t^{14} + a_2t^{16} + \cdots + a_5t^{22}) + t^{42}$.

Theorem 3.6. *Let $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}} (n \in \mathbb{N})$ be the elliptic surfaces associated to the elliptic curves $E_n : y^2 = x^3 + t^n x + t^n$ $(n \in \mathbb{N})$ over $\mathbb{P}^1_{\mathbb{C}}$. Then, for each $n \in \mathbb{N}$, $\mathrm{NS}(\mathcal{E}_n)$ has a $\mathbb{Q}$-basis $C_0, \infty, D_{d,j}, F_{t,a}(d \in \mathrm{Adm}_1, d|n,\ j \in (\mathbb{Z}/d\mathbb{Z})^\times, t \in \Sigma(\mathcal{E}_n), 1 \le a \le m_t - 1)$. Moreover if $\mathcal{E}_n$ is rational (i.e., $n = 1, 2, 3, 4, 6, 7, 8$ or 12), then these divisors form a $\mathbb{Z}$-basis.*

The argument of the proof is the same as the previous one, though the computation is very complicated.

j	1	3	5	7
$a_0(8, j)$	$2^{-1/2}$	$-\sqrt{2}$	$\sqrt{-1}+(-1)^{1/4}+1$	0
$a_1(8, j)$	0	$2^{3/4}$	0	$a_2(8, 7)^{1/4}(2^{5/2}+4)$
$a_2(8, j)$	0	-1	$\sqrt{-1}$	$\sqrt{-2\sqrt{2}-3}$
$b_0(8, j)$	$2^{-3/4}$	$-2^{3/4}\sqrt{-1}$	$(\sqrt{-1}+(-1)^{1/4}+1)^{3/2}$	0
$b_1(8, j)$	0	$3\sqrt{-1}$	0	1
$b_2(8, j)$	$2^{-1/4}$	$-2^{5/4}\sqrt{-1}$	0	$a_2(8, 7)^{3/4}\sqrt{2\sqrt{2}+2}(\sqrt{2}+1)^{1/4}$
$b_3(8, j)$	0	$\sqrt{-2}$	$\big(3(-1)^{1/4}(\sqrt{-1}-1)-4\big) \big/ \big(\sqrt{-1}+(-1)^{1/4}+1\big)^{3/2}$	$a_2(8, 7)^{3/2}\sqrt{2\sqrt{2}+2}(\sqrt{2}-1)^{1/4}$

j	1	5	7	11
$a_0(12, j)$	$-\frac{1}{12}a_1(12, 1)^2a_2(12, 1)(-33+5a_2(12, 1)^2)$	-1	$(1+\sqrt{-3})/2$	$\frac{1}{4}a_1(12, 11)^2a_2(12, 11)(5+3a_2(12, 11)^2)$
$a_1(12, j)$	a root of $z^6+180a_2(12, 1)+388a_2(12, 1)^3=0$	0	0	a root of $z^6-12a_2(12, 11)+36a_2(12, 11)^3=0$
$a_2(12, j)$	$\sqrt{2\sqrt{3}+3}$	$\sqrt{(2/3)\sqrt{3}-1}$	$\sqrt{(2/3)\sqrt{3}-1}$	$\sqrt{(2/3)\sqrt{3}-1}$

j	1, 5, 7, 11
$b_0(12, j)$	$\big(4a_0(12, j)a_1(12, j)a_2(12, j)(a_2(12, j)^2+1)(3a_2(12, j)^4+6a_2(12, j)^2-1) - a_1(12, j)^3(a_2(12, j)^2-1)(a_2(12, j)^4+6a_2(12, j)^2+1)\big)/16b_3(12, j)^5$
$b_1(12, j)$	$\big(4a_0(12, j)a_2(12, j)(a_2(12, j)^2+1)(3a_2(12, j)^2+1)+a_1(12, j)^2(3a_2(12, j)^4+6a_2(12, j)^2-1)\big)/8b_3(12, j)^3$
$b_2(12, j)$	$\big(a_1(12, j)(3a_2(12, j)^2+1)\big)/b_3(12, j)$
$b_3(12, j)$	$\sqrt{a_2(12, j)^3+a_2(12, j)}$

j	1, 11	5, 13
$a_0(18, j)$	0	$2b_1(18, j)b_2(18, j)-b_2(18, j)^6$
$a_1(18, j)$	$-2^{2/9}3^{-1/3}$	$b_2(18, j)^2$
$a_2(18, j)$	$2^{-2/9}3^{-2/3}$	0

j	1, 11	5, 13
$b_0(18, j)$	1	a root of $z^3 - 9z^2 - 9z + 9 = 0$
$b_1(18, j)$	0	a root of $8919936 - 8011872 b_0(18, j) - 9735552 b_0(18, j)^2 + z^9 = 0$
$b_2(18, j)$	0	$\frac{1}{36} b_1(18, j)^2(-45 - 15 b_0(18, j) + 2 b_0(18, j)^2)$
$b_3(18, j)$	$-2^{-1/3}3^{-1}$	0

j	7, 17
$a_0(18, j)$	$-4^{1/3}$
$a_1(18, j)$	a root of $4 + 3a_0(18, j)^2 z^3 + 6a_0(18, j) z^6 + z^9 = 0$
$a_2(18, j)$	$-\frac{1}{12} a_1(18, j)^2 (5a_0(18, j)^2 + 6a_0(18, j) a_1(18, j)^3 + a_1(18, j)^6)$
$b_0(18, j)$	$\big((12a_1(18, j)a_2(18, j)^4 - 4a_2(18, j)^3)a_0(18, j) - a_1(18, j)^3 a_2(18, j)^3 + 3a_1(18, j)^2 a_2(18, j)^2 + 5a_1(18, j)a_2(18, j) + 1\big)/16a_2(18, j)^{9/2}$
$b_1(18, j)$	$\big(12a_0(18, j)a_2(18, j)^3 + 3a_1(18, j)^2 a_2(18, j)^2 - 2a_1(18, j)a_2(18, j) - 1\big)/8a_2(18, j)^{5/2}$
$b_2(18, j)$	$\big(3a_1(18, j)a_2(18, j)^2 + a_2(18, j)\big)/2a_2(18, j)^{3/2}$
$b_3(18, j)$	$a_2(18, j)^{3/2}$

j	1, 3, 7, 9
$a_0(20, j)$	$\frac{1}{80} a_1(20, j)^2 a_2(20, j)(172 - 220a_2(20, j)^2 + 117a_1(20, j)^5 + 205a_1(20, j)^5 a_2(20, j)^2)$
$a_1(20, j)$	a root of $56 - 328a_2(20, j)^2 - 2500a_2(20, j)^2 z^5 + 625z^{10} = 0$
$a_2(20, j)$	a root of $5z^4 + 2z^2 + 1 = 0$

j	11, 13, 17, 19
$a_0(20, j)$	$-\frac{1}{16} a_1(20, j)^2 a_2(20, j)(-191 - 188a_2(20, j)^2 + 69a_2(20, j)^4 + 6a_2(20, j)^6)$
$a_1(20, j)$	a root of $z^5 + 1 - 45a_2(20, j)^2 - 15a_2(20, j)^4 + 15a_2(20, j)^6 = 0$
$a_2(20, j)$	a root of $z^8 + 12z^6 - 26z^4 - 52z^2 + 1 = 0$

j	1, 3, 7, 9, 11, 13, 17, 19
$b_0(20, j)$	$a_0(20, j)^{3/2}$
$b_1(20, j)$	$\big(3a_0(20, j)^2 a_1(20, j) + 1\big)/a_0(20, j)^{3/2}$
$b_2(20, j)$	$\big(12a_0(20, j)^5 a_2(20, j) + 3a_0(20, j)^4 a_1(20, j)^2 - 6a_0(20, j)^2 a_1(20, j) - 1\big)/8a_0(20, j)^{9/2}$
$b_3(20, j)$	$\big(12(a_0(20, j)^7 a_1(20, j) - a_0(20, j)^5)a_2(20, j) - a_0(20, j)^6 a_1(20, j)^3 + 15a_0(20, j)^4 a_1(20, j)^2 + 9a_0(20, j)^2 a_1(20, j) + 1\big)/16a_0(20, j)^{15/2}$

Table 2. Coefficients $a_k(d, j), b_k(d, j)$

Stiller's Example 2. This example is the minimal elliptic surface whose generic fiber is the elliptic curve defined by

$$Y^2 = 4X^3 - 3u^n X - u^{2n} \quad (u \in \mathbb{P}^1_{\mathbb{C}}, n \in \mathbb{N})$$

over $\mathbb{C}(u)$. By putting $X = -9x/4$, $Y = 27\sqrt{-1}y/4$, $u = \sqrt[n]{-27/4}t$, the defining equation becomes

$$y^2 = x^3 + t^n x + t^{2n} \quad (t \in \mathbb{P}^1_{\mathbb{C}}). \tag{11}$$

We denote by E_n^2 the elliptic curve defined by (11) and by $f : \mathcal{E}_n^2 \to \mathbb{P}^1_{\mathbb{C}}$ the associated elliptic surface. Similarly to Stiller's Example 1, the Mordell–Weil rank $r = \operatorname{rank}(E_n(\mathbb{C}(t)))$ is given by

$$r = \sum_{\substack{d \mid n \\ d \in \mathrm{Adm}_2}} \varphi(d),$$

where φ is the Euler function and $\mathrm{Adm}_2 = \{1, 2, 5, 6, 8, 9, 12, 14, 20, 21, 30\}$.

We now denote by $\mathcal{E}_n^1$ the elliptic surface of Stiller's Example 1 and by E_n^1 the associated elliptic curve. Recall that

$$E_n^1 : y^2 = x^3 + t^n x + t^n.$$

We assume that n is even and write $n = 2m$. By putting $\bar{x} = x/t^{2m}$, $\bar{y} = y/t^{3m}$, $\bar{t} = 1/t$, we obtain

$$\bar{y}^2 = \bar{x}^3 + \bar{t}^{2m}\bar{x} + \bar{t}^{4m}.$$

Since this equation is the defining equation of E_{2m}^2, we obtain an isomorphism

$$\begin{aligned} \mathcal{E}_{2m}^1 &\xrightarrow{\sim} \mathcal{E}_{2m}^2, \\ (x, y, t) &\longmapsto (x/t^{2m}, y/t^{3m}, 1/t). \end{aligned} \tag{12}$$

Therefore we define the $\varphi(d)$ rational points of E_d^2 for each $d (\geq 2) \in \mathrm{Adm}_2$ as the images of the points of Definition 3.5 via this isomorphism.

Definition 3.7. For $d \in \mathrm{Adm}_2$ and $j \in (\mathbb{Z}/d\mathbb{Z})^\times$, we define $\mathbb{C}(t)$-rational points $P_{d,j}$ of E_d^2 as follows.

$$\begin{aligned}
P_{1,1} &= (0, -t), \\
P_{2,1} &= \left(\sqrt{-1}\, t, -t^2\right), \\
P_{5,j} &= \left(2^{2/5}\zeta_5^{3j} t^3, -2^{1/5}\zeta_5^{4j} t^4 - t^5\right), \\
P_{6,j} &= \left(-\zeta_6^{4j} t^4, \sqrt{-1}\zeta_6^{5j} t^5\right), \\
P_{8,j} &= \left(\sum_{i=0}^{2} a_i(8, j) t^{6-i}, \sum_{i=0}^{3} b_i(8, j) t^{9-i}\right),
\end{aligned}$$

$$
\begin{aligned}
P_{9,j} &= \left(\sum_{i=0}^{2} a_i(18, \tilde{j})(\zeta_9^j t)^{6-i}, \sum_{i=0}^{3} b_i(18, \tilde{j})(\zeta_9^j t)^{9-i}\right),\\
P_{12,j} &= \left(\sum_{i=0}^{2} a_i(12, j)t^{8-i}, \sum_{i=0}^{3} b_i(12, j)t^{12-i}\right),\\
P_{14,j} &= \left(-\zeta_{14}^{8j}t^8 - \zeta_{14}^{10j}t^{10}, \sqrt{-1}(\zeta_{14}^{11j}t^{11} + \zeta_{14}^{13j}t^{13} + \zeta_{14}^{15j}t^{15})\right),\\
P_{20,j} &= \left(\sum_{i=0}^{2} a_i(20, j)(\zeta_{20}^j t)^{14-2i}, \sum_{i=0}^{3} b_i(20, j)(\zeta_{20}^j t)^{21-2i}\right),\\
P_{21,j} &= \left(\sum_{i=1}^{5} a_i(\zeta_{21}^j t)^{15-i}, \sum_{i=1}^{7} b_i(\zeta_{21}^j t)^{22-i}\right),\\
P_{30,j} &= \big(-3^{2/5}\zeta_{30}^{16j}t^{16} - 3^{1/5}\zeta_{30}^{18j}t^{18} - \zeta_{30}^{20j}t^{20},\\
&\qquad \sqrt{-1}(3^{1/5}\zeta_{30}^{23j}t^{23} + 2\zeta_{30}^{25j}t^{25} + 3^{4/5}\zeta_{30}^{27j}t^{27} + 3^{1/5}\zeta_{30}^{29j}t^{29})\big),
\end{aligned}
$$

where $\tilde{j}$ equals j if j is odd and equals $j+9$ if j is even, and the coefficients $a_k(d, j), b_k(d, j), a_1, \dots, a_5, b_1, \dots, b_7$ are same as them of Definition 3.5.

The following theorem follows from the isomorphism (12) and Theorem 3.6.

Theorem 3.8. *Let $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}} (n \in \mathbb{N})$ be the elliptic surfaces associated to the elliptic curves $E_n : y^2 = x^3 + t^n x + t^{2n} (n \in \mathbb{N})$ over $\mathbb{P}^1_{\mathbb{C}}$. Then, for each $n \in \mathbb{N}$, $\mathrm{NS}(\mathcal{E}_n)$ has a $\mathbb{Q}$-basis $C_0, \infty, D_{d,j}, F_{t,a} (d \in \mathrm{Adm}_2, d|n, j \in (\mathbb{Z}/d\mathbb{Z})^\times, t \in \Sigma(\mathcal{E}_n), 1 \le a \le m_t - 1)$. Moreover if $\mathcal{E}_n$ is rational* (*i.e.*, $n = 1, 2, 3, 4, 5, 6, 8, 9$ *or* 12), *then these divisors form a $\mathbb{Z}$-basis.*

Stiller's Example 3. Here we consider the minimal elliptic surface whose generic fiber is the elliptic curve defined by

$$Y^2 = 4X^3 - 3u^{3n}\big(u^n - \tfrac{8}{9}\big)X + u^{4n}\big(u^{2n} - \tfrac{4}{3}u^n + \tfrac{8}{27}\big) \quad (u \in \mathbb{P}^1_{\mathbb{C}}, n \in \mathbb{N})$$

over $\mathbb{C}(u)$. By changing the variables suitably, the defining equation becomes

$$y^2 = x^3 + x^2 + t^n x + \tfrac{1}{4}t^{2n} \quad (t \in \mathbb{P}^1_{\mathbb{C}}). \tag{13}$$

We denote by E_n the elliptic curve defined by (13) and by $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ the associated elliptic surface. By [Stiller 1987, Example 3], the Mordell–Weil rank $r = \mathrm{rank}(E_n(\mathbb{C}(t)))$ equals 1 if n is even and equals 0 if n is odd. Therefore similarly to Examples 1, 2 and 4, we obtain $r = \sum_{\substack{d \in \mathrm{Adm}_3 \\ d|n}} \varphi(d)$, $\mathrm{Adm}_3 = \{2\}$, and we can show:

Theorem 3.9. *Let $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}} (n \in \mathbb{N})$ be the elliptic surfaces associated to the elliptic curves $E_n : y^2 = x^3 + x^2 + t^n x + t^{2n}/4$ $(n \in \mathbb{N})$ over $\mathbb{P}^1_{\mathbb{C}}$. For each $n \in \mathbb{N}$, if n*

is odd, then $\mathrm{NS}(\mathcal{E}_n)$ *has a* $\mathbb{Z}$*-basis* $C_0, \infty, F_{t,a}$ $(t \in \Sigma(\mathcal{E}_n), 1 \le a \le m_t - 1)$, *and if* n *is even, then the group has a* $\mathbb{Q}$*-basis* $C_0, \infty, D_{2,1}, F_{t,a}$ $(t \in \Sigma(\mathcal{E}_n), 1 \le a \le m_t - 1)$, *where* $D_{2,1} = (P_{2,1}) - \infty$ *and* $\mathbb{C}(t)$*-rational point* $P_{2,1}$ *is defined by*

$$P_{2,1} = \left(-\tfrac{1}{2}t^n, \tfrac{1}{4}\sqrt{-2}t^{3n/2}\right).$$

Moreover if $\mathcal{E}_n$ *is rational* (*i.e.*, $n \le 3$), *then these divisors form a* $\mathbb{Z}$*-basis.*

Stiller's Example 5. We finally consider the minimal elliptic surface whose generic fiber is the elliptic curve defined by the equation†

$$Y^2 = 4X^3 - 3u^{12k+3}\left(u^{4k+1} - \tfrac{3}{4}\right)X - u^{20k+5}\left(u^{4k+1} - \tfrac{9}{8}\right) \quad (u \in \mathbb{P}^1_{\mathbb{C}}, k \in \mathbb{N})$$

over $\mathbb{C}(u)$. By changing the variables suitably, the defining equation becomes

$$y^2 = x^3 + x^2 + t^{4k+1}x \quad (t \in \mathbb{P}^1_{\mathbb{C}}).$$

Here we discuss a slightly more general equation:

$$y^2 = x^3 + x^2 + t^n x \quad (t \in \mathbb{P}^1_{\mathbb{C}}, n \in \mathbb{N}). \tag{14}$$

We denote by E_n the elliptic curve defined by (14) and by $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ the associated elliptic surface. The surface $\mathcal{E}_n$ has singular fibers of type I_{2n} over 0, type I_1 over $\zeta_n^i \sqrt[n]{1/4}$ $(0 \le i \le n-1)$ and III^* (resp. I_0^*, III, I_0) over ∞ as $n \equiv 1$ (resp. 2, 3, 0) modulo 4. Using Stiller's method, one can show that the Mordell–Weil rank $r = \operatorname{rank}(E_n(\mathbb{C}(t)))$ is given by

$$r = \sum_{\substack{d \mid n \\ d \in \mathrm{Adm}_5}} \varphi(d),$$

where φ is the Euler function and $\mathrm{Adm}_5 = \{2, 3\}$. We obtain the following theorem similarly to the other examples.

Theorem 3.10. *Let* $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ $(n \in \mathbb{N})$ *be the elliptic surfaces associated to the elliptic curves* $E_n : y^2 = x^3 + x^2 + t^n x$ $(n \in \mathbb{N})$ *over* $\mathbb{P}^1_{\mathbb{C}}$*. Then, for each* $n \in \mathbb{N}$, $\mathrm{NS}(\mathcal{E}_n)$ *has a* $\mathbb{Q}$*-basis* $C_0, \infty, D_{d,j}, F_{t,a}$ $(d \in \mathrm{Adm}_5, d \mid n, j \in (\mathbb{Z}/d\mathbb{Z})^\times, t \in \Sigma(\mathcal{E}_n), 1 \le a \le m_t - 1)$, *where* $D_{d,j} = (P_{d,j}) - \infty$ *and* $\mathbb{C}(t)$*-rational points* $P_{2,1}, P_{3,j}$ *are defined by*

$$\begin{aligned} P_{2,1} &= \left(\sqrt{-1}\, t^{n/2}, \sqrt{-1}\, t^{n/2}\right), \\ P_{3,j} &= \left(2^{2/3}\zeta_3^j t^{n/3}, 2^{2/3}\zeta_3^j t^{n/3} + 2^{1/3}\zeta_3^{2j} t^{2n/3}\right). \end{aligned}$$

Moreover if $\mathcal{E}_n$ *is rational* (*i.e.*, $n \le 4$), *then these divisors form a* $\mathbb{Z}$*-basis.*

†The equation in [Stiller 1987, page 188] is incorrect.

4. Alternative proof of Stiller's computations

Each surface in [Stiller 1987, Examples 1–5] has an automorphism such that it acts in multiplicity one (i.e., each eigenspace is one-dimensional) on the second de Rham cohomology modulo zero and fibral divisor classes. Stiller showed this by using the *inhomogeneous* de Rham cohomology, and this is essentially used in his argument on the computation of Picard numbers.

We gave the explicit $\mathbb{Q}$-bases of the Néron–Severi groups in the last section, where we used his results on the Picard numbers. However once one has the divisors as in the last section, one can conclude that they automatically form a $\mathbb{Q}$-basis of the Néron–Severi group. We show it in this section.

Let $f : \mathcal{E}_n \to \mathbb{P}^1_{\mathbb{C}}$ ($n \in \mathbb{N}$) be one of the families of elliptic surfaces of Examples 1–5. Let $\mathrm{NS}(\mathcal{E}_n)'$ be the subgroup of $\mathrm{NS}(\mathcal{E}_n)$ which is generated by all the divisors in Theorem 3.2, 3.6, 3.8, 3.9 or 3.10, as the case may be. Put $H^2_{\mathrm{tr}}(\mathcal{E}_n) = H^2(\mathcal{E}_n, \mathbb{Q})/\mathrm{NS}(\mathcal{E}_n)_{\mathbb{Q}}$, $V(\mathcal{E}_n) = H^2(\mathcal{E}_n, \mathbb{Q})/\mathrm{NS}(\mathcal{E}_n)'_{\mathbb{Q}}$. The goal is to show that $\mathrm{NS}(\mathcal{E}_n)_{\mathbb{Q}} = \mathrm{NS}(\mathcal{E}_n)'_{\mathbb{Q}}$, or equivalently

$$\dim V(\mathcal{E}_n) \le \dim H^2_{\mathrm{tr}}(\mathcal{E}_n). \tag{15}$$

We give a proof of (15) only for Stiller's Example 1 since the same argument works in the other cases. We already know the dimension of $V(\mathcal{E}_n)$. In the case at hand, the result can be written as

$$\dim V(\mathcal{E}_n) = \sum_{d \in S^1_n} \varphi(d), \tag{16}$$

where we put $S^1_n = \{d \in \mathbb{N} : d|n,\ d \notin \mathrm{Adm}_1 \cup \{4, 6\}\}$ and $\varphi(d)$ is the Euler function. In particular, when $n = 1, 2, 3, 4, 6, 7, 8$ or 12, the value of (16) is zero and there is nothing to prove. We assume $n \neq 1, 2, 3, 4, 6, 7, 8$ or 12.

Let $\sigma : \mathcal{E}_n \to \mathcal{E}_n$ be an automorphism which is defined by $(x, y, t) \mapsto (x, y, \zeta_n^{-1} t)$, and let σ^* be the automorphism on $H^2_{\mathrm{tr}}(\mathcal{E}_n)$ induced by σ. We denote by $f(T)$ the minimal polynomial of σ^* over $\mathbb{Q}$. If we have

$$f(\zeta_d) = 0 \ \text{ for each } d \in S^1_n, \tag{17}$$

then d-th cyclotomic polynomial divides into $f(T)$ and hence we have

$$\dim H^2_{\mathrm{tr}}(\mathcal{E}_n) \ge \deg f(T) \ge \sum_{d \in S^1_n} \varphi(d) = \dim V(\mathcal{E}_n)$$

and (15) follows. Let us prove (17).

Lemma 4.1. *Let $n = 12l + k$ with $l \ge 0$, $1 \le k \le 12$. If n equals* 1, 2, 3, 4, 6, 7, 8 *or* 12, *then $H^0(\mathcal{E}_n, \Omega^2_{\mathcal{E}_n}) = 0$. Otherwise, $H^0(\mathcal{E}_n, \Omega^2_{\mathcal{E}_n})$ has a basis*

$$t^{2n-a(n)-3} dt \frac{dx}{y}, \quad \ldots, \quad t^{2n-a(n)-b(n)-3} dt \frac{dx}{y},$$

k	1	2	3	4	5	6
$a(n)$	$9l-1$	$9l$	$9l+1$	$9l+2$	$9l+2$	$9l+3$
$b(n)$	$l-1$	$l-1$	$l-1$	$l-1$	l	$l-1$
k	7	8	9	10	11	12
$a(n)$	$9l+4$	$9l+5$	$9l+5$	$9l+6$	$9l+7$	$9l+8$
$b(n)$	$l-1$	$l-1$	l	l	l	$l-1$

Table 3. Definitions of $a(n)$ and $b(n)$ for Lemma 4.1.

where $a(n), b(n)$ are defined in Table 3.

Proof. The proof is left as an exercise (see, for example, [Stiller 1987, Proposition 3.3] for details). □

For an integer i with $0 \le i \le b(n)$, since we have

$$\sigma^*\Big(t^{2n-a(n)-i-3}dt\frac{dx}{y}\Big) = \zeta_n^{a(n)+i+2}\Big(t^{2n-a(n)-i-3}dt\frac{dx}{y}\Big),$$

the automorphism σ^* on $H^2_{\mathrm{tr}}(\mathcal{E}_n)$ over $\mathbb{Q}$ has eigenvalues $\zeta_n^{a(n)+i+2}$, so we have $f(\zeta_n^{a(n)+i+2}) = 0 \ \ (0 \le i \le b(n))$. On the other hand, since we have

$$J(n) := \{a(n)+2, \dots, a(n)+b(n)+2\} = \{j \in \mathbb{N} \mid 9n < 12j < 10n\},$$

we obtain $\{(a(d)+2)n/d, \dots, (a(d)+b(d)+2)n/d\} \subset J(n)$ for each d which divides into n. Then $J(d) = \varnothing$ if and only if $d = 1, 2, 3, 4, 6, 7, 8$ or 12. In addition, $d \in \mathrm{Adm}_1$ if and only if each $j \in \{j \in \mathbb{N} \mid 9d \le 12j \le 10d\}$ is not relatively prime to d (see Remark 3.4). Therefore for each $d \in S^1_n$, there exists a natural number j which is relatively prime to d such that $jn/d \in J(n)$, and we have $\zeta_n^{jn/d} = \zeta_d^j$. This implies (17).

Acknowledgements

The author would like to thank Professor Masanori Asakura for helpful comments and suggestions. He also thanks Professor Matthias Schütt whose comments made enormous contribution to this paper.

References

[Cox and Zucker 1979] D. A. Cox and S. Zucker, “Intersection numbers of sections of elliptic surfaces”, *Invent. Math.* **53**:1 (1979), 1–44. MR 81i:14023 Zbl 0444.14004

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Kodaira 1963a] K. Kodaira, “On compact analytic surfaces, II”, *Ann. of Math.* (2) **77** (1963), 563–626. MR 32 #1730 Zbl 0118.15802

[Kodaira 1963b] K. Kodaira, "On compact analytic surfaces, III", *Ann. of Math.* (2) **78** (1963), 1–40. MR 32 #1730 Zbl 0171.19601

[Shioda 1972] T. Shioda, "On elliptic modular surfaces", *J. Math. Soc. Japan* **24** (1972), 20–59. MR 55 #2927 Zbl 0226.14013

[Shioda 1990] T. Shioda, "On the Mordell–Weil lattices", *Comment. Math. Univ. St. Paul.* **39**:2 (1990), 211–240. MR 91m:14056 Zbl 0725.14017

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, New York, 1994. MR 96b:11074 Zbl 0911.14015

[Silverman 2000] J. H. Silverman, "A bound for the Mordell–Weil rank of an elliptic surface after a cyclic base extension", *J. Algebraic Geom.* **9**:2 (2000), 301–308. MR 2001a:11107 Zbl 0951.14023

[Stiller 1987] P. F. Stiller, "The Picard numbers of elliptic surfaces with many symmetries", *Pacific J. Math.* **128**:1 (1987), 157–189. MR 88c:14054 Zbl 0591.14022

Received September 23, 2013.

MASAMICHI KURODA
DEPARTMENT OF MATHEMATICS
HOKKAIDO UNIVERSITY
SAPPORO 060-0810
JAPAN
m-kuroda@math.sci.hokudai.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.123

ON A PRIME ZETA FUNCTION OF A GRAPH

TAKEHIRO HASEGAWA AND SEIKEN SAITO

In the first half of this paper, we introduce a prime zeta function associated with the Ihara zeta function, and study several properties of this function. In the last half, using results of the first half, we present graph-theoretic analogues to Mertens' theorems.

1. Introduction

Throughout this paper, we use the notation of [Stark and Terras 1996; Terras 2011] for graph theory and the theory of (Ihara) zeta functions $Z_X(u)$ of graphs, and the notation of [Hardy and Wright 2008] and [Titchmarsh 1958; 1986] for the theory of functions and the Riemann zeta function $\zeta(s)$.

In the analytic theory of the Riemann zeta function, the following theorems are well-known:

- Mertens' first theorem [1874, Equality (5)] (also see [Hardy and Wright 2008, Theorem 425], [Jameson 2003, Theorem 2.6.3], and [Titchmarsh 1986, Equality (3.14.3)]): as $x \to \infty$,

$$\sum_{p \le x} \frac{\log p}{p} = \log x + O(1).$$

- Mertens' second theorem [1874, Equality (13)] (also see [Hardy and Wright 2008, Theorem 427], [Jameson 2003, Theorem 2.6.4/Exercise 4, p. 191], and [Titchmarsh 1986, Equality (3.14.5)]): as $x \to \infty$,

$$\sum_{p \le x} \frac{1}{p} = \log(\log x) + B_1 + O\Big(\frac{1}{\log^k x}\Big)$$

for each $k \ge 1$, where $B_1 = 0.26149\ldots$ is the Mertens constant.

- Mertens' third theorem [1874, Equality (15)] (also see [Hardy and Wright 2008, Theorem 429], [Jameson 2003, Exercise 1, p. 96], and [Titchmarsh 1986,

MSC2010: primary 11N45; secondary 05C30, 05C38, 05C50.
Keywords: Ihara zeta functions, primes in graphs, Mertens' theorem.

Equality (3.15.2)]): as $x \to \infty$,

$$\prod_{p \le x} \Big(1 - \frac{1}{p}\Big) \sim \frac{e^{-\gamma}}{\log x},$$

where $\gamma = 0.57721\ldots$ is the Euler–Mascheroni constant.

- Prime number theorem (proved by de la Vallée Poussin and Hadamard in 1896; see, e.g., [Hardy and Wright 2008, Theorem 6], [Jameson 2003, Theorem 3.4.3], and [Titchmarsh 1986, Equality (3.7.1)]): as $x \to \infty$,

$$\pi(x) \sim \frac{x}{\log x},$$

where $\pi(x)$ denotes the number of rational prime numbers p less than x, that is,

$$\pi(x) := \big|\{p : p \text{ is a rational prime number with } p \le x\}\big|.$$

All proofs of the above formulae are related to the Riemann zeta function

$$\zeta(s) = \prod_{p \in \mathcal{P}} \Big(1 - \frac{1}{p^s}\Big)^{-1},$$

where $\mathcal{P}$ denotes the set of all rational prime numbers, that is,

$$\mathcal{P} := \{p \in \mathbb{Z} : p \text{ is a rational prime number}\},$$

and to the prime zeta function, defined first by Glaisher [1891],

$$P(s) = \sum_{p \in \mathcal{P}} \frac{1}{p^s}.$$

In graph theory, there exists an analogue of the Riemann zeta function, the so-called (Ihara) zeta function $Z_X(u)$ of a graph X (see [Ihara 1966]). Therefore, studying graph-theoretic analogues of these theorems is very interesting. Indeed, Terras and coworkers gave an analogue of the prime number theorem (see Theorem 2.10 in [Horton et al. 2006], and also Theorem 10.1 in [Terras 2011]):

If Δ_X divides n, then, as $n \to \infty$,

$$\pi_X(n) \sim \frac{\Delta_X}{n \cdot R_X^n},$$

and otherwise $\pi_X(n) \sim 0$. (For the definitions of $\pi_X(n)$ and R_X, see this section, and for that of Δ_X, see Section 3.) This is called the graph-theoretic prime number theorem.

In this paper, we define a prime zeta function of a graph, and investigate several properties of this function. In particular, we show that this has a natural boundary. Moreover, by using this function, we present graph-theoretic analogues of Mertens' theorems.

We shall note a relation between previous works and our works. A zeta function of a graph can be specialized from a dynamical zeta function for a flow (see Chapter 4 in [Terras 2011]), and dynamical-systemic analogues to the above formulae are already known (see, e.g., [Sharp 1991] for Mertens' theorems, and [Parry 1983; Parry and Pollicott 1983] for a prime number theorem). In that sense, our statements for Mertens' theorems are not new (see Remark 17). However, since our proofs are graph-theoretic and elementary, they are completely different from previous proofs.

In this section, we first recall the notation for graph theory and zeta functions of graphs, define a prime zeta function of a graph, and finally state the main theorem.

Now we recall the notation of graph theory. Throughout this paper, we always assume that X is a finite, connected, non-cycle and undirected graph without degree-one vertices. Let X be a graph with vertex set V, with $\nu := |V|$, and edge set E, with $\epsilon := |E|$. Simply, such a graph X is denoted by $X := (V, E)$. Note that ϵ is the number of edges of X.

An oriented edge (or an arc) a from a vertex u to a vertex v is denoted by $a = (u, v)$, and the inverse of a is denoted by $a^{-1} = (v, u)$. The origin and terminus of a are denoted by $o(a)$ and $t(a)$, respectively. We can now orient the edges of X, and label the edges as follows:

$$\vec{E} = \{e_1, e_2, \ldots, e_\epsilon, e_{\epsilon+1} = e_1^{-1}, e_{\epsilon+2} = e_2^{-1}, \ldots, e_{2\epsilon} = e_\epsilon^{-1}\}.$$

A path $C = a_1 \cdots a_s$, where the a_i are oriented edges, is said to have a backtrack (resp. tail) if $a_{j+1} = a_j^{-1}$ for some j (resp. $a_s = a_1^{-1}$), and a path C is called a cycle (or a closed path) if $o(a_1) = t(a_s)$. The length $\ell(C)$ of a path $C = a_1 \cdots a_s$ is defined by $\ell(C) = s$.

A cycle C is called prime (or primitive) if it satisfies the following:

- C does not have backtracks or a tail;
- no cycle D exists such that $C = D^f$ for some $f > 1$.

The equivalence class $[C]$ of a cycle $C = a_1 \cdots a_s$ is defined as the set of cycles

$$[C] := \{a_1 a_2 \cdots a_{s-1} a_s, a_2 \cdots a_{s-1} a_s a_1, \ldots, a_s a_1 a_2 \cdots a_{s-1}\},$$

and an equivalence class $[P]$ of a prime cycle P is called a prime in the graph X. Throughout this paper, we denote a prime by the symbol $[P]$. Two cycles C_1 and C_2 are called equivalent if $C_2 \in [C_1]$. Note that if $[C_1] = [C_2]$, then $\ell(C_1) = \ell(C_2)$, and thus $u^{\ell(C_1)} = u^{\ell(C_2)}$.

Next, we recall the zeta function of a graph $X = (V = \{v_1, \dots, v_\nu\}, E)$, and we define a prime zeta function associated with it. Let u be a complex variable, and let $f_X(u)$ denote

$$f_X(u) := \det(I_\nu - Au + Qu^2),$$

where I_ν is the $\nu \times \nu$ identity matrix, A is the adjacency matrix of X (see Definition 2.1 in [Terras 2011]), and

$$Q = \operatorname{diag}(\deg(v_1) - 1, \dots, \deg(v_\nu) - 1).$$

Let $\pi_X(n)$ denote

$$\pi(n) = \pi_X(n) := \bigl|\{[P] : [P] \text{ is a prime in } X \text{ with } \ell(P) = n\}\bigr|.$$

Throughout this paper, we fix an arbitrary real number $t > 1$ (that is, $\log t > 0$), and we set $u = t^{-s}$. The (Ihara) zeta function of X (see Definition 2.2 and Theorem 2.5 in [Terras 2011]) and the prime zeta function of X are defined as follows:

$$Z_X(u) := \prod_{[P]} (1 - u^{\ell(P)})^{-1} = \frac{1}{(1-u^2)^{\epsilon-\nu} f_X(u)}, \quad \mathscr{Z}_X(s) := Z_X(t^{-s}),$$

$$P_X(u) := \sum_{[P]} u^{\ell(P)} = \sum_{n=1}^{\infty} \pi_X(n) u^n, \qquad \mathscr{P}_X(s) := P_X(t^{-s}),$$

with $|u|$ sufficiently small, where $[P]$ runs through all primes in X. In this paper, we do not distinguish between the two functions $Z_X(u)$ and $\mathscr{Z}_X(s)$, or between $P_X(u)$ and $\mathscr{P}_X(s)$. The right-hand side of the first equality is called the Ihara–Bass formula (see [Bass 1992]). Note that, owing to our assumption for X, the zeta function $Z_X(u)$ is expressible like that.

Note that, for two finite connected graphs X_1 and X_2 without degree-one vertices, $P_{X_1}(u) = P_{X_2}(u)$ if and only if $Z_{X_1}(u) = Z_{X_2}(u)$ (see Proposition 7 in [Storm 2010]).

Let

$$T := \bigcup_{n=1}^{\infty} T_n \quad \text{and} \quad T_n := \{u \in \mathbb{C} : f_X(u^n) = 0\}$$

be the zeroes of the $f_X(u^n)$. Note that the elements of T_n are poles of $Z_X(u^n)$. The radius of convergence of $Z_X(u)$ is denoted by R_X. Note that $0 < R_X < 1$ since X is a non-cycle graph (see, e.g., [Terras 2011, p. 197]). It follows from the graph-theoretic prime number theorem (see Theorem 10.1 in [Terras 2011]) that the radius of convergence of the other function $P_X(u)$ is also equal to R_X. Note that the point $u = R_X$ is a singularity of $P_X(u)$, and that

$$P_X(u) \sim -\log(R_X - u)$$

as $u \uparrow R_X$, which is similar to

$$P(s) \sim -\log(s-1)$$

as $s \downarrow 1$ (see, e.g., [Fröberg 1968]), where $P(s) = \sum_p 1/p^s$ denotes the prime zeta function associated with the Riemann zeta function.

Euclid proved that the number of primes p is infinite. Euler showed that the prime zeta function $\sum_p 1/p$ diverges, and as an application he proved the infinitude of primes. In graph theory, it is also well known that the number of primes $[P]$ in X is infinite. We can give another proof "à la Euler" for this fact since $u = R_X$ is a singularity of $P_X(u)$.

Our main theorem is:

Main Theorem. *Suppose that $X = (V, E)$ is a finite, connected and non-cycle graph without degree-one vertices.*

(1) *Let $\mu(n)$ denote the Möbius function. If $|u| < R_X$, then*

$$P_X(u) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \log Z_X(u^n).$$

Furthermore, the right-hand side is absolutely convergent for u satisfying $|u| < 1$ and $u \notin T$, and so $P_X(u)$ has an analytic extension to the region $\{u \in \mathbb{C} : |u| < 1\} \setminus T$.

(2) *The imaginary axis* $\mathrm{Re}(s) = 0$ *is a natural boundary for the function $\mathcal{P}_X(s)$, that is, every point on this line can be realized as a limit point of singularities of $\mathcal{P}_X(s)$.*

(3) (*Graph-theoretic Mertens' first theorem*) *As $N \to \infty$,*

$$\sum_{n \le N} n \cdot \pi_X(n) R_X^n = N + O(1).$$

(4) (*Graph-theoretic Mertens' second theorem*) *There exists a constant B_X such that, as $N \to \infty$,*

$$\sum_{n \le N} \pi_X(n) R_X^n = \log N + B_X + O\Big(\frac{1}{N}\Big).$$

(5) (*Graph-theoretic Mertens' third theorem*) *Let $\gamma = 0.57721\ldots$ denote the Euler–Mascheroni constant. As $N \to \infty$,*

$$\prod_{\ell(P) \le N} (1 - R_X^{\ell(P)}) \sim \frac{e^{-\gamma}}{C_X} \cdot \frac{1}{N},$$

where

$$C_X = -\frac{1}{(1-R_X^2)^{\epsilon-\nu} R_X f_X'(R_X)}$$

(*for the definition*, *in detail*, *see Section 3 in this paper*).

The contents of this paper are as follows. In the next section, we prove the first two claims in the main theorem, that is, several properties of $P_X(u)$. In Section 3, we prove the remaining claims in the main theorem, namely, the graph-theoretic Mertens theorems.

2. Prime zeta function of a graph

In this section, we give a proof of parts (1) and (2) of the Main Theorem introduced in Section 1.

The following facts about $Z_X(u)$, etc., are known, and are often used in this paper.

Facts 1. (1) (Basic facts) *For an arbitrary real number $t > 1$, set $u = t^{-s}$. Then the function $\mathfrak{Z}_X(s)$ is absolutely convergent and holomorphic for all s satisfying* $\mathrm{Re}(s) > -\log R_X/\log t$ (≥ 0).

Since the function $Z_X(u)$ is the reciprocal of a polynomial by the Ihara–Bass formula, the function $Z_X(u)$ is meromorphic for all $u \in \mathbb{C}$, and therefore $\mathfrak{Z}_X(s)$ is also meromorphic for all $s \in \mathbb{C}$.

(2) [Kotani and Sunada 2000, Theorem 1.3(1)] *Let $q+1$ and $p+1$ be the maximum and minimum degrees of a graph X, respectively. Then $1/q \leq R_X \leq 1/p$, the point $u = R_X$ is a simple pole of $Z_X(u)$, and every pole of $Z_X(u)$ satisfies $R_X \leq |u| \leq 1$.*

(3) [Terras 2011, p. 197] *Suppose that X is a finite connected graph without degree-one vertices. Then $R_X = 1$ if and only if X is a cycle graph. This follows from the equation $p = q = 1$.*

(4) [Kotani and Sunada 2000, p. 8] *The leading coefficient of the polynomial f_X is given by*

$$c = \prod_{v \in V} (\deg(v) - 1),$$

and therefore that of the polynomial $1/Z_X$ is equal to $c_{2\epsilon} = (-1)^{\epsilon-\nu} c$.

In this section, the following lemma is important.

Key Lemma 2. *Let*

$$\phi(u) = 1 + \sum_{i=1}^{d} c_i u^i \in \mathbb{Z}[u]$$

be a polynomial function of degree $d \geq 0$, and let

$$T = \{u \in \mathbb{C} : \textit{there exists } n \geq 1 \textit{ such that } \phi(u^n) = 0\}$$

denote the zeroes of the $\phi(u^n)$. Suppose that r is an arbitrary real number, and assume that $\Phi(u)$ is a series defined by

$$\Phi(u) = \sum_{n=1}^{\infty} \frac{1}{n^r} \log \phi(u^n).$$

Then $\Phi(u)$ is absolutely convergent for u satisfying $|u| < 1$ and $u \notin T$.

Proof. First, we suppose that $d = 0$. Then the $\phi(u^n) = 1$ are constant, and therefore $\Phi(u) = 0$ is also constant. Hence, the claim is trivial. From now on, we assume that $d \geq 1$. Set $c := \max\{|c_i| : 1 \leq i \leq d\}$, choose a number C_0 with $C_0 \geq cd + 1$ (≥ 2), and fix it.

Let r_n ($n \geq 3$) be a number defined by

$$r_n := \left(\frac{1 - \exp(-1/n^{2-r})}{C_0} \right)^{1/n}.$$

Note that $r_n < (1/C_0)^{1/n}$, the sequence $\{r_n\}_{n\geq 3}$ is increasing, and $\lim_{n\to\infty} r_n = 1$.

Take u satisfying $|u| < 1$ and $u \notin T$, and fix it. Then there exists a number N such that $|u| \leq r_N$, and thus $|u| < r_n$ for all $n \geq N + 1$. Now we fix such numbers N and n.

Since $|u| < (1/C_0)^{1/n}$ and $|u^n| \leq |u| < 1$, we obtain, by the triangle inequality,

$$0 < 1 - C_0|u^n| \leq |\phi(u^n)|, \quad \text{and so} \quad -\log|\phi(u^n)| \leq -\log(1 - C_0|u^n|). \tag{1}$$

On the other hand, since $|u| < r_n$, then $C_0|u^n| < 1 - \exp(-1/n^{2-r})$, so we obtain the inequality $-\log(1 - C_0|u^n|) < 1/n^{2-r}$. Combining this result with (1), we obtain

$$\mathrm{Re}(-\log \phi(u^n)) = -\log|\phi(u^n)| < \frac{1}{n^{2-r}}. \tag{2}$$

The first inequality in (1) also shows that the function $\log \phi(u^n)$ is holomorphic in the closed disk $|u| \leq r_{N+1}$. By applying the Borel–Carathéodory theorem (see, e.g., [Titchmarsh 1958, §5.5]) to the function $\log \phi(u^n)$ and the two circles $|u| = r_{N+1}$, $|u| = r_N$, we obtain

$$|\log \phi(u^n)| \leq \max_{|u|=r_N} |\log \phi(u^n)| \leq K \max_{|u|=r_{N+1}} \mathrm{Re}(-\log \phi(u^n)) \leq K \frac{1}{n^{2-r}},$$

where $K := 2r_N/(r_{N+1} - r_N)$. Therefore, it follows that

$$\sum_{n=N+1}^{\infty} \frac{1}{n^r} |\log \phi(u^n)| \leq K \sum_{n=N+1}^{\infty} \frac{1}{n^2} < K \cdot \zeta(2) < \infty.$$

Hence, for u satisfying $|u| < 1$ and $u \notin T$, the series $\Phi(u)$ converges absolutely. $\square$

Using this lemma, we can prove the following proposition.

Proposition 3. *Let $\mu(n)$ denote the Möbius function. If $|u| < R_X$, then*

$$P_X(u) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \log Z_X(u^n). \tag{3}$$

Moreover, the right-hand side of (3) *is absolutely convergent for u satisfying $|u| < 1$ and $u \notin T$, and therefore $P_X(u)$ extends analytically to the region $\{u \in \mathbb{C} : |u| < 1\} \setminus T$.*

Equivalently, if $\mathrm{Re}(s) > -\log R_X / \log t$, *then*

$$\mathcal{P}_X(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \log \mathfrak{Z}_X(ns). \tag{4}$$

The right-hand side of (4) *is absolutely convergent for s satisfying* $\mathrm{Re}(s) > 0$ *and $t^{-s} \notin T$, and so* (4) *gives the analytic continuation of $\mathcal{P}_X(s)$ to the region.*

Proof. Note that $R_X \leq 1$ (from Fact 1(2)) and $\exp(z) = \prod_{n=1}^{\infty}(1 - z^n)^{-\mu(n)/n}$ for $|z| < 1$. Suppose that $|u| < R_X$. Since $|u^{\ell(P)}| \leq |u| < 1$, we obtain the equality

$$\exp(P_X(u)) = \prod_{[P]} \exp(u^{\ell(P)}) = \prod_{[P]} \prod_{n=1}^{\infty} (1 - u^{n\ell(P)})^{-\mu(n)/n} = \prod_{n=1}^{\infty} Z_X(u^n)^{\mu(n)/n},$$

and therefore (3) holds for u satisfying $|u| < R_X$.

Set

$$1/Z_X(u) = (1 - u^2)^{\epsilon - \nu} f_X(u) = 1 + c_1 u + \cdots + c_{2\epsilon} u^{2\epsilon} \in \mathbb{Z}[x],$$

$c = \max\{|c_i| : 1 \leq i \leq 2\epsilon\}$ and $C_0 = 2\epsilon c \geq 2$. By applying Key Lemma 2 to $\phi(u) = 1/Z_X(u)$ and $r = 1$, it follows that, for u satisfying $|u| < 1$ and $u \notin T$, the series $\sum_{n=1}^{\infty} \log Z_X(u^n)/n$ is absolutely convergent, and so the right-hand side of (3) is absolutely convergent. $\square$

Moreover, for a Ramanujan graph, we can prove the following.

Corollary 4. *Suppose that X is a finite connected Ramanujan graph with degree $q + 1$, that is, $Z_X(u)$ satisfies the Riemann hypothesis* (see Theorem 7.4 in [Terras 2011]). *Then the function $P_X(u)$ is absolutely convergent for u satisfying $|u| < 1$ and $|u| \neq (1/q)^{1/n}$ for all n.*

Equivalently, the function $\mathcal{P}_X(s)$ is absolutely convergent for s such that $\mathrm{Re}(s) > 0$ *and* $\mathrm{Re}(s) \neq \log q / \log t^n$ *for all n.*

Proof. Since X is a Ramanujan graph, by Theorem 1.3 in [Kotani and Sunada 2000], every real (resp. nonreal) zero of $f_X(u)$ satisfies $|u| = 1$ or $1/q$ (resp. $|u| = 1/\sqrt{q}$). Thus, every point $|u| \neq (1/q)^{1/n}$ is not zero of $f_X(u^n)$. Hence, the proof of the assertion follows from Proposition 3. $\square$

We can completely interchange the roles of the functions $P_X(u)$ and $\log Z_X(u)$.

Corollary 5. *If* $|u| < 1$ *and* $u \notin T$, *then*

$$\log Z_X(u) = \sum_{n=1}^{\infty} \frac{1}{n} P_X(u^n). \tag{5}$$

Equivalently, *if* $\mathrm{Re}(s) > 0$ *and* $t^{-s} \notin T$, *then*

$$\log \mathscr{Z}_X(s) = \sum_{n=1}^{\infty} \frac{1}{n} \mathscr{P}_X(ns). \tag{6}$$

Proof. By applying the Möbius inversion formula (see, e.g., Theorem 270 in [Hardy and Wright 2008], or Theorem 2.2.8 in [Jameson 2003]) to the equality (3) for $|u| < 1$, we obtain the equality (5). $\square$

Remark 6. The equalities (4) and (6) indicate that $\mathscr{P}_X(s)$ is a graph-theoretic analogue to the prime zeta function $P(s)$ for the Riemann zeta function $\zeta(s)$. The relations between $P(s)$ and $\zeta(s)$ are given as follows (see [Glaisher 1891], and also [Fröberg 1968] and Equality (1.6.1) in [Titchmarsh 1986]):

For $\mathrm{Re}(s) > 1$,

$$P(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \log \zeta(ns) \quad \text{and} \quad \log \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n} P(ns).$$

We can orient the edges of X, and label the edges as follows:

$$\vec{E} = \{a_1, a_2, \ldots, a_\epsilon, a_{\epsilon+1} = a_1^{-1}, a_{\epsilon+2} = a_2^{-1}, \ldots, a_{2\epsilon} = a_\epsilon^{-1}\}.$$

Let $W = W_X := (w_{ij})$ denote the edge adjacency matrix of a graph X, that is, a $2\epsilon \times 2\epsilon$ matrix defined by

$$w_{ij} := \begin{cases} 1 & \text{if } t(a_i) = o(a_j) \text{ and } a_j \neq a_i^{-1} \text{ for } a_i, a_j \in \vec{E}, \\ 0 & \text{otherwise} \end{cases}$$

(see p. 28 in [Terras 2011]). Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of W, and let $e_1, \ldots, e_k$ be their multiplicities. Note that $\sum_{i=1}^{k} e_i = 2\epsilon$. Let $e := \sum_{i=1, \lambda_i \neq \pm 1}^{k} e_i$. By the determinant formula given by Hashimoto [1989] and Bass [1992], the polynomial $1/Z_X(u)$ can be written as

$$1/Z_X(u) = \det(I_{2\epsilon} - Wu) = \prod_{i=1}^{k} (1 - \lambda_i u)^{e_i}.$$

Note that $f_X(1) = 0$. We now define a polynomial $g_X(u)$ by

$$g_X(u) := f_X(u)/(1-u).$$

Note that since $f_X'(1) = 2(\epsilon - \nu)\kappa$ by [Northshield 1998, Theorem],

$$g_X(1) = -f_X'(1) = -2(\epsilon - \nu)\kappa,$$

where κ is the complexity of X, that is, the number of spanning trees in X. Since X is a non-cycle graph, that is, $\epsilon \neq \nu$, the polynomial $g_X(u)$ can be also written as

$$g_X(u) = \frac{1/Z_X(u)}{(1-u^2)^{\epsilon-\nu}(1-u)} = (1+u)^{2\nu-1-e} \prod_{\substack{i=1 \\ \lambda_i \neq \pm 1}}^{k} (1-\lambda_i u)^{e_i}. \tag{7}$$

We can show that the function $\mathcal{P}_X(s)$ has a natural boundary.

Proposition 7. *Let $X = (V, E)$ be a finite, connected and non-cycle graph without degree-one vertices.*

(1) *There exists an eigenvalue λ of W such that $|\lambda| > 1$.*

(2) *The imaginary axis* $\mathrm{Re}(s) = 0$ *is a natural boundary for the function $\mathcal{P}_X(s)$, that is, every point on this line can be realized as a limit point of singularities of $\mathcal{P}_X(s)$.*

Proof. (1) The leading coefficient $c_{2\epsilon}$ of the polynomial $1/Z_X(u)$ is given by

$$(-1)^{\epsilon-\nu} \prod_{v \in V} (\deg(v) - 1) = c_{2\epsilon} = \prod_{i=1}^{k} \lambda_i^{e_i}$$

(from Fact 1(4)). By our assumption for X, the graph X is not a 2-regular graph. Thus $|c_{2\epsilon}| > 1$ and so there exists an eigenvalue λ_i with $|\lambda_i| \neq 1$. Note that every pole $1/\lambda_i$ of $Z_X(u)$ satisfies $|1/\lambda_i| \leq 1$ by Fact 1(2). So there exists an eigenvalue λ_i with $|\lambda_i| > 1$.

(2) Note that $\exp(z) = \prod_{n=1}^{\infty}(1-z^n)^{-\mu(n)/n}$ for $|z| < 1$. If $|u| < 1$ and $u \notin T$, then

$$\begin{aligned}
\exp(P_X(u)) &= \prod_{n=1}^{\infty} Z_X(u^n)^{\mu(n)/n} \\
&= \left(\prod_{n=1}^{\infty}(1-u^{2n})^{-\mu(n)/n}\right)^{\epsilon-\nu} \left(\prod_{n=1}^{\infty}(1-u^n)^{-\mu(n)/n}\right) \prod_{n=1}^{\infty} g_X(u^n)^{-\mu(n)/n} \\
&= \exp((\epsilon-\nu)u^2 + u) \prod_{n=1}^{\infty} g_X(u^n)^{-\mu(n)/n},
\end{aligned}$$

and therefore the equality

$$P_X(u) = (\epsilon-\nu)u^2 + u - \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \log g_X(u^n)$$

holds.

Note that $u = t^{-s}$. By using the equalities (7) and 2, the function $\mathcal{P}_X(s)$ can be written as

$$\mathcal{P}_X(s) = (\epsilon - \nu)t^{-2s} + t^{-s} \\ - \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \Bigg((2\nu - 1 - e)\log(1 + t^{-ns}) + \sum_{\substack{i=1 \\ \lambda_i \neq \pm 1}}^{k} e_i \log(1 - \lambda_i t^{-ns}) \Bigg)$$

for all s satisfying $\mathrm{Re}(s) > 0$. By part (1), there exists λ such that $|\lambda| > 1$ among the eigenvalues $\lambda_1, \ldots, \lambda_k$ of W. Note that $1 - \lambda t^{-ns} = 0$ if and only if $s = r(\lambda, n, m)$, where

$$r(\lambda, n, m) := \frac{\log|\lambda|}{n \log t} + i\frac{\mathrm{Arg}(\lambda) + 2\pi m}{n \log t},$$

and $\mathrm{Arg}(\lambda)$ is the argument of λ with $-\pi \leq \mathrm{Arg}(\lambda) < \pi$. Note that

$$\varepsilon_n := \frac{\log|\lambda|}{n \log t} \to 0$$

as $n \to \infty$. We now fix an arbitrary point $\alpha = ia$ on the imaginary axis $\mathrm{Re}(s) = 0$. Then, we can arrange a sequence of integers $\{m_n\}$ for each integer n so that

$$\frac{\mathrm{Arg}(\lambda) + 2\pi m_n}{n \log t} \to a$$

as $n \to \infty$. Hence, each point α on the boundary is a limit point of singularities of $\mathcal{P}_X(s)$. Since $\varepsilon_n > 0$ for all n, we cannot continue $\mathcal{P}_X(s)$ beyond the boundary at $\mathrm{Re}(s) = 0$. □

Remark 8. Proposition 7(2) is an analogue of the fact that the imaginary axis $\mathrm{Re}(s) = 0$ is a natural boundary for the prime zeta function $P(s)$ of the Riemann zeta function $\zeta(s)$ (see [Landau and Walfisz 1920]).

3. Graph-theoretic Mertens' theorem

In this section, we prove parts (3)–(5) of the Main Theorem introduced in Section 1.

Throughout this section, we always assume that $X = (V, E)$ is a finite, connected, non-cycle graph without degree-one vertices. Note in particular that $\nu \neq \epsilon$ and $0 < R_X < 1$.

First, we define the constants H_X, C_X and γ_X, and study their properties, which play important roles in this section. Let u be a complex variable. We define a function by

$$H_X(u) := \log Z_X(u) - P_X(u) = \sum_{n \geq 2} \frac{1}{n} P_X(u^n) = \sum_{[P]} \sum_{n \geq 2} \frac{1}{n} u^{n\ell(P)}.$$

Note that the point $u = R_X$ is a common pole of $Z_X(u)$ and $P_X(u)$ by Fact 1(2), and that the series $H_X(u)$ is absolutely convergent for u satisfying $|u| < 1$ and $u \notin T$, from Corollary 5.

Since $u = R_X$ is a simple pole of $Z_X(u)$, we can define constants c_X and C_X by

$$c_X := -\operatorname*{Res}_{u=R_X} Z_X(u) = \lim_{u \uparrow R_X} (R_X - u) Z_X(u) = \frac{-1}{(1 - R_X^2)^{\epsilon-\nu} f_X'(R_X)}$$

and $C_X := c_X / R_X$.

Lemma 9. (1) *The value $H_X := H_X(R_X)$ is finite.*

(2) *The constants c_X and C_X are positive.*

Proof. (1) Since $R_X^n < R_X < 1$ $(n \geq 2)$, the function $P_X(u)$ is holomorphic at $u = R_X^n$, and therefore $P_X(u^n)$ is holomorphic at $u = R_X$. We have

$$\begin{aligned} H_X(R_X) &= \sum_{[P]} \sum_{n \geq 2} \frac{1}{n} R_X^{n\ell(P)} \leq \sum_{[P]} \sum_{n \geq 2} R_X^{n\ell(P)} \\ &= \sum_{[P]} \frac{R_X^{2\ell(P)}}{1 - R_X^{\ell(P)}} \leq \frac{1}{1 - R_X} \sum_{[P]} R_X^{2\ell(P)} = \frac{P_X(R_X^2)}{1 - R_X} < +\infty, \end{aligned}$$

and the assertion follows.

(2) Note that the leading coefficient of the polynomial f_X is given by

$$c = \prod_{v \in V} (\deg(v) - 1) > 0$$

by Fact 1(4). Then f_X factors as the product of irreducible polynomials such that

$$f_X(u) = c \prod_{i=1}^{m_1} (u - \alpha_i) \cdot \prod_{j=1}^{m_2} f_j(u),$$

where the f_j are monic of $\deg f_j = 2$, and $\deg f_X = 2\nu = m_1 + 2m_2$. Note that m_1 is even. Since $u = R_X$ is a simple pole of $Z_X(u)$, it is a simple zero of f_X. We may assume that $\alpha_1 = R_X$. Since $\alpha_i > R_X$ $(2 \leq i \leq m_1)$ and the discriminants of the f_j are negative, the sign of

$$f_X'(R_X) = c \prod_{i=2}^{m_1} (R_X - \alpha_i) \prod_{j=1}^{m_2} f_j(R_X)$$

is equal to $(-1)^{m_1 - 1} = -1$, i.e., $f_X'(R_X) < 0$, so $c_X > 0$ and $C_X = c_X / R_X > 0$. □

Since the function $Z_X(u) - c_X/(R_X - u)$ is holomorphic at $u = R_X$, we can define a constant γ_X by

$$\gamma_X := \lim_{u \uparrow R_X} \left(Z_X(u) - \frac{c_X}{R_X - u} \right),$$

which is an analogue of the Euler–Mascheroni constant $\gamma = \lim_{s \downarrow 1}(\zeta(s) - 1/(s-1))$ for $\zeta(s)$.

In a neighborhood of $u = R_X$, the function $Z_X(u)$ can be expanded as

$$Z_X(u) = \frac{c_X}{R_X - u} + \gamma_X + O(R_X - u),$$

and so

$$\log Z_X(u) = \log \frac{c_X}{R_X - u} + O(R_X - u). \tag{8}$$

Similarly, in a neighborhood of $u = R_X$, the function $P_X(u)$ can be expanded as

$$P_X(u) = \log \frac{c_X}{R_X - u} - H_X(u) + O(R_X - u) = \log \frac{c_X}{R_X - u} - H_X(R_X) + O(R_X - u).$$

In this section, the following facts are used.

Facts 10. (1) (See, for example, Theorem 18.1 in [Korevaar 2002].) *Let x be a complex variable and let $F(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with $a_n \geq 0$ that converges for $|x| < 1$. Suppose that*

$$F(x) - \frac{C}{1-x} = O(1)$$

as $x \to 1$. Then the partial sum $A(N) = \sum_{n \leq N} a_n$ satisfies

$$A(N) = C \cdot N + O(\log N)$$

as $N \to \infty$.

(2) (See, for example, Exercises 9-6 in [Apostol 1974], and Theorem 1.3.6 in [Jameson 2003], the Abel partial summation formula). *Let $\{a_n\}$ be real numbers, and let $f(t)$ be a (real- or complex-valued) function with a continuous derivative in the interval $[1, N]$. Then*

$$\sum_{n \leq N} a_n f(n) = A(N) f(N) - \int_1^N A(t) f'(t)\, dt.$$

By using Fact 10, we can prove the following proposition.

Proposition 11. *Suppose that X is a finite, connected and non-cycle graph without degree-one vertices. In a neighborhood of $u = R_X$, expand $Z_X(u)$ into the* *power series*

$$Z_X(u) = \sum_{n=0}^{\infty} a'_n u^n.$$

Then, *as* $N \to \infty$,

$$\sum_{n \le N} a'_n R_X^n = C_X \cdot N + O(\log N).$$

Proof. First, for simplicity of arguments, we normalize the function $Z_X(u)$:

$$F(x) = Z_X(R_X x) = \sum_{n=0}^{\infty} a'_n R_X^n x^n = \sum_{n=0}^{\infty} a_n x^n,$$

where $a_n = a'_n R_X^n$. Note that the normalized function $F(x)$ converges for $|x| < 1$. Since all coefficients a'_n are nonnegative (by page 13 in [Terras 2011]), all coefficients a_n are also nonnegative. Since X is a non-cycle graph, the point $x = 1$ is a simple pole of $F(x)$. Hence, we obtain

$$F(x) - \frac{C_X}{1-x} = O(1)$$

as $x \to 1$. By applying Fact 10(1) to this equality, as $N \to \infty$,

$$\sum_{n \le N} a_n = C_X \cdot N + O(\log N), \quad \text{and so} \quad \sum_{n \le N} a'_n R_X^n = C_X \cdot N + O(\log N)$$

holds, and the assertion follows. □

Now, we compute the following example.

Example 12 [Terras 2011, Example 2.8, p. 18]. Consider the graph $X = K_4 - \{\text{one edge}\}$. Then

$$f_X(u) = (1-u)(1+u^2)(1+u+2u^2)(1-u^2-2u^3) \quad \text{and} \quad Z_X(u)^{-1} = (1-u^2) f_X(u).$$

Since the radius of convergence R_X of $Z_X(u)$ is the smallest positive real zero of $f_X(u)$,

$$R_X = \tfrac{1}{6}(\alpha - 1 + \alpha^{-1}) = 0.6572981\ldots, \quad \alpha = (53 + 6\sqrt{78})^{1/3}.$$

Then C_X is computed as $C_X = 0.5540954\ldots$. For example, if $N = 50000$, then

$$\frac{1}{N} \sum_{n \le N} a'_n R_X^n = 0.5540867\ldots \approx C_X.$$

Let $X = (V, E)$ be a graph, and set $|V| = \nu$ and $|E| = \epsilon$. Let $W = W_X$ be the edge adjacency matrix of X (see page 28 in [Terras 2011], or Section 2 in this paper), and let $\mathrm{Spec}(W)$ denote the spectrum of W, that is, the list of its eigenvalues together with their multiplicities. Note that $|\mathrm{Spec}(W)| = 2\epsilon$. The polynomial $1/Z_X(u)$ has an expression different from that in Section 2. In fact, this can be written as

$$1/Z_X(u) = \det(I_{2\epsilon} - Wu) = \prod_{\lambda \in \mathrm{Spec}(W)} (1 - \lambda u) \quad \left(= \prod_{i=1}^{k} (1 - \lambda_i u)^{e_i} \right).$$

Since the points $u = 1/\lambda$ are the poles of $Z_X(u)$, we obtain $1 \le |\lambda| \le 1/R_X$ by Fact 1(2).

The following lemma is used in the proof of Theorem 14 in this section.

Key Lemma 13. *Suppose that X is a finite, connected and non-cycle graph without degree-one vertices.*

(1) *As $N \to \infty$, we have*

$$\sum_{n=1}^{N} \sum_{\lambda \in \mathrm{Spec}(W)} (\lambda R_X)^n = N + O(1).$$

(2) *Let $0 < \alpha < \frac{1}{2}$ be a fixed real number. Then there exists a natural number N_0 such that, for any $n \ge N_0$,*

$$\left| n \cdot \pi(n) - \sum_{\lambda \in \mathrm{Spec}(W)} \lambda^n \right| < 2\epsilon \left(\frac{1}{R_X} \right)^{(1-\alpha)n}.$$

Proof. (1) Let Δ_X denote

$$\Delta = \Delta_X := \gcd\{\ell(P) : [P] \text{ is a prime in } X\}$$

(see Definition 2.12 in [Terras 2011]). It follows from Theorem 1.4 in [Kotani and Sunada 2000] that the poles of $Z_X(u)$ on the circle $|u| = R_X$ have the form $u = R_X e^{2\pi i a/\Delta}$ $(1 \le a \le \Delta)$. It is well known that

$$\sum_{a=1}^{\Delta} e^{2\pi i a n/\Delta} = \begin{cases} \Delta & \text{if } \Delta \mid n, \\ 0 & \text{otherwise} \end{cases}$$

(see, e.g., Exercise 10.1 in [Terras 2011]). Then we obtain

$$\left| N - \sum_{|\lambda| = 1/R_X} \sum_{n=1}^{N} (\lambda R_X)^n \right| = \left| N - \sum_{n=1}^{N} \sum_{a=1}^{\Delta} e^{2\pi i a n/\Delta} \right| = N - \left[\frac{N}{\Delta} \right] \Delta < \Delta,$$

where $[r]$ denotes the integer part of the real number r. On the other hand, we obtain

$$\left|\sum_{|\lambda|<1/R_X}\sum_{n=1}^{N}(\lambda R_X)^n\right| < 2\epsilon\sum_{n\geq 1}(\rho R_X)^n = \frac{2\epsilon\rho R_X}{1-\rho R_X},$$

where

$$\rho := \max\{|\lambda| : \lambda\in \mathrm{Spec}(W),\, |\lambda| < 1/R_X\}.$$

Combining these inequalities, by the triangle inequality we obtain

$$\left|N - \sum_{n=1}^{N}\sum_{\lambda\in\mathrm{Spec}(W)}(\lambda R_X)^n\right| < \Delta + \frac{2\epsilon\rho R_X}{1-\rho R_X}$$

as $N\to\infty$, and the assertion follows.

(2) Let $\mu(n)$ denote the Möbius function. Note that $\sum_{d|n}|\mu(d)| \leq n$. It is known that

$$\pi(n) = \frac{1}{n}\sum_{d|n}\mu(d)N_{n/d} \quad \text{and} \quad N_n = \sum_{\lambda\in\mathrm{Spec}(W)}\lambda^n$$

(see (10.3) and (10.4) in [Terras 2011]). Combining these equalities, we obtain

$$n\cdot\pi(n) = \sum_{\lambda\in\mathrm{Spec}(W)}\sum_{d|n}\mu(d)\lambda^{n/d},$$

and thus

$$\begin{aligned}\left|n\cdot\pi(n) - \sum_{\lambda\in\mathrm{Spec}(W)}\lambda^n\right| &= \left|\sum_{\lambda\in\mathrm{Spec}(W)}\sum_{\substack{d|n\\ d\geq 2}}\mu(d)\lambda^{n/d}\right| \\ &\leq \sum_{\lambda\in\mathrm{Spec}(W)}\sum_{\substack{d|n\\ d\geq 2}}|\mu(d)|\cdot|\lambda|^{n/d} \leq \sum_{\lambda\in\mathrm{Spec}(W)}\sum_{\substack{d|n\\ d\geq 2}}|\mu(d)|\cdot|\lambda|^{n/2} \\ &\leq n\sum_{\lambda\in\mathrm{Spec}(W)}\left(\frac{1}{R_X}\right)^{n/2} \leq 2\epsilon n\left(\frac{1}{R_X}\right)^{n/2}.\end{aligned}$$

On the other hand, since $R_X < 1$ and $0 < \alpha < \frac{1}{2}$ by our assumptions, there exists a natural number N_0 such that, for any $n \geq N_0$,

$$n \leq \left(\frac{1}{R_X}\right)^{(1/2-\alpha)n}, \quad \text{and so} \quad n\left(\frac{1}{R_X}\right)^{n/2} \leq \left(\frac{1}{R_X}\right)^{(1-\alpha)n}.$$

Hence, for any $n \geq N_0$,

$$\left|n\cdot\pi(n) - \sum_{\lambda\in\mathrm{Spec}(W)}\lambda^n\right| \leq 2\epsilon\left(\frac{1}{R_X}\right)^{(1-\alpha)n},$$

and the assertion follows. □

At last, we can prove the main theorem in this section.

Theorem 14. *Suppose that X is a finite, connected and non-cycle graph without degree-one vertices. Let $\gamma = 0.57721\ldots$ be the Euler–Mascheroni constant, and let $H_X = H_X(R_X)$ and C_X be the constants.*

(1) (*Graph-theoretic Mertens' first theorem*) *As $N \to \infty$,*

$$\sum_{n \le N} n \cdot \pi(n) R_X^n = N + O(1).$$

(2) (*Graph-theoretic Mertens' second theorem*) *There exists a constant B_X such that, as $N \to \infty$,*

$$\sum_{n \le N} \pi(n) R_X^n = \log N + B_X + O\Big(\frac{1}{N}\Big).$$

(3) *The equality $B_X = \gamma + \log C_X - H_X$ holds. Equivalently,*

$$\begin{aligned} B_X &= \gamma + \log C_X - \sum_{[P]} \sum_{n \ge 2} \frac{1}{n} R_X^{n\ell(P)} \\ &= \gamma + \log C_X + \prod_{[P]} \big(\log(1 - R_X^{\ell(P)}) + R_X^{\ell(P)}\big). \end{aligned}$$

(4) (*Graph-theoretic Mertens' third theorem*) *As $N \to \infty$,*

$$\prod_{\ell(P) \le N} \big(1 - R_X^{\ell(P)}\big) = \prod_{n \le N} (1 - R_X^n)^{\pi(n)} \sim \frac{e^{-\gamma}}{C_X} \cdot \frac{1}{N}.$$

Proof. (1) Let N_0 be a number as in the proof of Key Lemma 13(2), and let K denote the constant

$$K := \left| \sum_{n=1}^{N_0 - 1} n \cdot \pi(n) R_X^n - \sum_{n=1}^{N_0 - 1} \sum_{\lambda \in \mathrm{Spec}(W)} (\lambda R_X)^n \right|.$$

Assume that N is sufficiently large. Then it follows from Key Lemma 13(2) that

$$\begin{aligned} \left| \sum_{n=1}^{N} n \cdot \pi(n) R_X^n - \sum_{n=1}^{N} \sum_{\lambda \in \mathrm{Spec}(W)} (\lambda R_X)^n \right| &\le K + \left| \sum_{n=N_0}^{N} R_X^n \Big(n \cdot \pi(n) - \sum_{\lambda \in \mathrm{Spec}(W)} \lambda^n \Big) \right| \\ &\le K + 2\epsilon \sum_{n=N_0}^{N} R_X^{\alpha n} < K + \frac{2\epsilon}{1 - R_X^{\alpha}}, \end{aligned}$$

and therefore by Key Lemma 13(1) we have

$$\sum_{n=1}^{N} n \cdot \pi(n) R_X^n = \sum_{n=1}^{N} \sum_{\lambda \in \mathrm{Spec}(W)} (\lambda R_X)^n + O(1) = N + O(1) \quad \text{as } N \to \infty.$$

(2) We set $a_n = n \cdot \pi(n) R_X^n$. By part (1), we obtain $A(t) = t + O(1)$. By applying Fact 10(2) with $f(t) = 1/t$, we get

$$\begin{aligned}\sum_{n \le N} \pi(n) R_X^n &= \frac{A(N)}{N} + \int_1^N \frac{A(t)}{t^2}\,dt = \frac{N + O(1)}{N} + \int_1^N \frac{t + O(1)}{t^2}\,dt \\ &= 1 + O\Big(\frac{1}{N}\Big) + \int_1^N \Big(\frac{1}{t} + O\Big(\frac{1}{t^2}\Big)\Big)\,dt \\ &= 1 + O\Big(\frac{1}{N}\Big) + \Big[\log t + O\Big(\frac{1}{t}\Big)\Big]_1^N \\ &= 1 + O\Big(\frac{1}{N}\Big) + \log N + O\Big(\frac{1}{N}\Big) + O(1) = \log N + O(1) + O\Big(\frac{1}{N}\Big),\end{aligned}$$

and the assertion follows.

(3) Fix an arbitrary x satisfying $0 < x < 1$. By applying Fact 10(2) with $a_n = \pi(n) R_X^n$ and $f(t) = x^t$,

$$\sum_{n \le N} \pi(n) R_X^n x^n = A(N) x^N - \log x \int_1^N x^t A(t)\,dt$$

holds. It follows from part (2) that

$$\sum_{n \le N} \pi(n) R_X^n x^n = \Big(\log N + B_X + O\Big(\frac{1}{N}\Big)\Big) x^N - \log x \int_1^N x^t \Big(\log t + B_X + O\Big(\frac{1}{t}\Big)\Big)\,dt,$$

and, moreover, as $N \to \infty$,

$$P_X(R_X x) = -\log x \int_1^\infty x^t \Big(\log t + B_X + O\Big(\frac{1}{t}\Big)\Big)\,dt. \tag{9}$$

In order to calculate the right-hand side of this equality, for simplicity of arguments, we define the functions $I_n = I_n(x)$:

$$-\log x \int_1^\infty x^t \Big(\log t + B_X + O\Big(\frac{1}{t}\Big)\Big)\,dt = I_1 + I_2 + O(I_3),$$

where

$$I_1 = -\log x \int_1^\infty x^t \log t \, dt,$$
$$I_2 = -B_X \cdot \log x \int_1^\infty x^t \, dt = B_X \cdot x, \quad \text{and}$$
$$I_3 = -\log x \int_1^\infty \frac{x^t}{t} \, dt.$$

First, we compute the function I_1:

$$I_1 = -\int_1^\infty (x^t)' \log t \, dt = \int_1^\infty \frac{x^t}{t} \, dt.$$

Now we take $r = -t \log x$. Note that $\log x < 0$. Then we obtain

$$I_1 = \int_{-\log x}^\infty \frac{e^{-r}}{r} \, dr = -\mathrm{Ei}(\log x),$$

where $\mathrm{Ei}(z)$ ($z \in \mathbb{C}$ and $|\mathrm{Arg}(-z)| < \pi$) is the exponential integral

$$-\mathrm{Ei}(-z) = \int_z^\infty \frac{e^{-r}}{r} \, dr$$

(see, e.g., Equality (3.1.3) in [Lebedev 1972]). Since the function $\mathrm{Ei}(z)$ expands as

$$\mathrm{Ei}(z) = \gamma + \log(-z) + \sum_{k=1}^\infty \frac{z^k}{k \cdot k!}$$

(see Equality (3.1.6) in [ibid.]),

$$I_1 = -\gamma - \log(-\log x) + O(\log x) = -\gamma - \log(-\log x) + O(1-x).$$

Next we calculate the function I_3. It follows from the above result that

$$I_3 = -\log x \int_1^\infty \frac{x^t}{t} \, dt = (-\log x) I_1 = O(1-x)$$

as $x \uparrow 1$.

By combining the above results, the equality (9) is written as follows:

$$P_X(R_X x) = -\gamma - \log(-\log x) + B_X x + O(1-x),$$

and, moreover, as $x \uparrow 1$,

$$P_X(R_X x) + \log(-\log x) \to B_X - \gamma. \tag{10}$$

On the other hand, since

$$\log Z_X(R_X x) = \log \frac{1}{1-x} + \log C_X + O(1-x)$$

from the equality (8), as $x \uparrow 1$,

$$\log Z_X(R_X x) + \log(-\log x) = \log\left(\frac{-\log x}{1-x}\right) + \log C_X \to \log C_X. \tag{11}$$

Combining (10) with (11), we obtain

$$\begin{aligned} H_X &= \lim_{x\uparrow 1} H_X(R_X x) = \lim_{x\uparrow 1}\bigl(\log Z_X(R_X x) - P_X(R_X x)\bigr) \\ &= \lim_{x\uparrow 1}\bigl((\log Z_X(R_X x) + \log(-\log x)) - (P_X(R_X x) + \log(-\log x))\bigr) \\ &= \log C_X + \gamma - B_X. \end{aligned}$$

(4) Fix an arbitrary positive real number N. We define the following functions:

$$H_X^{\leq N} = \sum_{n\leq N} \pi(n) \sum_{m=2}^{\infty} \frac{1}{m} R_X^{mn} \quad \text{and} \quad H_X^{>N} = \sum_{n>N} \pi(n) \sum_{m=2}^{\infty} \frac{1}{m} R_X^{mn}.$$

Note that $H_X = H_X^{\leq N} + H_X^{>N}$. From parts (2) and (3), we obtain

$$\sum_{n\leq N} \pi(n) R_X^n + H_X^{\leq N} = \log N + \gamma + \log C_X - H_X^{>N} + O\Bigl(\frac{1}{N}\Bigr).$$

Since the left-hand side of this equality is equal to

$$\begin{aligned} \sum_{n\leq N} \pi(n) R_X^n + H_X^{\leq N} &= \sum_{n\leq N} \pi(n) \sum_{m=1}^{\infty} \frac{1}{m} R_X^{mn} \\ &= -\sum_{n\leq N} \pi(n) \log(1 - R_X^n) = -\log\Bigl(\prod_{n\leq N} (1 - R_X^n)^{\pi(n)}\Bigr), \end{aligned}$$

we obtain

$$\prod_{n\leq N} (1 - R_X^n)^{\pi(n)} = \frac{e^{-\gamma}}{C_X} \cdot \frac{1}{N} \exp\Bigl(H_X^{>N} + O\Bigl(\frac{1}{N}\Bigr)\Bigr).$$

Since $H_X^{>N} \to 0$ and $1/N \to 0$ as $N \to \infty$, the assertion follows. □

Last, we compute the following example.

Example 15 (continued from Example 12). Consider the graph $X = K_4 - \{\text{one edge}\}$. Then

$$H_X = 0.25613\ldots, \quad B_X = \gamma + \log C_X - H_X = -0.26933\ldots.$$

For example, if $N = 550$, then

$$\sum_{n\le N} \pi(n) R_X^n - \log N = -0.26842\cdots \approx B_X,$$

$$\prod_{n\le N} (1 - R_X^n)^{\pi(n)} = 0.18447\cdots \approx \frac{e^{-\gamma}}{C_X} \cdot \frac{1}{N} = 0.18457\ldots.$$

Remark 16. (See [Mertens 1874, Equation (17)], or [Hardy and Wright 2008, Theorem 428].) A number-theoretic analogue to part (3) in the preceding theorem is

$$B_1 = \gamma - H = \gamma + \sum_p \Bigl(\log\Bigl(1 - \frac{1}{p}\Bigr) + \frac{1}{p}\Bigr),$$

where $H = \sum_{n\ge 2} P(n)/n$ is a constant, and $P(s)$ is the prime zeta function.

Remark 17. We now compare parts (2)–(4) of our Theorem 14 with Theorem 1 in [Sharp 1991]. We define

$$h_X := -\log R_X, \quad N(P) = e^{h_X \ell(P)} \quad \text{and} \quad x = e^{h_X N}.$$

The quantity h_X is called the topological entropy of a flow in ergodic theory (see [Sharp 1991]), which is a constant in our setting. Note that $\ell(P) \le N$ if and only if $N(P) \le x$. Note that $R_X^{\ell(P)} = 1/N(P)$. Then our Mertens' second theorem can be rewritten as

$$\sum_{N(P)\le x} \frac{1}{N(P)} = \log(\log x) + B + O\Bigl(\frac{1}{\log x}\Bigr),$$

where $B := -\log h_X + B_X$, and, similarly, our Mertens' third theorem becomes

$$\prod_{N(P)\le x} \Bigl(1 - \frac{1}{N(P)}\Bigr) \sim \frac{1}{C_X/h_X} \cdot \frac{e^{-\gamma}}{\log x}.$$

In Theorem 1 in [Sharp 1991], our constant C_X/h_X, which is equal to a residue (up to sign) of the Ihara zeta function, corresponds with that of a dynamical zeta function for a flow.

Moreover, our Theorem 14(3) becomes

$$B = \gamma + \log(C_X/h_X) + \sum_{[P]} \Bigl(\log\Bigl(1 - \frac{1}{N(P)}\Bigr) + \frac{1}{N(P)}\Bigr).$$

Remark 18. Let $X = (V, E)$ be a finite, connected, non-cycle graph without degree-one vertices, and let $S = (V', E')$ be its k-subdivision (that is, let S be the graph obtained from X by adding k new vertices to each edge of X) (see Examples 6.4 and 8.5 in [Terras 2011]). Then

$$H_X = H_S, \quad C_X = (k+1)C_S, \quad \text{and} \quad B_X = B_S + \log(k+1).$$

This is proved as follows: note that $\Delta_S = (k+1)\Delta_X$, $R_S^{k+1} = R_X$, and

$$\pi_S(n) = \begin{cases} \pi_X(n/(k+1)) & \text{if } (k+1) \mid n, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\begin{aligned} H_S &= \sum_{m\geq 2} \frac{1}{m} P_S(R_S^m) = \sum_{m\geq 2} \frac{1}{m} \sum_{n=1}^{\infty} \pi_S(n) R_S^{mn} = \sum_{m\geq 2} \frac{1}{m} \sum_{n=1}^{\infty} \pi_X(n) R_S^{(k+1)mn} \\ &= \sum_{m\geq 2} \frac{1}{m} \sum_{n=1}^{\infty} \pi_X(n) R_X^{mn} = \sum_{m\geq 2} \frac{1}{m} P_X(R_X^m) = H_X. \end{aligned}$$

Note that $\nu' = \nu + k\epsilon$, $\epsilon' = (k+1)\epsilon$, and $Z_S(u) = Z_X(u^{k+1})$, and so

$$(1-u^2)^{\epsilon-\nu} f_S(u) = (1-u^{2(k+1)})^{\epsilon-\nu} f_X(u^{k+1}),$$
$$(1-R_S^2)^{\epsilon-\nu} R_S f_S'(R_S) = (k+1)(1-R_X^2)^{\epsilon-\nu} R_X f_X'(R_X).$$

Therefore,

$$(k+1)C_S = \frac{-(k+1)}{(1-R_S^2)^{\epsilon'-\nu'} R_S f_S'(R_S)} = C_X,$$

and so

$$B_X = \gamma + \log C_X - H_X = \gamma + \log C_S - H_S + \log(k+1) = B_S + \log(k+1).$$

Acknowledgements

Saito would like to thank Professor Toyokazu Hiramatsu for his encouragement and thoughtful suggestions. The authors thank the referee for his or her valuable comments and careful review of this paper.

References

[Apostol 1974] T. M. Apostol, *Mathematical analysis*, 2nd ed., Addison-Wesley, Reading, MA, 1974. MR 49 #9123 Zbl 0309.26002

[Bass 1992] H. Bass, "The Ihara–Selberg zeta function of a tree lattice", *Internat. J. Math.* **3**:6 (1992), 717–797. MR 94a:11072 Zbl 0767.11025

[Fröberg 1968] C.-E. Fröberg, "On the prime zeta function", *Nordisk Tidskr. Informationsbehandling* (*BIT*) **8** (1968), 187–202. MR 38 #4421 Zbl 0167.04201

[Glaisher 1891] J. W. L. Glaisher, "On the sums of the inverse powers of the prime numbers", *Quart. J. Pure. Appl. Math.* **25** (1891), 347–362. JFM 23.0275.02

[Hardy and Wright 2008] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 6th ed., Oxford University Press, 2008. Revised by D. R. Heath-Brown and J. H. Silverman. MR 2009i:11001 Zbl 1159.11001

[Hashimoto 1989] K.-I. Hashimoto, "Zeta functions of finite graphs and representations of p-adic groups", pp. 211–280 in *Automorphic forms and geometry of arithmetic varieties*, edited by K.-I. Hashimoto and Y. Namikawa, Adv. Stud. Pure Math. **15**, Academic Press, Boston, 1989. MR 91i:11057 Zbl 0709.22005

[Horton et al. 2006] M. D. Horton, H. M. Stark, and A. A. Terras, "What are zeta functions of graphs and what are they good for?", pp. 173–189 in *Quantum graphs and their applications*, edited by G. Berkolaiko et al., Contemp. Math. **415**, Amer. Math. Soc., Providence, RI, 2006. MR 2007i:05088 Zbl 1222.11109

[Ihara 1966] Y. Ihara, "On discrete subgroups of the two by two projective linear group over $\mathfrak{p}$-adic fields", *J. Math. Soc. Japan* **18** (1966), 219–235. MR 36 #6511 Zbl 0158.27702

[Jameson 2003] G. J. O. Jameson, *The prime number theorem*, London Mathematical Society Student Texts **53**, Cambridge University Press, 2003. MR 2004c:11002 Zbl 1033.11001

[Korevaar 2002] J. Korevaar, "A century of complex Tauberian theory", *Bull. Amer. Math. Soc.* (*N.S.*) **39**:4 (2002), 475–531. MR 2003g:40004 Zbl 1001.40007

[Kotani and Sunada 2000] M. Kotani and T. Sunada, "Zeta functions of finite graphs", *J. Math. Sci. Univ. Tokyo* **7**:1 (2000), 7–25. MR 2001f:68110 Zbl 0978.05051

[Landau and Walfisz 1920] E. Landau and A. Walfisz, "Über die Nichtfortsetzbarkeit einiger durch Dirichletsche Reihen definierter Funktionen", *Rend. Circ. Mat. Palermo* **44** (1920), 82–86. JFM 47.0287.02

[Lebedev 1972] N. N. Lebedev, *Special functions and their applications*, 2nd ed., Dover, New York, 1972. MR 50 #2568 Zbl 0271.33001

[Mertens 1874] F. Mertens, "Ein Beitrag zur analytischen Zahlentheorie", *J. Reine Angew. Math.* **78** (1874), 46–63. JFM 06.0116.01

[Northshield 1998] S. Northshield, "A note on the zeta function of a graph", *J. Combin. Theory Ser. B* **74**:2 (1998), 408–410. MR 99g:05106 Zbl 1027.05048

[Parry 1983] W. Parry, "An analogue of the prime number theorem for closed orbits of shifts of finite type and their suspensions", *Israel J. Math.* **45**:1 (1983), 41–52. MR 85c:58089 Zbl 0552.28020

[Parry and Pollicott 1983] W. Parry and M. Pollicott, "An analogue of the prime number theorem for closed orbits of Axiom A flows", *Ann. of Math.* (2) **118**:3 (1983), 573–591. MR 85i:58105 Zbl 0537.58038

[Sharp 1991] R. Sharp, "An analogue of Mertens' theorem for closed orbits of Axiom A flows", *Bol. Soc. Brasil. Mat.* (*N.S.*) **21**:2 (1991), 205–229. MR 93a:58142 Zbl 0761.58041

[Stark and Terras 1996] H. M. Stark and A. A. Terras, "Zeta functions of finite graphs and coverings", *Adv. Math.* **121**:1 (1996), 124–165. MR 98b:11094 Zbl 0874.11064

[Storm 2010] C. Storm, "An infinite family of graphs with the same Ihara zeta function", *Electron. J. Combin.* **17**:1 (2010), Research Paper 82. MR 2011g:05156 Zbl 1215.05096

[Terras 2011] A. A. Terras, *Zeta functions of graphs: a stroll through the garden*, Cambridge Studies in Advanced Mathematics **128**, Cambridge University Press, 2011. MR 2012d:05016 Zbl 1206.05003

[Titchmarsh 1958] E. C. Titchmarsh, *The theory of functions*, 2nd ed., Oxford University Press, 1958. MR 3155290 Zbl 0336.30001

[Titchmarsh 1986] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., Clarendon, New York, 1986. MR 88c:11049 Zbl 0601.10026

Received October 3, 2013. Revised May 29, 2014.

TAKEHIRO HASEGAWA
SHIGA UNIVERSITY
OTSU
SHIGA 520-0862
JAPAN
thasegawa3141592@yahoo.co.jp

SEIKEN SAITO
WASEDA UNIVERSITY
SHINJUKU
TOKYO 169-8050
JAPAN
seiken.saitou@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.147

ON WHITTAKER MODULES FOR A LIE ALGEBRA ARISING FROM THE 2-DIMENSIONAL TORUS

SHAOBIN TAN, QING WANG AND CHENGKANG XU

Let A be the ring of Laurent polynomials in two variables and B be the set of skew derivations of A. We denote by $\tilde{L}$ the semidirect product of A and B, and by L the universal central extension of the derived Lie algebra of $\tilde{L}$. We study the Whittaker modules for the Lie algebra L. The irreducibilities for the universal Whittaker modules are given. Moreover, a $\mathbb{Z}$-gradation is defined on the universal Whittaker modules and we determine all $\mathbb{Z}$-graded irreducible quotients of the reducible universal Whittaker modules.

1. Introduction

The Lie algebra we considered in this paper can be seen as a generalization of the rank one Heisenberg–Virasoro algebra. The rank one Heisenberg–Virasoro algebra HVir was first given in [Arbarello et al. 1988]; it is the universal central extension of the Lie algebra $\mathfrak{D}$ of differential operators on a circle of order at most one; $\mathfrak{D}$ has a basis $\{t^n, d_n = t^{n+1}d/dt \mid n \in \mathbb{Z}\}$ with Lie bracket relations

$$[t^n, t^m] = 0, \quad [d_i, t^n] = nt^{i+n}, \quad [d_i, d_j] = (j-i)d_{i+j},$$

and HVir has the Lie bracket relations

$$\begin{aligned}
&[d_m, d_n] = (n-m)d_{m+n} + \delta_{m+n,0}\frac{m^3-m}{12}c_1,\\
&[d_m, t^n] = nt^{m+n} + (m^2-m)\delta_{m+n,0}c_2,\\
&[t^m, t^n] = m\delta_{m+n,0}c_3,\\
&[c_i, \mathrm{HVir}] = 0, \quad i = 1, 2, 3.
\end{aligned}$$

One can see that HVir contains a Heisenberg subalgebra and a Virasoro subalgebra. In [Xue et al. 2006], the authors generalized the rank one Heisenberg–Virasoro

Tan was partially supported by NSF of China (grant number 11471268). Wang was partially supported by NSF of China (grant number 11371024), Natural Science Foundation of Fujian Province (grant number 2013J01018) and Fundamental Research Funds for the Central University (grant number 2013121001).
MSC2010: primary 06B15; secondary 17B65, 17B66.
Keywords: Whittaker module, infinite-dimensional Lie algebra, torus.

algebra to the rank two case. More precisely, let $A = \mathbb{C}[t_1^{\pm 1}, t_2^{\pm 1}]$ be the ring of Laurent polynomials and B be the set of skew derivations of A spanned by elements of the form

$$E(\alpha) = t^\alpha(\alpha(2)d_1 - \alpha(1)d_2),$$

where $\alpha = (\alpha(1), \alpha(2)) \in \mathbb{Z}^2$, $t^\alpha = t_1^{\alpha(1)} t_2^{\alpha(2)}$ and d_1, d_2 are degree derivations of A. Set $\tilde{L} = A \oplus B$. Then $\tilde{L}$ becomes a Lie algebra under the Lie bracket relations

$$[t^\alpha, t^\beta] = 0, \quad [t^\alpha, E(\beta)] = \det\begin{pmatrix}\beta\\ \alpha\end{pmatrix} t^{\alpha+\beta}, \quad [E(\alpha), E(\beta)] = \det\begin{pmatrix}\beta\\ \alpha\end{pmatrix} E(\alpha+\beta),$$

where $\alpha, \beta \in \mathbb{Z}^2$, and

$$\det\begin{pmatrix}\beta\\ \alpha\end{pmatrix} = \beta(1)\alpha(2) - \alpha(1)\beta(2).$$

Let $\tilde{L}'$ be the derived Lie subalgebra of $\tilde{L}$. Then $\tilde{L}'$ is perfect and has a universal central extension L with the following Lie bracket relations [Xue et al. 2006]:

$$\begin{aligned} &[t^\alpha, t^\beta] = 0, \quad [K_i, L] = 0 \quad \text{for } i = 1, 2, 3, 4, \\ &[t^\alpha, E(\beta)] = \det\begin{pmatrix}\beta\\ \alpha\end{pmatrix} t^{\alpha+\beta} + \delta_{\alpha+\beta,0} h(\alpha), \\ &[E(\alpha), E(\beta)] = \det\begin{pmatrix}\beta\\ \alpha\end{pmatrix} E(\alpha+\beta) + \delta_{\alpha+\beta,0} f(\alpha), \end{aligned} \tag{1-1}$$

where $\alpha, \beta \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, K_1, K_2, K_3, K_4 are central elements, and

$$h(\alpha) = \alpha(1)K_1 + \alpha(2)K_2 \quad \text{and} \quad f(\alpha) = \alpha(1)K_3 + \alpha(2)K_4. \tag{1-2}$$

One can see that L contains a Virasoro-like subalgebra spanned by

$$\{E(\alpha), K_3, K_4 \mid \alpha \in \mathbb{Z}^2 \setminus \{(0, 0)\}\},$$

which was introduced by Kirkman, Procesi and Small [Kirkman et al. 1994]. In this paper, we study Whittaker modules for the Lie algebra L.

Whittaker modules were first discovered by Arnal and Pinczon [1974] in the study of the irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$. Kostant [1978] introduced the term "Whittaker module" and studied Whittaker modules for a complex semisimple Lie algebra $\mathfrak{g}$. In particular, he built up a one-to-one correspondence between the set of all equivalence classes of Whittaker modules and the set of all ideals in the center of the universal enveloping algebra of $\mathfrak{g}$. Moreover, Whittaker modules were shown to be one important class in the classification of the irreducible modules for the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ [Block 1981]. Since Kostant's definition of Whittaker module for finite-dimensional semisimple Lie algebras is based on a triangular decomposition [Kostant 1978], it is natural to consider Whittaker modules for other algebras with a triangular decomposition, such as Heisenberg algebras, affine Lie algebras, generalized Weyl algebras and the Virasoro algebra, which were studied in [Christodoulopoulou 2008; Benkart and Ondrus 2009; Ondrus and Wiesner 2009], respectively. Recently, Whittaker modules for other infinite-dimensional Lie algebras related to the Virasoro algebra were also studied, such as the rank one Heisenberg–Virasoro algebra [Lu and Zhao 2013], the Schrödinger–Witt algebra [Zhang et al. 2010], and so on. Note that these algebras are of rank one, that is, they are graded by $\mathbb{Z}$.

Motivated by these works, Batra and Mazorchuk [2011] defined a Whittaker pair $(\mathfrak{g}, \mathfrak{n})$ for a Lie algebra $\mathfrak{g}$ and a quasinilpotent subalgebra $\mathfrak{n}$ such that $\mathfrak{n}$ acts locally nilpotent on the adjoint module $\mathfrak{g}/\mathfrak{n}$. They obtained a general setup for the study of Whittaker modules, which includes Lie algebras with triangular decomposition and simple Lie algebras of Cartan type. However, this general theory doesn't work for many exceptions such as the generalized Virasoro algebras [Guo and Liu 2011a], the Virasoro-like algebra $\mathcal{V}$ [Guo and Liu 2011b] and the Lie algebra L considered in this paper. Note that the Virasoro-like algebra $\mathcal{V}$ is of rank two, that is, it is graded by $\mathbb{Z}^2$. Therefore, Guo and Liu used a different technique to deal with the Whittaker modules for the Lie algebra $\mathcal{V}$ [ibid.]. We note that the Lie algebra L considered in this paper contains the Virasoro-like Lie algebra $\mathcal{V}$ as a Lie subalgebra, and we will see that the study of Whittaker modules for L is more complicated than that for $\mathcal{V}$.

The paper is organized as follows. In Section 2, we state some facts about total orders on $\mathbb{Z}^2$ and give the definition of Whittaker modules for the Lie algebra L. In Section 3, we determine all the Whittaker vectors for the universal Whittaker modules. In Section 4, we study irreducibility for the universal Whittaker modules. We define a $\mathbb{Z}$-gradation on the universal Whittaker modules and determine all $\mathbb{Z}$-graded irreducible quotients for the reducible universal Whittaker modules. Finally, we prove some more properties of these $\mathbb{Z}$-graded irreducible quotients.

Throughout this paper, we denote the sets of complex numbers, nonzero complex numbers, integers, nonnegative integers and positive integers by $\mathbb{C}$, $\mathbb{C}^\times$, $\mathbb{Z}$, $\mathbb{Z}_+$ and $\mathbb{N}$, respectively. All Lie algebras mentioned in this paper are over the complex field $\mathbb{C}$. The universal enveloping algebra for a Lie algebra $\mathfrak{g}$ is denoted by $\mathcal{U}(\mathfrak{g})$.

2. Whittaker modules for the Lie algebra L

In this section, we recall the definition of the Lie algebra L given in [Xue et al. 2006] and the definition of Whittaker module. We also present some facts about them.

For an element α in $G = \mathbb{Z}^2$, we denote $\alpha = (\alpha(1), \alpha(2))$. For any $\alpha,\ \beta \in \mathbb{Z}^2$, we set

$$\det\binom{\beta}{\alpha} = \beta(1)\alpha(2) - \alpha(1)\beta(2).$$

For any $\alpha \in G \setminus \{\mathbf{0} = (0, 0)\}$, let $X(\alpha)$ denote t^α or $E(\alpha)$ if it has no special explanation. The Lie algebra L is spanned by the elements of the form

$$\{t^\alpha, E(\alpha), K_i \mid \alpha \in \mathbb{Z}^2 \setminus \{\mathbf{0}\},\ i = 1, 2, 3, 4\},$$

with Lie bracket relations defined by (1-1). Clearly, L is $\mathbb{Z}^2$-graded and contains a Virasoro-like algebra $\mathcal{V}$ as the Lie subalgebra spanned by

$$\{E(\alpha), K_3, K_4 \mid \alpha \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}\}.$$

Fix a total order $\prec$ on $G = \mathbb{Z}^2$ which is compatible with the addition of G, i.e., $\alpha \prec \beta$ implies $\alpha + \gamma \prec \beta + \gamma$ for all $\gamma \in G$. We have the obvious meanings for $\preceq$, $\succ$, and $\succeq$. Then we have a decomposition $G = G_+ \uplus \{\mathbf{0}\} \uplus G_-$, where $G_\pm = \{\alpha \in G \mid \pm\alpha \succ \mathbf{0}\}$.

We say that $\prec$ is dense if for any $\alpha \in G_+$, there is some $\beta \in G_+$ such that $\beta \prec \alpha$; $\prec$ is discrete if there exists a smallest element in G_+. For example, the lexicographical order is discrete, since (0, 1) is the smallest element in G_+. Dense total orders on G exist. For example, let $\alpha = (\alpha(1), \alpha(2))$, $\beta = (\beta(1), \beta(2))$. We say $\alpha \prec \beta$ if $\alpha(1) + \alpha(2)\pi < \beta(1) + \beta(2)\pi$. One can check that this is a dense compatible total order on G. The following lemma is from [Guo and Liu 2011b].

Lemma 2.1. (1) *Nonzero elements $\alpha, \beta \in G$ form a basis of G if and only if* $\det\binom{\alpha}{\beta} = \pm 1$.

(2) *If $\prec$ is dense, then for any $\alpha \succ \mathbf{0}$, there is some $\mathbf{0} \prec \beta \prec \alpha$ such that* $\det\binom{\alpha}{\beta} \neq 0$.

(3) *If $\prec$ is discrete, let ϵ denote the smallest positive element in G. Then there exists $\epsilon' \succ \mathbf{0}$ such that ϵ, ϵ' form a basis of G.*

According to the total order on G fixed above, L has a triangular decomposition

$$L = L_- \oplus L_0 \oplus L_+,$$

where $L_\pm = \mathrm{Span}_{\mathbb{C}}\{t^\alpha, E(\alpha) \mid \pm\alpha \succ \mathbf{0}\}$ and $L_0 = \mathrm{Span}_{\mathbb{C}}\{K_i \mid i = 1, 2, 3, 4\}$.

Recall from [Batra and Mazorchuk 2011] that a Lie algebra $\mathfrak{g}$ is called quasi-nilpotent if

$$\bigcap_{k\in\mathbb{N}} \mathfrak{g}^k = 0,$$

where $\mathfrak{g}^{k+1} = [\mathfrak{g}^k, \mathfrak{g}]$ is defined by induction. Now we claim that L_+ is not quasinilpotent. Indeed, L_+ contains $\mathcal{V}_+ = \bigoplus_{\alpha\in G_+} \mathbb{C}E(\alpha)$ as a subalgebra, which is proved to be not quasinilpotent in [Guo and Liu 2011b], so that

$$\bigcap_{k\in\mathbb{N}} L_+^k \supseteq \bigcap_{k\in\mathbb{N}} \mathcal{V}_+^k \neq 0.$$

So $(L,\ L_+)$ is not a Whittaker pair in the sense of [Batra and Mazorchuk 2011], and the general theory for Whittaker modules there does not apply to the Lie algebra L. Thus we treat it as follows.

Fix any nonzero Lie algebra homomorphism $\varphi : L_+ \to \mathbb{C}$ and let $k_1, k_2, k_3, k_4 \in \mathbb{C}$. Given an L-module V, a vector $v \in V$ is called a Whittaker vector of type $(\varphi, k_1, k_2, k_3, k_4)$ if $xv = \varphi(x)v$ for all $x \in L_+$, and $K_i v = k_i v$ for $i = 1, 2, 3, 4$. V is called a Whittaker module of type $(\varphi, k_1, k_2, k_3, k_4)$ if $V = \mathcal{U}(L)v$ for some Whittaker vector v of type $(\varphi, k_1, k_2, k_3, k_4)$. In this paper, all Whittaker modules and Whittaker vectors are of type $(\varphi, k_1, k_2, k_3, k_4)$ if not specified. Clearly, u is a Whittaker vector if and only if $(X(\alpha) - \varphi(X(\alpha)))u = 0$ for all $\alpha \in G_+$, $X(\alpha) = t^\alpha$ and $E(\alpha)$. Notice that $\varphi(L_+^2) = [\varphi(L_+), \varphi(L_+)] = 0$. We have the following facts.

Proposition 2.2. *Let $\prec$ be a total order on G.*

(1) *If $\prec$ is dense, then any Lie algebra homomorphism $\varphi : L_+ \to \mathbb{C}$ is the zero homomorphism.*

(2) *If $\prec$ is discrete and ϵ denotes the smallest positive element in G, then $\varphi(t^\alpha) = \varphi(E(\alpha)) = 0$ for all $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$.*

Proof. (1) Suppose $\prec$ is dense. Then by Lemma 2.1, for any $\alpha \in G_+$ there is some $\beta \in G_+$ such that $\beta \prec \alpha$ and $\det\binom{\alpha}{\beta} \neq 0$. Thus

$$X(\alpha) = \frac{1}{\det\binom{\beta}{\alpha-\beta}}[E(\alpha-\beta), X(\beta)] = \frac{1}{\det\binom{\beta}{\alpha}}[E(\alpha-\beta), X(\beta)] \in L_+^2.$$

So $\varphi(t^\alpha) = \varphi(E(\alpha)) = 0$, and this shows that $\varphi = 0$.

(2) Let $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$. We have $\alpha - i\epsilon \in G_+ \setminus \mathbb{Z}\epsilon$ for all $i \in \mathbb{Z}$, and $\det\binom{\alpha}{\epsilon} \neq 0$. Thus

$$X(\alpha) = \frac{1}{\det\binom{\alpha-\epsilon}{\epsilon}}[X(\epsilon), E(\alpha-\epsilon)] = \frac{1}{\det\binom{\alpha}{\epsilon}}[X(\epsilon), E(\alpha-\epsilon)] \in L_+^2,$$

which shows that $\varphi(t^\alpha) = \varphi(E(\alpha)) = 0$ for all $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$. □

Thus, we assume that $\prec$ is discrete with smallest positive element ϵ in G throughout the rest of this paper.

3. Whittaker vectors in universal Whittaker modules

In this section we study the universal Whittaker module and determine all its Whittaker vectors. By Lemma 2.1, we know that there exists $\epsilon' \in G_+$ such that $\{\epsilon, \epsilon'\}$ is a basis of G. We will always use this basis for G from now on.

We construct the universal Whittaker module of type $(\varphi, k_1, k_2, k_3, k_4)$ over L, denoted $M_{\varphi,k_1,k_2,k_3,k_4}$, as follows: let $\mathbb{C}\tilde{v}$ be the one-dimensional $(L_0 \oplus L_+)$-module defined by $x\tilde{v} = \varphi(x)\tilde{v}$ for any $x \in L_+$ and $K_i\tilde{v} = k_i\tilde{v}$ for $i = 1, 2, 3, 4$. Set

$$M_{\varphi,k_1,k_2,k_3,k_4} = \mathcal{U}(L) \otimes_{\mathcal{U}(L_0 \oplus L_+)} \mathbb{C}\tilde{v}.$$

This is a left $\mathcal{U}(L)$-module under left multiplication. Set $v = 1 \otimes \tilde{v}$. We have $M_{\varphi,k_1,k_2,k_3,k_4} = \mathcal{U}(L)v$. It is obvious that $M_{\varphi,k_1,k_2,k_3,k_4}$ has the following universal property: for any Whittaker module W of type $(\varphi, k_1, k_2, k_3, k_4)$ generated by a Whittaker vector w, there is an L-module epimorphism ϕ from $M_{\varphi,k_1,k_2,k_3,k_4}$ to W which maps v to w.

By the Poincaré–Birkhoff–Witt (PBW) theorem, $M_{\varphi,k_1,k_2,k_3,k_4}$ is isomorphic to $\mathcal{U}(L_-)$ as a vector space. Let $L_- = L_-^t \oplus L_-^E$, where

$$L_-^t = \mathrm{Span}_{\mathbb{C}}\{t^{-\alpha} \mid \alpha \succ 0\}, \quad L_-^E = \mathrm{Span}_{\mathbb{C}}\{E(-\alpha) \mid \alpha \succ 0\}.$$

Since $\mathcal{U}(L_-^t)$ and $\mathcal{U}(L_-^E)$ have $\mathbb{C}$-bases

$$B^t = \{1, t^{-\beta_m} \cdots t^{-\beta_1} \mid m \in \mathbb{N}, \beta_m \succeq \cdots \succeq \beta_1 \succ 0\}$$

and

$$B^E = \{1, E(-\alpha_n) \cdots E(-\alpha_1) \mid n \in \mathbb{N}, \alpha_n \succeq \cdots \succeq \alpha_1 \succ 0\},$$

respectively, $\mathcal{U}(L_-)$ has a $\mathbb{C}$-basis

$$\begin{aligned} B = B^t B^E = B^t \cup B^E \cup \{&t^{-\beta_m} \cdots t^{-\beta_1} E(-\alpha_n) \cdots E(-\alpha_1) \\ &\mid m, n \in \mathbb{N}, \alpha_n \succeq \cdots \succeq \alpha_1 \succ 0, \beta_m \succeq \cdots \succeq \beta_1 \succ 0\} \end{aligned}$$

and $M_{\varphi,k_1,k_2,k_3,k_4}$ has a $\mathbb{C}$-basis Bv. For convenience, we set $M = M_{\varphi,k_1,k_2,k_3,k_4}$ and

$$\begin{aligned} E_\pm &= \bigoplus_{k \in \mathbb{N}} \mathbb{C}E(\pm k\epsilon), & T_\pm &= \bigoplus_{k \in \mathbb{N}} \mathbb{C}t^{\pm k\epsilon}, \\ H_\pm &= E_\pm \oplus T_\pm, & H &= H_- \oplus L_0 \oplus H_+, \\ E &= E_- \oplus E_+, & T &= T_- \oplus T_+. \end{aligned}$$

Set

$$M(H) = \mathcal{U}(H)v = \mathcal{U}(H_-)v, \quad M(T) = \mathcal{U}(T)v = \mathcal{U}(T_-)v.$$

For $\alpha \in G$, set $\alpha = \alpha[1]\epsilon + \alpha[2]\epsilon'$, where $\alpha[1], \alpha[2] \in \mathbb{Z}$.

Lemma 3.1. (1) *If $\alpha \in G_+$, then $\alpha[2] \geq 0$. In particular, if $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$, then $\alpha[2] > 0$.*

(2) *If $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$, then for any $u \in M(H)$, $x \in \mathcal{U}(L_-)$, we have*

$$(X(\alpha) - \varphi(X(\alpha)))xu = [X(\alpha), x]u. \tag{3-1}$$

(3) *Let $\alpha_1, \ldots, \alpha_n \in G_+$, $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$. If $\alpha - \sum_{i=1}^n \alpha_i \in G_+ \setminus \mathbb{Z}\epsilon$, then we have*

$$X(\alpha)X(-\alpha_n) \cdots X(-\alpha_1)w = 0 \quad \textit{for } w \in M(H), \tag{3-2}$$

where all $X(\beta)$ denote t^β or $E(\beta)$.

Proof. (1) Suppose $\alpha[2] < 0$. Then we have $-\alpha[2]\epsilon' \succeq \epsilon' \succ \alpha[1]\epsilon$, which implies $\alpha = \alpha[1]\epsilon + \alpha[2]\epsilon' \prec 0$. This is a contradiction with $\alpha \in G_+$.

(2) We may assume $u = X(-n_1\epsilon)\cdots X(-n_s\epsilon)v$, where $s \in \mathbb{Z}_+$, $n_1, \ldots n_s \in \mathbb{N}$. Then

$$\begin{aligned}(X(\alpha) - \varphi(X(\alpha)))xu &= X(\alpha)xu - xX(-n_1\epsilon)\cdots X(-n_s\epsilon)X(\alpha)v \\ &= [X(\alpha), x]u + x[X(\alpha), X(-n_1\epsilon)\cdots X(-n_s\epsilon)]v.\end{aligned}$$

Notice that

$$x[X(\alpha), X(-n_1\epsilon)\cdots X(-n_s\epsilon)]v \in x \sum_{\eta \in G_+ \setminus \mathbb{Z}\epsilon} \mathcal{U}(H)X(\eta)v = 0$$

by Proposition 2.2, and thus (3-1) holds.

(3) Now we prove (3-2) by induction on n. For $n = 0$, since $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$, we have that $\varphi(X(\alpha)) = 0$ by Proposition 2.2. Hence, by (2),

$$X(\alpha)w = (X(\alpha) - \varphi(X(\alpha)))w = [X(\alpha), 1]w = 0$$

for $w \in M(H)$. Suppose that $n > 0$ and that the result holds for any positive integer $k < n$. Then for $w \in M(H)$, by applying the induction hypothesis, we have

$$\begin{aligned}X(\alpha)X(-\alpha_n)\cdots X(-\alpha_1)w &= \det\begin{pmatrix}\alpha \\ \alpha_n\end{pmatrix} X(\alpha - \alpha_n)X(-\alpha_{n-1})\cdots X(-\alpha_1)w \\ &\qquad + X(-\alpha_n)X(\alpha)X(-\alpha_{n-1})\cdots X(-\alpha_1)w \\ &= 0.\end{aligned}$$

□

For later use, we define some subsets of B. Set

$$B(0) = \{1, t^{-\beta_m}\cdots t^{-\beta_1} \in B \mid m \in \mathbb{N}, \beta_m \succeq \cdots \succeq \beta_1 \in G_+ \setminus \mathbb{Z}\epsilon\}.$$

For $h \in \mathbb{N}$, set

$$B_E(h) = \left\{ E(-\alpha_n)\cdots E(-\alpha_1) \in B \,\middle|\, n \in \mathbb{N}, \alpha_m \succeq \cdots \succeq \alpha_1 \in G_+ \setminus \mathbb{Z}\epsilon, \sum_{i=1}^n \alpha_i[2] = h \right\},$$

$$B(h) = B(0)B_E(h), \quad B(-h) = \varnothing, \quad B'(h) = \bigcup_{h' < h} B(h'), \quad \bar{B}(h) = B(h) \cup B'(h).$$

For $h \in \mathbb{N}$, $\beta \in G_+ \setminus \mathbb{Z}\epsilon$, set

$$B_T(h, \beta) = \left\{ t^{-\beta_m}\cdots t^{-\beta_1} \,\middle|\, \beta_m \succeq \cdots \succeq \beta_1 = \beta, \sum_{i=1}^m \beta_i[2] = h \right\}$$

and

$$B_T(h) = \bigcup_{\beta \in G_+ \setminus \mathbb{Z}\epsilon} B_T(h, \beta).$$

Let $\mathcal{H} = \{(h, \beta) \mid B_T(h, \beta) \neq \varnothing\}$, and define a total order $\gg$ on $\mathcal{H}$ by setting

$$(h, \beta) \gg (h', \beta') \quad \text{if } h > h', \text{ or } h = h' \text{ and } \beta \prec \beta'.$$

Moreover, we denote

$$B'_T(h, \beta) = \bigcup_{(h,\beta)\gg(h',\beta')} B_T(h', \beta'), \qquad B'_T(h) = \bigcup_{h'<h} B_T(h'),$$
$$\bar{B}_T(h, \beta) = B_T(h, \beta) \cup B'_T(h, \beta), \quad \bar{B}_T(h) = \bigcup_{h'\leq h} B_T(h'),$$

and we set $B_T(0) = \bar{B}_T(0) = \{1\}$, $B'_T(0) = \varnothing$. Then one can see that $B(0) = \bigcup_{h\in\mathbb{Z}_+} B_T(h)$.

Lemma 3.2. (1) *For any $u \in M \setminus B(0)M(H)$, there exists $\eta \in G_+ \setminus \mathbb{Z}\epsilon$ such that*

$$(t^\eta - \varphi(t^\eta))u \in B(0)M(H) \setminus \mathbb{C}v.$$

(2) *For any $u' \in B(0)M(H) \setminus M(H)$, there exist $\gamma_1, \ldots, \gamma_s \in G_+$ such that*

$$(E(\gamma_s) - \varphi(E(\gamma_s))) \cdots (E(\gamma_1) - \varphi(E(\gamma_1)))u' \in M(H) \setminus \mathbb{C}v.$$

Proof. (1) Since $u \in M \setminus B(0)M(H)$, there exists $h \in \mathbb{N}$ such that $u \in \bar{B}(h)M(H)$ and $u \notin B'(h)M(H)$. Thus, we may write

$$u = \sum_{x\in B(h)} x v_x + \sum_{y\in B'(h)} y v_y,$$

where $v_x, v_y \in M(H)$ and both sums are finite, and the elements x are of the form

$$x = t^{-\beta_m} \cdots t^{-\beta_1} E(-\alpha_n) \cdots E(-\alpha_1), \tag{3-3}$$

with $m \in \mathbb{Z}_+$, $n \in \mathbb{N}$, $\beta_m \succeq \cdots \succeq \beta_1 \in G_+ \setminus \mathbb{Z}\epsilon$, $\alpha_n \succeq \cdots \succeq \alpha_1 \in G_+ \setminus \mathbb{Z}\epsilon$ and $\sum_{i=1}^n \alpha_i[2] = h$, where $m = 0$ means that $x = E(-\alpha_n) \cdots E(-\alpha_1)$.

Let $\eta \in G$ be such that $\eta[2] = h$, $\eta - \sum_{i=1}^n \alpha_i \in -\mathbb{N}\epsilon$ and, for each $(\alpha_n, \ldots, \alpha_1)$ in (3-3) associated to an element x appearing in $\sum_{x\in B(h)} x v_x$,

$$\det\begin{pmatrix} \eta - \sum_{j=i+1}^n \alpha_j \\ \alpha_i \end{pmatrix} \neq 0 \quad \text{for } 1 \leq i \leq n-1 \quad \text{and} \quad \det\begin{pmatrix} \eta \\ \alpha_n \end{pmatrix} \neq 0.$$

Since the sum $\sum_{x\in B(h)} x v_x$ is finite, it is obvious that such an η exists. It follows that $\eta \in G_+ \setminus \mathbb{Z}\epsilon$, thus we have $\bar{u} = (t^\eta - \varphi(t^\eta))u = t^\eta u$ by Proposition 2.2. Note that $t^\eta\big(\sum_{y\in B'(h)} y v_y\big) = 0$ by Lemma 3.1(3), thus we have

$$\bar{u} = \sum_{x\in B(h)} t^{-\beta_m} \cdots t^{-\beta_1} t^\eta E(-\alpha_n) \cdots E(-\alpha_1) v_x.$$

For every summand in $\bar{u}$, by using Lemma 3.1(3), we have

$$\begin{aligned}
&t^{-\beta_m}\cdots t^{-\beta_1}t^{\eta}E(-\alpha_n)\cdots E(-\alpha_1)v_x\\
&\quad= t^{-\beta_m}\cdots t^{-\beta_1}[t^{\eta},E(-\alpha_n)]E(-\alpha_{n-1})\cdots E(-\alpha_1)v_x\\
&\qquad\qquad + t^{-\beta_m}\cdots t^{-\beta_1}E(-\alpha_n)t^{\eta}E(-\alpha_{n-1})\cdots E(-\alpha_1)v_x\\
&\quad= \det\binom{\eta}{\alpha_n}t^{-\beta_m}\cdots t^{-\beta_1}t^{\eta-\alpha_n}E(-\alpha_{n-1})\cdots E(-\alpha_1)v_x\\
&\quad= \det\binom{\eta}{\alpha_n}\prod_{i=1}^{n-1}\det\binom{\eta-\sum_{j=i+1}^{n}\alpha_j}{\alpha_i}t^{-\beta_m}\cdots t^{-\beta_1}t^{\eta-\sum_{j=1}^{n}\alpha_j}v_x\\
&\quad\in B(0)M(H)\setminus\mathbb{C}v.
\end{aligned}$$

This implies that $\bar{u}\in B(0)M(H)\setminus\mathbb{C}v$.

(2) Since $u'\in B(0)M(H)\setminus M(H)$, there exists $h\in\mathbb{N}$ and $\beta\in G_+\setminus\mathbb{Z}\epsilon$ such that $u'\in\bar{B}_T(h,\beta)M(H)$ and $u'\notin B'_T(h,\beta)M(H)$. Thus we may write

$$u'=\sum_{x\in B_T(h,\beta)}xv_x+\sum_{y\in B'_T(h,\beta)}yv_y,$$

where $v_x,v_y\in M(H)$ and both sums are finite. Then there exists some $n_0\in\mathbb{N}$ such that all v_x,v_y appeared above lie in $\mathcal{U}\bigl(\sum_{i<n_0}\mathbb{C}t^{-i\epsilon}\oplus\mathbb{C}E(-i\epsilon)\bigr)v$.

Take $\gamma_1=\beta-n_0\epsilon$, so $\det\binom{\gamma_1}{\beta}\neq 0$. We consider $(E(\gamma_1)-\varphi(E(\gamma_1)))u'$. First we consider the term $(E(\gamma_1)-\varphi(E(\gamma_1)))xv_x$ for $x\in B_T(h,\beta)$. We may write

$$x=t^{-\beta_m}\cdots t^{-\beta_1}(t^{-\beta})^k,$$

where $m\in\mathbb{Z}_+,\ \beta_m\succeq\cdots\succeq\beta_1\succ\beta,\ k\in\mathbb{N}$ and $\bigl(k\beta+\sum_{i=1}^{m}\beta_i\bigr)[2]=h$. Then, by Lemma 3.1(2), we have

$$\begin{aligned}
&(E(\gamma_1)-\varphi(E(\gamma_1)))xv_x\\
&\quad=[E(\gamma_1),t^{-\beta_m}\cdots t^{-\beta_1}(t^{-\beta})^k]v_x\\
&\quad=\sum_{i=1}^{m}\det\binom{\gamma_1}{\beta_i}t^{-\beta_m}\cdots t^{-\beta_i+\beta-n_0\epsilon}\cdots t^{-\beta_1}(t^{-\beta})^kv_x\\
&\qquad\qquad+\det\binom{\gamma_1}{\beta}x't^{-n_0\epsilon}v_x,
\end{aligned}\tag{3-4}$$

where $x'=kt^{-\beta_m}\cdots t^{-\beta_1}(t^{-\beta})^{k-1}$. Set $M(T_{-n_0})=\mathcal{U}\bigl(\bigoplus_{k\neq n_0}t^{-k\epsilon}\bigr)\mathcal{U}(E_-)v$. Then we see that the first sum in (3-4) lies in $B_T(h-\beta[2])M(T_{-n_0})$. Thus we have

$$(E(\gamma_1)-\varphi(E(\gamma_1)))xv_x\equiv\det\binom{\gamma_1}{\beta}x't^{-n_0\epsilon}v_x\pmod{B_T(h-\beta[2])M(T_{-n_0})}.$$

Now we consider the term $(E(\gamma_1) - \varphi(E(\gamma_1)))yv_y$ for $y \in B'_T(h,\beta)$. We may write

$$y = t^{-\beta_m} \cdots t^{-\beta_1},$$

where $\beta_m \succeq \cdots \succeq \beta_1 \in G_+ \setminus \mathbb{Z}\epsilon$, $\sum_{i=1}^m \beta_i[2] = h' \le h$. It is clear that $\beta_i \succ \beta$ for all i when $h' = h$. By Lemma 3.1(2), we have

$$\begin{aligned}(E(\gamma_1) - \varphi(E(\gamma_1)))yv_y &= [E(\gamma_1), t^{-\beta_m} \cdots t^{-\beta_1}]v_y \\ &= \sum_{i=1}^m \det\binom{\gamma_1}{\beta_i} t^{-\beta_m} \cdots t^{-\beta_i+\beta-n_0\epsilon} \cdots t^{-\beta_1} v_y.\end{aligned}$$

If $h' < h$, it is obvious that

$$(E(\gamma_1) - \varphi(E(\gamma_1)))yv_y \in B_T(h' - \beta[2])M(H) \subseteq B'_T(h - \beta[2])M(H).$$

If $h' = h$, then we have $\beta_i - \beta \in G_+$ for all $1 \le i \le m$. If $\beta_i - \beta \in G_+ \setminus \mathbb{Z}\epsilon$, then it is clear that

$$t^{-\beta_m} \cdots t^{-\beta_i+\beta-n_0\epsilon} \cdots t^{-\beta_1} v_y \in B_T(h - \beta[2])M(T_{-n_0}).$$

If $\beta_i - \beta = n_i\epsilon \in \mathbb{N}\epsilon$, then

$$t^{-\beta_m} \cdots t^{-\beta_i+\beta-n_0\epsilon} \cdots t^{-\beta_1} v_y = t^{-\beta_m} \cdots t^{-\beta_{i+1}} t^{-\beta_{i-1}} \cdots t^{-\beta_1} t^{(n_i - n_0)\epsilon} v_y,$$

which also lies in $B_T(h - \beta[2])M(T_{-n_0})$. Thus we have

$$(E(\gamma_1) - \varphi(E(\gamma_1)))yv_y \in B'_T(h - \beta[2])M(H) + B_T(h - \beta[2])M(T_{-n_0}).$$

From this discussion, we see that

$$(E(\gamma_1) - \varphi(E(\gamma_1)))u' = \det\binom{\gamma_1}{\beta} \sum_{x \in B_T(h,\beta)} x' t^{-n_0\epsilon} v_x + u''$$

for some $u'' \in B'_T(h - \beta[2])M(H) + B_T(h - \beta[2])M(T_{-n_0})$. Note that

$$\sum_{x \in B_T(h,\beta)} x' t^{-n_0\epsilon} v_x \in B_T(h - \beta[2]) t^{-n_0\epsilon} M(H)$$

is linearly independent from u''. This, together with the facts that

$$\sum_{x \in B_T(h,\beta)} x v_x \neq 0 \quad \text{and} \quad \det\binom{\gamma_1}{\beta} \neq 0,$$

imply

$$\det\binom{\gamma_1}{\beta} \sum_{x \in B_T(h,\beta)} x' t^{-n_0\epsilon} v_x \neq 0.$$

In particular, we have $(E(\gamma_1)-\varphi(E(\gamma_1)))u' \notin \mathbb{C}v$. Clearly, we have

$$(E(\gamma_1)-\varphi(E(\gamma_1)))u' \in \bar{B}_T(h-\beta[2])M(H).$$

Then, repeating this process finitely many times, we can take some $\gamma_2, \dots, \gamma_s \in G_+ \setminus \mathbb{Z}\epsilon$, $s \in \mathbb{N}$ such that

$$(E(\gamma_s)-\varphi(E(\gamma_s)))\cdots(E(\gamma_1)-\varphi(E(\gamma_1)))u' \in \bar{B}_T(0)M(H)\setminus\mathbb{C}v = M(H)\setminus\mathbb{C}v.$$

This completes the proof. $\square$

Lemma 3.3. *For $u \in M(H)\setminus\mathbb{C}v$, there exist $r, s \in \mathbb{Z}_+$, $n_1, \dots, n_s, m_1, \dots, m_r \in \mathbb{N}$ and $A \in \mathbb{C}^\times$ such that*

$$\begin{aligned}(E(n_s\epsilon)-\varphi(E(n_s\epsilon)))\cdots(E(n_1\epsilon)-\varphi(E(n_1\epsilon)))\\ \cdot(t^{m_r\epsilon}-\varphi(t^{m_r\epsilon}))\cdots(t^{m_1\epsilon}-\varphi(t^{m_1\epsilon}))u = Ah(\epsilon)^{r+s}v.\end{aligned} \tag{3-5}$$

Proof. First, we may assume $u \notin M(T)$, and we write

$$u = \sum_{i=1}^{n} a_i f_i v_i,$$

where all $a_i \neq 0$, $v_i \in M(T)$ and f_i for $1 \le i \le n$ are monic monomials with variables from the set $\{E(-j\epsilon) \mid j \in \mathbb{N}\}$. Without loss of generality, we may assume that f_1 has the maximal degree, and write

$$f_1 = E(-\epsilon)^{m_1}\cdots E(-r\epsilon)^{m_r}, \quad m_i \in \mathbb{Z}_+,$$

where $m_1, \dots, m_r$ are not all zero. For any monomial g with variables from $\{E(-j\epsilon) \mid j \in \mathbb{N}\}$, note that $[t^{i\epsilon}, E(j\epsilon)] = \delta_{i+j,0}\, i h(\epsilon)$ for any $i, j \in \mathbb{Z}$. Then for any $w \in M(T)$, we have

$$(t^{i\epsilon}-\varphi(t^{i\epsilon}))gw = ih(\epsilon)\partial_i'(g)w,$$

where $\partial_i'(g)$ is the partial derivative of g with respect to $E(-i\epsilon)$. Then by induction on r it is easy to see that

$$(t^{\epsilon}-\varphi(t^{\epsilon}))^{m_1}\cdots(t^{r\epsilon}-\varphi(t^{r\epsilon}))^{m_r} f_i v_i = \delta_{1,i}\prod_{j=1}^{r} m_j!(jh(\epsilon))^{m_j} v_1,$$

where $\delta_{1,i}$ is the Kronecker delta function and $m_j!$ is the factorial of m_j. So we get

$$(t^{\epsilon}-\varphi(t^{\epsilon}))^{m_1}\cdots(t^{r\epsilon}-\varphi(t^{r\epsilon}))^{m_r} u = A_1 h(\epsilon)^{m_1+\cdots+m_r} v_1,$$

where $A_1 = a_1 \prod_{j=1}^{r} m_j! j^{m_j} \neq 0$. If $v_1 \in \mathbb{C}v$, the lemma is clear. Otherwise, $v_1 \in M(T)$ and $v_1 \notin \mathbb{C}v$, and we write

$$v_1 = \sum_{i=1}^{n} b_i g_i v \in M(T),$$

where all $b_i \neq 0$ and g_i for $1 \leq i \leq n$ are monic monomials with variables from the set $\{t^{-j\epsilon} \mid j \in \mathbb{N}\}$. Without loss of generality, we may assume that g_1 has the maximal degree, and write

$$g_1 = (t^{-\epsilon})^{n_1} \cdots (t^{-s\epsilon})^{n_s}, n_i \in \mathbb{Z}_+,$$

where $n_1, \ldots, n_s$ are not all zero. For any monomial g with variables from the set $\{t^{-j\epsilon} \mid j \in \mathbb{N}\}$, we have

$$(E(i\epsilon) - \varphi(E(i\epsilon)))gv = ih(\epsilon)\partial_i''(g)v,$$

where $\partial_i''(g)$ is the partial derivative of g with respect to $t^{-i\epsilon}$. Then by induction on s it is easy to see that

$$(E(\epsilon) - \varphi(E(\epsilon)))^{n_1} \cdots (E(s\epsilon) - \varphi(E(s\epsilon)))^{n_s} g_i v = \delta_{1,i} \prod_{j=1}^{s} n_j!(jh(\epsilon))^{n_j} v.$$

Thus we get

$$(E(\epsilon) - \varphi(E(\epsilon)))^{n_1} \cdots (E(s\epsilon) - \varphi(E(s\epsilon)))^{n_s} v_1 = A_2 h(\epsilon)^{n_1+\cdots+n_s} v,$$

where $A_2 = b_1 \prod_{j=1}^{s} n_j! j^{n_j} \neq 0$.

Now we take $A = A_1 A_2 \neq 0$, and obtain the identity (3-1). If $u \in M(T)$, by the same discussion, we obtain the lemma. Thus the proof is completed. □

Let $\mathrm{Wh}(V)$ denote the set of Whittaker vectors for any Whittaker module V. In what follows, we determine the set $\mathrm{Wh}(M)$.

Proposition 3.4. (1) *If $h(\epsilon)$, $f(\epsilon)$ act on M as 0, then* $\mathrm{Wh}(M) = M(H)$.

(2) *If $h(\epsilon)$ acts on M as 0 and $f(\epsilon)$ does not act as 0, then* $\mathrm{Wh}(M) = M(T)$.

(3) *If $h(\epsilon)$ does not act as 0, then* $\mathrm{Wh}(M) = \mathbb{C}v$.

Proof. From Lemma 3.2, we see that any element in $M \setminus M(H)$ is not a Whittaker vector, thus we have $\mathrm{Wh}(M) \subseteq M(H)$.

(1) Suppose $f(\epsilon) = h(\epsilon) = 0$ on M. For any nonzero element $u \in M(H)$, we prove that u is a Whittaker vector. Write

$$u = \sum_{i=1}^{n} f_i v,$$

where f_i are monomials with variables from $\{t^{-j\epsilon}, E(-j\epsilon) \mid j \in \mathbb{N}\}$. Then for any $j \in \mathbb{N}$, we have

$$(E(j\epsilon) - \varphi(E(j\epsilon)))u = j \sum_{i=1}^{n} \big(f(\epsilon)\partial'_j(f_i)v + h(\epsilon)\partial''_j(f_i)v \big) = 0,$$

where ∂'_j and ∂''_j have the same meaning as in the proof of Lemma 3.3. Since $h(\epsilon) = 0$, we have $[t^{i\epsilon}, E(-j\epsilon)] = 0$ on M for any $i, j \in \mathbb{N}$, and $t^{i\epsilon}$ commutes with all f_k for $1 \le k \le n$ on M. This implies $t^{i\epsilon}u = \varphi(t^{i\epsilon})u$ for all $i \in \mathbb{N}$. Moreover, for all $\alpha \in G_+ \setminus \mathbb{Z}\epsilon$, note that $E(\alpha)u = \varphi(E(\alpha))u$, and $t^{\alpha}u = \varphi(t^{\alpha})u$ by Lemma 3.1(2). Thus $u \in \mathrm{Wh}(M)$ and we have $\mathrm{Wh}(M) = M(H)$.

(2) Suppose $h(\epsilon) = 0$, $f(\epsilon) \neq 0$ on M and $u \in M(H) \setminus M(T)$, then there exist some $m, p \in \mathbb{N}$ such that

$$u = \sum_{r=0}^{m} \sum c_{rkn}\, t^{-k_s\epsilon} \cdots t^{-k_1\epsilon} E(-n_l\epsilon) \cdots E(-n_1\epsilon)(E(-p\epsilon))^r v, \quad c_{rkn} \in \mathbb{C}^{\times},$$

where the second sum is finite and ranges over $s, l \in \mathbb{Z}_+$, $k_s \ge \cdots \ge k_1 \in \mathbb{N}$, $n_l \ge \cdots \ge n_1 \in \mathbb{N}$, and $n_1 > p$. Then we have

$$\begin{aligned}
&(E(p\epsilon) - \varphi(E(p\epsilon)))u \\
&\quad = \sum_{r=0}^{m} \sum c_{rkn} [E(p\epsilon), t^{-k_s\epsilon} \cdots t^{-k_1\epsilon} E(-n_l\epsilon) \cdots E(-n_1\epsilon)(E(-p\epsilon))^r] v \\
&\quad = pf(\epsilon) \sum_{r=1}^{m} \sum r c_{rkn}\, t^{-k_s\epsilon} \cdots t^{-k_1\epsilon} E(-n_l\epsilon) \cdots E(-n_1\epsilon)(E(-p\epsilon))^{r-1} v \neq 0,
\end{aligned}$$

which implies that u is not a Whittaker vector and $\mathrm{Wh}(M) \subseteq M(T)$. For any $u \in M(T)$, it is easy to check that u is a Whittaker vector as in the discussion in (1).

(3) Suppose $h(\epsilon) \neq 0$ on M and $u \in M(H) \setminus \mathbb{C}v$. Since $h(\epsilon) \neq 0$, Lemma 3.3 shows that u is not a Whittaker vector. Thus $\mathrm{Wh}(M) = \mathbb{C}v$. $\square$

4. Irreducible quotients of the universal Whittaker modules

In this section we study irreducibility for the universal Whittaker modules and we define a $\mathbb{Z}$-gradation on them, then we determine all $\mathbb{Z}$-graded irreducible quotients for the reducible universal Whittaker modules.

The Lie algebra L has a $\mathbb{Z}$-gradation $L = \bigoplus_{n\in\mathbb{Z}} L(n)$, where

$$L(-n) = \begin{cases} \bigoplus_{m\in\mathbb{Z}} (\mathbb{C}t^{m\epsilon+n\epsilon'} + \mathbb{C}E(m\epsilon+n\epsilon')) & \text{if } n \neq 0, \\ \bigoplus_{m\in\mathbb{Z}\setminus\{0\}} (\mathbb{C}t^{m\epsilon} + \mathbb{C}E(m\epsilon)) \oplus L_0 & \text{if } n = 0. \end{cases}$$

Set $B_H(0)=\{1\}$, and for $h\in\mathbb{N}$, recall that B is a basis of $\mathcal{U}(L_-)$ and we set

$$B_H(h)=\left\{X(-\alpha_n)\cdots X(-\alpha_1)\in B\;\middle|\; n\in\mathbb{N}, \alpha_i\in G_+\setminus\mathbb{Z}\epsilon, \sum_{i=1}^{n}\alpha_i[2]=h\right\}.$$

Let $M(h)=B_H(h)M(H)$ for $h\in\mathbb{Z}_+$. We have that $M=\bigoplus_{h\in\mathbb{Z}_+}M(h)$ and $L(n)M(h)\subseteq M(n+h)$. Hence M is $\mathbb{Z}$-graded. Note that $M(0)=M(H)=\mathcal{U}(H)v$ is a $\mathcal{U}(H)$-module.

Recall from (1-2) the definition of $h(\epsilon)$. The following theorem determines when the universal Whittaker module is irreducible.

Theorem 4.1. *The universal Whittaker module M is irreducible if and only if $h(\epsilon)\neq 0$.*

Proof. Suppose that $h(\epsilon)=0$. By Proposition 3.4(1) and (2), we see that M has a nonzero Whittaker vector $w\notin\mathbb{C}v$. It is easy to see that $\mathcal{U}(L)w$ is a proper submodule of M.

Conversely, suppose V is a nonzero submodule of M, and take $0\neq w\in V\setminus\mathbb{C}v$. Lemma 3.2 and Lemma 3.3 imply that $h(\epsilon)^k v\in\mathcal{U}(L)w\subseteq V$ for some $k\in\mathbb{N}$. Since $h(\epsilon)\neq 0$, we have $v\in V$. Thus $V=M$. □

Now we determine all $\mathbb{Z}$-graded irreducible quotients for the universal Whittaker modules on which $h(\epsilon)$ acts as 0 by constructing all maximal $\mathbb{Z}$-graded submodules. The main idea is that we first construct all maximal $\mathcal{U}(H)$-submodules of $M(H)$, then we build up maximal $\mathbb{Z}$-graded $\mathcal{U}(L)$-submodules of M. We divide the construction into two cases: $f(\epsilon)=h(\epsilon)=0$ on M, and $f(\epsilon)\neq 0$, $h(\epsilon)=0$ on M. Let $\mathcal{M}$ denote the set of all maximal $\mathbb{Z}$-graded $\mathcal{U}(L)$-submodules of M and $\mathcal{M}_H$ denote the set of all maximal $\mathcal{U}(H)$-submodules of $M(H)$.

First we consider the case where $f(\epsilon)=h(\epsilon)=0$ on M. For any pair

$$(a,b)=((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}})\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}},$$

let I_{ab} denote the ideal of $\mathcal{U}(H_-)$ generated by $\{t^{-i\epsilon}-a_i, E(-i\epsilon)-b_i \mid i\in\mathbb{N}\}$. Clearly I_{ab} is maximal.

Lemma 4.2. *The set $\{I_{ab}\mid (a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}\}$ exhausts all maximal ideals of $\mathcal{U}(H_-)$.*

Proof. Suppose K is a maximal ideal of $\mathcal{U}(H_-)$. Since $\mathcal{U}(H_-)$ is an integral domain, $\mathcal{U}(H_-)/K$ is a field extension of $\mathbb{C}$. Notice that any nontrivial field extension of $\mathbb{C}$ is of uncountable dimension over $\mathbb{C}$, but $\mathcal{U}(H_-)/K$ is of countable dimension by the PBW theorem, so $\mathcal{U}(H_-)/K\cong\mathbb{C}$. Then we have an algebra epimorphism $\pi:\mathcal{U}(H_-)\to\mathcal{U}(H_-)/K\cong\mathbb{C}$ with kernel K. Set $a_i=\pi(t^{-i\epsilon})$ and $b_i=\pi(E(-i\epsilon))$ for all $i\in\mathbb{N}$ and $(a,b)=((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}})$. Clearly, $t^{-i\epsilon}-a_i, E(-i\epsilon)-b_i\in\ker\pi=K$ for all $i\in\mathbb{N}$. That is, $I_{ab}\subseteq K$. Since I_{ab} is maximal, we have $I_{ab}=K$. □

From Lemma 4.2, we see that any maximal $\mathcal{U}(H_-)$-submodule of $M(H)$ is of the form $I_{ab}v$ for some $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$. Thus $I_{ab}v$ for $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$ are maximal $\mathcal{U}(H)$-submodules of $M(H)$. Furthermore, we claim that any maximal $\mathcal{U}(H)$-submodule of $M(H)$ is of the form $I_{ab}v$ for some $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$. Indeed, suppose that V is a maximal $\mathcal{U}(H)$-submodule of $M(H)$. Then V is a $\mathcal{U}(H_-)$-submodule of $M(H)$. Thus there exists some $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$ such that $V\subseteq I_{ab}v$. So $V=I_{ab}v$. That is, $\mathcal{M}_H=\{I_{ab}v\mid(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}\}$.

Let $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$. For $h\in\mathbb{Z}_+$, we define

$$M_{ab}(h)=\{u\in M(h)\mid X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u\in I_{ab}v\ \forall i_1,\dots,i_h\in\mathbb{Z}\}.$$

Set $M_{ab}=\sum_{h\in\mathbb{Z}_+}M_{ab}(h)$. We claim that M_{ab} is a proper submodule of M. Indeed, since $v\notin M_{ab}$, we see that $M_{ab}\subsetneqq M$. To prove that M_{ab} is an L-submodule of M, note that $\{X(\pm\epsilon'+i\epsilon)\mid i\in\mathbb{Z}\}$ generates L, thus we only need to prove the two inclusions

$$X(\epsilon'+i\epsilon)M_{ab}(h)\subseteq M_{ab}(h-1)\quad\text{and}\quad X(-\epsilon'+i\epsilon)M_{ab}(h)\subseteq M_{ab}(h+1)$$

for any $i\in\mathbb{Z}$ and $h\in\mathbb{Z}_+$. The first one is obvious. For the second one, let $u\in M_{ab}(h)$, and note that for $\alpha\in G_+\setminus\mathbb{Z}\epsilon$ we have $X(\alpha)M(H)=0$ by Lemma 3.1. Then for any $i,i_1,\dots,i_{h+1}\in\mathbb{Z}$, we have

$$\begin{aligned}
&X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_{h+1}\epsilon)X(-\epsilon'+i\epsilon)u\\
&\quad=X(-\epsilon'+i\epsilon)X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_{h+1}\epsilon)u\\
&\qquad\qquad+\big[X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_{h+1}\epsilon),X(-\epsilon'+i\epsilon)\big]u\\
&\quad\in X(-\epsilon'+i\epsilon)X(\epsilon'+i_1\epsilon)I_{ab}v\\
&\qquad\qquad+\sum_{k,j_1,\cdots,j_h\in\mathbb{Z}}\mathbb{C}X(k\epsilon)X(\epsilon'+j_1\epsilon)\cdots X(\epsilon'+j_h\epsilon)u\\
&\quad\subseteq\sum_{k\in\mathbb{Z}}X(k\epsilon)I_{ab}v\subseteq I_{ab}v,
\end{aligned}$$

where $X(0)=1$. Thus the second inclusion is obtained. Moreover, it is easy to see that M_{ab} is $\mathbb{Z}$-graded.

In what follows, we prove that the M_{ab} for $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$ exhaust all maximal $\mathbb{Z}$-graded submodules of M. The following result gives the characterization of all maximal $\mathbb{Z}$-graded $\mathcal{U}(L)$-submodules of M for the case $f(\epsilon)=h(\epsilon)=0$.

Proposition 4.3. $\mathcal{M}=\{M_{ab}\mid(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}\}$. *Moreover, all M_{ab} for $(a,b)\in\mathbb{C}^{\mathbb{N}}\times\mathbb{C}^{\mathbb{N}}$ are maximal L-submodules of M.*

Proof. First we prove that M_{ab} is a maximal L-submodule of M for any $(a, b) \in \mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}$. Note that for any $u \in M$, we may write

$$u = u_h + u' \tag{4-1}$$

for some $h \in \mathbb{Z}_+$, where $0 \neq u_h \in M(h)$ and $u' \in \sum_{h'<h} M(h')$. Then we have

$$X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u = X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u_h \in M(0) = M(H) \tag{4-2}$$

for any $i_1, \ldots, i_h \in \mathbb{Z}$. Now for any $(a, b) \in \mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}$ and $u \in M \setminus M_{ab}$, write $u = u_h + u'$ as in (4-1). We claim that there exists some $w \in (M(H) \setminus I_{ab}v) \cap (\mathcal{U}(L)u)$. In fact, if $h = 0$, then the claim holds for $w = u$. If $h > 0$, we may assume that $u_h \notin M_{ab}$; otherwise, if $u_h \in M_{ab}$, then $u' = u - u_h \in M \setminus M_{ab}$, thus we may consider u' instead of u. Then by the definition of M_{ab} and (4-2), we have

$$X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u = X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u_h \in M(H) \setminus I_{ab}v \tag{4-3}$$

for some $i_1, \ldots, i_h \in \mathbb{Z}$. Take $w = X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u$ for $i_1, \ldots, i_h \in \mathbb{Z}$ satisfying (4-3). Obviously, $w \in (M(H) \setminus I_{ab}v) \cap (\mathcal{U}(L)u)$.

Since $I_{ab}v$ is a maximal $\mathcal{U}(H)$-submodule of $M(H)$, we have

$$v \in M(H) = I_{ab}v + \mathcal{U}(H)w \subseteq M_{ab} + \mathcal{U}(L)u.$$

Since v generates M, it follows that $M = M_{ab} + \mathcal{U}(L)u$ for any $u \in M \setminus M_{ab}$. Thus M_{ab} is maximal. Since all M_{ab} are $\mathbb{Z}$-graded, we have $\mathcal{M} \supseteq \{M_{ab} \mid (a, b) \in \mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}\}$.

On the other hand, let $N \in \mathcal{M}$. Note that $N \cap M(H)$ is a proper $\mathcal{U}(H)$-submodule of $M(H)$. Then there exists $(a, b) \in \mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}$ such that $N \cap M(H) \subseteq I_{ab}v$.

Take any $u \in N$ and write $u = u_h + u'$ as in (4-1). If $h = 0$, we see that $u = u_0 \in N \cap M(H) \subseteq I_{ab}v \subseteq M_{ab}$. If $h > 0$, then we have

$$X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u = X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u_h \in N \cap M(H) \subseteq I_{ab}v.$$

It follows that $u_h \in M_{ab}$. Since N is $\mathbb{Z}$-graded, we have $u_h \in N$. So $u' = u - u_h \in N$. Now by induction on h we get that $u \in M_{ab}$. So $N \subseteq M_{ab}$ and therefore $N = M_{ab}$. This completes the proof. □

Now we consider the second case, when $h(\epsilon) = 0$ and $f(\epsilon) \neq 0$ on M. For any $\xi = (\xi_i)_{i\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$, let J_ξ denote the ideal of $\mathcal{U}(T_-)$ generated by $\{t^{-i\epsilon} - \xi_i \mid i \in \mathbb{N}\}$. Since $\mathcal{U}(T_-)$ is commutative, a similar argument as in Lemma 4.2 shows that $\{J_\xi \mid \xi \in \mathbb{C}^{\mathbb{N}}\}$ exhausts all maximal ideals of $\mathcal{U}(T_-)$, so $\{J_\xi v \mid \xi \in \mathbb{C}^{\mathbb{N}}\}$ exhausts all maximal $\mathcal{U}(T_-)$-submodules of $\mathcal{U}(T_-)v = M(T)$. Note that $M(H) = \mathcal{U}(E)M(T)$. We give all maximal $\mathcal{U}(H)$-submodules of $M(H)$ in the following proposition.

Proposition 4.4. $\qquad \mathcal{M}_H = \{\mathcal{U}(E)J_\xi v \mid \xi \in \mathbb{C}^{\mathbb{N}}\}.$

Proof. Since $h(\epsilon) = 0$, we have

$$T\mathcal{U}(E)J_\xi v = \mathcal{U}(E)TJ_\xi v \subseteq \mathcal{U}(E)J_\xi v.$$

Moreover, for any $k \in \mathbb{Z} \setminus \{0\}$, it is obvious that $E(k\epsilon)\mathcal{U}(E)J_\xi v \subseteq \mathcal{U}(E)J_\xi v$. So $\mathcal{U}(E)J_\xi v$ is a proper $\mathcal{U}(H)$-submodule of $M(H)$. For any $u \in M(H) \setminus \mathcal{U}(E)J_\xi v$, we may write

$$u = \sum_{i=1}^{n} a_i f_i v_i, \quad n \in \mathbb{N},$$

where $a_i \neq 0$, $v_i \in \mathcal{U}(T_-)v = M(T)$, f_i for $1 \leq i \leq n$ are monic monomials with variables from $\{E(-j\epsilon) \mid j \in \mathbb{N}\}$. We remark that, since $u \in M(H) \setminus \mathcal{U}(E)J_\xi v$, at least one $v_i \notin J_\xi v$. Set $J = \{i \in \{1, \ldots, n\} \mid v_i \notin J_\xi v\} \neq \varnothing$ and

$$u' = u - \sum_{i \notin J} a_i f_i v_i = \sum_{i \in J} a_i f_i v_i.$$

Since $f_i v_i \in \mathcal{U}(E)J_\xi v$ for $i \notin J$, and since $\mathcal{U}(E)J_\xi v$ is a $\mathcal{U}(H)$-module, it follows that $\mathcal{U}(H)u' \subseteq \mathcal{U}(H)u + \mathcal{U}(E)J_\xi v$. Without loss of generality, we may assume that $1 \in J$ and f_1 has the maximal degree among $\{f_i \mid i \in J\}$. Write

$$f_1 = E(-\epsilon)^{m_1} \cdots E(-r\epsilon)^{m_r}$$

for some $m_i \in \mathbb{Z}_+$. For any monomial g with variables from $\{E(-j\epsilon) \mid j \in \mathbb{N}\}$, note that $[E(i\epsilon), E(-j\epsilon)] = \delta_{ij} i f(\epsilon)$ and $[E(i\epsilon), t^{-j\epsilon}] = \delta_{ij} i h(\epsilon)$ for any $i, j \in \mathbb{N}$. Thus for any $w \in \mathcal{U}(T_-)v = M(T)$ we have

$$(E(i\epsilon) - \varphi(E(i\epsilon)))gw = if(\epsilon)\partial_i'(g)w,$$

where $\partial_i'(g)$ is the partial derivative of g with respect to $E(-i\epsilon)$. Then by induction on r, it is easy to see that, for $i \in J$,

$$(E(\epsilon) - \varphi(E(\epsilon)))^{m_1} \cdots (E(r\epsilon) - \varphi(E(r\epsilon)))^{m_r} f_i v_i = \delta_{1,i} \prod_{j=1}^{r} m_j!(jf(\epsilon))^{m_j} v_1.$$

So we get

$$(E(\epsilon) - \varphi(E(\epsilon)))^{m_1} \cdots (E(r\epsilon) - \varphi(E(r\epsilon)))^{m_r} u' = Af(\epsilon)^{m_1 + \cdots + m_r} v_1,$$

where $A = a_1 \prod_{j=1}^{r} m_j! j^{m_j} \neq 0$. Since $f(\epsilon) \neq 0$, it follows that $v_1 \in \mathcal{U}(H)u' \cap M(T)$.

Since $v_1 \notin J_\xi v$ and $J_\xi v$ is a maximal $\mathcal{U}(T_-)$-submodule of $M(T)$, we have

$$v \in M(T) = \mathcal{U}(T_-)v_1 + J_\xi v \subseteq \mathcal{U}(H)u' + \mathcal{U}(E)J_\xi v \subseteq \mathcal{U}(H)u + \mathcal{U}(E)J_\xi v.$$

This implies that $M(H) = \mathcal{U}(H)u + \mathcal{U}(E)J_\xi v$ for any $u \in M(H) \setminus \mathcal{U}(E)J_\xi v$, and thus $\mathcal{U}(E)J_\xi v$ is a maximal $\mathcal{U}(H)$-submodule of $M(H)$. That is,

$$\mathcal{M}_H \supseteq \{\mathcal{U}(E)J_\xi v \mid \xi \in \mathbb{C}^{\mathbb{N}}\}.$$

On the other hand, we note that any $\mathcal{U}(H)$-module $W \in \mathcal{M}_H$ is also a $\mathcal{U}(T_-)$-submodule of $M(H)$. Thus $W \cap M(T)$ is a proper $\mathcal{U}(T_-)$-submodule of $M(T)$. It follows that $W \cap M(T) \subseteq J_\xi v$ for some $\xi \in \mathbb{C}^{\mathbb{N}}$.

For any nonzero element $u \in W \subseteq M(H)$, we write

$$u = \sum_{i=1}^{k} g_i v_i, \quad k \in \mathbb{N},$$

where $v_i \in M(T)$ and g_i for $1 \le i \le k$ are monomials with variables from $\{E(-j\epsilon) \mid j \in \mathbb{N}\}$. Without loss of generality, we may assume that g_1 has the maximal degree. If $\deg g_1 = 0$, we have $u \in M(T) \cap W \subseteq J_\xi v \subseteq \mathcal{U}(E) J_\xi v$ for some $\xi \in \mathbb{C}^{\mathbb{N}}$. If $\deg g_1 > 0$, write

$$g_1 = a_1 E(-\epsilon)^{m_1} \cdots E(-r\epsilon)^{m_r}$$

for $a_1 \in \mathbb{C}^\times$, $m_i \in \mathbb{Z}_+$, $i = 1, \dots, r$. Then by induction on r it is easy to see that

$$(E(\epsilon) - \varphi(E(\epsilon)))^{m_1} \cdots (E(r\epsilon) - \varphi(E(r\epsilon)))^{m_r} g_i v_i = \delta_{1,i} a_1 \prod_{j=1}^{r} m_j! (j f(\epsilon))^{m_j} v_1.$$

So we get

$$(E(\epsilon) - \varphi(E(\epsilon)))^{m_1} \cdots (E(r\epsilon) - \varphi(E(r\epsilon)))^{m_r} u = A f(\epsilon)^{m_1 + \cdots + m_r} v_1,$$

where $A = a_1 \prod_{j=1}^{r} m_j! j^{m_j} \ne 0$. Since $f(\epsilon) \ne 0$, we have $v_1 \in \mathcal{U}(E) u \subseteq W$. Thus $v_1 \in W \cap M(T) \subseteq J_\xi v$ for some $\xi \in \mathbb{C}^{\mathbb{N}}$. It follows that $g_1 v_1 \in W \cap \mathcal{U}(E) J_\xi v$. So $u - g_1 v_1 \in W$. By iteration, we can get $u \in W \cap \mathcal{U}(E) J_\xi v \subseteq \mathcal{U}(E) J_\xi v$. Thus $W \subseteq \mathcal{U}(E) J_\xi v$. Then, by the maximality of W as a $\mathcal{U}(H)$-submodule of $M(H)$ and since $\mathcal{U}(E) J_\xi v$ is a proper $\mathcal{U}(H)$-submodule of $M(H)$, we have $W = \mathcal{U}(E) J_\xi v$. Thus $\mathcal{M}_H \subseteq \{\mathcal{U}(E) J_\xi v \mid \xi \in \mathbb{C}^{\mathbb{N}}\}$. This completes the proof. □

In what follows, we construct certain submodules of M, and prove that these submodules exhaust all the maximal $\mathbb{Z}$-graded submodules of M in the case that $h(\epsilon) = 0$, $f(\epsilon) \ne 0$ on M.

For $\xi \in \mathbb{C}^{\mathbb{N}}$, let $M_\xi(0) = \mathcal{U}(E) J_\xi v$. For any $h \in \mathbb{N}$, set

$$M_\xi(h) = \{u \in M(h) \mid X(\epsilon' + i_1 \epsilon) \cdots X(\epsilon' + i_h \epsilon) u \in \mathcal{U}(E) J_\xi v \ \forall i_1, \dots, i_h \in \mathbb{Z}\}.$$

Set $M_\xi = \bigoplus_{h \in \mathbb{Z}_+} M_\xi(h)$. By a similar argument as in the previous case where $f(\epsilon) = h(\epsilon) = 0$ on M, we can prove that M_ξ is a proper $\mathbb{Z}$-graded L-submodule of M. Then by a similar proof as in Proposition 4.3, we obtain the following result.

Proposition 4.5. *$\mathcal{M} = \{M_\xi \mid \xi \in \mathbb{C}^{\mathbb{N}}\}$. Moreover, for any $\xi \in \mathbb{C}^{\mathbb{N}}$, M_ξ is a maximal L-submodule of M.*

By Theorem 4.1, Proposition 4.3 and Proposition 4.5, we obtain the main result of the paper.

Theorem 4.6. *Suppose that V is a $\mathbb{Z}$-graded irreducible quotient of the universal Whittaker module $M_{\varphi,k_1,k_2,k_3,k_4}$ of type $(\varphi, k_1, k_2, k_3, k_4)$ over L. Then V is irreducible as an L-module. Furthermore*:

(1) *If $h(\epsilon) \neq 0$ on V, then $V = M_{\varphi,k_1,k_2,k_3,k_4}$.*

(2) *If $h(\epsilon) = f(\epsilon) = 0$ on V, then*

$$V = V_{ab} =: M_{\varphi,k_1,k_2,k_3,k_4}/M_{ab}$$

for some $(a, b) \in \mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}$. Moreover, $V_{ab} \cong V_{a'b'}$ if and only if $(a, b) = (a', b')$.

(3) *If $h(\epsilon) = 0$, $f(\epsilon) \neq 0$ on V, then*

$$V = V_\xi =: M_{\varphi,k_1,k_2,k_3,k_4}/M_\xi$$

for some $\xi \in \mathbb{C}^{\mathbb{N}}$. Moreover, $V_\xi \cong V_{\xi'}$ if and only if $\xi = \xi'$.

Remark. For the Virasoro-like algebra $\mathcal{V}$, the notion of Whittaker module of type (φ, k_3, k_4) was given in [Guo and Liu 2011b]. We can similarly define a $\mathbb{Z}$-gradation on the universal Whittaker module M_{φ,k_3,k_4}. For any $\alpha \in \mathbb{Z}^2 \setminus \{(0, 0)\}$, we define the action of t^α on M_{φ,k_3,k_4} trivially; then it is easy to see that K_1, K_2, hence $h(\epsilon)$, act as 0 on M_{φ,k_3,k_4}, and M_{φ,k_3,k_4} becomes a Whittaker module of type $(\varphi, 0, 0, k_3, k_4)$ for L. Therefore, Theorem 4.6 gives all $\mathbb{Z}$-graded irreducible quotients of the universal Whittaker modules for $\mathcal{V}$ and also proves that all $\mathbb{Z}$-graded irreducible quotients of the universal Whittaker modules are actually irreducible.

Furthermore, we prove that any $\mathbb{Z}$-graded irreducible quotient of a universal Whittaker module for L admits a unique Whittaker vector up to scalars. This result also applies to the Virasoro-like algebra.

Corollary 4.7. *Suppose that V is a $\mathbb{Z}$-graded irreducible quotient of a universal Whittaker module. Then* $\dim \mathrm{Wh}(V) = 1$.

Proof. Using Theorem 4.6, we prove this corollary in three cases.

Case 1: $h(\epsilon) \neq 0$ on V. Notice that $V = M$ and $\mathrm{Wh}(M) = \mathbb{C}v$. So $\dim \mathrm{Wh}(V) = 1$.

Case 2: $h(\epsilon) = f(\epsilon) = 0$ on V. We have $V = M/M_{ab}$ for some $(a, b) \in \mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}$. Let $u + M_{ab}$ be a Whittaker vector for some $u \in M \setminus M_{ab}$. Write $u = u_h + u'$, where $0 \neq u_h \in M(h)$ and $u' \in \sum_{i<h} M(i)$. If $u_h \in M_{ab}$, then $u' + M_{ab} = u + M_{ab}$. We consider u' instead. Hence we may assume that h is the smallest nonnegative integer such that $u_h \notin M_{ab}$. We claim that $h = 0$.

On the contrary, suppose that $h > 0$. Note that $u_h \notin M_{ab}$. By the definition of M_{ab} we have

$$X(\epsilon' + i_1\epsilon) \cdots X(\epsilon' + i_h\epsilon)u_h \in M(H) \setminus I_{ab}v \quad \text{for some } i_1, \ldots, i_h \in \mathbb{Z}.$$

So, by Proposition 2.2(2) and (4-3), we have

$$
\begin{aligned}
&(X(\epsilon'+i_1\epsilon)-\varphi(X(\epsilon'+i_1\epsilon)))\cdots(X(\epsilon'+i_h\epsilon)-\varphi(X(\epsilon'+i_h\epsilon)))u\\
&\quad = X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u\\
&\quad = X(\epsilon'+i_1\epsilon)\cdots X(\epsilon'+i_h\epsilon)u_h \notin M_{ab}.
\end{aligned}
$$

This contradicts that $u+M_{ab}$ is a Whittaker vector, so $h=0$. Therefore $u\in M(H)$. Since $M(H)=\mathbb{C}v+I_{ab}v$, we have $u\in\mathbb{C}v+M_{ab}$. So

$$\dim \mathrm{Wh}(V)=\dim \mathrm{Wh}(M/M_{ab})=1.$$

Case 3: $h(\epsilon)=0$, $f(\epsilon)\neq 0$ on V. The proof is similar to that of Case 2. □

References

[Arbarello et al. 1988] E. Arbarello, C. De Concini, V. G. Kac, and C. Procesi, "Moduli spaces of curves and representation theory", *Comm. Math. Phys.* **117**:1 (1988), 1–36. MR 89i:14019 Zbl 0647.17010

[Arnal and Pinczon 1974] D. Arnal and G. Pinczon, "On algebraically irreducible representations of the Lie algebra sl(2)", *J. Mathematical Phys.* **15** (1974), 350–359. MR 50 #9995 Zbl 0298.17003

[Batra and Mazorchuk 2011] P. Batra and V. Mazorchuk, "Blocks and modules for Whittaker pairs", *J. Pure Appl. Algebra* **215**:7 (2011), 1552–1568. MR 2012c:17013 Zbl 1228.17008

[Benkart and Ondrus 2009] G. Benkart and M. Ondrus, "Whittaker modules for generalized Weyl algebras", *Represent. Theory* **13** (2009), 141–164. MR 2010e:16039 Zbl 1251.16020

[Block 1981] R. E. Block, "The irreducible representations of the Lie algebra $\mathfrak{sl}(2)$ and of the Weyl algebra", *Adv. in Math.* **39**:1 (1981), 69–110. MR 83c:17010 Zbl 0454.17005

[Christodoulopoulou 2008] K. Christodoulopoulou, "Whittaker modules for Heisenberg algebras and imaginary Whittaker modules for affine Lie algebras", *J. Algebra* **320**:7 (2008), 2871–2890. MR 2009i:17009 Zbl 1221.17009

[Guo and Liu 2011a] X. Guo and X. Liu, "Whittaker modules over generalized Virasoro algebras", *Comm. Algebra* **39**:9 (2011), 3222–3231. MR 2012m:17038 Zbl 1255.17005

[Guo and Liu 2011b] X. Guo and X. Liu, "Whittaker modules over Virasoro-like algebra", *J. Math. Phys.* **52**:9 (2011), 093504. MR 2012j:17032 Zbl 1272.17028

[Kirkman et al. 1994] E. Kirkman, C. Procesi, and L. Small, "A q-analog for the Virasoro algebra", *Comm. Algebra* **22**:10 (1994), 3755–3774. MR 96b:17016 Zbl 0813.17009

[Kostant 1978] B. Kostant, "On Whittaker vectors and representation theory", *Invent. Math.* **48**:2 (1978), 101–184. MR 80b:22020 Zbl 0405.22013

[Lu and Zhao 2013] R. Lu and K. Zhao, "Generalized oscillator representations of the twisted Heisenberg–Virasoro algebra", preprint, 2013. arXiv 1308.6023v1

[Ondrus and Wiesner 2009] M. Ondrus and E. Wiesner, "Whittaker modules for the Virasoro algebra", *J. Algebra Appl.* **8**:3 (2009), 363–377. MR 2010f:17040 Zbl 1220.17019

[Xue et al. 2006] M. Xue, W. Lin, and S. Tan, "Central extension, derivations and automorphism group for Lie algebras arising from the 2-dimensional torus", *J. Lie Theory* **16**:1 (2006), 139–153. MR 2006i:17038 Zbl 1105.17006

[Zhang et al. 2010] X. Zhang, S. Tan, and H. Lian, "Whittaker modules for the Schrödinger–Witt algebra", *J. Math. Phys.* **51**:8 (2010), 083524. MR 2011m:81152

Received December 7, 2013.

SHAOBIN TAN
SCHOOL OF MATHEMATICAL SCIENCES
XIAMEN UNIVERSITY
XIAMEN 361005
CHINA
tans@xmu.edu.cn

QING WANG
SCHOOL OF MATHEMATICAL SCIENCES
XIAMEN UNIVERSITY
XIAMEN 361005
CHINA
qingwang@xmu.edu.cn

CHENGKANG XU
SCHOOL OF MATHEMATICAL SCIENCES
XIAMEN UNIVERSITY
XIAMEN 361005
CHINA
xiaoxiongxu@126.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.169

FRÉCHET QUANTUM SUPERGROUPS

AXEL DE GOURSAC

We introduce Fréchet quantum supergroups and their representations. By using the universal deformation formula of the abelian supergroups $\mathbb{R}^{m|n}$ we construct various classes of Fréchet quantum supergroups that are deformation of classical ones. For such quantum supergroups, we find an analog of Kac–Takesaki operators that are superunitary and satisfy the pentagonal relation.

1. Introduction

Noncommutative geometry [Connes 1994] is a vibrant field of mathematics whose essential principle lies in the duality between spaces and commutative algebras, so that the properties of spaces can be algebraically characterized. Then, a noncommutative algebra can be seen as corresponding to some "noncommutative space". This very rich way of thinking allows generalizing classical notions and theorems of usual geometry; and it is sometimes possible to prove new results for differential geometry in this more general noncommutative framework (for instance the classification of foliations of the torus [Rieffel 1981]). In this point of view, the noncommutative analogs of groups are quantum groups [Woronowicz 1987; Majid 1995].

As productive examples of noncommutative algebras, deformation quantization [Bayen et al. 1978a; 1978b] consists in introducing a deformed product on the space of smooth functions $\mathscr{C}^\infty(M)$ on a Poisson manifold M. This product depends on a deformation parameter θ so that $\theta = 0$ yields the usual commutative product on $\mathscr{C}^\infty(M)$. There is thus possibility of studying deformations with a formal deformation parameter (see in particular [Kontsevich 2003]) or a nonformal one ($\theta \in \mathbb{R}$).

In the case of a symplectic Lie group G, to any left-invariant formal deformation on $\mathscr{C}^\infty(G)$ is associated a Drinfeld twist [Drinfeld 1989] on the universal enveloping Hopf algebra $\mathscr{U}(\mathfrak{g})$ of the Lie algebra of G. Then, such a twist $F \in \mathscr{U}(\mathfrak{g}) \otimes \mathscr{U}(\mathfrak{g})[[\theta]]$

Work supported by the Belgian Interuniversity Attraction Pole (IAP) within the framework "Dynamics, Geometry and Statistical Physics" (DYGEST).
MSC2010: primary 16T05, 46E10; secondary 46L65, 58A50.
Keywords: Hopf algebra, quantum group, noncommutative supergeometry, Fréchet spaces, deformation quantization, multiplicative unitary.

deforms also any $\mathcal{U}(\mathfrak{g})$-module-algebra $\boldsymbol{A}$; this is called a universal deformation formula (UDF). The external symmetries of the UDF correspond thus to the twisted Hopf algebra on which the deformation of the algebras $\boldsymbol{A}$ are module-algebras (see [Giaquinto and Zhang 1998]).

For nonformal deformation quantization of Lie groups in the smooth setting, there are only few available examples. Rieffel [1989] built the deformation of abelian groups and the associated UDF. This was also recently extended to nonabelian Kählerian Lie groups [Bieliavsky and Gayral 2013; Bieliavsky et al. 2014].

Coming from another direction, supergeometry [Kostant 1977; Tuynman 2005] is a mathematical theory in which the objects are supermanifolds involving, besides the usual commuting coordinates, also anticommuting coordinates (Grassmann variables). The algebra of smooth functions of a supermanifold is then $\mathbb{Z}_2$-graded commutative. Supergeometry was applied to various domains of mathematics and in physics.

It is then natural to ask whether a noncommutative supergeometry corresponding to noncommutative geometry with $\mathbb{Z}_2$-grading does exist and possess nice properties. Noncommutative algebraic geometry developed fruitfully this graded approach with projective schemes [Artin and Zhang 1994]. A work in the direction of noncommutative Q-manifolds was also achieved in [Schwarz 1999]. In [de Goursac et al. 2012] we built some geometric tools such as noncommutative differential calculi, connections, for algebras with more general grading and interpreted as "noncommutative graded spaces". More recently, we constructed a nonformal deformation quantization of abelian Lie supergroups in [Bieliavsky et al. 2012]. It was initially motivated by physics since a renormalizable scalar quantum field theory on the Moyal space can be interpreted with the star product of the superspace $\mathbb{R}^{m|1}$ (see [de Goursac 2010; Bieliavsky et al. 2012]), as well as its associated gauge theory [de Goursac et al. 2007; 2012]. In this deformation, we had to introduce the notion of C*-superalgebra in order to implement the UDF associated to the Heisenberg supergroup. This notion has nice properties and should be the natural object of noncommutative supergeometry at the topological level.

The corresponding notion of quantum group in noncommutative supergeometry should be called "quantum supergroup". Some algebraic definitions of quantum supergroups were already given (see e.g., [Majid 1995]). In this paper, we introduce this notion in the context of topological Hopf superalgebras.

To this aim, we first look at the external symmetries of the UDF associated to the deformation of the abelian Lie supergroups. We indeed find a nonnuclear Fréchet–Hopf superalgebra H whose comodule algebras are deformed by the twist of the UDF and which corresponds to the external symmetries.

As external symmetries form quantum groups in general, properties of H lead us to a Fréchet definition of quantum supergroups and of their representations. This definition is actually a direct extension of Kostant's definition [1977] of supergroups without the supercommutativity condition.

We then study three examples of Fréchet quantum supergroups. First, the Clifford algebra that is topologically trivial as finite-dimensional. The second example uses the UDF of the abelian Lie supergroups to deform a class of solvable (nonnilpotent) Lie supergroups into Fréchet quantum supergroups. We introduce an analog of Kac–Takesaki operator for such quantum supergroups and show that it satisfies the pentagonal equation, but it is superunitary and not unitary. Finally, we construct Fréchet quantum supergroups with supertoral subgroups and exhibit their multiplicative superunitary operators.

Note that the definition and properties of C*-quantum supergroups are currently under study, but the Fréchet framework presented here — even though not nuclear — is much less constrained and could be useful in some cases where the C* notion is not available.

2. Nonformal deformation of superspaces

Supergeometric setting. We start with some recalls about the concrete approach of supergeometry developed in [DeWitt 1984; Tuynman 2005; Rogers 2007]. The essence of this approach consists of replacing the basis field $\mathbb{R}$ by a real supercommutative superalgebra $\mathcal{A}$ in all the geometric constructions.

Let $\mathcal{A} = \bigwedge V$, where V is a real infinite-dimensional vector space. Then, $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{A}_1$ is a $\mathbb{Z}_2$-graded commutative algebra with

$$ab = (-1)^{|a||b|} ba \quad \text{for all } a, b \in \mathcal{A},$$

where $|a| \in \mathbb{Z}_2$ denotes the degree of the homogeneous element a, and the expression is extended by linearity to inhomogeneous elements of $\mathcal{A}$. Moreover, it satisfies $\mathcal{A}/\mathcal{N}_{\mathcal{A}} \simeq \mathbb{R}$, where $\mathcal{N}_{\mathcal{A}}$ denotes the ideal of nilpotent elements of $\mathcal{A}$. We denote by $\mathbb{B} : \mathcal{A} \to \mathbb{R}$ the quotient map by $\mathcal{N}_{\mathcal{A}}$, and call it the body map. Actually, the explicit form of the algebra $\mathcal{A}$ is not important here; only its above properties play a role. Moreover, no topology is needed for $\mathcal{A}$ here, the Fréchet topology will appear at the level of the superfunctions on the involved supermanifolds.

Definition 2.1 (superspace). The superspace of (graded) dimension $m \mid n$ is defined as $\mathbb{R}^{m|n} := (\mathcal{A}_0)^m \times (\mathcal{A}_1)^n$. It involves m even (commuting) coordinates and n odd (anticommuting) coordinates in the canonical basis. The body map can be applied on each even coordinate and is also denoted by $\mathbb{B} : \mathbb{R}^{m|n} \to \mathbb{R}^m$.

Moreover, if m is even, this superspace can be endowed with the even symplectic structure associated to the $(m+n) \times (m+n)$ matrix given in the canonical basis by

$$\omega = \begin{pmatrix} \omega_0 & 0 \\ 0 & 2\mathbb{1} \end{pmatrix},$$

where $\omega_0 = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right)$ of size m.

The DeWitt topology of $\mathbb{R}^{m|n}$ can be constructed as follows. A subset U of $\mathbb{R}^{m|n}$ is called open if $\mathbb{B}U$ is an open subset of $\mathbb{R}^m$ and $U = \mathbb{B}^{-1}(\mathbb{B}U)$, namely U is saturated with nilpotent elements. It is of course not a Hausdorff topology.

The smooth functions on $\mathbb{R}^{m|0} = (\mathcal{A}_0)^m$ can be defined as associated to elements of $\mathscr{C}^\infty(\mathbb{R}^m)$.

Definition 2.2. To any smooth function $f \in \mathscr{C}^\infty(\mathbb{R}^m)$ one can associate the function $\tilde{f} : \mathbb{R}^{m|0} \to \mathcal{A}_0$ defined by

$$\tilde{f}(x) = \sum_{\alpha \in \mathbb{N}^m} \frac{1}{\alpha!} \partial^\alpha f(x_0) n^\alpha$$

for all $x \in \mathbb{R}^{m|0} = (\mathcal{A}_0)^m$ of the form $x = x_0 + n$, with $x_0 = \mathbb{B}(x) \in \mathbb{R}^m$ and $n \in \mathbb{R}^{m|0}$ a nilpotent element. (The usual conventions for the multiindex α apply.) Note that the sum over α is finite due to the nilpotency of n.

Definition 2.3 (smooth superfunctions). Let U be an open subset of $\mathbb{R}^{m|n}$. A map $f : U \to \mathcal{A}$ is said to be smooth on U, and written $f \in \mathscr{C}^\infty(U)$; if there exist unique functions $f_I \in \mathscr{C}^\infty(\mathbb{B}U)$ for all ordered subsets I of $\{1, \dots, n\}$, such that for all $(x, \xi) \in \mathbb{R}^{m|n}$ ($x \in \mathbb{R}^{m|0}$ and $\xi \in \mathbb{R}^{0|n}$),

$$f(x, \xi) = \sum_I \tilde{f}_I(x) \xi^I,$$

where ξ^I denotes the ordered product of the corresponding coefficients. This means that, if $I = \{i_1, \dots, i_k\}$ with $1 < i_1 < \cdots < i_k \leq n$, then $\xi^I := \prod_{i \in I} \xi^i = \xi^{i_1} \xi^{i_2} \cdots \xi^{i_k}$, and we take as a convention: $\xi^\varnothing = 1$. We extend this definition in the usual way to functions with values in a superspace.

For any two (ordered) subsets $I = \{i_1, \dots, i_l\}$ and $J = \{j_1, \dots, j_\ell\}$ of $\{1, \dots, n\}$, we define $\varepsilon(I, J)$ to be zero if I and J overlap; if $I \cap J = \varnothing$, we set $\varepsilon(I, J)$ to the parity of the list $(i_1, \dots, i_k, j_1, \dots, j_\ell)$, defined as -1 raised to the number of transpositions needed to put it in increasing order. This function satisfies

$$\begin{aligned} \varepsilon(I, J) &= (-1)^{|I||J|} \varepsilon(J, I), \\ \varepsilon(I, J \cup K) &= \varepsilon(I, J)\varepsilon(I, K) \quad \text{if } J \cap K = \varnothing. \end{aligned} \tag{2-1}$$

As a consequence, we have $\xi^I \cdot \xi^J = \varepsilon(I, J) \xi^{I \cup J}$. The smooth superfunctions then satisfy $\mathscr{C}^\infty(\mathbb{R}^{m|n}) \simeq \mathscr{C}^\infty(\mathbb{R}^m) \otimes \bigwedge \mathbb{R}^n$. We recall the Lebesgue–Berezin integration for superfunctions

$$\int_{\mathbb{R}^{m|n}} \mathrm{d}z \; f(z) = \int_{\mathbb{R}^m} \mathrm{d}x \; f_{\{1,\dots,n\}}(x).$$

With this definition of smooth superfunctions and the DeWitt topology, it is possible to define supermanifolds and Lie supergroups (see [DeWitt 1984; Rogers 2007; Tuynman 2005]).

Definition 2.4 (supermanifold, Lie supergroup). Let M be a topological space.

- A chart of M is a homeomorphism $\varphi : U \to W$, with U an open subset of M and W an open subset of $\mathbb{R}^{m|n}$, for $m, n \in \mathbb{N}$.
- An atlas of M is a collection of charts
$$\mathcal{S} = \{\varphi_i : U_i \to W_i,\ i \in I\},$$
where $\bigcup_{i \in I} U_i = M$ and $\varphi_i \circ \varphi_j^{-1} \in \mathscr{C}^\infty(\varphi_j(U_i \cap U_j), W_i)_0$ for all $i,\ j \in I$.
- If M is endowed with an atlas, we define its body as
$$\mathbb{B}M = \{y \in M : \text{for some } i,\ y \in U_i \text{ and } \varphi_i(y) \in \mathbb{B}W_i\},$$
and the body map $\mathbb{B} : M \to \mathbb{B}M$ on each subset U_i by $\mathbb{B}_{|U_i} = \varphi_i^{-1} \circ \mathbb{B} \circ \varphi_i$.
- M is called a supermanifold if it is endowed with an atlas such that $\mathbb{B}M$ is a real manifold.
- Let M be a supermanifold. A function f on M is called smooth, and we write $f \in \mathscr{C}^\infty(M)$, if $f \circ \varphi_i^{-1} \in \mathscr{C}^\infty(W_i)$ for every chart φ_i in some atlas for M.
- A Lie supergroup is a supermanifold G which has a group structure for which the multiplication is a smooth map. Consequently, the identity element of the supergroup has real coordinates (it lies in $\mathbb{B}G$), and the inverse map is smooth.

The algebra $\mathscr{C}^\infty(M)$ of smooth superfunctions on a supermanifold M carries a structure of $\mathbb{Z}_2$-graded Fréchet superalgebra for the pointwise product (see [Bieliavsky et al. 2012, Lemma 2.18]). A supermanifold M of dimension $m \mid n$ is called trivial if there exists a supermanifold M_0 of dimension $m \mid 0$ such that $M \simeq M_0 \times \mathbb{R}^{0|n}$. Note that $\mathbb{B}M_0 = \mathbb{B}M$ and that M_0 is totally determined by $\mathbb{B}M$. In particular, it can be showed (see [Tuynman 2005]) that every Lie supergroup has an underlying structure of trivial supermanifold.

Note that the superspace $\mathbb{R}^{m|n}$ has a structure of abelian supergroup. Its law can be expressed as
$$(x, \xi) \cdot (y, \eta) = (x + y, \xi + \eta) \quad \text{for all } (x, \xi), (y, \eta) \in \mathbb{R}^{m|n}.$$

The star product. The construction of the deformation quantization of the symplectic superspace $\mathbb{R}^{m|n}$ (see Definition 2.1) has been performed in [Bieliavsky et al. 2012] if m is an even integer. Let us recall here the corresponding $\mathbb{R}^{m|n}$-invariant star product. Its expression is given by the von Neumann formula extended to the graded setting: for $x \in \mathbb{R}^{m|0}$, $\xi \in \mathbb{R}^{0|n}$ (we write $(x,\xi) \in \mathbb{R}^{m|n}$),

$$
\begin{aligned}
(2\text{-}2)\quad & (f_1 \star f_2)(x,\xi) \\
&= \kappa \int dx_1\, d\xi_1\, dx_2\, d\xi_2\; f_1(x_1,\xi_1) f_2(x_2,\xi_2) \\
&\quad \times \exp\Big(\frac{-2i}{\theta}(\omega_0(x_1,x_2)+\omega_0(x_2,x)+\omega_0(x,x_1)+2\xi_1\xi_2+2\xi_2\xi+2\xi\xi_1)\Big),
\end{aligned}
$$

where $\kappa = (-1)^{n(n+1)/2}(i\theta)^n/(4^n(\pi\theta)^m)$ is a normalization factor while θ is the deformation parameter.

This product is defined on smooth superfunctions with compact support (i.e., its body support is compact), but it is possible to extend it to a larger algebra by using the method of oscillatory integrals. Let us introduce the space

$$\mathcal{B}(\mathbb{R}^{m|n}) = \mathcal{B}(\mathbb{R}^m) \otimes \bigwedge \mathbb{R}^n$$

of complex-valued bounded smooth superfunctions with every derivative bounded. It is a generalization of the space $\mathcal{B}(\mathbb{R}^m)$ of Schwartz to the graded setting. Endowed with the seminorms

$$|f|_\alpha = \sup_{x\in\mathbb{R}^m} \Big\{ \sum_I |D_x^\alpha f_I(x)| \Big\} \tag{2-3}$$

and the pointwise product, this space is a Fréchet superalgebra. See for example [Inoue and Maeda 1991; 2003] for close examples of Fréchet superalgebras and related analysis.

The oscillatory integrals give a meaning to expressions like[1]

$$\int dx_i\, d\xi_i\; e^{i\omega_0(x_1,x_2)} f(x_1,\xi_1,x_2,\xi_2)$$

for a (nonintegrable) function $f \in \mathcal{B}(\mathbb{R}^{2m|2n})$. Let us define the operator O by

$$(O\cdot f)(x_1,\xi_1,x_2,\xi_2) = (1-\Delta_{(x_1,x_2)})\Big(\frac{1}{1+x_1^2+x_2^2} f(x_1,\xi_1,x_2,\xi_2)\Big),$$

for a smooth superfunction f with compact support and where $\Delta_{(x_1,x_2)}$ denotes the Laplacian with respect to the variables $(x_1,x_2) \in \mathbb{R}^m \times \mathbb{R}^m$. An integration by parts shows that

$$
\begin{aligned}
(2\text{-}4)\quad & \int dx_i\, d\xi_i\; e^{i\omega_0(x_1,x_2)} f(x_1,\xi_1,x_2,\xi_2) \\
&\qquad = \int dx_i\, d\xi_i\; e^{i\omega_0(x_1,x_2)} (O^k\cdot f)(x_1,\xi_1,x_2,\xi_2),
\end{aligned}
$$

[1] We adopt the notation $dx_i\, d\xi_i := dx_1\, d\xi_1\, dx_2\, d\xi_2 \cdots$.

for any $k \in \mathbb{N}$. Moreover, there exist (bounded) functions $b^\alpha \in \mathcal{B}(\mathbb{R}^{2m})$ such that

$$(O^k \cdot f)(x_1, \xi_1, x_2, \xi_2) = \frac{1}{(1 + x_1^2 + x_2^2)^k} \sum_{\substack{\alpha \in \mathbb{N}^{2m} \\ |\alpha| \leq 2k}} b^\alpha(x_1, x_2) D^\alpha f(x_1, \xi_1, x_2, \xi_2). \tag{2-5}$$

As a consequence, for any $f \in \mathcal{B}(\mathbb{R}^{2m|2n})$, there exists an integer k such that $(O^k \cdot f) \in L^1(\mathbb{R}^{2m|2n})$. Thus, the oscillatory integral of f is given by the RHS member of (2-4). With this notion, the formula (2-2) defines an associative product on $\mathcal{B}(\mathbb{R}^{m|n})$.

Universal deformation formula. In this subsection, we consider an action of the supergroup $\mathbb{R}^{m|n}$ on a Fréchet algebra $(\boldsymbol{A}, |\cdot|_j)$,

$$\rho : \mathbb{R}^{m|n} \times (\boldsymbol{A} \otimes \mathcal{A}) \to (\boldsymbol{A} \otimes \mathcal{A}),$$

satisfying the conditions:

- $\rho_0 = \mathrm{id}$; for all $z_1, z_2 \in \mathbb{R}^{m|n}$, $\rho_{z_1+z_2} = \rho_{z_1}\rho_{z_2}$.
- For all $z \in \mathbb{R}^{m|n}$, $\rho_z : (\boldsymbol{A} \otimes \mathcal{A}) \to (\boldsymbol{A} \otimes \mathcal{A})$ is an $\mathcal{A}$-linear automorphism of algebras.
- By writing $z = (x, \xi) \in \mathbb{R}^{m|n}$, we can expand the action as $\rho_{(x,\xi)}(a) = \sum_I \rho_x(a)_I \xi^I$; for all $a \in \boldsymbol{A}$ and all I, $x \mapsto \rho_x(a)_I$ is $\boldsymbol{A}$-valued and continuous.
- There exists a constant $C > 0$ such that, for all $a \in \boldsymbol{A}$ and all I, j, there exists k such that
$$|\rho_x(a)_I|_j \leq C|a|_k \quad \text{for all } x \in \mathbb{B}M.$$

We notice that the star product (2-2) can be trivially extended to $\boldsymbol{A}$-valued superfunctions $\mathcal{B}_{\boldsymbol{A}}(\mathbb{R}^{m|n})$ that are bounded with every derivative bounded. Note that this space is also Fréchet for the seminorms $|f|_{j,\alpha} = \sup_{x \in \mathbb{R}^m}\{\sum_I |D^\alpha f_I(x)|_j\}$.

With the action ρ, we can deform the product of $\boldsymbol{A}$ by this extended star product.

Definition 2.5 (smooth vectors). The set of smooth vectors of $\boldsymbol{A}$ for the action ρ is defined as

$$\boldsymbol{A}^\infty = \{a \in \boldsymbol{A},\ \rho^a := z \mapsto \rho_z(a) \text{ is smooth on } \mathbb{R}^{m|n}\}.$$

Lemma 2.6 [Bieliavsky et al. 2012]. *The set of smooth vectors $\boldsymbol{A}^\infty$ is dense in $\boldsymbol{A}$. Moreover, for any $a \in \boldsymbol{A}^\infty$, the map ρ^a lies in $\mathcal{B}_{\boldsymbol{A}^\infty}(\mathbb{R}^{m|n})$.*

This means that we can now form the star product of ρ^a and ρ^b, for a and b smooth vectors.

Proposition 2.7 [Bieliavsky et al. 2012]. *The expression* $a \star_\rho b := (\rho^a \star \rho^b)(0)$, *for* $a, b \in \boldsymbol{A}^\infty$, *yields an associative product on* $\boldsymbol{A}^\infty$. *Endowed with the seminorms*

$$|a|_{j,\alpha} := |\rho^a|_{j,\alpha} = \sup_{x\in\mathbb{R}^m}\Big\{\sum_I |D^\alpha \rho_x(a)_I|_j\Big\},$$

$(\boldsymbol{A}^\infty, \star_\rho)$ *is a* (*noncommutative*) *Fréchet algebra.*

It turns out that the star product (2-2) can be rewritten as

$$(f_1 \star f_2)(x,\xi) \\ = \kappa \int \mathrm{d}x_1\, \mathrm{d}\xi_1\, \mathrm{d}x_2\, \mathrm{d}\xi_2\ f_1(x_1+x, \xi_1+\xi) f_2(x_2+x, \xi_2+\xi) e^{\frac{-2i}{\theta}(\omega_0(x_1,x_2)+2\xi_1\xi_2)}.$$

Then, we can write directly the twist $F : \boldsymbol{A}^\infty \otimes \boldsymbol{A}^\infty \to \boldsymbol{A}^\infty \otimes \boldsymbol{A}^\infty$ associated to the deformation

$$F = \kappa \int_{\mathbb{R}^{m|n}\times\mathbb{R}^{m|n}} \mathrm{d}z_1\, \mathrm{d}z_2\ e^{-\frac{2i}{\theta}\omega(z_1,z_2)} \rho_{z_1} \otimes \rho_{z_2}, \tag{2-6}$$

with $z = (x, \xi) \in \mathbb{R}^{m|n}$ and where ρ replaces the translation for a general action ρ on an algebra $\boldsymbol{A}$. Denoting by $\mu_0 : \boldsymbol{A} \otimes \boldsymbol{A} \to \boldsymbol{A}$ the undeformed product of $\boldsymbol{A}$, we can express the deformed product of Proposition 2.7 as $\mu_F := \mu_0 \circ F$, namely, $\mu_F(a \otimes b) = a \star_\rho b$. The expression (2-6) is also called the universal deformation formula of the supergroup $\mathbb{R}^{m|n}$, as it can deform a dense subspace $\boldsymbol{A}^\infty$ of every algebra $\boldsymbol{A}$ on which $\mathbb{R}^{m|n}$ acts (with some regularity assumed at the beginning of this section).

We can now show new properties regarding the twist of this deformation. Let us recall the definition of the projective tensor product [Grothendieck 1955] of two Fréchet algebras $(\boldsymbol{A}, |\cdot|_j)$ and $(\boldsymbol{B}, |\cdot|_k)$. It is the completion of the algebraic tensor product $\boldsymbol{A} \otimes \boldsymbol{B}$ for the family of seminorms: for all $c \in \boldsymbol{A} \otimes \boldsymbol{B}$,

$$\pi_{j,k}(c) = \inf\Big\{\sum_i |a_i|_j |b_i|_k,\ c = \sum_i a_i \otimes b_i\Big\}, \tag{2-7}$$

where the infimum is taken over all decompositions $c = \sum_i a_i \otimes b_i$. This completion is denoted by $\boldsymbol{A}\widehat{\otimes}_\pi \boldsymbol{B}$.

Proposition 2.8. *The twist* F *is a continuous endomorphism on the projective tensor product of* $\boldsymbol{A}^\infty$ *with itself*: $F \in \mathcal{L}(\boldsymbol{A}^\infty \widehat{\otimes}_\pi \boldsymbol{A}^\infty)$.

Proof. Let $c \in \boldsymbol{A}^\infty \otimes \boldsymbol{A}^\infty$. Then,

$$\pi_{j,\alpha;k,\beta}(F(c)) = \inf\Big\{\Big|\sum_i F(a_i \otimes b_i)\Big|_{j,\alpha,k,\beta}\Big\},$$

where c can be written as $\sum_i a_i \otimes b_i$, and the infimum is taken over all such decompositions. By using the definition of oscillatory integral (2-4), and defining the partial operators

$$
\begin{aligned}
(O_{z_1} \cdot f)(z_1, z_2) &= \frac{1}{1+x_1^2}(1-\Delta_{x_2}) f(z_1, z_2), \\
(O_{z_2} \cdot f)(z_1, z_2) &= \frac{1}{1+x_2^2}(1-\Delta_{x_1}) f(z_1, z_2),
\end{aligned}
\tag{2-8}
$$

with $z_i = (x_i, \xi_i) \in \mathbb{R}^{m|n}$, we obtain

$$
\begin{aligned}
&\pi_{j,\alpha;k,\beta}(F(c)) \\
&\quad = \inf \left| \kappa' \int \mathrm{d}z_1 \, \mathrm{d}z_2 \; e^{-\frac{2i}{\theta}\omega(z_1,z_2)} O_{z_1}^{k_1} O_{z_2}^{k_2} \sum_i \rho_{z_1}(a_i) \otimes \rho_{z_2}(b_i) \right|_{j,\alpha,k,\beta} \\
&\quad \le \inf |\kappa'| \sum_{i,I,J} \int \mathrm{d}x_1 \, \mathrm{d}x_2 \; \frac{1}{(1+x_1^2)^{k_1}(1+x_2^2)^{k_2}} \\
&\qquad\qquad \times \sum_{\gamma,\delta} |b_1^{\gamma}(x_1) b_2^{\delta}(x_2)| \, |D^{\gamma} \rho_{x_1}(a_i)_I|_{j,\alpha} |D^{\delta} \rho_{x_2}(b_i)_J|_{k,\beta}
\end{aligned}
$$

in the notation of (2-5), if I, J are summed over $\{1, \dots, n\}$ with some conditions, and for κ' a constant. By definition of the seminorm,

$$
|D^{\gamma} \rho_{x_1}(a_i)_I|_{j,\alpha} = \sup_{x_3 \in \mathbb{R}^m} \left\{ \sum_K |D_{x_3}^{\alpha} \rho_{x_3}(D_{x_1}^{\gamma} \rho_{x_1}(a_i)_I)_K|_j \right\}.
$$

Since ρ is a group action, we can deduce that

$$
\rho_{x_3}(D_{x_1}^{\gamma} \rho_{x_1}(a_i)_I)_K = (-1)^{|I||K|} \varepsilon(I, K) D_{x_1}^{\gamma} \rho_{x_1+x_3}(a_i)_{I \cup K}. \tag{2-9}
$$

We then choose sufficiently large numbers k_1 and k_2 such that there exists a constant $C > 0$ with

$$
\pi_{j,\alpha;k,\beta}(F(c)) \le C \inf \sum_{i,\gamma,\delta} |a_i|_{j,\alpha+\gamma} |b_i|_{k,\beta+\delta} = C \sum_{\gamma,\delta} |c|_{j,\alpha+\gamma,k,\beta+\delta},
$$

where the sum on multiindices $\gamma, \delta \in \mathbb{N}^m$ satisfies the constraint $|\gamma| \le 2k_1$ and $|\delta| \le 2k_2$. The last inequality shows that F is continuous on $\boldsymbol{A}^{\infty} \widehat{\otimes}_{\pi} \boldsymbol{A}^{\infty}$. □

Example 2.9. If we take $\boldsymbol{A} = \mathcal{B}(\mathbb{R}^{m|n})$ and $\rho_z(f)(z') = f(z+z')$, then the space of smooth vectors is $\boldsymbol{A}^{\infty} = \mathcal{B}(\mathbb{R}^{m|n})$ and the product μ_F corresponds to (2-2).

There are a lot of other examples, like the actions of $\mathbb{R}^{m|n}$ over a certain class of continuous superfunctions on the trivial supermanifolds on which $\mathbb{R}^{m|n}$ is acting (see [Bieliavsky et al. 2012]).

External symmetries of the deformation. To introduce the external symmetries of the deformation or of the twist F, we need the notion of topological Hopf algebra, endowed with a Fréchet topology.

Definition 2.10. A Fréchet–Hopf algebra is a Hopf algebra H endowed with a Fréchet topology, such that the algebraic operations - product, unit, coproduct, counit and antipode - are continuous maps for the Fréchet structure and for a given topological tensor product on H.

Given a Fréchet–Hopf algebra H with topological tensor product $\widehat{\otimes}_{HH}$, as well as a topological tensor product $\widehat{\otimes}_{AH}$ between H and a Fréchet algebra $\boldsymbol{A}$ that has itself another topological tensor product $\widehat{\otimes}_{\boldsymbol{AA}}$; we say that $\boldsymbol{A}$ is a comodule algebra of H if it is an algebraic comodule algebra of H, if the coaction can be continuously extended to

$$\boldsymbol{A} \to \boldsymbol{A} \widehat{\otimes}_{AH} H$$

and if the three topological tensor are compatible, i.e., if the flips involved in the axioms of a comodule algebra are continuous for the Fréchet structures (see Lemma 2.13 for an example).

In the context of superspaces, we can introduce the following Fréchet–Hopf algebra. Let $H := \mathcal{B}(\mathbb{R}^{m|n})$ with its Fréchet topology (2-3). We introduce a topological tensor product different from the projective one, denoted by τ, as follows. We define $\boldsymbol{A} \widehat{\otimes}_\tau H$ to be the completion of the algebraic tensor product for the family of seminorms of $\mathcal{B}_{\boldsymbol{A}}(\mathbb{R}^{m|n})$:

$$\tau_{j,\alpha}(f) = |f|_{j,\alpha} = \sup_{x \in \mathbb{R}^m} \Big\{ \sum_I |D^\alpha f_I(x)|_j \Big\}. \tag{2-10}$$

One can then see that $H \widehat{\otimes}_\tau H \simeq \mathcal{B}(\mathbb{R}^{m|n} \times \mathbb{R}^{m|n})$ and by definition, $\boldsymbol{A} \widehat{\otimes}_\tau H \simeq \mathcal{B}_{\boldsymbol{A}}(\mathbb{R}^{m|n})$. On H we consider the standard Hopf algebra structure, whose algebraic operations can be continuously extended for the tensor product τ:

- the product $\mu : H \widehat{\otimes}_\tau H \to H$ defined by $\mu(f_1 \otimes f_2)(z) = f_1(z) f_2(z)$,
- the unit $\mathbb{1} : \mathbb{C} \to H$ defined by $\mathbb{1}(\lambda)(z) = \lambda$,
- the coproduct $\Delta : H \to H \widehat{\otimes}_\tau H$ defined by $\Delta f(z_1, z_2) = f(z_1 z_2)$,
- the counit $\varepsilon : H \to \mathbb{C}$ defined by $\varepsilon(f) = f(0)$,
- the antipode $S : H \to H$ defined by $Sf(z) = f(-z)$,

where $f_i \in H$, $z_i \in \mathbb{R}^{m|n}$, $\lambda \in \mathbb{C}$. These operations satisfy the useful axioms of Hopf algebra, taking into account that the flip $\sigma_{12} : H \otimes H \to H \otimes H$ is defined by

$$\sigma_{12}(f_1 \otimes f_2) = (-1)^{|f_1||f_2|} f_2 \otimes f_1 \tag{2-11}$$

because of the grading. This means that for $f \in H \widehat{\otimes}_\tau H$, $\sigma_{12} f(z, z') = f(z', z)$.

Proposition 2.11. *$H = \mathcal{B}(\mathbb{R}^{m|n})$ is a $\mathbb{Z}_2$-graded supercommutative Fréchet–Hopf algebra for the topological tensor product τ.*

Proof. Due to the explicit expression of the coproduct

$$\Delta(f)(x_1,\xi_1;x_2,\xi_2)=\sum_{I,J}\varepsilon(I,J)f_{I\cup J}(x_1+x_2)\xi_1^I\xi_2^J$$

obtained by an expansion on the odd variables and by (2-1), we have, for all $f\in\mathcal{B}(\mathbb{R}^{m|n})$,

$$\tau_{\alpha,\beta}(\Delta(f))=\sup_{x_1,x_2\in\mathbb{R}^m}\left\{\sum_{I,J}|\varepsilon(I,J)D_{x_1}^{\alpha}D_{x_2}^{\beta}f_{I\cup J}(x_1+x_2)|\right\}\le 2^n|f|_{\alpha+\beta},$$

which shows the continuity of $\Delta : H\to H\widehat{\otimes}_\tau H$. The continuity of the other operations can be proved in the same way. The algebraic properties between operations are the same as in the nongraded setting except $(S\otimes S)\Delta=\sigma_{12}\Delta S$ involving the flip (2-11). It can be showed that

$$\sigma_{12}\Delta(f)(x_1,\xi_1;x_2,\xi_2)=\Delta(f)(x_2,\xi_2;x_1,\xi_1)=\Delta(f)(x_1,\xi_1;x_2,\xi_2),$$

because $\mathbb{R}^{m|n}$ is abelian. Then, we have

$$\begin{aligned}(S\otimes S)\Delta(f)(x_1,\xi_1;x_2,\xi_2)&=f(-x_1-x_2,-\xi_1-\xi_2)\\&=\sigma_{12}\Delta S(f)(x_1,\xi_1;x_2,\xi_2).\end{aligned}$$ □

Remark 2.12. Note that $\mathscr{C}^\infty(\mathbb{R}^{m|n})$ is also a Fréchet–Hopf algebra (see [Bonneau and Sternheimer 2005] in the nongraded setting). Since it is nuclear contrary to $\mathcal{B}(\mathbb{R}^{m|n})$, this structure is independent of the choice of the topological tensor product. In this paper, we consider $\mathcal{B}(\mathbb{R}^{m|n})$ for the deformation quantization since $\mathscr{C}^\infty(\mathbb{R}^{m|n})$ is too large for the star product to be defined on it (see Section 2). $\mathcal{B}(\mathbb{R}^{m|n})$ is not nuclear but we will see that the tensor products τ and π (needed for representations) are compatible in a certain sense. We could of course have considered a smaller nuclear subalgebra like the Schwartz algebra $\mathscr{S}(\mathbb{R}^{m|n})$ - see [Bieliavsky et al. 2010] in the nongraded setting - but then the coproduct does not stabilize this algebra and we have to see it as valued in (the tensor product of) the multiplier algebra of $\mathscr{S}(\mathbb{R}^{m|n})$. See also [Voigt 2008] for another framework (bornological vector spaces) adapted to quantum groups.

Let us present the dual version of the universal deformation formula studied in Section 2, which will lead to the external symmetries. As before, we consider the Fréchet–Hopf algebra $H=\mathcal{B}(\mathbb{R}^{m|n})$ associated to the supergroup $\mathbb{R}^{m|n}$. The reformulation of the action ρ in this context will be done by the notion of H-comodule algebras (see Definition 2.10). To this aim, we need the following intermediate result.

Lemma 2.13. *The topological tensor product τ is compatible with the projective one π, in the sense that the flip*

$$\sigma_{23} : (\boldsymbol{A} \widehat{\otimes}_\tau H) \widehat{\otimes}_\pi (\boldsymbol{A} \widehat{\otimes}_\tau H) \to (\boldsymbol{A} \widehat{\otimes}_\pi \boldsymbol{A}) \widehat{\otimes}_\tau (H \widehat{\otimes}_\tau H),$$

defined by $\sigma_{23}(a_1 \otimes f_1 \otimes a_2 \otimes f_2) = a_1 \otimes a_2 \otimes f_1 \otimes f_2$, is continuous, for any Fréchet algebra $(\boldsymbol{A}, |\cdot|_j)$.

Proof. For $a_i, b_i \in \boldsymbol{A}$ and $f_i, g_i \in H$, due to the expressions (2-7) and (2-10) of the seminorms of π and τ, one has

$$\begin{aligned}\pi_{j,\alpha;k,\beta}\Bigl(\sum_i a_i \otimes f_i \otimes b_i \otimes g_i\Bigr) \\ &= \inf \sum_i \tau_{j,\alpha}(a_i \otimes f_i)\tau_{k,\beta}(b_i \otimes g_i) \\ &= \inf \sum_i \sup_{x,y\in\mathbb{R}^m} \sum_{I,J} |a_i|_j |D^\alpha f_{i,I}(x)| |b_i|_k |D^\beta g_{i,J}(y)|.\end{aligned}$$

Moreover,

$$\begin{aligned}\tau_{j,k;\alpha,\beta}\Bigl(\sigma_{23}\Bigl(\sum_i a_i \otimes f_i \otimes b_i \otimes g_i\Bigr)\Bigr) \\ &= \sup_{x,y\in\mathbb{R}^m} \sum_{I,J} \pi_{j,k}\Bigl(\sum_i (a_i \otimes b_i) D^\alpha f_{i,I}(x) D^\beta g_{i,J}(y)\Bigr) \\ &= \sup_{x,y\in\mathbb{R}^m} \sum_{I,J} \inf \sum_i |a_i|_j |b_i|_k |D^\alpha f_{i,I}(x)| |D^\beta g_{i,J}(y)|.\end{aligned}$$

Since for all $x, y \in \mathbb{R}^m$,

$$\begin{aligned}\inf \sum_i |a_i|_j |b_i|_k |D^\alpha f_{i,I}(x)| |D^\beta g_{i,J}(y)| \\ &\le \inf \sum_i \sup_{x,y\in\mathbb{R}^m} |a_i|_j |b_i|_k |D^\alpha f_{i,I}(x)| |D^\beta g_{i,J}(y)|,\end{aligned}$$

there exists a constant $1 \le C \le 2^{n+1}$ such that

$$\tau_{j,k;\alpha,\beta}\Bigl(\sigma_{23}\Bigl(\sum_i a_i \otimes f_i \otimes b_i \otimes g_i\Bigr)\Bigr) \le C\, \pi_{j,\alpha;k,\beta}\Bigl(\sum_i a_i \otimes f_i \otimes b_i \otimes g_i\Bigr),$$

which proves the continuity of σ_{23}. □

Proposition 2.14. *The action ρ of $\mathbb{R}^{m|n}$ on a Fréchet algebra $(\boldsymbol{A}, \mu_0)$ with axioms of* Section 2, *generates the continuous coaction $\chi : \boldsymbol{A}^\infty \to \boldsymbol{A}^\infty \widehat{\otimes}_\tau H$ defined by*

$$\chi(a)(z) := \rho_z(a) \quad \text{for all } a \in \boldsymbol{A}^\infty \text{ and } z \in \mathbb{R}^{m|n}.$$

Then $(\boldsymbol{A}^\infty, \mu_0)$ is an H-comodule algebra, with $\widehat{\otimes}_{\boldsymbol{A}H} := \widehat{\otimes}_\tau$ and $\widehat{\otimes}_{\boldsymbol{A}\boldsymbol{A}} := \widehat{\otimes}_\pi$.

Proof. Since ρ is a group action and $\rho_z : (\boldsymbol{A}\otimes\mathscr{A}) \to (\boldsymbol{A}\otimes\mathscr{A})$ is an algebra morphism for all $z \in \mathbb{R}^{m|n}$, we deduce that χ satisfies the axioms of a coaction:

$$(\mathrm{id}\otimes\Delta)\chi = (\chi\otimes\mathrm{id})\chi, \quad (\mathrm{id}\otimes\varepsilon)\chi = \mathrm{id}.$$

Thus, $\boldsymbol{A}^\infty$ is an algebraic H-comodule algebra

$$(\mu_0\otimes\mu)\sigma_{23}(\chi\otimes\chi) = \chi\mu_0, \tag{2-12}$$

where $\mu_0 : \boldsymbol{A}^\infty\widehat{\otimes}_\pi\boldsymbol{A}^\infty \to \boldsymbol{A}^\infty$ corresponds to the undeformed product of $\boldsymbol{A}$ and σ_{23} is the flip of Lemma 2.13 for the algebra $\boldsymbol{A}^\infty$. Let a be in $\boldsymbol{A}^\infty$; we then have $\chi(a) \in \boldsymbol{A}^\infty\widehat{\otimes}_\tau\mathscr{B}(\mathbb{R}^{m|n}) \simeq \mathscr{B}_{\boldsymbol{A}^\infty}(\mathbb{R}^{m|n})$, so

$$\tau_{j,\alpha;\beta}(\chi(a)) = \sup_{y\in\mathbb{R}^m}\sum_I |D^\beta\rho_y(a)_I|_{j,\alpha} = \sup_{y,y'}\sum_{I,J}|D^\alpha_{y'}\rho_{y'}(D^\beta_y\rho_y(a)_I)_J|_j.$$

By using (2-9), we obtain

$$\tau_{j,\alpha;\beta}(\chi(a)) = \sup_{y,y'}\sum_{I,J}|\varepsilon(I,J)D^\alpha_{y'}D^\beta_y\rho_{y+y'}(a)_{I\cup J}|_j,$$

which shows that there exists $C > 0$ such that $\tau_{j,\alpha;\beta}(\chi(a)) \le |a|_{j,\alpha+\beta}$, i.e., χ is continuous. Note that the flip σ_{23} is continuous due to the compatibility of the topological tensor products τ and π showed in Lemma 2.13. Indeed, all the maps involved in (2-12) have to be continuous in order for $\boldsymbol{A}^\infty$ to be a comodule algebra of the Fréchet–Hopf algebra H. □

Now, the algebra $(\boldsymbol{A}^\infty, \mu_0)$ can be deformed by the twist F defined in (2-6) in such a way $(\boldsymbol{A}^\infty, \mu_F = \mu_0 F)$ is a Fréchet algebra. The universal deformation formula constructed before provides therefore a deformation of the category of the H-comodule algebras. Of course, once deformed, there is a priori no reason for $(\boldsymbol{A}^\infty, \mu_F)$ to be again an H-comodule algebra.

Definition 2.15. Given a twist F which deforms the category of comodule algebras $(\boldsymbol{A}, \mu_0)$ of a given Fréchet–Hopf algebra H, we call external symmetries of the twist the Fréchet–Hopf algebras H_F for which any deformed algebra $(\boldsymbol{A}, \mu_F)$ is an H_F-comodule algebra.

In the nongraded setting and formally in the deformation parameter, there is a way to obtain the external symmetries H_F from H and the twist F [Drinfeld 1989; Giaquinto and Zhang 1998]. This has been extended to nonformal deformations of a large class of solvable Lie groups in [Bieliavsky et al. 2010]. Let us describe this process for such a Lie group G and where H denotes (a closed subclass of) $\mathscr{C}^\infty(G)$ with its Hopf algebra structure. If the nonformal twist of G on algebras $\boldsymbol{A}$, where G acts by ρ, has the form

$$F = \int_{G\times G} \mathrm{d}x_1\,\mathrm{d}x_2\; e^{-\frac{2i}{\theta}S(x_1,x_2)}A(x_1,x_2)\rho_{x_1}\otimes\rho_{x_2},$$

where S and A are the phase and amplitude of the deformation quantization, then we can consider the left L and right R actions of G on itself to obtain maps $\mathscr{C}^\infty(G)\widehat{\otimes}\mathscr{C}^\infty(G) \to \mathscr{C}^\infty(G)\widehat{\otimes}\mathscr{C}^\infty(G)$:

$$
\begin{aligned}
F_L &= \int_{G\times G} \mathrm{d}x_1\, \mathrm{d}x_2\, e^{-\frac{2i}{\theta}S(x_1,x_2)} A(x_1,x_2) R^*_{x_1} \otimes R^*_{x_2},\\
F_R &= \int_{G\times G} \mathrm{d}x_1\, \mathrm{d}x_2\, e^{-\frac{2i}{\theta}S(x_1,x_2)} A(x_1,x_2) L^*_{(x_1)^{-1}} \otimes L_{(x_2)^{-1}}.
\end{aligned}
\tag{2-13}
$$

To obtain the external symmetries of F, the product μ of H has to be twisted [Bieliavsky et al. 2010] into $\mu_{QG} := F_L \circ F_R \circ \mu$, which is compatible with the undeformed coproduct Δ. Thus, (a subclass of) $\mathscr{C}^\infty(G)$ with μ_{QG} and Δ is the topological Hopf algebra corresponding to the external symmetries.

In the graded setting, the construction has not been provided in general. However, for the supergroup $\mathbb{R}^{m|n}$, we can see that the external symmetries of the deformation of $\mathbb{R}^{m|n}$ are $H = (\mathscr{B}(\mathbb{R}^{m|n}), \mu, \Delta)$ without twisting its product.

Proposition 2.16. *$(\boldsymbol{A}^\infty, \mu_F)$ is an H-comodule algebra.*

Proof. The only remaining condition to check is $(\mu_F \otimes \mu)\sigma_{23}(\chi \otimes \chi) = \chi\mu_F$. For $a, b \in \boldsymbol{A}^\infty$ and $z \in \mathbb{R}^{m|n}$,

$$
\begin{aligned}
\chi\mu_F(a\otimes b)(z) &= \kappa \int \mathrm{d}z_1\, \mathrm{d}z_2\, e^{-\frac{2i}{\theta}\omega(z_1,z_2)} \rho_z(\rho_{z_1}(a)\rho_{z_2}(b))\\
(\mu_F \otimes \mu)\sigma_{23}(\chi\otimes\chi)(a\otimes b)(z) &= \kappa \int \mathrm{d}z_1\, \mathrm{d}z_2\, e^{-\frac{2i}{\theta}\omega(z_1,z_2)} \rho_{z_1}\rho_z(a)\rho_{z_2}\rho_z(b).
\end{aligned}
$$

Since ρ_z is an algebra morphism, ρ a group action and $\mathbb{R}^{m|n}$ an abelian supergroup, we obtain that $\chi\mu_F(a\otimes b)(z) = (\mu_F \otimes \mu)\sigma_{23}(\chi\otimes\chi)(a\otimes b)(z)$. □

Note that $(\mathscr{B}(\mathbb{R}^{m|n}), \mu_F = \mu\circ F, \Delta)$ is not a Hopf algebra anymore: the deformed product μ_F is not compatible with the undeformed coproduct Δ.

3. Construction of quantum supergroups

Definition of a Fréchet quantum supergroup. In Definition 2.15, we saw that external symmetries of the deformation quantization of actions of a Lie group on Fréchet algebras correspond to a deformation of the Fréchet–Hopf algebra associated to the Lie group by using (2-13). External symmetries form a quantum group.

In the case of $\mathbb{R}^{m|n}$, we saw in Proposition 2.16 that the external symmetries of the twist F correspond to the group $\mathbb{R}^{m|n}$ itself (i.e., the undeformed Hopf algebra $H = \mathscr{B}(\mathbb{R}^{m|n})$), because $\mathbb{R}^{m|n}$ is abelian. However, to anticipate what could be the external symmetries of a more general supergroup, we have to introduce the new notion of quantum supergroup. Taking into account the nature of external symmetries, we see that this notion has to correspond to a topological graded Hopf algebra, but is not supercommutative in general.

Definition 3.1. We define a Fréchet quantum supergroup to be a Fréchet–Hopf algebra (see Definition 2.10), for a given topological tensor product, with a $\mathbb{Z}_2$-grading and for which the algebraic operations - product, unit, coproduct, counit and antipode - respect this grading, i.e., are homogeneous maps of degree 0.

There exist in the literature other definitions of quantum supergroups, as there are different notions of quantum groups — related to topological Hopf algebras or using deformations of universal enveloping algebras of Lie algebras. In particular, the purely algebraic version of Definition 3.1 corresponds exactly to the notion of quantum supergroup in [Majid 1995]. But here, we place ourselves in the context of topological Hopf (super)algebras. Note also that we do not assume that the Fréchet–Hopf algebra has to be nuclear (see Remark 2.12).

Remark 3.2. In the case of $\mathbb{R}^{m|n}$, the definition of a supergroup given by Kostant [1977] is equivalent to the data of the sheaf $\mathscr{C}^\infty$ or $\mathscr{B}$ assuming that $\mathscr{C}^\infty(\mathbb{R}^{m|n})$ or $\mathscr{B}(\mathbb{R}^{m|n})$ is a $\mathbb{Z}_2$-graded commutative Fréchet–Hopf algebra. We can notice indeed that the conditions in [ibid.] of smoothness on the coproduct and the antipode are equivalent to continuity conditions for the Fréchet structure. This is why Definition 3.1 is an extension of Kostant's definition of a supergroup to the quantum level, omitting the supercommutativity condition.

Following again the analogy with external symmetries of the deformation quantization of $\mathbb{R}^{m|n}$, we introduce the representations of a Fréchet quantum supergroup.

Definition 3.3. A representation of a given Fréchet quantum supergroup H is a $\mathbb{Z}_2$-graded comodule algebra $\boldsymbol{A}$ of H (see Definition 2.10) such that the continuous coaction $\boldsymbol{A} \to \boldsymbol{A} \widehat{\otimes}_{AH} H$ is homogeneous of degree 0.

The Clifford algebra. In this section, we consider the simplest example of Clifford algebra, for which we present the structure of (Fréchet) quantum supergroup. The Clifford algebra can be seen as a deformation quantization of the superspace $\mathbb{R}^{0|n}$: for all $f_1, f_2 \in \mathscr{C}^\infty(\mathbb{R}^{0|n})$,

$$(f_1 \star f_2)(\xi) = \kappa \int \mathrm{d}\xi_1 \, \mathrm{d}\xi_2 \; f_1(\xi_1 + \xi) f_2(\xi_2 + \xi) e^{-\frac{4i}{\theta}\xi_1\xi_2},$$

for $\xi \in \mathbb{R}^{0|n}$. The star product above corresponds actually to (2-2) for $m = 0$. We set $H := \mathscr{C}^\infty(\mathbb{R}^{0|n})$ and we can endow it with the norm $\|f\| := \sum_I |f_I|$ for I to be summed over the parts of $\{1, \dots, n\}$. On this finite-dimensional space, any other norm would have been equivalent, so that we do not look anymore at the topology of this example. H is associative, with unit 1. As generators, we take $e_i := \xi^i$ with

$\xi = (\xi^1, \dots, \xi^n) \in \mathbb{R}^{0|n}$. Since

$$e_i \star e_j = e_i e_j + \frac{i\theta}{4}\delta_{ij},$$

we have the following relations of $Cl(n, \mathbb{C})$,

$$e_i \star e_j + e_j \star e_i = \frac{i\theta}{2}\delta_{ij}.$$

If $\theta = -4i$, we can endow H [Albuquerque and Majid 2002] with a structure of quantum supergroup:

- coproduct $\Delta(e_i) := e_i \otimes e_i$,
- counit $\varepsilon(e_i) := 1$,
- antipode $S(e_i) := e_i$,
- product on tensors $(e_i \otimes e_j) \star (e_k \otimes e_l) := \sigma_{jk}(e_i \star e_k) \otimes (e_j \star e_l)$,

with $\sigma_{ij} = 1$ if $i \leq j$ and $\sigma_{ij} = -1$ if $i > j$. Note that σ_{ij} is a Schur multiplier of the group $\mathbb{Z}_2^n$ for which the algebra $Cl(n, \mathbb{C})$ is $\mathbb{Z}_2^n$-graded commutative [de Goursac et al. 2012]. A corresponding Kac–Takesaki operator would be given by $W(e_i \otimes e_j) := e_i \otimes (e_i \star e_j)$.

Examples of solvable Fréchet quantum supergroups. Let us now construct other examples of Fréchet quantum supergroups, which are deformation of solvable Lie supergroups. These are consistent extensions of [Rieffel 1992] to the graded setting. We consider a $(1|0)$-dimensional split extension of the symplectic superspace $(\mathbb{R}^{m|n}, \omega)$ of Definition 2.1. Let indeed $\pi : \mathbb{R}^{1|0} \to \mathrm{Sp}(\mathbb{R}^{m|n}, \omega)$ be a symplectic representation of $\mathbb{R}^{1|0}$ on $\mathbb{R}^{m|n}$, homogeneous of degree 0. It can be written as

$$\pi = \begin{pmatrix} \pi_0 & 0 \\ 0 & \pi_1 \end{pmatrix}$$

(square matrix of size $m+n$). We also assume each matrix coefficient of π to be smooth with respect to the variable $a \in \mathbb{R}^{1|0}$. Then, the split extension is of the form $G := \mathbb{R}^{1|0} \ltimes_\pi \mathbb{R}^{m|n}$ with supergroup law,

$$(a, x, \xi) \cdot (a', x', \xi') = (a + a', \pi_0(a')x + x', \pi_1(a')\xi + \xi'). \tag{3-1}$$

Here $a \in \mathbb{R}^{1|0}$, $x \in \mathbb{R}^{m|0}$ and $\xi \in \mathbb{R}^{0|n}$. We use the natural action of $\mathbb{R}^{m|n}$ on G together with the universal deformation formula of Proposition 2.7 to deform the product of functions on G as

$$\begin{aligned}(f_1 \star f_2)(a, x, \xi) &\\ = \kappa(a) \int \mathrm{d}x_1\, \mathrm{d}\xi_1\, \mathrm{d}x_2\, \mathrm{d}\xi_2\ & f_1(a, x_1 + x, \xi_1 + \xi) f_2(a, x_2 + x, \xi_2 + \xi) \\ &\times e^{-\frac{2i}{a}(\omega_0(x_1, x_2) + 2\xi_1\xi_2)}\end{aligned} \tag{3-2}$$

with $\kappa(a) = (-1)^{n(n+1)/2}(ia)^n/(4^n(\pi a)^m)$. Note that we used the extension variable a as the deformation parameter. This will be crucial to define a consistent coproduct. We define H to be the space of smooth superfunctions on G that are bounded with every derivative bounded in the variables $(x, \xi) \in \mathbb{R}^{m|n}$,

$$H := \mathscr{C}^\infty(\mathbb{R}^{1|0}) \widehat{\otimes} \mathscr{B}(\mathbb{R}^{m|n}).$$

The standard Fréchet structure of H is defined by the seminorms

$$|f|_{\alpha,K,\beta} = \sup_{\substack{a\in K\\ x\in\mathbb{R}^m}} \left\{ \sum_I |D_a^\alpha D_x^\beta f_I(a,x)| \right\} \tag{3-3}$$

for K compact of $\mathbb{R} = \mathbb{B}(\mathbb{R}^{1|0})$, $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}^m$.

Proposition 3.4. *Endowed with the star product* (3-2) *and the seminorms* (3-3), *H is a unital associative Fréchet superalgebra.*

Proof. What remains to prove here is the continuity of the star product (3-2). Let $f_1, f_2 \in H$, K compact of $\mathbb{R}$, $\alpha \in \mathbb{N}$ and $\beta \in \mathbb{N}^m$. First we perform a change of variable: $x_1 \mapsto ax_1$ in the expression of $|f_1 \star f_2|_{\alpha,K,\beta}$. Then, we can estimate this expression by expanding the superfunctions f_1 and f_2 along the odd variables in (3-2) and integrate over these odd variables, and also apply operators (2-8) inside the integrals. Thus for $k_1, k_2 \in \mathbb{N}$, there exist functions $b_1^\gamma, b_2^\delta \in \mathscr{B}(\mathbb{R}^m)$ such that

$$\begin{aligned} &|f_1 \star f_2|_{\alpha,K,\beta} \\ &\quad\le \frac{1}{4^n\pi^m} \sup_{\substack{a\in K\\ x\in\mathbb{R}^m}} \sum_{\substack{I,J,\gamma,\delta,\\ \tau,\nu,\mu}} \int \mathrm{d}x_1\,\mathrm{d}x_2\, \frac{1}{(1+x_1^2)^{k_1}(1+x_2^2)^{k_2}} |b_1^\gamma(x_1) b_2^\delta(x_2)|\,|a|^{|\mu|} \\ &\qquad\qquad \times |D_a^\tau D_x^\gamma (f_1)_I(a, x+ax_1)|\,|D_a^\nu D_x^\delta (f_2)_J(a, x+x_2)| \end{aligned}$$

where I, J are summed over $\{1, \dots, n\}$ with some conditions; τ, ν, μ over $\mathbb{N}$ with $\tau + \nu \le \alpha$ and $\mu \le |\gamma|$; γ, δ over $\mathbb{N}^m$ with $|\gamma| \le |\beta| + 2k_1$ and $|\delta| \le |\beta| + 2k_2$. For an adapted choice of k_1, k_2, it means that there exists a constant $C > 0$ such that

$$|f_1 \star f_2|_{\alpha,K,\beta} \le C \sum_{\gamma,\delta,\tau,\nu} |f_1|_{\tau,K,\gamma} |f_2|_{\nu,K,\delta}$$

where the sum is finite. This proves the continuity of the star product. □

We then consider the coproduct, counit and antipode coming from the (undeformed) supergroup structure of G:

- the coproduct $\Delta : H \to H \widehat{\otimes} H$ defined by $\Delta f(g, g') = f(g \cdot g')$ for $g, g' \in G$ and the supergroup law (3-1),
- the counit $\varepsilon : H \to \mathbb{C}$ defined by $\varepsilon(f) = f(0, 0, 0)$,

- the antipode $S : H \to H$ defined by $Sf(g) = f(g^{-1})$, with $f \in H$ and $(a, x, \xi)^{-1} = (-a, -\pi_0(-a)x, -\pi_1(-a)\xi)$.

We note $\mu : H\widehat{\otimes}H \to H$ the star product: $\mu(f_1 \otimes f_2) := f_1 \star f_2$.

Theorem 3.5. *$(H, \mu, 1, \Delta, \varepsilon, S)$ is a Fréchet quantum supergroup.*

Proof. We know from Proposition 3.4 that $(H, \star, 1)$ is a Fréchet superalgebra. Let us show first that the coproduct is continuous. For $f \in H$, (a, x, ξ), $(a', x', \xi') \in G$, the coproduct takes the form

$$\Delta(f)(a, x, \xi, a', x', \xi') = \sum_{I,J,L} \varepsilon(I, J) f_{I\cup J}(a+a', \pi_0(a')x+x')(\pi_1(a'))_{IL}\xi^L(\xi')^J$$

with some constraints on I, J, L, and $\varepsilon(I, J)$ given by (2-1). Then,

$$\begin{aligned}
|\Delta(f)|_{\alpha,K,\beta;\alpha',K',\beta'} &\leq \sup_{\substack{a\in K\\ a'\in K'\\ x,x'\in\mathbb{R}^m}} \sum_{I,J,L} |D_a^\alpha D_{a'}^{\alpha'} D_x^\beta D_{x'}^{\beta'} f_{I\cup J}(a+a', \pi_0(a')x+x')(\pi_1(a'))_{IL}| \\
&\leq C \sum_{\tau,\gamma} |f|_{\tau,K'',\gamma}
\end{aligned}$$

where K'' is a compact containing $\{a+a', a \in K, a' \in K'\}$, $\tau \leq \alpha+\alpha'$, $|\gamma| \leq |\beta|+|\beta'|$, and C a constant depending in particular on the smooth matrix coefficients of π and their derivatives. This proves that Δ is continuous. In the same way, the counit ε and the antipode S are continuous.

Let us show that Δ is an algebra morphism for the star product. For $f_1, f_2 \in H$, we have

$$\begin{aligned}
&\Delta(f_1 \star f_2)(a, x, \xi, a', x', \xi') \\
&= \kappa(a+a') \int \mathrm{d}x_1\, \mathrm{d}\xi_1\, \mathrm{d}x_2\, \mathrm{d}\xi_2\; f_1\big(a+a', x_1+\pi_0(a')x+x', \xi_1+\pi_1(a')\xi+\xi'\big) \\
&\qquad \times f_2\big(a+a', x_2+\pi_0(a')x+x', \xi_2+\pi_1(a')\xi+\xi'\big) e^{-\frac{2i}{a+a'}(\omega_0(x_1,x_2)+2\xi_1\xi_2)}
\end{aligned}$$

Besides,

$$\begin{aligned}
&\Delta(f_1) \star \Delta(f_2)(a, x, \xi, a', x', \xi') \\
&= \kappa(a)\kappa(a') \int \mathrm{d}x_1\, \mathrm{d}\xi_1\, \mathrm{d}x_2\, \mathrm{d}\xi_2\, \mathrm{d}x_1'\, \mathrm{d}\xi_1'\, \mathrm{d}x_2'\, \mathrm{d}\xi_2' \\
&\quad \times f_1\big(a+a', \pi_0(a')(x_1+x)+x_1'+x', \pi_1(a')(\xi_1+\xi)+\xi_1'+\xi'\big) e^{-\frac{2i}{a}(\omega_0(x_1,x_2)+2\xi_1\xi_2)} \\
&\quad \times f_2\big(a+a', \pi_0(a')(x_2+x)+x_2'+x', \pi_1(a')(\xi_2+\xi)+\xi_2'+\xi'\big) e^{-\frac{2i}{a'}(\omega_0(x_1',x_2')+2\xi_1'\xi_2')}.
\end{aligned}$$

The sign of the star product of elements of $H\widehat{\otimes}H$ coming from the flip (2-11) has been taken into account. We perform the change of variables $x_i'' = x_i' + \pi_0(a')x_i$,

$\xi_i'' = \xi_i' + \pi_1(a')\xi_i$. Using the identity $\int \mathrm{d}\xi\, e^{c\xi\xi'} = (-1)^{\frac{1}{2}n(n-1)} c^n (\xi')^{\{1,\dots,n\}}$, we can integrate over x_1, ξ_1, obtaining

$$\begin{aligned}
&\Delta(f_1) \star \Delta(f_2)(a, x, \xi, a', x', \xi') \\
&\quad = (-4i)^n (-1)^{\frac{n(n+1)}{2}} \pi^m \kappa(a)\kappa(a') \int \mathrm{d}x_2\, \mathrm{d}\xi_2\, \mathrm{d}x_1''\, \mathrm{d}\xi_1''\, \mathrm{d}x_2''\, \mathrm{d}\xi_2'' \\
&\quad\quad \times \delta\biggl(\frac{a+a'}{aa'} x_2 - \frac{1}{a'}\pi_0(a')^* x_2''\biggr)\biggl(\frac{a+a'}{aa'}\xi_2 - \frac{1}{a'}\pi_1(a')^*\xi_2''\biggr)^{\{1,\dots,n\}} \\
&\quad\quad \times f_1(a+a', x_1'' + \pi_0(a')x + x', \xi_1''\pi_1(a')\xi + \xi') \\
&\quad\quad \times f_2(a+a', x_2'' + \pi_0(a')x + x', \xi_2'' + \pi_1(a')\xi) + \xi') \\
&\quad\quad \times \exp\biggl(-\frac{2i}{a'}(\omega_0(x_1'', x_2'' - \pi_0(a')x_2) + 2\xi_1''(\xi_2'' - \pi_1(a')\xi_2))\biggr).
\end{aligned}$$

In the previous step, we used the fact that π is a symplectic representation, i.e., $\omega(\pi_0(a)x, \pi_0(a)y) = \omega_0(x, y)$ and $(\pi_1(a)\xi)(\pi_1(a)\eta) = \xi\eta$. Moreover we denote $\pi_0(a)^* := \omega_0^{-1}\pi_0(a)^T\omega_0$ and $\pi_1(a)^* := \pi_1(a)^T$. If we now perform the Dirac integration on x_2, ξ_2, we obtain

$$\Delta(f_1) \star \Delta(f_2)(a, x, \xi, a', x', \xi') = \Delta(f_1 \star f_2)(a, x, \xi, a', x', \xi').$$

All the other algebraic identities are the same as in the undeformed case except $\mu(\mathrm{id} \otimes S)\Delta = 1 \otimes \varepsilon = \mu(S \otimes \mathrm{id})\Delta$. For this, we compute

$$\begin{aligned}
&\mu(\mathrm{id} \otimes S)\Delta(f)(a, x, \xi) \\
&= \kappa(a) \int \mathrm{d}x_1\, \mathrm{d}\xi_1\, \mathrm{d}x_2\, \mathrm{d}\xi_2\, f(0, \pi_0(-a)(x_1 - x_2), \pi_1(-a)(\xi_1 - \xi_2)) e^{-\frac{2i}{a}(\omega_0(x_1, x_2) + 2\xi_1\xi_2)} \\
&= \varepsilon(f).
\end{aligned}$$

□

We can now exhibit the analog of Kac–Takesaki operator $W : H\widehat{\otimes}H \to H\widehat{\otimes}H$ associated to this quantum supergroup, also called multiplicative unitary in the nongraded context. It is defined in [Baaj and Skandalis 1993; Woronowicz 1996] by

$$W(a \otimes b) := (\Delta a) \star (1 \otimes b) = a_{(1)} \otimes (a_{(2)} \star b), \tag{3-4}$$

for all $a, b \in H$, using the Sweedler notation for the coproduct. Its explicit expression is given by, for all $f \in H\widehat{\otimes}H$,

$$\begin{aligned}
&W(f)(a, x, \xi, a', x', \xi') \\
&\quad = \kappa(a') \int \mathrm{d}x_1\, \mathrm{d}\xi_1\, \mathrm{d}x_2\, \mathrm{d}\xi_2\; e^{-\frac{2i}{a'}(\omega_0(x_1, x_2) + 2\xi_1\xi_2)} \\
&\quad\quad \times f(a+a', x_1 + \pi_0(a')x + x', \xi_1 + \pi_1(a')\xi + \xi', a', x_2 + x', \xi_2 + \xi').
\end{aligned} \tag{3-5}$$

Proposition 3.6. *The Kac–Takesaki operator* (3-5) *is a continuous operator* $W : H\widehat{\otimes}H \to H\widehat{\otimes}H$ *homogeneous of degree* 0, *and it satisfies the pentagonal relation*

$$W_{12}W_{13}W_{23} = W_{23}W_{12}.$$

Proof. Indeed, as $W = (\mu \otimes \mu)\sigma_{23}(\Delta \otimes 1 \otimes \mathrm{id})$, it is continuous. To prove the pentagonal relation where involved signs are different from the nongraded case, we use the Sweedler notation for the coproduct since its coassociativity has been showed in Theorem 3.5. On the left side,

$$W_{12}W_{13}W_{23}(a \otimes b \otimes c) = (-1)^{|a_{(3)}||b_{(1)}|}a_{(1)} \otimes (a_{(2)} \star b_{(1)}) \otimes (a_{(3)} \star b_{(2)} \star c),$$

where the sign appears because of the action of

$$W_{13} = (\mu \otimes \mathrm{id} \otimes \mu)\sigma_{24}(\Delta \otimes \mathrm{id} \otimes 1 \otimes \mathrm{id})$$

and with

$$\sigma_{24}(a_1 \otimes a_2 \otimes a_3 \otimes a_4 \otimes a_5) = (-1)^{|a_2|(|a_3|+|a_4|)+|a_3||a_4|}a_1 \otimes a_4 \otimes a_3 \otimes a_2 \otimes a_5.$$

On the right side,

$$W_{23}W_{12}(a \otimes b \otimes c) = (-1)^{|a_{(3)}||b_{(1)}|}a_{(1)} \otimes (a_{(2)} \star b_{(1)}) \otimes (a_{(3)} \star b_{(2)} \star c),$$

where we used $\Delta(a \star b) = \Delta(a) \star \Delta(b) = (-1)^{|a_{(2)}||b_{(1)}|}(a_{(1)} \star b_{(1)}) \otimes (a_{(2)} \star b_{(2)})$ due to Theorem 3.5. □

Remark 3.7. For the Lebesgue–Berezin measure on $\mathbb{R}^{m|n}$, we can define a "natural" superhermitian (not positive definite) scalar product

$$\langle f_1, f_2\rangle := \int \mathrm{d}x\,\mathrm{d}\xi\; \overline{f_1(x, \xi)}\, f_2(x, \xi)$$

and a hermitian positive definite one $(f_1, f_2) := \langle f_1, *f_2\rangle$ via the Hodge operation

$$*\sum_I f_I(x)\xi^I := \sum_I \varepsilon(I, \complement I) f_I(x)\xi^{\complement I}.$$

Account being taken of the right-invariant measure

$$\mathrm{d}^R(a, x, \xi) = \frac{1}{\mathrm{sdet}(\pi(a))}\mathrm{d}(a, x, \xi) = \frac{\det \pi_1(a)}{\det \pi_0(a)}\mathrm{d}a\,\mathrm{d}x\,\mathrm{d}\xi$$

on G, a straightforward computation using (3-5) shows that

$$\int \mathrm{d}^R(a, x, \xi)\, \mathrm{d}^R(a', x', \xi')\; \overline{W(f_1)(a, x, \xi, a', x', \xi')}\, W(f_2)(a, x, \xi, a', x', \xi')$$
$$= \int \mathrm{d}^R(a, x, \xi)\, \mathrm{d}^R(a', x', \xi')\; \overline{f_1(a, x, \xi, a', x', \xi')}\, f_2(a, x, \xi, a', x', \xi')$$

for $f_1, f_2 \in (H \widehat{\otimes} H) \cap L^2(G \times G)$. This means that the operator W is superunitary for the superhermitian scalar product associated to $L^2(G \times G, \mathrm{d}^R g \otimes \mathrm{d}^R g')$,

$$\langle W(f_1), W(f_2)\rangle = \langle f_1, f_2\rangle,$$

which is not true for the positive definite scalar product $(\,\cdot\,,\,\cdot\,)$. W is a "multiplicative superunitary" rather than a multiplicative unitary.

Fréchet quantum supergroups with supertoral subgroups. In this section, we will follow the philosophy of [Rieffel 1993] to construct deformation of compact Lie supergroups with supertoral subgroups. Let G be a compact connected Lie supergroup (i.e., its body $\mathbb{B}G$ is compact connected) with $\Gamma := \mathbb{T}^{m|n}$ a supertoral subgroup of G. We assume that m is even so that the symplectic superspace $(\mathbb{R}^{m|n}, \omega)$ (see Definition 2.1) is the Lie algebra of Γ. We note

$$z = (x, \xi) \in \mathbb{R}^{m|n} \quad \mapsto \quad e^z = e^{(x,\xi)} \in G$$

the exponential restricted to this Lie algebra. Note that $\mathscr{C}^\infty(G) \simeq \mathscr{C}^\infty(\mathbb{B}G) \otimes \bigwedge \mathbb{R}^n$ is a Fréchet superalgebra for the supercommutative pointwise product (see below Definition 2.4) and the seminorms

$$|f|_{\alpha,K} = \sup_{\substack{g\in K\\ |\nu|\le\alpha}} \left\{ \sum_I |D^\nu f_I(g)| \right\}, \tag{3-6}$$

for K compact subset of a coordinate chart of $\mathbb{B}G$, $\alpha \in \mathbb{N}$ and D^ν a multiderivation of order $|\nu|$ for even coordinates. The action $\rho : \mathbb{R}^{m|n} \times \mathbb{R}^{m|n} \times \mathscr{C}^\infty(G) \to \mathscr{C}^\infty(G)$, defined by

$$\forall z, z' \in \mathbb{R}^{m|n}, \forall g \in G, \quad \rho_{(z,z')} f(g) := f(e^{-z} g e^{z'}),$$

allows to deform the pointwise product into the star product

$$\begin{aligned}(f_1 \star f_2)(g) \\ = \kappa^2 \int \mathrm{d}z_1\, \mathrm{d}z_3\, \mathrm{d}z_2\, \mathrm{d}z_4\; f_1(e^{-z_1} g e^{z_3}) f_2(e^{-z_2} g e^{z_4}) e^{-\frac{2i}{\theta}(\omega(z_1,z_2)-\omega(z_3,z_4))},\end{aligned} \tag{3-7}$$

with $\kappa = (-1)^{\frac{1}{2}n(n+1)}(i\theta/4)^n/(\pi\theta)^m$, for $g \in G$ and for any $f_1, f_2 \in \mathscr{C}^\infty(G)$. Note that the underlying symplectic space is $(\mathbb{R}^{m|n}, \omega) \oplus (\mathbb{R}^{m|n}, -\omega)$, where the minus sign, which can also be found in the phase of the star product, will be crucial. We note $H := \mathscr{C}^\infty(G)$.

Proposition 3.8. *Endowed with the star product* (3-7) *and the seminorms* (3-6), *H is a unital associative Fréchet superalgebra.*

Proof. Associativity is a consequence of the universal deformation formula. Let us check that the star product is continuous. Then, we use the same method as in the proof of Proposition 3.4 and we get that for $k_i \in \mathbb{N}$, there exist functions $b_i^{\gamma_i} \in \mathcal{B}(\mathbb{R}^m)$ and a constant $C > 0$ (depending on θ) such that

$$\begin{aligned}
&|f_1 \star f_2|_{\alpha,K}\\
&\quad \leq C \sup_{g\in\mathbb{B}K} \sum_{I,J,\gamma_i,\nu_i} \int \mathrm{d}x_1\, \mathrm{d}x_3\, \mathrm{d}x_2\, \mathrm{d}x_4 \; \frac{|c_{\nu_1,\nu_2}||b_1^{\gamma_1}(x_1) b_2^{\gamma_2}(x_2) b_3^{\gamma_3}(x_3) b_4^{\gamma_4}(x_4)|}{(1+x_1^2)^{k_1}(1+x_2^2)^{k_2}(1+x_3^2)^{k_3}(1+x_4^2)^{k_4}}\\
&\qquad\qquad \times |D_g^{\nu_1} D_{x_1}^{\gamma_1} D_{x_3}^{\gamma_3}((f_1)_I(e^{-\theta x_1} g e^{x_3}))||D_g^{\nu_2} D_{x_2}^{\gamma_2} D_{x_4}^{\gamma_4}((f_2)_J(e^{-\theta x_2} g e^{x_4}))|,
\end{aligned}$$

where I, J are summed over $\{1, \ldots, n\}$ with some conditions; and ν_i, γ_i are such that $\nu_1 + \nu_2 \leq \alpha$ and $|\gamma_i| \leq 2k_i$. It follows that there exists a constant $C' > 0$ and a compact K' of $\mathbb{B}G$ containing $\{\mathbb{B}(e^{-z_1} g e^{z_2}),\, g \in K,\, z_i \in \Gamma\}$ such that

$$|f_1 \star f_2|_{\alpha,K} \leq C' \sum_{\tau,\nu} |f_1|_{\tau,K'} |f_2|_{\nu,K'},$$

where the sum is finite. Therefore, the star product is continuous. □

Let us endow H with the following (undeformed) operations:

- the coproduct $\Delta : H \to H \widehat{\otimes} H$ defined by $\Delta f(g, g') = f(g \cdot g')$ for $g, g' \in G$,
- the counit $\varepsilon : H \to \mathbb{C}$ defined by $\varepsilon(f) = f(e_G)$, with e_G the neutral element of G,
- the antipode $S : H \to H$ defined by $Sf(g) = f(g^{-1})$, with $f \in H$.

We denote by $\mu : H \widehat{\otimes} H \to H$ the star product: $\mu(f_1 \otimes f_2) := f_1 \star f_2$.

Theorem 3.9. $(H, \mu, 1, \Delta, \varepsilon, S)$ *is a Fréchet quantum supergroup.*

Proof. First, we check the compatibility between the coproduct and the product. Set $f_1, f_2 \in H$.

$$\begin{aligned}
&\Delta(f_1 \star f_2)(g, g')\\
&\qquad = \kappa^2 \int \mathrm{d}z_1\, \mathrm{d}z_3\, \mathrm{d}z_2\, \mathrm{d}z_4 \; f_1(e^{-z_1} g g' e^{z_3}) f_2(e^{-z_2} g g' e^{z_4}) e^{-\frac{2i}{\theta}(\omega(z_1,z_2)-\omega(z_3,z_4))}.
\end{aligned}$$

Then, as in the previous section, we want to compute

$$\begin{aligned}
&\Delta(f_1) \star \Delta(f_2)(g, g')\\
&= \kappa^4 \int \mathrm{d}z_1\, \mathrm{d}z_3\, \mathrm{d}z_2\, \mathrm{d}z_4\, \mathrm{d}z_1'\, \mathrm{d}z_3'\, \mathrm{d}z_2'\, \mathrm{d}z_4' \; f_1(e^{-z_1} g e^{z_3 - z_1'} g' e^{z_3'}) f_2(e^{-z_2} g e^{z_4 - z_2'} g' e^{z_4'})\\
&\qquad\qquad \times \exp\Big(-\frac{2i}{\theta}(\omega(z_1, z_2) - \omega(z_3, z_4) + \omega(z_1', z_2') - \omega(z_3', z_4'))\Big).
\end{aligned}$$

For this, we change the variables $z_3'' = z_3 - z_1'$, $z_4'' = z_4 - z_2'$ and we perform the integration on z_1', z_2'. After simplification, it gives the compatibility

$$\Delta(f_1) \star \Delta(f_2)(g, g') = \Delta(f_1 \star f_2)(g, g').$$

Let us show for example the identity $\mu(\mathrm{id}\otimes S)\Delta = 1\otimes\varepsilon$. Indeed,

$$\begin{aligned}\mu(\mathrm{id}\otimes S)\Delta(f)(g) &= \kappa^2\int \mathrm{d}z_1\,\mathrm{d}z_3\,\mathrm{d}z_2\,\mathrm{d}z_4\; f(e^{-z_1}ge^{z_3}(e^{-z_2}ge^{z_4})^{-1})e^{-\frac{2i}{\theta}(\omega(z_1,z_2)-\omega(z_3,z_4))}\\ &= \kappa(-1)^n\int \mathrm{d}z_1\,\mathrm{d}z_2\; f(e^{-z_1}gg^{-1}e^{z_2})e^{-\frac{2i}{\theta}\omega(z_1,z_2)} = f(e_G).\end{aligned}$$

For the continuity of the coproduct, we need to choose a global odd coordinate system $\{\eta\}$ on G since it is a trivial supermanifold (see Definition 2.4). Then, the coproduct can be expressed as

$$\Delta(f)(g,g') = f(g\cdot g') = \sum_{I,J,L} c_{I,J,L}\,f_L((\mathbb{B}g)(\mathbb{B}g'))\eta^I(\eta')^J$$

if $g=(\mathbb{B}g,\eta)$, $g'=(\mathbb{B}g',\eta')$, and by denoting $c_{I,J,L}$ some coefficients related to the group law of G and independent of f. Thus, we have the estimate

$$|\Delta(f)|_{\alpha,K;\alpha',K'} \le \sup_{\substack{g\in K,\,g'\in K'\\|\nu|\le\alpha,\,|\nu'|\le\alpha'}} \sum_{I,J,L}|c_{I,J,L}||D^{\nu}_g D^{\nu'}_{g'} f_L(g\cdot g')| \le C\sum_{\tau}|f|_{\tau,K''},$$

where K'' is a compact subset of $\mathbb{B}G$ containing $\{g\cdot g',\, g\in K,\, g'\in K'\}$, $\tau\le\alpha+\alpha'$, and C a constant depending in particular on $c_{I,J,L}$ and on the (smooth) modular function of G and its derivatives. So, the coproduct is continuous, as well as the other operations. □

Proposition 3.10. *The subgroup $\Gamma\subset G$ is not deformed in H. This means that $\Gamma=\mathbb{T}^{m|n}$ is a subgroup of the quantum supergroup $(H,\mu,1,\Delta,\varepsilon,S)$.*

Proof. The coproduct is indeed not deformed. For the product, we can see that for all $g\in\Gamma$ and all $f_1, f_2\in H$,

$$(f_1\star f_2)(g) = \kappa^2\int \mathrm{d}z_1\,\mathrm{d}z_3\,\mathrm{d}z_2\,\mathrm{d}z_4\; f_1(ge^{z_3-z_1})f_2(ge^{z_4-z_2})e^{-\frac{2i}{\theta}(\omega(z_1,z_2)-\omega(z_3,z_4))}$$

since Γ is abelian. By performing the change of variables $z_1\mapsto z_1+z_3$, $z_2\mapsto z_2+z_4$ and integrating over z_3, z_4, we find $(f_1\star f_2)(g)=f_1(g)f_2(g)$. So, $\mathscr{C}^\infty(\Gamma)$ is not deformed in H. □

The analog of the Kac–Takesaki operator defined in (3-4) has in this context the expression

$$\begin{aligned}W(f)(g,g') &\\ &= \kappa^2\int \mathrm{d}z_1\,\mathrm{d}z_3\,\mathrm{d}z_2\,\mathrm{d}z_4\; f(e^{-z_1}gg'e^{z_3}, e^{-z_2}g'e^{z_4})e^{-\frac{2i}{\theta}(\omega(z_1,z_2)-\omega(z_3,z_4))},\end{aligned}\tag{3-8}$$

for $f\in H\widehat{\otimes}H$. As in Proposition 3.6, we can show that it is a continuous operator

$W : H\widehat{\otimes}H \to H\widehat{\otimes}H$ homogeneous of degree 0 and that it satisfies the pentagonal equation

$$W_{12}W_{13}W_{23} = W_{23}W_{12}.$$

Moreover if G is unimodular, a computation analog as in Remark 3.7 proves that W is superunitary for the superhermitian scalar product canonically associated to $L^2(G \times G)$: for all $f_1, f_2 \in (H\widehat{\otimes}H) \cap L^2(G \times G)$,

$$\langle W(f_1), W(f_2)\rangle = \langle f_1, f_2\rangle.$$

We finally give an explicit example of Fréchet quantum supergroup with a supertoral subgroup. For this, we will present the special linear supergroup in low dimension. We need to recall what supermatrices are.

Definition 3.11. A square supermatrix A of size $(m\,|\,n)$ is a matrix with coefficients in $\mathscr{A}$ (see Section 2) and of the form

$$A = \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix}$$

where A_{00} is an $m \times m$ matrix with entries in $\mathscr{A}_0$, A_{01} is an $m \times n$ matrix with entries in $\mathscr{A}_1$, A_{10} is an $n \times m$ matrix with entries in $\mathscr{A}_1$, and A_{11} is an $n \times n$ matrix with entries in $\mathscr{A}_0$.

The set of square supermatrices of size $(m\,|\,n)$ is a superalgebra for the standard addition and multiplication. We denote by $\mathrm{GL}(m\,|\,n)$ the supergroup of invertible square supermatrices of size $(m\,|\,n)$. Finally, we define the Berezinian (or superdeterminant) of a supermatrix A by

$$\mathrm{Ber}(A) = \det(A_{00} - A_{01}A_{11}^{-1}A_{10})\det(A_{11}^{-1}).$$

Now the special linear supergroup is defined as

$$\mathrm{SL}(m \mid n) := \{A \in \mathrm{GL}(m \mid n),\ \mathrm{Ber}(A) = 1\}.$$

Restricted to dimension $m\,|\,n = 1\,|\,1$, this supergroup contains the elements

$$g = \begin{pmatrix} a & \beta \\ \gamma & d \end{pmatrix}$$

with $a, d \in \mathscr{A}_0$, $\beta, \gamma \in \mathscr{A}_1$ such that $a = d + \frac{1}{d}\beta\gamma$. We see directly that $\mathrm{SL}(1\,|\,1)$ contains two supertoral subgroups generated by $\beta \in \mathbb{T}^{0|1}$ and $\gamma \in \mathbb{T}^{0|1}$. We can choose for example to consider the deformation using the supertorus generated by β, and we want to see if this deformation is not trivial. For this, we compute the explicit expression of the star product (3-7): for any $f_1, f_2 \in \mathscr{C}^\infty(\mathrm{SL}(1 \mid 1))$,

$$\begin{aligned}(f_1 \star f_2)(g) \\ = f_1(g) f_2(g) + \frac{i\theta}{2} a\gamma (\partial_\beta f_1(g)\partial_d f_2(g) - \partial_d f_1(g)\partial_\beta f_2(g)) \\ + \frac{i\theta}{2} d\gamma(\partial_\beta f_1(g)\partial_a f_2(g) - \partial_a f_1(g)\partial_\beta f_2(g)) + \frac{i\theta}{2}(d^2 - a^2)\partial_\beta f_1(g)\partial_\beta f_2(g).\end{aligned}$$

We see that already by taking a supertoral subgroup of dimension $0|1$ we can produce a nontrivial Fréchet quantum supergroup, deformation of $\mathrm{SL}(1|1)$. Note that this associative star product stops at the finite level θ because only odd variables are involved in the deformation. This is a simple example that shows how such a construction can be useful in concrete cases. Of course, it applies on a large class of supergroups for which explicit expressions can be much more complicated.

Acknowledgements

The author thanks Pierre Bieliavsky, Philippe Bonneau and Gijs Tuynman for interesting discussions on this work.

References

[Albuquerque and Majid 2002] H. Albuquerque and S. Majid, “Clifford algebras obtained by twisting of group algebras”, *J. Pure Appl. Algebra* **171**:2-3 (2002), 133–148. MR 2003f:15037 Zbl 1054.15024

[Artin and Zhang 1994] M. Artin and J. J. Zhang, “Noncommutative projective schemes”, *Adv. Math.* **109**:2 (1994), 228–287. MR 96a:14004 Zbl 0833.14002

[Baaj and Skandalis 1993] S. Baaj and G. Skandalis, “Unitaires multiplicatifs et dualité pour les produits croisés de C^*-algèbres”, *Ann. Sci. École Norm. Sup.* (4) **26**:4 (1993), 425–488. MR 94e:46127 Zbl 0804.46078

[Bayen et al. 1978a] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, “Deformation theory and quantization, I: Deformations of symplectic structures”, *Ann. Physics* **111**:1 (1978), 61–110. MR 58 #14737a Zbl 0377.53024

[Bayen et al. 1978b] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, “Deformation theory and quantization, II: Physical applications”, *Ann. Physics* **111**:1 (1978), 111–151. MR 58 #14737b Zbl 0377.53025

[Bieliavsky and Gayral 2013] P. Bieliavsky and V. Gayral, “Deformation quantization for actions of Kählerian Lie groups”, preprint, 2013. To appear in *Mem. Amer. Math. Soc.* arXiv 1109.3419

[Bieliavsky et al. 2010] P. Bieliavsky, P. Bonneau, F. D’Andrea, V. Gayral, Y. Maeda, and Y. Voglaire, “Non-formal Drinfeld twists for Kählerian Lie groups”, 2010. In progress.

[Bieliavsky et al. 2012] P. Bieliavsky, A. de Goursac, and G. Tuynman, “Deformation quantization for Heisenberg supergroup”, *J. Funct. Anal.* **263**:3 (2012), 549–603. MR 2925914 Zbl 1247.53091

[Bieliavsky et al. 2014] P. Bieliavsky, V. Gayral, A. de Goursac, and F. Spinnler, “Harmonic analysis on homogeneous complex bounded domains and noncommutative geometry”, in *Developments and retrospectives in Lie theory: geometric and analytic methods*, edited by G. Mason et al., Developments in Mathematics **37**, Springer, New York, 2014. arXiv 1311.1871

[Bonneau and Sternheimer 2005] P. Bonneau and D. Sternheimer, "Topological Hopf algebras, quantum groups and deformation quantization", pp. 55–70 in *Hopf algebras in noncommutative geometry and physics* (Brussels, 2002), edited by S. Caenepeel and F. Van Oystaeyen, Lecture Notes in Pure and Appl. Math. **239**, Dekker, New York, 2005. MR 2006d:16062 Zbl 1080.16037

[Connes 1994] A. Connes, *Noncommutative geometry*, Academic Press, San Diego, CA, 1994. MR 95j:46063 Zbl 0818.46076

[DeWitt 1984] B. DeWitt, *Supermanifolds*, Cambridge University Press, 1984. MR 87b:58007 Zbl 0551.53002

[Drinfeld 1989] V. G. Drinfeld, "Quasi–Hopf algebras", *Algebra i Analiz* **1**:6 (1989), 114–148. In Russian, translated in *Leningrad Math. J.* **1**:6 (1990), 1419–1457. MR 91b:17016 Zbl 0718.16033

[Giaquinto and Zhang 1998] A. Giaquinto and J. J. Zhang, "Bialgebra actions, twists, and universal deformation formulas", *J. Pure Appl. Algebra* **128**:2 (1998), 133–151. MR 2000a:16072 Zbl 0938.17015

[de Goursac 2010] A. de Goursac, "On the origin of the harmonic term in noncommutative quantum field theory", *SIGMA Symmetry Integrability Geom. Methods Appl.* **6** (2010), 048. MR 2011k:81270 Zbl 1217.81137

[de Goursac et al. 2007] A. de Goursac, J.-C. Wallet, and R. Wulkenhaar, "Noncommutative induced gauge theory", *Eur. Phys. J. C Part. Fields* **51**:4 (2007), 977–987. MR 2009b:81201 Zbl 1189.81215

[de Goursac et al. 2012] A. de Goursac, T. Masson, and J.-C. Wallet, "Noncommutative ϵ-graded connections", *J. Noncommut. Geom.* **6**:2 (2012), 343–387. MR 2914869 Zbl 1275.58003

[Grothendieck 1955] A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc. **16**, Amer. Math. Soc., Providence, RI, 1955. MR 17,763c Zbl 0064.35501

[Inoue and Maeda 1991] A. Inoue and Y. Maeda, "Foundations of calculus on super Euclidean space $\mathbf{R}^{m|n}$ based on a Fréchet–Grassmann algebra", *Kodai Math. J.* **14**:1 (1991), 72–112. MR 93c:58034 Zbl 0737.58009

[Inoue and Maeda 2003] A. Inoue and Y. Maeda, "On a construction of a good parametrix for the Pauli equation by Hamiltonian path-integral method: An application of superanalysis", *Japan. J. Math.* (*N.S.*) **29**:1 (2003), 27–107. MR 2004e:37135 Zbl 1057.81029

[Kontsevich 2003] M. Kontsevich, "Deformation quantization of Poisson manifolds", *Lett. Math. Phys.* **66**:3 (2003), 157–216. MR 2005i:53122 Zbl 1058.53065

[Kostant 1977] B. Kostant, "Graded manifolds, graded Lie theory, and prequantization", pp. 177–306 in *Differential geometrical methods in mathematical physics* (Bonn, 1975), edited by K. Bleuler and A. Reetz, Lecture Notes in Math. **570**, Springer, Berlin, 1977. MR 58 #28326 Zbl 0358.53024

[Majid 1995] S. Majid, *Foundations of quantum group theory*, Cambridge University Press, 1995. MR 97g:17016 Zbl 0857.17009

[Rieffel 1981] M. A. Rieffel, "C^*-algebras associated with irrational rotations", *Pacific J. Math.* **93**:2 (1981), 415–429. MR 83b:46087 Zbl 0499.46039

[Rieffel 1989] M. A. Rieffel, "Deformation quantization of Heisenberg manifolds", *Comm. Math. Phys.* **122**:4 (1989), 531–562. MR 90e:46060 Zbl 0679.46055

[Rieffel 1992] M. A. Rieffel, "Some solvable quantum groups", pp. 146–159 in *Operator algebras and topology* (Craiova, 1989), edited by W. B. Arveson et al., Pitman Res. Notes Math. Ser. **270**, Longman Sci. Tech., Harlow, 1992. MR 93i:46132 Zbl 0804.17010

[Rieffel 1993] M. A. Rieffel, "Compact quantum groups associated with toral subgroups", pp. 465–491 in *Representation theory of groups and algebras*, edited by J. Adams et al., Contemp. Math. **145**, Amer. Math. Soc., Providence, RI, 1993. MR 94i:22022 Zbl 0795.17017

[Rogers 2007] A. Rogers, *Supermanifolds: Theory and applications*, World Scientific, Hackensack, NJ, 2007. MR 2008h:58008 Zbl 1135.58004

[Schwarz 1999] A. Schwarz, "Noncommutative supergeometry and duality", *Lett. Math. Phys.* **50**:4 (1999), 309–321. MR 2001g:58014 Zbl 0967.58004

[Tuynman 2005] G. M. Tuynman, *Supermanifolds and supergroups: Basic theory*, Mathematics and its Applications **570**, Kluwer, Dordrecht, 2005. MR 2005k:58007 Zbl 1083.58001

[Voigt 2008] C. Voigt, "Bornological quantum groups", *Pacific J. Math.* **235**:1 (2008), 93–135. MR 2009c:16129 Zbl 1157.46041

[Woronowicz 1987] S. L. Woronowicz, "Compact matrix pseudogroups", *Comm. Math. Phys.* **111**:4 (1987), 613–665. MR 88m:46079 Zbl 0627.58034

[Woronowicz 1996] S. L. Woronowicz, "From multiplicative unitaries to quantum groups", *Internat. J. Math.* **7**:1 (1996), 127–149. MR 96k:46136 Zbl 0876.46044

Received December 10, 2013. Revised April 7, 2014.

AXEL DE GOURSAC
CHARGÉ DE RECHERCHES AU FRS-FNRS, IRMP
UNIVERSITÉ CATHOLIQUE DE LOUVAIN
CHEMIN DU CYCLOTRON 2
1348 LOUVAIN-LA-NEUVE
BELGIUM

axelmg@melix.net

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.197

GENERATORS OF THE GAUSS–PICARD MODULAR GROUP IN THREE COMPLEX DIMENSIONS

BAOHUA XIE, JIEYAN WANG AND YUEPING JIANG

In this paper, we will describe a method to obtain the generators system of Gauss–Picard modular group PU(3, 1; $\mathbb{Z}[i]$). More precisely, we will show that PU(3, 1; $\mathbb{Z}[i]$) can be generated by five given transformations: two Heisenberg translations, two Heisenberg rotations and one involution. Indeed, the same method works for the other higher-dimensional Euclidean Picard modular groups.

1. Introduction

There are some natural algebraic generalizations of the classical modular group $\mathrm{PSL}(2, \mathbb{Z})$. For example, a Bianchi group is a group of the form $\mathrm{PSL}(2, \mathbb{O}_d)$, where d is a positive square-free integer. Here, PSL denotes the projective special linear group and $\mathbb{O}_d$ is the ring of integers in the imaginary quadratic number field $\mathbb{Q}(\sqrt{-d})$. These groups were first studied by Bianchi [1892] as a natural class of discrete subgroups of $\mathrm{PSL}(2, \mathbb{C})$. A general method to determine finite presentation for each $\mathrm{PSL}(2, \mathbb{O}_d)$ was developed by Swan [1971] based on the geometrical work of Bianchi, while a separate purely algebraic method was given by Cohn [1968]. As another generalization of the modular group, the construction was generalized by Picard [1883; 1884]. Suppose H is a Hermitian matrix of signature (2, 1) with entries in $\mathbb{O}_d$, and let $\mathrm{SU}(\mathrm{H}; \mathbb{O}_d)$ denote the subgroup of SU(H) consisting of those matrices whose entries lie in $\mathbb{O}_d$. Picard studied the group $\mathrm{PU}(\mathrm{H}; \mathbb{O}_d)$ acting on the complex hyperbolic plane $\mathbb{H}^2_{\mathbb{C}}$. Now, Picard modular groups $\mathrm{PU}(\mathrm{H}; \mathbb{O}_d)$ have attracted a great deal of attention both for their intrinsic interest as discrete groups and also for their applications in complex hyperbolic geometry.

One can view the modular group or a Bianchi group acting discontinuously on hyperbolic spaces. Then Poincaré's polyhedra theorem provides a geometric method to obtain their generators from their fundamental polyhedra. But [Mostow 1980] told us that the explicit construction of fundamental domains for lattices in complex hyperbolic spaces was particularly difficult. Until recently, the geometry of $\mathrm{SU}(\mathrm{H}; \mathbb{O}_3)$ had been studied by Falbel and Parker [2006], while the geometry of

MSC2010: primary 32M05, 22E40; secondary 32M15.
Keywords: complex hyperbolic space, Picard modular groups, generators.

SU(H; $\mathbb{O}_1$) had been studied by Francsics and Lax [2005a; 2005b; 2006] and Falbel, Francsics and Parker [Falbel et al. 2011b]. By applying similar ideas to those of [Falbel and Parker 2006; Falbel et al. 2011b], Zhao [2012] obtained generators of the Euclidean Picard groups PU(2, 1; $\mathbb{O}_d$) for $d = 2, 7, 11$.

There are some simple algorithms to obtain the generators of the modular group or some Picard modular groups. For example, the continued fraction algorithm may be applied to any element of the modular group PSL(2, $\mathbb{Z}$). This shows that $S(z) = -1/z$ and $T(z) = z + 1$ generate PSL(2, $\mathbb{Z}$). This algorithm was extended to PU(2, 1; $\mathbb{O}_1$) in [Falbel et al. 2011a], which provided a different system of generators from those obtained via a fundamental domain in [Falbel et al. 2011b]. In [Wang et al. 2011], the authors applied the continued fraction algorithm to PU(2, 1; $\mathbb{O}_3$) and produced a different system of generators from that obtained in [Falbel and Parker 2006].

There is an obvious generalization of Picard modular groups to higher complex dimensions. We observe that very little is known about the geometry and algebraic properties, e.g., explicit fundamental domain or generating system of the higher-dimensional Picard modular groups PU(n, 1; $\mathbb{O}_d$). In [Xie et al. 2013], the continued fraction algorithm was generalized to Picard modular groups in higher complex dimensions. It contained the first generalization that we were aware of to a group of 4×4 matrices. However, it seems very difficult to extend the continued fraction algorithm to other higher-dimensional Picard modular groups. Using a combination of the ideas from [Falbel et al. 2011a; Xie et al. 2013] and [Falbel and Parker 2006; Zhao 2012], we will present a method to obtain the generating system of the Gauss–Picard modular group PU(3, 1; $\mathbb{Z}[i]$). We first get the generators of the stabilizer of infinity of PU(3, 1; $\mathbb{Z}[i]$) by applying a similar argument as in our previous paper [Xie et al. 2013]. Then we will construct a subset in the boundary of complex hyperbolic space which contains the fundamental domain for the stabilizers of infinity in PU(3, 1; $\mathbb{Z}[i]$). Finally, we will show the boundaries of some isometric spheres that contain this subset. This method works for the other higher-dimensional Euclidean Picard modular groups.

2. Preliminaries

2.1. ***The Siegel domain.*** We recall some basic notions of complex hyperbolic geometry. For more details we refer the reader to [Goldman 1999; Parker 2010].

Let $\mathbb{C}^{n,1}$ denote the vector space $\mathbb{C}^{n+1}$ equipped with the Hermitian form of signature $(n, 1)$ given by the matrix

$$J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & I_{n-1} & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The Hermitian product of two vectors $\boldsymbol{z}$ and $\boldsymbol{w}$ is given by $\langle \boldsymbol{z}, \boldsymbol{w}\rangle = \boldsymbol{w}^* J \boldsymbol{z}$, where $\boldsymbol{w}^*$ denotes the Hermitian transpose of $\boldsymbol{w}$.

We denote by V_- and V_0 the negative and null cones associated to the Hermitian form, respectively. The complex hyperbolic n-space $\mathbb{H}^n_{\mathbb{C}}$ is the projectivization of V_-, and its boundary is the projectivization of V_0. The model of $\mathbb{H}^n_{\mathbb{C}}$ associated to the Hermitian form given above is often referred to as the Siegel model of $\mathbb{H}^n_{\mathbb{C}}$.

We define the Siegel domain $\mathfrak{S}$ of the complex hyperbolic n-space $\mathbb{H}^n_{\mathbb{C}}$ by identifying points of $\mathfrak{S}$ with their horospherical coordinates,

$$z = (\zeta, v, u) \in \mathbb{C}^{n-1} \times \mathbb{R} \times \mathbb{R}^+.$$

The boundary of $\mathfrak{S}$ is given by $H_0 \cup \{q_\infty\}$, where q_∞ is a distinguished point at infinity and $H_0 = \mathbb{C}^{n-1} \times \mathbb{R} \times \{0\}$.

2.2. ***Heisenberg group.*** The boundary of a complex hyperbolic space is identified with the one-point compactification of the Heisenberg group. The $(2n-1)$-dimensional Heisenberg group $\mathcal{H}_{2n-1}$ is $\mathbb{C}^{n-1} \times \mathbb{R}$ with the group law

$$(\xi, \nu) \cdot (z, u) = (\xi + z, \nu + u + 2\Im\langle\langle \xi, z\rangle\rangle).$$

Here $\langle\langle \xi, z\rangle\rangle = z^*\xi$ is the standard positive definite Hermitian form on $\mathbb{C}^{n-1}$. In particular, we write $\|\xi\|^2 = \xi^*\xi$.

The Heisenberg group acts on itself by Heisenberg translation. For $(\tau, t) \in \mathcal{H}_{2n-1}$, this translation is

$$N_{(\tau,t)}(\xi, \nu) = (\tau + \xi, t + \nu + 2\Im\langle\langle \tau, \xi\rangle\rangle).$$

The unitary group $U(n-1)$ acts on the Heisenberg group by Heisenberg rotation.

2.3. ***Holomorphic isometries.*** Define a map $\mathfrak{S} \to \mathbb{CP}^n$ by

$$\psi : (\xi, \nu, u) \mapsto \begin{pmatrix} \frac{1}{2}(-\|\xi\|^2 - u + i\nu) \\ \xi \\ 1 \end{pmatrix}, \quad \psi : q_\infty \mapsto \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then ψ maps the set of points $z \in \mathfrak{S}$ homeomorphically to the set of points $\boldsymbol{z} \in \mathbb{CP}^n$ with $\langle \boldsymbol{z}, \boldsymbol{z}\rangle < 0$, and maps the set of points in $\partial\mathfrak{S}$ homeomorphically to the set of points $\boldsymbol{z} \in \mathbb{CP}^n$ with $\langle \boldsymbol{z}, \boldsymbol{z}\rangle = 0$. We write $\psi(z) = \boldsymbol{z}$.

The Bergman metric on $\mathfrak{S}$ is given by the distance formula

$$\cosh^2 \frac{\rho(z, w)}{2} = \frac{\langle \boldsymbol{z}, \boldsymbol{w}\rangle\langle \boldsymbol{w}, \boldsymbol{z}\rangle}{\langle \boldsymbol{z}, \boldsymbol{z}\rangle\langle \boldsymbol{w}, \boldsymbol{w}\rangle}.$$

The holomorphic isometry group of $\mathfrak{S}$ with respect to the Bergman metric is the projective unitary group $\mathrm{PU}(n, 1)$, and it acts on $\mathbb{CP}^n$ by matrix multiplication.

2.4. ***Picard modular groups.*** Let $\mathcal{O}_d$ be the ring of integers in the imaginary quadratic number field $\mathbb{Q}(i\sqrt{d})$, where d is a positive square-free integer. If $d \equiv 1, 2 \bmod 4$, then $\mathcal{O}_d = \mathbb{Z}[\sqrt{d}i]$, and if $d \equiv 3 \bmod 4$, then $\mathcal{O}_d = \mathbb{Z}[(1+\sqrt{d}i)/2]$. The subgroup of $\mathrm{PU}(n, 1)$ with entries in $\mathcal{O}_d$ is called the Picard modular group for $\mathcal{O}_d$ and is written as $\mathrm{PU}(n, 1; \mathcal{O}_d)$. Obviously, if $d = 1$, then the ring $\mathcal{O}_d$ can be written as $\mathbb{Z}[i]$.

Remark 1. The matrices corresponding to the generators obtained in this paper belong to the group $U(3, 1; \mathbb{Z}[i])$. In relation to complex hyperbolic isometries, the relevant group is $\mathrm{PU}(3, 1; \mathbb{Z}[i]) = \mathrm{SU}(3, 1; \mathbb{Z}[i])/\mathbb{Z}_4$. The center of $\mathrm{SU}(3, 1)$ is isomorphic to $\mathbb{Z}_4$, the group of fourth roots of unity. By abuse of notation, we will denote the Gauss–Picard modular group in three complex dimensions by $U(3, 1; \mathbb{Z}[i])$.

2.5. ***Heisenberg automorphism groups.*** The action of Heisenberg isometries extends to the Siegel domain, fixing q_∞. Some examples of Heisenberg isometries are as follows: for $U \in U(n-1)$ and $(\tau, t) \in \mathcal{H}_{2n-1}$, the Heisenberg rotation and Heisenberg translation correspond to the matrices

$$M_U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & U & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad N_{(\tau,t)} = \begin{pmatrix} 1 & -\tau^* & \frac{1}{2}(-\|\tau\|^2 + it) \\ 0 & I_{n-1} & \tau \\ 0 & 0 & 1 \end{pmatrix}$$

in $\mathrm{SU}(n, 1)$, respectively. The Heisenberg dilation by r fixing q_∞ and 0 corresponds to the matrix $A_r \in \mathrm{SU}(n, 1)$, where

$$A_r = \begin{pmatrix} r & 0 & 0 \\ 0 & I_{n-1} & 0 \\ 0 & 0 & 1/r \end{pmatrix}.$$

Finally, the Heisenberg inversion interchanging q_∞ and 0 corresponds to the matrix $R \in \mathrm{SU}(n, 1)$, where

$$R = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -I_{n-1} & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

2.6. ***Isometric spheres.*** Given an element $G \in \mathrm{PU}(3, 1)$ such that $G(q_\infty) \neq q_\infty$, we define the isometric sphere of G to be the hypersurface

$$\{z \in \mathbb{H}^3_{\mathbb{C}} : |\langle z, q_\infty\rangle| = |\langle z, G^{-1}(q_\infty)\rangle|\}.$$

For example, the isometric sphere of

$$R = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

is

$$\mathcal{B}_0 = \big\{(\zeta_1, \zeta_2, t, u) \in \mathfrak{S} : \big||\zeta_1|^2 + |\zeta_2|^2 + u + it\big| = 2\big\}$$

in horospherical coordinates.

All other isometric spheres are images of $\mathcal{B}_0$ by Heisenberg dilations, rotations and translations. Thus, the isometric sphere with radius r and center $(\zeta_1^0, \zeta_2^0, t^0, 0)$ is given by

$$\big\{(\zeta_1, \zeta_2, t, u) : \big||\zeta_1 - \zeta_1^0|^2 + |\zeta_2 - \zeta_2^0|^2 + u + it - it^0 + 2i\Im(\zeta_1\overline{\zeta_1^0} + \zeta_2\overline{\zeta_2^0})\big| = r^2\big\}.$$

If G has the matrix form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix},$$

then $G(q_\infty) \neq q_\infty$ if and only if $a_{41} \neq 0$. The isometric sphere of G has radius $r = \sqrt{2/|a_{41}|}$ and center $G^{-1}(q_\infty)$, which in horospherical coordinates is

$$(\zeta_1^0, \zeta_2^0, t^0, 0) = (\overline{a_{42}}/\overline{a_{41}}, \overline{a_{43}}/\overline{a_{41}}, 2\Im(\overline{a_{44}}/\overline{a_{41}}), 0).$$

3. The generators of the stabilizer

Let Γ_∞ be the stabilizer subgroup of q_∞ in $\mathrm{PU}(n, 1)$. That is,

$$\Gamma_\infty \equiv \{g \in \mathrm{PU}(n, 1) : g(q_\infty) = q_\infty\}.$$

We recall from [Falbel et al. 2011a; Francsics and Lax 2005a; 2005b; Xie et al. 2013] that the Langlands decomposition can be used to parametrize a transformation in the stabilizer subgroup of q_∞.

Lemma 2 (Langlands decomposition). *Any element $P \in \Gamma_\infty$ can be decomposed as a product of a Heisenberg translation, dilation, and a rotation*:

$$P = N_{(\tau,t)} A_r M_U = \begin{pmatrix} r & -\tau^* U & (-\|\tau\|^2 + it)/2r \\ 0 & U & \tau/r \\ 0 & 0 & 1/r \end{pmatrix}.$$

The parameters satisfy the corresponding conditions. That is, $U \in U(n-1)$, $r \in \mathbb{R}^+$ and $(\tau, t) \in \mathcal{H}_{2n-1}$.

First, we describe the Heisenberg rotations in the Gauss–Picard modular group $U(3,1;\mathbb{Z}[i])$. Let $U(2;\mathbb{Z}[i])$ be the unitary group $U(2)$ over the ring $\mathbb{Z}[i]$. Then we have the following result.

Lemma 3. *$U(2;\mathbb{Z}[i])$ can be generated by the two unitary matrices*

$$U_1=\begin{pmatrix}0&1\\1&0\end{pmatrix}\quad \textit{and}\quad U_2=\begin{pmatrix}i&0\\0&1\end{pmatrix}.$$

Remark 4. A similar lemma was proved for $U(2;\mathbb{Z}[(1+\sqrt{-3})/2])$ in [Xie et al. 2013].

Next, we characterize the elements of the stabilizer subgroup Γ_∞ of infinity in the Picard modular group $U(3,1;\mathbb{Z}[i])$. We denote this stabilizer by $\Gamma_\infty(3,1;\mathbb{Z}[i])$.

Lemma 5. *An element $P\in U(3,1;\mathbb{Z}[i])$ lies in $\Gamma_\infty(3,1;\mathbb{Z}[i])$ if and only if the parameters in the Langlands decomposition of P satisfy the conditions*

$$r=1,\quad t\in 2\mathbb{Z},\quad \tau=(\tau_1,\tau_2)^T\in\mathbb{Z}[i]^2,\quad U\in U(2;\mathbb{Z}[i]),\quad \|\tau\|\in 2\mathbb{Z}.$$

Proof. The proof of this lemma follows from the Langlands decomposition form of $P\in\Gamma_\infty(3,1;\mathbb{Z}[i])$. □

We are now in a position to determine the generators of the stabilizer subgroup of q_∞.

Proposition 6. *Let $\Gamma_\infty(3,1;\mathbb{Z}[i])$ be stated as above. Then $\Gamma_\infty(3,1;\mathbb{Z}[i])$ is generated by the Heisenberg translations $N_{((1,1)^T,0)}$, $N_{((0,0)^T,2)}$ and the Heisenberg rotations M_{U_i} $(i=1,2)$.*

Proof. Our proof starts with the observation that there is no dilation component of $P\in\Gamma_\infty(3,1;\mathbb{Z}[i])$ in its Langlands decomposition. That is, P must have the form

$$P=N_{(\tau,t)}M_U=\begin{pmatrix}1&-\tau^*&(-||\tau||^2+it)/2\\0&I_2&\tau\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&U&0\\0&0&1\end{pmatrix}.$$

Since the unitary matrix U lies in $U(2;\mathbb{Z}[i])$, the rotation component of P in the Langlands decomposition is generated by M_{U_i} $(i=1,2)$ by Lemma 3.

What is left is to consider the Heisenberg translation part $N_{(\tau,t)}$ of P. Let

$$\tau=(m_1+n_1i,m_2+n_2i)^T,$$

where $m_1,n_1,m_2,n_2\in\mathbb{Z}$. Since $|\tau|^2=m_1^2+n_1^2+m_2^2+n_2^2\in 2\mathbb{Z}$, there are two cases:

(1) $m_1^2+n_1^2\in 2\mathbb{Z}$ and $m_2^2+n_2^2\in 2\mathbb{Z}$;

(2) $m_1^2+n_1^2\in 2\mathbb{Z}+1$ and $m_2^2+n_2^2\in 2\mathbb{Z}+1$.

We first consider the case (1). We can write τ as

$$\tau = \big(k_1(1+i)+l_1(1-i),\, k_2(1+i)+l_2(1-i)\big),$$

where $k_1, l_1, k_2, l_2 \in \mathbb{Z}$. $N_{(\tau,t)}$ splits as

$$N_{(\tau,t)} = N_{((0,0)^T,t)} \circ N_{(\tau,0)}.$$

Since $t = 2k \in 2\mathbb{Z}$, $N_{((0,0)^T,t)} = N^k_{((0,0)^T,2)}$. We also have

$$N_{(\tau,0)} = N^{k_1}_{((1+i,0)^T,0)} \circ N^{l_1}_{((i-1,0)^T,0)} \circ N^{2k_1l_1}_{((0,0)^T,2)} \circ N^{k_2}_{((0,1+i)^T,0)} \circ N^{l_2}_{((0,1+i)^T,0)} \circ N^{-2k_2l_2}_{((0,0)^T,2)}.$$

We observe that

$$\begin{aligned} N_{((1+i,0)^T,0)} &= N_{((1,1)^T,0)} \circ N_{((i,-1)^T,0)} \circ N_{((0,0)^T,2)},\\ N_{((i-1,0)^T,0)} &= N_{((i,1)^T,0)} \circ N^{-1}_{((1,1)^T,0)} \circ N^{-1}_{((0,0)^T,2)},\\ N_{((0,1+i)^T,0)} &= N_{((1,1)^T,0)} \circ N_{((-1,i)^T,0)} \circ N_{((0,0)^T,2)},\\ N_{((0,i-1)^T,0)} &= N_{((1,i)^T,0)} \circ N^{-1}_{((1,1)^T,0)} \circ N^{-1}_{((0,0)^T,2)}. \end{aligned}$$

It is easy to see that

$$\begin{aligned} N_{((i,1)^T,0)} &= M_{U_2} N_{((1,1)^T,0)} M^{-1}_{U_2},\\ N_{((i,-1)^T,0)} &= M_{U_1} M^2_{U_2} M_{U_1} M_{U_2} N_{((1,1)^T,0)} M^3_{U_2} (M_{U_1} M_{U_2} M_{U_1})^2,\\ N_{((-1,i)^T,0)} &= M^2_{U_2} M_{U_1} M_{U_2} M_{U_1} N_{((1,1)^T,0)} M^2_{U_2} (M_{U_1} M_{U_2} M_{U_1})^3,\\ N_{((1,i)^T,0)} &= M_{U_1} M_{U_2} M_{U_1} N_{((1,1)^T,0)} (M_{U_1} M_{U_2} M_{U_1})^3. \end{aligned}$$

In case (2), similar considerations apply to the translation $N_{(\tau,0)} \circ N_{((1,1)^T,0)}$, where $N_{(\tau,0)}$ belongs to case (1). □

4. Fundamental domain for the stabilizer in PU(2, 1; $\mathbb{Z}[i]$)

In [Falbel et al. 2011b], the authors described a method to find the fundamental domain for the stabilizer of q_∞ in the Gauss–Picard modular group PU(2, 1; $\mathbb{Z}[i]$) in two complex dimensions. We review it now.

Let Γ be PU(2, 1; $\mathbb{Z}[i]$) and Γ_∞ be the stabilizer of q_∞. Every element of Γ_∞ is upper triangular, and its diagonal entries are units in $\mathbb{Z}[i]$. Recall that the units of $\mathbb{Z}[i]$ are $\pm 1, \pm i$. Therefore Γ_∞ contains no dilations and fits into the exact sequence

$$0 \longrightarrow \mathbb{Z} \longrightarrow \Gamma_\infty \xrightarrow{\Pi_*} \Delta \longrightarrow 1,$$

where $\Delta \subset \operatorname{Isom} \mathbb{Z}[i]$ is of index 2 and Π is the vertical projection defined by $\Pi : (z, t) \in \mathcal{H} \mapsto z \in \mathbb{C}$.

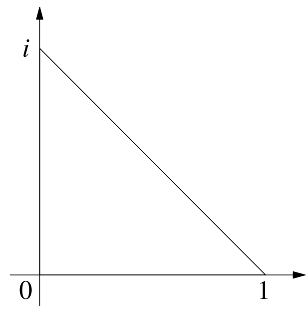

Figure 1. A fundamental domain for the index two subgroup $\Delta \subset \operatorname{Isom} \mathbb{Z}[i]$ is a triangle $\triangle$ with vertices $0, 1, i$.

As the first step toward the construction of a fundamental domain for the action of Γ_∞ on $\mathscr{H}$, one should construct a fundamental domain in $\mathbb{C}$ of $\Delta \subseteq \operatorname{Isom} \mathbb{Z}[i]$. From the generators of Δ, one finds that a fundamental domain for $\Delta \subseteq \operatorname{Isom} \mathbb{Z}[i]$ is the triangle $\triangle$ with vertices $0, 1, i$; see Figure 1.

In order to produce a fundamental domain for Γ_∞, we look at all the preimages of the triangle (that is, a fundamental domain of $\Pi_*(\Gamma_\infty)$) under the vertical projection Π and we intersect this with a fundamental domain for $\ker(\Pi_*)$. The inverse image of the triangle under Π is an infinite prism. The kernel of Π_* is the infinite cyclic group generated by T, the vertical translation by $(0, 2)$. Hence, a fundamental domain for Γ_∞ is the prism in $\mathscr{H}$ with vertices $(0, \pm 1)$, $(1, \pm 1)$, $(i, \pm 1)$.

5. Statement of the results

In this section, we recall the geometric method used in [Falbel and Parker 2006; Falbel et al. 2011b] to determine the generators of the Euclidean Picard groups, and then state our method and results.

The geometric method is based on the special feature that the Euclidean Picard modular orbifold has only one cusp for $d = 1, 2, 3, 7, 11$. The basic idea of the proof can be described easily. Analogously to Theorem 3.5 of [Falbel and Parker 2006], it can be proved that $\langle \Gamma_\infty, R \rangle$ has only one cusp. The fact that $\mathrm{PU}(2, 1; \mathbb{O}_d)$ has the same cusp and the stabilizer of infinity as the group generated by $\langle \Gamma_\infty, R \rangle$ shows that they are the same. The key step is to find a union of isometric spheres so that a fundamental domain for Γ_∞ is contained in the intersection of their exteriors and a fundamental domain for the stabilizer, which implies that the group $\langle \Gamma_\infty, R \rangle$ has only one cusp. In other words, one should show that the union of the boundaries of these isometric spheres in the Heisenberg group contains a fundamental domain for the stabilizer of infinity.

We will prove our result by using a similar idea. The main observation is that there is no need to know the exact fundamental domain for the stabilizer of infinity.

We will construct a set in the Heisenberg group which contains a fundamental set for the stabilizer of infinity as a subset. Then we show that the union of the boundaries of some isometric spheres in the Heisenberg group covers this set. This also show that the group $\langle \Gamma_\infty, R\rangle$ has only one cusp.

More precisely, let Σ be the set

$$\{(\xi_1, \xi_2, t) : \xi_i \in \triangle, -1 \le t \le 1\}.$$

Here $\triangle$ is the fundamental domain of $\Delta \subset \operatorname{Isom} \mathbb{Z}[i]$.

Note that Σ is not a fundamental domain for the stabilizer of infinity because this set is preserved by some Heisenberg rotations.

Proposition 7. *Σ contains a fundamental domain for the stabilizer of infinity.*

Proof. The restriction of the action of the stabilizer of infinity on each copy of $\mathbb{C}$ has the same fundamental domain $\triangle$ as $\Delta \subset \operatorname{Isom} \mathbb{Z}[i]$. Then Σ is the preimage of $\triangle \times \triangle$ under vertical projection intersected with a fundamental domain for the vertical translation by $((0, 0)^T, 2)$. It is clear that Σ is preserved by the Heisenberg rotations M_{U_1}. Hence, a fundamental domain for Γ_∞ lies inside Σ. □

In next section we will prove our main theorem. Our main step is to show that Σ lies inside the boundaries of some isometric spheres in the Heisenberg group. It is obvious that the geodesic cone from q_∞ over Σ contains a fundamental domain for the Gauss–Picard modular group $U(3, 1; \mathbb{Z}[i])$.

Theorem 8. *The Picard modular group $U(3, 1; \mathbb{Z}[i])$ is generated by the Heisenberg translations*

$$N_{((1,1)^T,0)} = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad N_{((0,0)^T,2)} = \begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

the Heisenberg rotations

$$M_{U_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad M_{U_2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and the involution

$$R = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

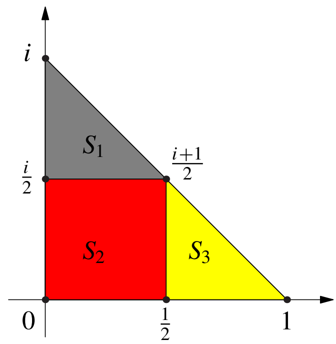

Figure 2. The decomposition of the fundamental domain $\triangle$ for $\Delta \subset \operatorname{Isom} \mathbb{Z}[i]$ into three parts.

6. Proof of Theorem 8

In this section we will prove that the generators of Picard modular groups consist of the generators of the stabilizer and the involution.

Recall that the Cygan sphere $\mathcal{B}_0$ is the isometric sphere of R. The boundary $\mathcal{S}_0$ of $\mathcal{B}_0$ is called the Heisenberg sphere in the Heisenberg group. $\mathcal{S}_0$ is defined by

$$\mathcal{S}_0 = \left\{ \left| |\xi_1|^2 + |\xi_2|^2 + ti \right| = 2 \right\}.$$

Indeed, we only need to consider the boundaries of isometric spheres in the Heisenberg group because two isometric spheres have a nonempty interior intersection if and only if the boundaries have a nonempty interior intersection.

It is not hard to see that parts of Σ lie outside $\mathcal{S}_0$. Therefore we need to find more isometric spheres whose boundaries together with $\mathcal{S}_0$ contain the set Σ.

Note that Σ has the form

$$\Sigma = \{(\xi_1, \xi_2, t) : \xi_1 \in \triangle, \xi_2 \in \triangle, -1 \le t \le 1\}.$$

First, we decompose $\triangle$ into three parts. We write $\triangle = S_1 \cup S_2 \cup S_3$, where S_1 is a triangle with vertices i, $\frac{i}{2}$, $\frac{1}{2}(1+i)$, S_2 is a square with vertices 0, $\frac{i}{2}$, $\frac{1}{2}$, $\frac{1}{2}(1+i)$, and S_3 is a triangle with vertices 0, 1, $\frac{1}{2}(1+i)$; see Figure 2.

Therefore, Σ will be decomposed into nine subsets:

- $\Sigma_1 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_1, \xi_2 \in S_1, -1 \le t \le 1\}$,
- $\Sigma_2 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_1, \xi_2 \in S_2, -1 \le t \le 1\}$,
- $\Sigma_3 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_1, \xi_2 \in S_3, -1 \le t \le 1\}$,
- $\Sigma_4 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_2, \xi_2 \in S_1, -1 \le t \le 1\}$,
- $\Sigma_5 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_2, \xi_2 \in S_2, -1 \le t \le 1\}$,
- $\Sigma_6 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_2, \xi_2 \in S_3, -1 \le t \le 1\}$,

- $\Sigma_7 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, -1 \le t \le 1\}$,
- $\Sigma_8 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_2, -1 \le t \le 1\}$,
- $\Sigma_9 = \{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_3, -1 \le t \le 1\}$.

We first prove that $\mathcal{S}_0$ covers the subsets Σ_2, Σ_4, Σ_5, Σ_6, Σ_8.

If $(\xi_1, \xi_2, t) \in \Sigma_5$, then

$$|\xi_1|^2 + |\xi_2|^2 \le \left(\tfrac{\sqrt{2}}{2}\right)^2 + \left(\tfrac{\sqrt{2}}{2}\right)^2 = 1,$$

so

$$\big||\xi_1|^2 + |\xi_2|^2 + it\big| \le \sqrt{1+1} = \sqrt{2} < 2.$$

Hence $\Sigma_5 \subset \mathcal{S}_0$.

If $(\xi_1, \xi_2, t) \in \Sigma_2$, then

$$|\xi_1|^2 + |\xi_2|^2 \le 1 + \tfrac{\sqrt{2}}{2} = \tfrac{3}{2},$$

so

$$\big||\xi_1|^2 + |\xi_2|^2 + it\big| \le \sqrt{\left(\tfrac{3}{2}\right)^2 + 1} = \sqrt{\tfrac{13}{4}} < 2.$$

Therefore $\Sigma_2 \subset \mathcal{S}_0$.

Similarly, we can show that Σ_4, Σ_6, Σ_8 are included in $\mathcal{S}_0$.

In order to prove this theorem, it is sufficient to prove that the remaining four subsets are covered by some Heisenberg spheres.

For the set Σ_9, we consider the map $N_{((1,1)^T,0)} R N^{-1}_{((1,1)^T,0)}$. The isometric sphere $\mathcal{B}_1$ of this map is the Cygan sphere centered at the point $((1, 1)^T, 0, 0)$ (in horospherical coordinates) with radius 1. The boundary of $\mathcal{B}_1$ is a Heisenberg sphere given by

$$\mathcal{S}_1 = \Big\{\big||\xi_1 - 1|^2 + |\xi_2 - 1|^2 + i(t + 2\Im(\xi_1 + \xi_2))\big| = 2\Big\}.$$

If $(\xi_1, \xi_2, t) \in \Sigma_9$, then $\xi_1 \in S_3$, $\xi_2 \in S_3$, $-1 \le t \le 1$. We get that

$$0 \le \Im\xi_i \le \tfrac{1}{2}, \quad |\xi_i - 1|^2 \le \tfrac{1}{2},$$

so

$$-1 \le t + 2\Im(\xi_1 + \xi_2) \le 3.$$

Let $T = N_{((0,0)^T,2)}$. It is easy to see that the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_3, -1 \le t + 2\Im(\xi_1 + \xi_2) \le 1\}$$

lies inside $\mathcal{S}_1$ and the set

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_3, 1 \le t + 2\Im(\xi_1 + \xi_2) \le 3\}$$

lies inside

$$T^{-1}(\mathcal{S}_1) = \big\{ \big| |\xi_1 - 1|^2 + |\xi_2 - 1|^2 + i(t - 2 + 2\Im(\xi_1 + \xi_2)) \big| = 2 \big\}.$$

Therefore, $\mathcal{S}_1$ and $T^{-1}(\mathcal{S}_1)$ cover the set Σ_9.

For the set Σ_7, we consider the map $N_{((1,i)^T,0)} R N^{-1}_{((1,i)^T,0)}$. The isometric sphere $\mathcal{B}_2$ of this map is the Cygan sphere centered at the point $((1,i)^T, 0, 0)$. The boundary of $\mathcal{B}_2$ is given by

$$\mathcal{S}_2 = \big\{ \big| |\xi_1 - 1|^2 + |\xi_2 - i|^2 + i(t + 2\Im(\xi_1) + 2\Re(\xi_2)) \big| = 2 \big\}.$$

If $(\xi_1, \xi_2, t) \in \Sigma_7$, then $\xi_1 \in S_3$, $\xi_2 \in S_1$, $-1 \le t \le 1$. We get that

$$0 \le \Im\xi_1 \le \tfrac{1}{2}, \quad 0 \le \Re\xi_2 \le \tfrac{1}{2}, \quad |\xi_1 - 1|^2 \le \tfrac{1}{2}, \quad |\xi_2 - i|^2 \le \tfrac{1}{2},$$

so

$$-2 \le t + 2\Im(\xi_1 + \xi_2) \le 2.$$

If $-1 \le t + 2\Im(\xi_1 + \xi_2) \le 1$, then the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, -1 \le t + 2\Im(\xi_1 + \xi_2) \le 1\}$$

lies inside $\mathcal{S}_2$.

If $-2 \le t + 2\Im(\xi_1 + \xi_2) \le -1$, then the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, -2 \le t + 2\Im(\xi_1 + \xi_2) \le -1\}$$

lies inside $T(\mathcal{S}_2)$.

If $1 \le t + 2\Im(\xi_1 + \xi_2) \le 2$, then the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, 1 \le t + 2\Im(\xi_1 + \xi_2) \le 2\}$$

lies inside $T^{-1}(\mathcal{S}_2)$.

For the set Σ_1, we consider the map $N_{((i,i)^T,0)} R N^{-1}_{((i,i)^T,0)}$. The isometric sphere $\mathcal{B}_3$ of this map is the Cygan sphere centered at the point $((i,i)^T, 0, 0)$. The boundary of $\mathcal{B}_3$ is a Heisenberg sphere given by

$$\mathcal{S}_3 = \big\{ \big| |\xi_1 - 1|^2 + |\xi_2 - 1|^2 + i(t + 2\Im(\xi_1 + \xi_2)) \big| = 2 \big\}.$$

If $(\xi_1, \xi_2, t) \in \Sigma_1$, then $\xi_1 \in S_1$, $\xi_2 \in S_1$, $-1 \le t \le 1$. We get that

$$0 \le \Re\xi_i \le \tfrac{1}{2}, |\xi_i - 1|^2 \le \tfrac{1}{2},$$

so

$$-3 \le t + 2\Im(\xi_1 + \xi_2) \le 1.$$

As before, we can see that Σ_1 is covered by the Heisenberg spheres corresponding to the maps

$$N_{((i,i)^T,0)} R N^{-1}_{((i,i)^T,0)} \quad \text{and} \quad T N_{((i,i)^T,0)} R N^{-1}_{((i,i)^T,0)} T^{-1}.$$

It remains to consider the set Σ_3. We consider the map $N_{((i,1)^T,0)} R N^{-1}_{((i,1)^T,0)}$. The isometric sphere $\mathcal{B}_4$ of this map is the Cygan sphere centered at the point $((i, 1)^T, 0, 0)$. The boundary of $\mathcal{B}_4$ is a Heisenberg sphere given by

$$\mathcal{S}_4 = \left\{ \left| |\xi_1 - i|^2 + |\xi_2 - 1|^2 + i(t - 2\Re(\xi_1) + 2\Im(\xi_2)) \right| = 2 \right\}.$$

If $(\xi_1, \xi_2, t) \in \Sigma_1$, then $\xi_1 \in S_1$, $\xi_2 \in S_3$, $-1 \le t \le 1$. We get that

$$0 \le \Re\xi_1 \le \tfrac{1}{2}, \quad 0 \le \Im\xi_1 \le \tfrac{1}{2}, \quad |\xi_1 - i|^2 \le \tfrac{1}{2}, \quad |\xi_2 - 1|^2 \le \tfrac{1}{2},$$

so

$$-2 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le 2.$$

If $-1 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le 1$, then the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, -1 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le 1\}$$

lies inside $\mathcal{S}_4$.

If $-2 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le -1$, then the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, -2 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le -1\}$$

lies inside $T(\mathcal{S}_4)$.

If $1 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le 2$, then the subset

$$\{(\xi_1, \xi_2, t) : \xi_1 \in S_3, \xi_2 \in S_1, 1 \le t - 2\Re(\xi_1) + 2\Im(\xi_2) \le 2\}$$

lies inside $T^{-1}(\mathcal{S}_4)$. Thus Σ_3 is covered by the Heisenberg spheres corresponding to the maps

$$N_{((i,1)^T,0)} R N^{-1}_{((i,1)^T,0)}, \quad T N_{((i,1)^T,0)} R N^{-1}_{((i,1)^T,0)} T^{-1},$$

and

$$T^{-1} N_{((i,1)^T,0)} R N^{-1}_{((i,1)^T,0)} T.$$

Remark 9. This method works for the other higher-dimensional Euclidean Picard modular groups $\mathrm{PU}(n, 1; \mathcal{O}_d)$ for $d = 2, 7, 11$ and $n \ge 3$. But the calculation will be more complicated. For example, the set Σ will be decomposed into smaller parts. Then one needs more Heisenberg spheres to cover the set Σ which contains the fundamental set.

Acknowledgements

We would like to thank D. Allcock for his interest in our work and for explaining to us some results in [Allcock 1999; 2000a; 2000b]. This work was supported by NSF (grant number 11071059) and NSF (grant number 11371126). B. Xie was also supported by NSF (grant number 11201134) and the young teachers support program of Hunan University.

References

[Allcock 1999] D. Allcock, "Reflection groups on the octave hyperbolic plane", *J. Algebra* **213**:2 (1999), 467–498. MR 2000e:17028 Zbl 0932.51002

[Allcock 2000a] D. Allcock, "New complex- and quaternion-hyperbolic reflection groups", *Duke Math. J.* **103**:2 (2000), 303–333. MR 2001f:11105 Zbl 0962.22007

[Allcock 2000b] D. Allcock, "The Leech lattice and complex hyperbolic reflections", *Invent. Math.* **140**:2 (2000), 283–301. MR 2002b:11091 Zbl 1012.11053

[Bianchi 1892] L. Bianchi, "Sui gruppi di sostituzioni lineari con coefficienti appartenenti a corpi quadratici immaginarî", *Math. Ann.* **40**:3 (1892), 332–412. MR 1510727 JFM 24.0188.02

[Cohn 1968] P. M. Cohn, "A presentation of SL_2 for Euclidean imaginary quadratic number fields", *Mathematika* **15** (1968), 156–163. MR 38 #4568 Zbl 0169.34501

[Falbel and Parker 2006] E. Falbel and J. R. Parker, "The geometry of the Eisenstein–Picard modular group", *Duke Math. J.* **131**:2 (2006), 249–289. MR 2007f:22011 Zbl 1109.22007

[Falbel et al. 2011a] E. Falbel, G. Francsics, P. D. Lax, and J. R. Parker, "Generators of a Picard modular group in two complex dimensions", *Proc. Amer. Math. Soc.* **139**:7 (2011), 2439–2447. MR 2012d:22014 Zbl 1222.32038

[Falbel et al. 2011b] E. Falbel, G. Francsics, and J. R. Parker, "The geometry of the Gauss–Picard modular group", *Math. Ann.* **349**:2 (2011), 459–508. MR 2011k:22009 Zbl 1213.14049

[Francsics and Lax 2005a] G. Francsics and P. D. Lax, "A semi-explicit fundamental domain for a Picard modular group in complex hyperbolic space", pp. 211–226 in *Geometric analysis of PDE and several complex variables*, edited by S. Chanillo et al., Contemp. Math. **368**, Amer. Math. Soc., Providence, RI, 2005. MR 2006b:22011 Zbl 1065.22007

[Francsics and Lax 2005b] G. Francsics and P. Lax, "An explicit fundamental domain for the Picard modular group in two complex dimensions", preprint, 2005. arXiv math/0509708

[Francsics and Lax 2006] G. Francsics and P. D. Lax, "Analysis of a Picard modular group", *Proc. Natl. Acad. Sci. USA* **103**:30 (2006), 11103–11105. MR 2007h:11048 Zbl 1206.11048

[Goldman 1999] W. M. Goldman, *Complex hyperbolic geometry*, Clarendon Press, Oxford, 1999. MR 2000g:32029 Zbl 0939.32024

[Mostow 1980] G. D. Mostow, "On a remarkable class of polyhedra in complex hyperbolic space", *Pacific J. Math.* **86**:1 (1980), 171–276. MR 82a:22011 Zbl 0456.22012

[Parker 2010] J. R. Parker, "Notes on complex hyperbolic geometry", preprint, 2010, Available at http://www.icts.res.in/media/uploads/Program/Files/NCHG.pdf.

[Picard 1883] E. Picard, "Sur des fonctions de deux variables indépendantes analogues aux fonctions modulaires", *Acta Math.* **2**:1 (1883), 114–135. MR 1554595 JFM 15.0432.01

[Picard 1884] E. Picard, "Sur les formes quadratiques ternaires indéfinies á indéterminées conjuguées et sur les fonctions hyperfuchsiennes correspondantes", *Acta Math.* **5**:1 (1884), 121–182. MR 1554651 JFM 16.0385.01

[Swan 1971] R. G. Swan, "Generators and relations for certain special linear groups", *Advances in Math.* **6** (1971), 1–77. MR 44 #1741 Zbl 0221.20060

[Wang et al. 2011] J. Wang, Y. Xiao, and B. Xie, "Generators of the Eisenstein–Picard modular group", *J. Aust. Math. Soc.* **91**:3 (2011), 421–429. MR 2900617 Zbl 1246.32023

[Xie et al. 2013] B. Xie, J. Wang, and Y. Jiang, "Generators of the Eisenstein–Picard modular group in three complex dimensions", *Glasg. Math. J.* **55**:3 (2013), 645–654. MR 3084667 Zbl 1273.32027

[Zhao 2012] T. Zhao, "Generators for the Euclidean Picard modular groups", *Trans. Amer. Math. Soc.* **364**:6 (2012), 3241–3263. MR 2888244 Zbl 1250.22014

Received January 11, 2014. Revised April 27, 2014.

BAOHUA XIE
COLLEGE OF MATHEMATICS AND ECONOMETRICS
HUNAN UNIVERSITY
CHANGSHA, 410082
CHINA
xiexbh@hnu.edu.cn

JIEYAN WANG
ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
BEIJING, 100190
CHINA
jywang@hnu.edu.cn

YUEPING JIANG
COLLEGE OF MATHEMATICS AND ECONOMETRICS
HUNAN UNIVERSITY
CHANGSHA, 410082
CHINA
ypjiang@hnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.213

COMPLETE CHARACTERIZATION OF ISOLATED HOMOGENEOUS HYPERSURFACE SINGULARITIES

STEPHEN YAU AND HUAIQING ZUO

Dedicated to Professor Michael Artin on the occasion of his 80th birthday

Let X be a nonsingular projective variety in $\mathbb{CP}^{n-1}$. Then the cone over X in $\mathbb{C}^n$ is an affine variety V with an isolated singularity at the origin. It is a very natural and important question to ask when an affine variety with an isolated singularity at the origin is a cone over nonsingular projective variety.

This problem is very hard in general. In this paper we shall treat the hypersurface case. Given a function f with an isolated singularity at the origin, we can ask whether f is a weighted homogeneous polynomial or a homogeneous polynomial after a biholomorphic change of coordinates. The former question was answered in a celebrated 1971 paper by Saito. However, the latter question had remained open for 40 years until Xu and Yau solved it for f with three variables. Recently, Yau and Zuo solved it for f with up to six variables. However, the methods they used are hard to generalize. In this paper, we solve the latter question for general n completely; i.e., we show that f is a homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau = (\nu - 1)^n$, where μ, τ and ν are the Milnor number, Tjurina number and multiplicity of the singularity respectively. We also prove that there are at most $\mu^{1/n} + 1$ multiplicities within the same topological type of the isolated hypersurface singularity, while the famous Zariski multiplicity problem asserts that there is only one multiplicity.

1. Introduction

Let $w = (w_1, \dots, w_n)$ be an n-tuple of positive rational numbers. A polynomial $f(z_1, \dots, z_n)$ is said to be a weighted homogeneous polynomial with weight w if each monomial $\alpha z_1^{a_1} z_2^{a_2} \cdots z_n^{a_n}$ of f satisfies $a_1 w_1 + \cdots + a_n w_n = 1$. It has an isolated critical point at $0 \in \mathbb{C}^n$ if grad $f = (\partial f/\partial z_1, \dots, \partial f/\partial z_n)$ is zero at 0 but grad $f(z) \neq 0$ for all z in a neighborhood of 0.

Yau was partially supported by the start-up fund from Tsinghua University, and Zuo was partially supported by NSFC no. 11401335 and the start-up fund from Tsinghua University.
MSC2010: primary 32S25; secondary 32S10.
Keywords: homogeneous singularities, Milnor number, multiplicity.

Recall that a polynomial $f(z_1, \dots, z_n)$ is called quasihomogeneous if f is in the Jacobian ideal of f, i.e., $f \in (\partial f/\partial z_1, \dots, \partial f/\partial z_n)$.

By a theorem of Saito (see Theorem 2.7), if f is quasihomogeneous with an isolated critical point at 0, then after a biholomorphic change of coordinates, f becomes a weighted homogeneous polynomial.

Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be the germ of a complex analytic function with an isolated critical point at the origin. Let $V = \{z \in \mathbb{C}^n : f(z) = 0\}$. It is a natural question to ask when V is defined by a weighted homogeneous polynomial up to biholomorphic change of coordinates. Saito [1971] solved this question. He gave a necessary and sufficient condition for V to be defined by a weighted homogeneous polynomial. It is a natural and important question to characterize homogeneous polynomial with an isolated critical point at the origin. This question has remained open for 40 years. In fact it is the first important case of the following interesting problem. Let X be a nonsingular projective variety in $\mathbb{CP}^{n-1}$. Then the cone over X in $\mathbb{C}^n$ is an affine variety V with an isolated singularity at the origin. It is then natural to ask when an affine variety with an isolated singularity at the origin is a cone over nonsingular projective variety.

For a two-dimensional isolated hypersurface singularity V, Xu and Yau [1992; 1993] found a coordinate-free characterization when V is defined by a homogeneous polynomial. Recently, necessary and sufficient conditions were given for three-dimensional isolated hypersurface singularities with $p_g \geq 0$ [Lin and Yau 2004; Lin et al. 2006a; Xu and Yau 1996] and four-dimensional isolated hypersurface singularities with $p_g > 0$ [Chen et al. 2011], where p_g is the geometric genus of the singularity. Based on the classification of weighted homogeneous singularities, Yau and Zuo [2012] solved the problem for f with up to six variables. However, it is quite difficult to generalize their methods to characterize the homogeneous polynomials for general n. Ten years ago, Yau formulated the Yau homogeneous characterization conjecture: (1) Let μ and ν be the Milnor number and multiplicity of $(V, 0)$ respectively. Then $\mu \geq (\nu - 1)^n$, and equality holds if and only if f is a semihomogeneous function (i.e., $f = f_\nu + g$, where f_ν is a homogeneous polynomial of degree ν defining an isolated singularity at the origin and g consists of terms of degree at least $\nu + 1$) after a biholomorphic change of coordinates. (2) Moreover, if f is a weighted homogeneous function, then $\mu = (\nu - 1)^n$ if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates. In this paper we verify the Yau homogeneous characterization conjecture affirmatively. As a result, we have solved the characterization problem of homogeneous polynomials with an isolated critical point at the origin, i.e., we have shown that f is a homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau = (\nu - 1)^n$.

Recall that the multiplicity of the singularity V is defined to be the order of the lowest nonvanishing term in the power series expansion of f at 0. The Milnor number μ and the Tjurina number τ of the singularity $(V, 0)$ are defined by

$$\mu = \dim \mathbb{C}\{z_1, z_2, \ldots, z_n\}/(f_{z_1}, \ldots, f_{z_n}),$$
$$\tau = \dim \mathbb{C}\{z_1, z_2, \ldots, z_n\}/(f, f_{z_1}, \ldots, f_{z_n}).$$

They are numerical invariants of $(V, 0)$.

Let $\pi : (M, A) \to (V, 0)$ be a resolution of singularity of dimension n with exceptional set $A = \pi^{-1}(0)$. The geometric genus p_g of the singularity $(V, 0)$ is the dimension of $H^{n-1}(M, \mathbb{O})$ and is independent of the resolution M.

Using p_g, μ and ν, Yau made another conjecture in 1995 (see [Lin and Yau 2004; Chen et al. 2011]) describing when a weighted homogeneous singularity is a homogeneous singularity. Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a weighted homogeneous polynomial with an isolated singularity at the origin. Let μ, p_g and ν be the Milnor number, geometric genus and multiplicity of the singularity $V = \{z : f(z) = 0\}$; then

$$\mu - p(\nu) \geq n! p_g,$$

where $p(\nu) = (\nu - 1)^n - \nu(\nu - 1) \cdots (\nu - n + 1)$, and equality holds if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates.

In fact, we shall prove in this paper that if $p_g = 0$, then the Yau homogeneous characterization conjecture implies the 1995 Yau conjecture.

These conjectures are sharp estimates and have some important applications in geometry. However, they were proved only for low-dimensional singularities. For the Yau homogeneous characterization conjecture, Lin, Wu, Yau and Luk [Lin et al. 2006b] proved the following two theorems.

Theorem 1.1. *Let $f : (\mathbb{C}^2, 0) \to (\mathbb{C}, 0)$ be the germ of a holomorphic function defining an isolated plane curve singularity $V = \{z \in \mathbb{C}^2 : f(z) = 0\}$ at the origin. Let μ and ν be the Milnor number and multiplicity of $(V, 0)$, respectively. Then*

$$\mu \geq (\nu - 1)^2. \tag{1-1}$$

Furthermore, if V has at most two irreducible branches at the origin, or if f is a quasihomogeneous function, then equality holds in (1-1) *if and only if f is a homogeneous polynomial* (*after a biholomorphic change of coordinates*).

Theorem 1.2. *Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be the germ of a holomorphic function defining an isolated hypersurface singularity $V = \{z \in \mathbb{C}^n : f(z) = 0\}$ at the origin. Let μ, ν and $\tau = \dim \mathbb{C}\{z_1, \ldots, z_n\}/(f, \partial f/\partial z_1, \ldots, \partial f/\partial z_n)$ be the Milnor number, multiplicity and Tjurina number of $(V, 0)$, respectively. Suppose $\mu = \tau$ and n is either* 3 *or* 4. *Then*

$$\mu \geq (\nu - 1)^n, \tag{1-2}$$

and equality holds in (1-2) *if and only if f is a homogeneous polynomial* (*after a biholomorphic change of coordinates*).

For the 1995 Yau conjecture, Lin, Tu and Yau [Lin and Yau 2004; Lin et al. 2006a] have the following theorem:

Theorem 1.3. *Let $(V, 0)$ be a three-dimensional isolated hypersurface singularity defined by a weighted homogeneous polynomial $f(x, y, z, w) = 0$. Let μ, ν and p_g be the Milnor number, multiplicity and geometric genus of the singularity, respectively. Then*

$$\mu - (2\nu^3 - 5\nu^2 + 2\nu + 1) \geq 4!p_g \tag{1-3}$$

and equality holds in (1-3) *if and only if f is a homogeneous polynomial.*

Remark. The above theorem is proved in [Lin and Yau 2004] with $p_g > 0$. For $p_g = 0$, the theorem is proved in [Lin et al. 2006a].

An immediate corollary of Theorem 1.3 is the following:

Corollary 1.4 [Lin et al. 2006a]. *Let $(V, 0)$ be a three-dimensional isolated hypersurface singularity defined by a polynomial $f(x, y, z, w) = 0$. Let μ, ν, p_g and τ be the Milnor number, multiplicity, geometric genus and Tjurina number of the singularity, respectively. Then f is a homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau$ and $\mu - (2\nu^3 - 5\nu^2 + 2\nu + 1) = 4!p_g$.*

Chen, Lin, Yau and Zuo generalized the above theorem to four-dimensional isolated hypersurface singularities with the additional assumption $p_g > 0$.

Theorem 1.5 [Chen et al. 2011]. *Let $(V, 0)$ be a four-dimensional isolated hypersurface singularity defined by a weighted homogeneous polynomial $f(x, y, z, w, t) = 0$. Let μ, ν and p_g be the Milnor number, multiplicity and geometric genus of the singularity, respectively. If $p_g > 0$ then*

$$\mu - [(\nu - 1)^5 + \nu(\nu - 1)(\nu - 2)(\nu - 3)(\nu - 4)] \geq 5!p_g, \tag{1-4}$$

and equality holds in (1-4) *if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates.*

Corollary 1.6 [Chen et al. 2011]. *Let $(V, 0)$ be a four-dimensional isolated hypersurface singularity defined by a polynomial $f(x, y, z, w, t) = 0$. Let μ, ν, p_g and τ be the Milnor number, multiplicity, geometric genus and Tjurina number of the singularity, respectively. Moreover, if $p_g > 0$ then f is a homogeneous polynomial after a biholomorphic change of coordinates if and only if $\mu = \tau$ and*

$$\mu - [(\nu - 1)^5 + \nu(\nu - 1)(\nu - 2)(\nu - 3)(\nu - 4)] = 5!p_g.$$

Yau and Zuo [2012] proved the following theorem:

Theorem 1.7. *Let* $f : (\mathbb{C}^k, 0) \to (\mathbb{C}, 0)$, *where* k *is either* 5 *or* 6, *be a polynomial with an isolated singularity at the origin. Let* μ, ν *and* τ *be the Milnor number, multiplicity and Tjurina number of the singularity* $V = \{z :\ f(z) = 0\}$ *respectively. Then* f *is a homogeneous polynomial after a biholomorphic change of coordinates if and only if* $\mu = \tau = (\nu - 1)^k$.

The purpose of this paper is to prove the following results:

Theorem A. *Let* $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ *be a weighted homogeneous polynomial with an isolated singularity at the origin. Let* μ *and* ν *be the Milnor number and multiplicity of the singularity* $V = \{z :\ f(z) = 0\}$ *respectively. Then*

$$\mu \geq (\nu - 1)^n, \tag{1-5}$$

and equality holds in (1-5) *if and only if* f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

Corollary B. *Let* $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ *be a polynomial with an isolated singularity at the origin. Let* μ, ν *and* τ *be the Milnor number, multiplicity and Tjurina number of the singularity* $V = \{z :\ f(z) = 0\}$ *respectively. Then* f *is a homogeneous polynomial after a biholomorphic change of coordinates if and only if* $\mu = \tau = (\nu - 1)^n$.

Proposition 3.1 and Theorem A answer the Yau homogeneous characterization conjecture affirmatively, and Corollary B gives a complete characterization of isolated homogeneous hypersurface singularities. Let $(V, 0)$ and $(W, 0)$ be two isolated hypersurface singularities in $\mathbb{C}^n$. We say that $(V, 0)$ and $(W, 0)$ have the same topological type if $(\mathbb{C}^n, V, 0)$ is homeomorphically equivalent to $(\mathbb{C}^n, W, 0)$. The famous Zariski multiplicity question asks whether $(V, 0)$ and $(W, 0)$ have the same multiplicity if they have the same topological type, i.e., whether there is only one multiplicity within the same topological type. For two-dimensional isolated quasihomogeneous singularities, the Zariski multiplicity question was solved [Yau 1988; Xu and Yau 1989]. Proposition 3.1 says that there are at most $\mu^{1/n} + 1$ multiplicities within the same topological type. On the other hand, Theorem C below confirms that the 1995 Yau conjecture is true for the case of $p_g = 0$.

Theorem C. *Let* $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ *be a weighted homogeneous polynomial with an isolated singularity at the origin. Let* μ, p_g *and* ν *be the Milnor number, geometric genus and multiplicity of the singularity* $V = \{z :\ f(z) = 0\}$. *If* $p_g = 0$, *then*

$$\mu - p(\nu) \geq n!\, p_g,$$

where $p(\nu) = (\nu-1)^n - \nu(\nu-1)\cdots(\nu-n+1)$ $(=0)$, *and equality holds if and only if* f *is a homogeneous polynomial after a biholomorphic change of coordinates.*

In Section 2, we recall the material which is necessary to prove the main theorems. In Section 3, we prove the main theorems.

2. Preliminaries

In this section, we recall some known results which are needed to prove the main theorems. Let $f(z_1, \dots, z_n)$ be a germ of an analytic function at the origin such that $f(0) = 0$. Suppose f has an isolated critical point at the origin, and suppose f can be developed in a convergent Taylor series $f(z_1, \dots, z_n) = \sum a_\lambda z^\lambda$, where $z^\lambda = z_1^{\lambda_1} \cdots z_n^{\lambda_n}$. Recall that the Newton boundary $\Gamma(f)$ is the union of compact faces of $\Gamma_+(f)$, where $\Gamma_+(f)$ is the convex hull of the union of subsets $\{\lambda + \mathbb{R}_+^n\}$ for λ such that $a_\lambda \neq 0$. Let $\Gamma_-(f)$, the Newton polyhedron of f, be the cone over $\Gamma(f)$ with cone point at 0. For any closed face Δ of $\Gamma(f)$, we associate the polynomial $f_\Delta(z) = \sum_{\lambda \in \Delta} a_\lambda z^\lambda$. We say that f is nondegenerate if f_Δ has no critical point in $(\mathbb{C}^*)^n$ for any $\Delta \in \Gamma(f)$, where $\mathbb{C}^* = \mathbb{C} - \{0\}$. We say that a point p of the integral lattice $\mathbb{Z}^n$ in $\mathbb{R}^n$ is positive if all coordinates of p are positive.

Theorem 2.1 [Merle and Teissier 1980]. *Let $(V, 0)$ be an isolated hypersurface singularity defined by a nondegenerate holomorphic function $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$. Then the geometric genus $p_g = \#\{p \in \mathbb{Z}^n \cap \Gamma_-(f) : p \text{ is positive}\}$.*

Recall that a polynomial $f(z_1, \dots, z_n)$ is a weighted homogeneous polynomial of type $(w_1, \dots, w_n)$, where $w_1, \dots, w_n$ are fixed positive rational numbers, if it can be expressed as a linear combination of monomials $z_1^{i_1} \cdots z_n^{i_n}$ for which $i_1 w_1 + \cdots + i_n w_n = 1$. As a consequence of Theorem 2.1, for an isolated singularity defined by a weighted homogeneous polynomial, computing the geometric genus is equivalent to counting the number of positive integral points in the tetrahedron defined by $x_1 w_1 + \cdots + x_n w_n \leq 1$, $x_1 \geq 0, \dots, x_n \geq 0$. We also need the following result:

Theorem 2.2 [Milnor and Orlik 1970]. *Let $f(z_1, \dots, z_n)$ be a weighted homogeneous polynomial of type $(w_1, \dots, w_n)$ with an isolated singularity at the origin. Then the Milnor number μ is equal to $(1/w_1 - 1) \cdots (1/w_n - 1)$.*

Yau [1977] gave a lower bound for p_g of a hypersurface singularity.

Theorem 2.3 [Yau 1977]. *Let*

$$f(z_1, \dots, z_{n-1}, z_n) = z_n^m + a_1(z_1, \dots, z_{n-1}) z_n^{m-1} + \cdots + a_m(z_1, \dots, z_{n-1})$$

be holomorphic near $(0, \dots, 0)$. Let d_i be the order of the zero of $a_i(z_1, \dots, z_{n-1})$ at $(0, \dots, 0)$, with $d_i \geq i$. Let $d = \min_{1 \leq i \leq m}(d_i / i)$. Suppose that

$$V = \{(z_1, \dots, z_n) : f(z_1, \dots, z_n) = 0\},$$

defined in a suitably small polydisc, has $p = (0, \dots, 0)$ as its only singularity. Let $\pi : M \to V$ be a resolution of V. Then $\dim H^{n-2}(M, \mathbb{O}) > (m-1)d - (n-1)$.

Remark. Here, the singularity is $(n-1)$-dimensional, so $\dim H^{n-2}(M, \mathbb{O}) = p_g$.

Let $f \in \mathbb{C}\{z_1, \dots, z_n\}$ define an isolated singularity at the origin. Let $w = (w_1, \dots, w_n)$ be a weight on the coordinates $(z_1, \dots, z_n)$ for positive integers w_i, $i = 1, \dots, n$. We have the weighted Taylor expansion $f = f_\rho + f_{\rho+1} + \cdots$ with respect to w, where $f_\rho \neq 0$ and f_k is a weighted homogeneous polynomial of type $(w_1, \dots, w_n; k)$ for $k \geq \rho$, i.e., f_k is a linear combination of monomials $z_1^{i_1} \cdots z_n^{i_n}$ for which $i_1 w_1 + \cdots + i_n w_n = k$.

Theorem 2.4 [Furuya and Tomari 2004]. *Let $f \in \mathbb{C}\{z_1, \dots, z_n\}$ define an isolated singularity at the origin. With the above situation, then*:

(1) *The following inequality holds*:

$$\mu(f) \geq \left(\frac{\rho}{w_1} - 1\right) \cdots \left(\frac{\rho}{w_n} - 1\right).$$

(2) *Equality holds in* (1) *if and only if f_ρ defines an isolated singularity at the origin.*

Here we recall that f is called a semiquasihomogeneous function if the initial term f_ρ defines an isolated singularity at the origin.

Definition 2.1. Let $f, g : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be germs of holomorphic functions defining isolated hypersurface singularities $V_f = \{z : f(z) = 0\}$ and $V_g = \{z : g(z) = 0\}$. Let $\phi : (\mathbb{C}^n, 0) \to (\mathbb{C}^n, 0)$ be the germ of a biholomorphic map.

(1) f is contact-equivalent to g if $\phi(V_f) = V_g$.

(2) f is right-equivalent to g if $g = f \circ \phi$.

The Milnor number is an invariant under right-equivalence and the Tjurina number is an invariant under contact equivalence. It is a nontrivial theorem that the Milnor number is indeed an invariant under contact equivalence:

Theorem 2.5 [Greuel 1975]. *Let $\mathbb{K}$ be an algebraically closed field of characteristic zero and $f, g \in \mathbb{K}\{z_1, z_2, \dots, z_n\}$. If f is contact-equivalent to g, then $\mu(f) = \mu(g)$.*

Theorem 2.6 [Shoshitaishvili 1976; Benson and Yau 1990]. *If f and g are germs of isolated weighted homogeneous singularities at the origin in $\mathbb{C}^n$, then f and g are right-equivalent if and only if f and g are contact-equivalent.*

Theorem 2.7 [Saito 1971]. *Let $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be the germ of a complex analytic function with an isolated critical point at the origin.*

(a) *f is right-equivalent to a weighted homogeneous polynomial if and only if $\mu = \tau$ or*

$$f \in J_f := \left(\frac{\partial f}{\partial z_1}, \dots, \frac{\partial f}{\partial z_n}\right).$$

(b) *If f is weighted homogeneous with normalized weight system* $(w_1, \dots, w_n, 1)$ *with* $0 < w_1 \le \cdots \le w_n < 1$ *and if* $f \in \boldsymbol{m}^3_{\mathbb{C}^n,0}$, *then the weight system is unique and* $0 < w_1 \le \cdots \le w_n < \frac{1}{2}$.

(c) *If* $f \in J_f$, *then f is right-equivalent to a weighted homogeneous polynomial* $g(z_1, \dots, z_k) + z_{k+1}^2 + \cdots + z_n^2$ *with* $g \in \boldsymbol{m}^3_{\mathbb{C}^n,0}$. *Specifically, its normalized weight system satisfies* $0 < w_1 \le \cdots \le w_k < w_{k+1} = \cdots = w_n = \frac{1}{2}$.

(d) *If f and* $\bar{f} \in \mathbb{O}_{\mathbb{C}^n,0}$ *are right-equivalent and weighted homogeneous with normalized weight systems* $(w_1, \dots, w_n, 1)$ *and* $(\overline{w}_1, \dots, \overline{w}_n, 1)$ *with* $w_1 \le \cdots \le w_n \le \frac{1}{2}$ *and* $\overline{w}_1 \le \cdots \le \overline{w}_n \le \frac{1}{2}$, *then* $w_i = \overline{w}_i$.

3. Proof of the main theorems

The following statement is well known.

Proposition 3.1 [Teissier 1973]. *Let* $f : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ *be a holomorphic germ defining an isolated hypersurface singularity* $V = \{z \colon f(z) = 0\}$ *at the origin. Let μ and ν be the Milnor number and multiplicity of* $(V, 0)$ *respectively. Then*

$$\mu \ge (\nu - 1)^n, \tag{3-1}$$

and equality holds in (3-1) *if and only if f is a semihomogeneous function* (*i.e.*, $f = f_\nu + g$, *where f_ν is a nondegenerate homogeneous polynomial of degree ν and g consists of terms of degree at least* $\nu + 1$) *after a biholomorphic change of coordinates.*

Proof. Let $f(z_1, \dots, z_n) : (\mathbb{C}^n, 0) \to (\mathbb{C}, 0)$ be a holomorphic function with an isolated singularity at the origin. Let μ and ν be the Milnor number and multiplicity of the singularity $V = \{z \colon f(z) = 0\}$. By an analytic change of coordinates, one can assume that the z_n-axis is not contained in the tangent cones of V, so that $f(0, \dots, 0, z_n) \ne 0$. By the Weierstrass preparation theorem, near 0, the germ f can be represented as a product $f(z_1, \dots, z_n) = u(z_1, \dots, z_n) g(z_1, \dots, z_n)$, where $u(0, \dots, 0) \ne 0$ and

$$g(z_1, \dots, z_{n-1}, z_n) = z_n^\nu + a_1(z_1, \dots z_{n-1}) z_n^{\nu-1} + \cdots + a_\nu(z_1, \dots, z_{n-1}),$$

where ν is the multiplicity of $f(z_1, \dots, z_n)$ and $a_i \in (z_1, \dots, z_{n-1})^i$ for $i = 1, \dots, \nu$. Therefore $f(z_1, \dots, z_n)$ is contact-equivalent to $g(z_1, \dots, z_n)$.

Let d_i be the order of the zero of $a_i(z_1, \dots, z_{n-1})$ at $(0, \dots, 0)$, with $d_i \ge i$. Let $d = \min_{1 \le i \le \nu}(d_i / i)$; then $d \ge 1$. We define a new weight w on the coordinate system: $w(z_n) = d$, $w(z_i) = 1$ for $1 \le i \le n - 1$. With respect to the new weights, z_n^ν has degree $d\nu$ and $a_i(z_1, \dots, z_{n-1}) z_n^{\nu-i}$ has degree at least $d(\nu - i) + d_i \ge d\nu - di + di = d\nu$. Thus the initial term of $f(z_1, \dots, z_n)$ has degree $\rho = d\nu$. By

Theorem 2.5, the Milnor number is an invariant under the contact equivalence. By Theorem 2.4(1), we have

$$\mu = \mu(g) \geq \Big(\frac{d\nu}{d} - 1\Big)\Big(\frac{d\nu}{1} - 1\Big)\cdots\Big(\frac{d\nu}{1} - 1\Big) = (\nu-1)(d\nu-1)^{n-1} \geq (\nu-1)^n.$$

Thus, we proved the inequality (3-1).

We need to show that equality in (3-1) holds if and only if f is a semihomogeneous function after a biholomorphic change of coordinates.

$\Rightarrow$: If $\mu = (\nu-1)^n$, then by the fact that $\mu \geq (\nu-1)(d\nu-1)^{n-1} \geq (\nu-1)^n$, we have $d = 1$, and by Theorem 2.4(2), $g_{d\nu}(z_1, \dots, z_n) = g_\nu(z_1, \dots, z_n)$ is a homogeneous polynomial of degree ν defining an isolated singularity. Hence $f(z_1, \dots, z_n)$ is contact-equivalent to a semihomogeneous singularity g; i.e., f is a semihomogeneous function after a biholomorphic change of coordinates.

$\Leftarrow$: Suppose f is a semihomogeneous polynomial after a biholomorphic change of coordinates. Since the Milnor number of f is the same as the Milnor number of its initial part (see [Arnold 1974]) which is a homogeneous polynomial with degree ν, so $\mu = (\nu-1)^n$ is obvious. □

Proof of Theorem A. By Proposition 3.1, it is sufficient to show that if f is a weighted homogeneous singularity, then $\mu = (\nu-1)^n$ if and only if f is equivalent to a homogeneous singularity.

The "if" part is trivial. We only need to consider the "only if" part. By Saito's theorem (see Theorem 2.7(c)), we can choose normalized weights for f, which means that these weights satisfy $0 < w_i \leq \frac{1}{2}$, $1 \leq i \leq n$. By what we have proved above, we know that there exists a $g(z_1, \dots, z_n)$ so that f is contact equivalent to g; moreover the initial part of g is g_ν, a homogeneous polynomial with degree ν, and g_ν also defines an isolated singularity at the origin. We can rewrite f and g as

$$f(z_1, \dots, z_n) = f_\nu(z_1, \dots, z_n) + f_{\nu+1}(z_1, \dots, z_n) + \cdots,$$
$$g(z_1, \dots, z_n) = g_\nu(z_1, \dots, z_n) + g_{\nu+1}(z_1, \dots, z_n) + \cdots,$$

where f_i and g_i, $i \geq \nu$, are the homogeneous parts of f and g respectively and g_ν defines an isolated singularity at the origin. Since for weighted homogeneous singularities contact equivalence is the same as right equivalence (see Theorem 2.6), there exists a biholomorphism at the origin:

$$\phi : (\mathbb{C}^n, 0) \to (\mathbb{C}^n, 0),$$
$$(z_1, \dots, z_n) \mapsto (\phi_1(z_1, \dots, z_n), \dots, \phi_n(z_1, \dots, z_n)),$$

such that $f(z_1,\dots,z_n) = g(\phi_1(z_1,\dots,z_n),\dots,\phi_n(z_1,\dots,z_n))$ and

$$\begin{aligned}\phi_1(z_1,\dots,z_n) &= a_{11}z_1+\cdots+a_{1n}z_n+H_1^2+H_1^3+\cdots,\\ &\vdots\\ \phi_n(z_1,\dots,z_n) &= a_{n1}z_1+\cdots+a_{nn}z_n+H_n^2+H_n^3+\cdots,\end{aligned}$$

where $H_i^j = \sum_{\alpha_1+\cdots+\alpha_n=j} c_i(\alpha_1,\dots,\alpha_n)z_1^{\alpha_1}\cdots z_n^{\alpha_n}$. Since ϕ is a biholomorphism at the origin, we have $|(a_{ij})| = \det(a_{ij}) \neq 0$. It follows from the equality $f(z_1,\dots,z_n) = g(\phi_1(z_1,\dots,z_n),\dots,\phi_n(z_1,\dots,z_n))$ that

$$\begin{aligned}&g_\nu(\phi_1(z_1,\dots,z_n),\dots,\phi_n(z_1,\dots,z_n))\\ &\qquad + g_{\nu+1}(\phi_1(z_1,\dots,z_n),\dots,\phi_n(z_1,\dots,z_n))+\cdots\\ &\qquad\qquad = f_\nu(z_1,\dots,z_n)+f_{\nu+1}(z_1,\dots,z_n)+\cdots.\end{aligned}$$

Comparing the degree of each side, we have

$$g_\nu(\bar\phi_1(z_1,\dots,z_n),\dots,\bar\phi_n(z_1,\dots,z_n)) = f_\nu(z_1,\dots,z_n),$$

where $\bar\phi_i = a_{i1}z_1+\cdots+a_{in}z_n$, $1\le i\le n$. Since $\det(a_{ij})\neq 0$, f_ν is right-equivalent to g_ν. Therefore f_ν also defines an isolated singularity. Now we have two normalized weights for f_ν: one is $(w_1,\dots,w_n)$, because each monomial in f_ν comes from f, and the other is $(1/\nu,\dots,1/\nu)$ which follows from the fact that f_ν is a homogeneous polynomial with degree ν. By Theorem 2.7, we have $w_1 = w_2 = \cdots = w_n = 1/\nu$. Therefore $f(z_1,\dots,z_n) = f_\nu(z_1,\dots,z_n)$ is a homogeneous polynomial. □

Proof of Corollary B. This follows from Theorem A and Theorem 2.7(a). □

Proof of Theorem C. Since $p_g = 0$, by Theorem 2.3, we have $0 > (\nu-1)d-(n-1)$, where $d = \min_{1\le i\le\nu}(d_i/i)$, and d_i is the order of the zero of $a_i(x_1,\dots,x_n)$ at $(0,\dots,0)$, with $d_i\ge i$. Then $\nu < (n-1)/d+1$. Since $d\ge 1$, ν is an integer at least 2 for an isolated hypersurface singularity, so we have $2\le\nu\le n-1$. Therefore $p(\nu) = (\nu-1)^n - \nu(\nu-1)\cdots(\nu-n+1) = (\nu-1)^n$. The theorem is reduced to proving that

$$\mu \ge (\nu-1)^n,$$

where equality holds if and only if f is a homogeneous polynomial after a biholomorphic change of coordinates. The proof follows from Theorem A immediately. □

References

[Arnold 1974] V. I. Arnold, "Нормальные формы функций в окрестности вырожденных критических точек", *Uspekhi Mat. Nauk* **29**:2(176) (1974), 11–49. Translated as "Normal forms of functions in the neighborhood of degenerate critical points" in *Russian Math. Surveys* **29**:2 (1974), 10–50. MR 58 #24324 Zbl 0298.57022

[Benson and Yau 1990] M. Benson and S. S.-T. Yau, "Equivalences between isolated hypersurface singularities", *Math. Ann.* **287**:1 (1990), 107–134. MR 91h:58006 Zbl 0673.32016

[Chen et al. 2011] I. Chen, K.-P. Lin, S. S.-T. Yau, and H.-Q. Zuo, "Coordinate-free characterization of homogeneous polynomials with isolated singularities", *Comm. Anal. Geom.* **19**:4 (2011), 661–704. MR 2880212 Zbl 1246.32030

[Furuya and Tomari 2004] M. Furuya and M. Tomari, "A characterization of semi-quasihomogeneous functions in terms of the Milnor number", *Proc. Amer. Math. Soc.* **132**:7 (2004), 1885–1890. MR 2005f:14007 Zbl 1052.32023

[Greuel 1975] G.-M. Greuel, "Der Gauss–Manin-Zusammenhang isolierter Singularitäten von vollständigen Durchschnitten", *Math. Ann.* **214** (1975), 235–266. MR 53 #417 Zbl 0285.14002

[Lin and Yau 2004] K.-P. Lin and S. S.-T. Yau, "Classification of affine varieties being cones over nonsingular projective varieties: hypersurface case", *Comm. Anal. Geom.* **12**:5 (2004), 1201–1219. MR 2006a:14102 Zbl 1072.32022

[Lin et al. 2006a] K.-P. Lin, Z.-H. Tu, and S. S.-T. Yau, "Characterization of isolated homogeneous hypersurface singularities in $\mathbb{C}^4$", *Sci. China Ser. A* **49**:11 (2006), 1576–1592. MR 2008c:32037 Zbl 1115.32017

[Lin et al. 2006b] K.-P. Lin, X. Wu, S. S.-T. Yau, and H.-S. Luk, "A remark on lower bound of Milnor number and characterization of homogeneous hypersurface singularities", *Comm. Anal. Geom.* **14**:4 (2006), 625–632. MR 2008a:32026 Zbl 1111.32025

[Merle and Teissier 1980] M. Merle and B. Teissier, "Conditions d'adjonction: d'après Du Val", pp. 229–245 in *Seminaire sur les singularités des surfaces* (Palaiseau, 1976–1977), edited by M. Demazure et al., Lecture Notes in Math. **777**, Springer, Berlin, 1980. Zbl 0461.14009

[Milnor and Orlik 1970] J. Milnor and P. Orlik, "Isolated singularities defined by weighted homogeneous polynomials", *Topology* **9** (1970), 385–393. MR 45 #2757 Zbl 0204.56503

[Saito 1971] K. Saito, "Quasihomogene isolierte Singularitäten von Hyperflächen", *Invent. Math.* **14** (1971), 123–142. MR 45 #3767 Zbl 0224.32011

[Shoshitaishvili 1976] A. N. Shoshitaishvili, "О функциях с изоморфными Якобиевыми идеалами", *Funkcional. Anal. i Priložen.* **10**:2 (1976), 57–62. Translated as "Functions with isomorphic Jacobian ideals" in *Funct. Anal. Appl.* **10**:2 (1976), 128–133. MR 54 #5492 Zbl 0346.32023

[Teissier 1973] B. Teissier, "Cycles évanescents, sections planes et conditions de Whitney", pp. 285–362 in *Singularités à Cargèse* (Cargèse, 1972), Astérisque **7–8**, Société Mathématique de France, Paris, 1973. MR 51 #10682 Zbl 0295.14003

[Xu and Yau 1989] Y.-J. Xu and S. S.-T. Yau, "Classification of topological types of isolated quasi-homogeneous two-dimensional hypersurface singularities", *Manuscripta Math.* **64**:4 (1989), 445–469. MR 91e:32034 Zbl 0681.32008

[Xu and Yau 1992] Y.-J. Xu and S. S.-T. Yau, "A sharp estimate of the number of integral points in a tetrahedron", *J. Reine Angew. Math.* **423** (1992), 199–219. MR 93d:11067 Zbl 0734.11048

[Xu and Yau 1993] Y.-J. Xu and S. S.-T. Yau, "Durfee conjecture and coordinate free characterization of homogeneous singularities", *J. Differential Geom.* **37**:2 (1993), 375–396. MR 94f:32073 Zbl 0793.32016

[Xu and Yau 1996] Y.-J. Xu and S. S.-T. Yau, "A sharp estimate of the number of integral points in a 4-dimensional tetrahedra", *J. Reine Angew. Math.* **473** (1996), 1–23. MR 97d:11151 Zbl 0844.11063

[Yau 1977] S. S.-T. Yau, "Two theorems on higher dimensional singularities", *Math. Ann.* **231**:1 (1977), 55–59. MR 58 #11511 Zbl 0343.32010

[Yau 1988] S. S.-T. Yau, "Topological types and multiplicities of isolated quasihomogeneous surface singularities", *Bull. Amer. Math. Soc.* (*N.S.*) **19**:2 (1988), 447–454. MR 92b:32042 Zbl 0659.32013

[Yau and Zuo 2012] S. S.-T. Yau and H.-Q. Zuo, "Lower estimate of Milnor number and characterization of isolated homogeneous hypersurface singularities", *Pacific J. Math.* **260**:1 (2012), 245–255. MR 3001794 Zbl 1276.32022

Received February 8, 2014.

STEPHEN YAU
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING, 100084
CHINA
yau@uic.edu

HUAIQING ZUO
MATHEMATICAL SCIENCES CENTER
TSINGHUA UNIVERSITY
BEIJING, 100084
CHINA
hqzuo@math.tsinghua.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.225

A THEOREM OF MŒGLIN AND WALDSPURGER FOR COVERING GROUPS

SHIV PRAKASH PATEL

Let E be a nonarchimedean local field of characteristic zero and residual characteristic p. Let G be a connected reductive group defined over E and π an irreducible admissible representation of $G(E)$. A result of C. Mœglin and J.-L. Waldspurger (for $p \neq 2$) and S. Varma (for $p = 2$) states that the leading coefficient in the character expansion of π at the identity element of $G(E)$ gives the dimension of a certain space of degenerate Whittaker forms. In this paper we generalize this result of Mœglin and Waldspurger to the setting of covering groups of $G(E)$.

1. Introduction

Let E be a nonarchimedean local field of characteristic zero, $\boldsymbol{G}$ a connected split reductive group defined over E, and $G = \boldsymbol{G}(E)$. Let $\boldsymbol{\mathfrak{g}} = \mathrm{Lie}(\boldsymbol{G})$ be the Lie algebra of $\boldsymbol{G}$ and $\mathfrak{g} = \boldsymbol{\mathfrak{g}}(E)$. Let (π, W) be an irreducible admissible representation of G. A theorem of F. Rodier [1975] relates the dimension of the space of nondegenerate Whittaker forms of π to coefficients in the character expansion of π around the identity. More precisely, Rodier proved that if the residual characteristic of E is large enough and the group $\boldsymbol{G}$ is split then the dimension of any space of nondegenerate Whittaker functionals for (π, W) equals the coefficient in the character expansion of π at the identity corresponding to an appropriate maximal nilpotent orbit in the Lie algebra $\mathfrak{g}$. Rodier proved his theorem assuming that the residual characteristic of E is greater than a constant which depends only on the root datum of $\boldsymbol{G}$. C. Mœglin and J.-L. Waldspurger [1987] generalized this theorem of Rodier in several directions: in particular, they proved the result for fields E whose residual characteristic is odd, and they removed the assumption that $\boldsymbol{G}$ is split. The Mœglin–Waldspurger theorem is a more precise statement about the coefficients appearing in the character expansion around the identity and certain spaces of "degenerate" Whittaker forms. In [Varma 2014], this theorem has been proved for fields with even residual characteristic by modifying certain constructions in [Mœglin and Waldspurger 1987] (see the remark at the end of the introduction). So the Mœglin–Waldspurger theorem is

MSC2010: primary 22E50; secondary 11F70, 11S37.
Keywords: covering groups, character expansion, degenerate Whittaker forms.

true for all connected reductive groups without any restriction on the residual characteristic of the field E. We now recall the theorem. To state it we need to introduce some notation. Let Y be a nilpotent element in $\mathfrak{g}$ and suppose $\varphi : \mathbb{G}_m \to \boldsymbol{G}$ is a one-parameter subgroup satisfying

$$\operatorname{Ad}(\varphi(t))Y = t^{-2}Y. \tag{1}$$

Associated to such a pair (Y, φ) one can define a certain space $\mathcal{W}_{(Y,\varphi)}$, called the space of degenerate Whittaker forms of (π, W) relative to (Y, φ) (see Section 4 for the definition).

Define $\mathcal{N}_{\mathrm{Wh}}(\pi)$ to be the set of nilpotent orbits $\mathbb{O}$ of $\mathfrak{g}$ for which there exists an element $Y \in \mathbb{O}$ and a φ satisfying (1) such that the space $\mathcal{W}_{(Y,\varphi)}$ of degenerate Whittaker forms relative to the pair (Y, φ) is nonzero.

Recall that the character expansion of (π, W) around the identity is a sum $\sum_{\mathbb{O}} c_{\mathbb{O}}\widehat{\mu_{\mathbb{O}}}$, where $\mathbb{O}$ varies over the set of nilpotent orbits of $\mathfrak{g}$, $c_{\mathbb{O}} \in \mathbb{C}$ and $\widehat{\mu_{\mathbb{O}}}$ is the Fourier transform of a suitably chosen measure $\mu_{\mathbb{O}}$ on $\mathbb{O}$. One defines $\mathcal{N}_{\mathrm{tr}}(\pi)$ to be the set of nilpotent orbits $\mathbb{O}$ of $\mathfrak{g}$ such that the corresponding coefficient $c_{\mathbb{O}}$ in the character expansion of π around the identity is nonzero.

We have the standard partial order on the set of nilpotent orbits in $\mathfrak{g}$: $\mathbb{O}_1 \le \mathbb{O}_2$ if $\mathbb{O}_1 \subset \overline{\mathbb{O}_2}$. Let $\operatorname{Max}\mathcal{N}_{\mathrm{Wh}}(\pi)$ and $\operatorname{Max}\mathcal{N}_{\mathrm{tr}}(\pi)$ denote the sets of maximal elements in $\mathcal{N}_{\mathrm{Wh}}(\pi)$ and $\mathcal{N}_{\mathrm{tr}}(\pi)$, respectively, with respect to this partial order.

Theorem 1 [Mœglin and Waldspurger 1987, Chapter I]. *Let $\boldsymbol{G}$ be a connected reductive group defined over E. Let π be an irreducible admissible representation of $G = \boldsymbol{G}(E)$. Then*

$$\operatorname{Max}\mathcal{N}_{\mathrm{Wh}}(\pi) = \operatorname{Max}\mathcal{N}_{\mathrm{tr}}(\pi).$$

Moreover, if $\mathbb{O}$ is an element in either of these sets, then for any (Y, φ) as above with $Y \in \mathbb{O}$ we have

$$c_{\mathbb{O}} = \dim \mathcal{W}_{(Y,\varphi)}.$$

If one considers the case of the pair (Y, φ) with Y a regular nilpotent element, this theorem specializes to Rodier's theorem.

In this paper we generalize the Mœglin–Waldspurger theorem to the setting of a covering group $\tilde{G}$ of G. Let μ_r be the group of r-th roots of unity in $\mathbb{C}^\times$. An r-fold covering group $\tilde{G}$ of G is a central extension of locally compact groups by $\mu_r := \{z \in \mathbb{C} : z^r = 1\}$ giving rise to the short exact sequence

$$1 \longrightarrow \mu_r \longrightarrow \tilde{G} \longrightarrow G \longrightarrow 1. \tag{2}$$

The representations of $\tilde{G}$ on which μ_r acts by the natural embedding $\mu_r \hookrightarrow \mathbb{C}^\times$ are called genuine representations. The definition of the space of degenerate Whittaker forms of a representation of G involves only unipotent groups. Since the covering

$\tilde{G} \to G$ splits over any unipotent subgroup of G in a unique way (see [Mœglin and Waldspurger 1995]) this makes it possible to define the space of degenerate Whittaker forms for any genuine smooth representation (π, W) of $\tilde{G}$. In particular, it makes sense to talk of the set $\mathcal{N}_{\mathrm{Wh}}(\pi)$.

The existence of the character expansion of an admissible genuine representation (π, W) of $\tilde{G}$ has been proved by Wen-Wei Li [2012]. At the identity, the Harish-Chandra–Howe character expansion of an irreducible genuine representation has the same form, and therefore we have $\mathcal{N}_{\mathrm{tr}}(\pi)$. This makes it possible to have an analogue of Theorem 1 in the setting of covering groups. The main aim of this paper is to prove the following.

Theorem 2. *Let π be an irreducible admissible genuine representation of $\tilde{G}$. Then*

$$\operatorname{Max} \mathcal{N}_{\mathrm{Wh}}(\pi) = \operatorname{Max} \mathcal{N}_{\mathrm{tr}}(\pi).$$

Moreover, if $\mathbb{O}$ is an element in either of these sets, then for any (Y, φ) as above with $Y \in \mathbb{O}$ we have

$$c_{\mathbb{O}} = \dim \mathcal{W}_{(Y,\varphi)}.$$

We will use the results in [Mœglin and Waldspurger 1987], and to accommodate the even residual characteristic case we follow [Varma 2014]. Let us describe some ideas involved in the proof. Let Y be a nilpotent element in $\mathfrak{g}$ and φ a one-parameter subgroup as above. Let $\mathfrak{g}_i$ be the eigenspace of weight i under the action of $\mathbb{G}_m$ on $\mathfrak{g}$ via $\mathrm{Ad} \circ \varphi$. One can attach a parabolic subgroup $\boldsymbol{P}$ with unipotent radical $\boldsymbol{N}$ whose Lie algebras are $\mathfrak{p} := \bigoplus_{i \geq 0} \mathfrak{g}_i$ and $\mathfrak{n} = \bigoplus_{i>0} \mathfrak{g}_i$ respectively. The one-parameter subgroup φ also determines a parabolic subgroup $\boldsymbol{P}^-$ opposite to $\boldsymbol{P}$ with Lie algebra $\mathfrak{p}^- = \bigoplus_{i \leq 0} \mathfrak{g}_i$. For simplicity, assume $\mathfrak{g}_1 = 0$ for the purpose of the introduction. Then $\mathfrak{n} = \bigoplus_{i \geq 2} \mathfrak{g}_i$ and $\chi : \gamma \mapsto \psi(B(Y, \log \gamma))$ defines a character of $N = \boldsymbol{N}(E)$, where B is an $\mathrm{Ad}(G)$-invariant nondegenerate symmetric bilinear form on $\mathfrak{g}$ and ψ is an additive character of E. In this case (i.e., $\mathfrak{g}_1 = 0$), the space of degenerate Whittaker forms $\mathcal{W}_{(Y,\varphi)}$ is defined to be the twisted Jacquet module of π with respect to (N, χ). In the case where $\mathfrak{g}_1 \neq 0$, the definition of $\mathcal{W}_{(Y,\varphi)}$ needs to be appropriately modified (see Section 4).

On the other hand, to the pair (Y, φ) one attaches certain open compact subgroups G_n of G for large n and certain characters χ_n of G_n. One then proves that the covering $\tilde{G} \to G$ splits over G_n for large n, so that the G_n can be seen as subgroups of $\tilde{G}$ as well. Let ϖ be a uniformizer of E. Let $t := \varphi(\varpi)$ and $\tilde{t}$ be any lift of t in $\tilde{G}$. It turns out that $\tilde{t}^{-n} G_n \tilde{t}^n \cap N$ becomes an "arbitrarily large" subgroup of N and $\tilde{t}^{-n} G_n \tilde{t}^n \cap P^-$ becomes an "arbitrarily small" subgroup of P^-, as n becomes large. For large n, the characters χ_n have been so defined that the character $\chi_n' := \chi_n \circ \mathrm{Int}(\tilde{t}^n)$ restricted to $\tilde{t}^{-n} G_n \tilde{t}^n \cap N$ agrees with χ. Using the Harish-Chandra–Howe character expansion, one proves that the dimension of the

(G_n, χ_n)-isotypic component of W is equal to $c_{\mathbb{O}}$ for large enough n, where $\mathbb{O}$ is the nilpotent orbit of Y in $\mathfrak{g}$. Finally, one proves that there is a natural isomorphism between the $(\tilde{t}^{-n} G_n \tilde{t}^n, \chi_n \circ \mathrm{Int}(\tilde{t}^n))$-isotypic component of W and $\mathcal{W}_{(Y,\varphi)}$.

Remark 3. The definition of $\mathcal{W}_{(Y,\varphi)}$ (hence that of $\mathcal{N}_{\mathrm{Wh}}(\pi)$) depends on a choice of an additive character ψ of E and a choice of $\mathrm{Ad}(G)$-invariant nondegenerate bilinear form B on $\mathfrak{g}$. On the other hand, in the character expansion, the $c_{\mathbb{O}}$ (hence $\mathcal{N}_{\mathrm{tr}}(\pi)$) depend on ψ, B, a measure on G and a measure on $\mathfrak{g}$. However, by choosing a compatible measure on G and $\mathfrak{g}$ via the exponential map, one gets rid of the dependency of $c_{\mathbb{O}}$ on these measures, and therefore the $c_{\mathbb{O}}$ depend only on ψ and B. For more detailed discussion about how our results depend on B and ψ, see Remark 4 in [Varma 2014].

Remark 4. One aspect in Varma's proof for $p = 2$ which does not obviously generalize from the proof for $p \neq 2$ is the prescription of the character χ_n of G_n given in [Mœglin and Waldspurger 1987], which is due to the somewhat bad behavior of the Campbell–Hausdorff formula in the $p = 2$ case. Using Kirillov's theory of compact p-adic groups, Varma prescribed a χ_n (although not unique) which will serve our purpose. On the other hand, the definition of degenerate Whittaker forms of W has also been modified by Varma to accommodate the case $p = 2$.

Although the methods used in the paper are not new and heavily depend on the proofs in the linear case, the result is useful in the study of representation theory of covering groups. The author himself has made use of this result in his thesis, where he attempts to generalize a result of D. Prasad [1992] in the setting of covering groups, namely, in the harmonic analysis relating the pairs $(\mathrm{GL}_2(E)^{\sim}, \mathrm{GL}_2(F))$ and $(\mathrm{GL}_2(E)^{\sim}, D_F^{\times})$, where E/F is a quadratic extension of nonarchimedean local field, D_F is the quaternion division algebra with center F and $\mathrm{GL}_2(E)^{\sim}$ is a certain twofold cover of $\mathrm{GL}_2(E)$.

Let us briefly give an outline of the organization of the paper. In Section 2, we recall the definition of the subgroups G_n and state some properties of the character χ_n. In Section 3, we recall splitting of the covering groups over G_n and describe an appropriate choice of the splitting over the subgroup G_n for large n. In Section 4 we give the definition of the space of degenerate Whittaker forms and describe important setup to prove the main theorem. In Section 5, we transfer some results from linear groups to covering groups in a few lemmas, and, based on these lemmas, we prove the main theorem.

2. Subgroups G_n and characters χ_n

In this section, we recall a certain sequence of subgroups G_n of G, which form a basis of neighborhoods at the identity, and certain characters $\chi_n : G_n \to \mathbb{C}^{\times}$. The objects involved in this section were defined for linear groups in [Mœglin and

Waldspurger 1987; Varma 2014], and we will lift them to covering groups in a suitable way in Section 3 and work with these lifts in this paper.

Let $\mathfrak{O}_E$ denote the ring of integers in E. We fix an additive character ψ of E with conductor $\mathfrak{O}_E$. Fix an $\mathrm{Ad}(G)$-invariant nondegenerate symmetric bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to E$.

Let $Y \in \mathfrak{g}$ be nilpotent. Choose a one-parameter subgroup $\varphi : \mathbb{G}_m \to \boldsymbol{G}$ satisfying

$$\mathrm{Ad}(\varphi(s))Y = s^{-2}Y \quad \text{for all } s \in \mathbb{G}_m. \tag{3}$$

For a given nilpotent element Y, the existence of such a φ is known from the theory of $\mathfrak{sl}_2$-triplets, but there are examples which do not come from this theory.

For $i \in \mathbb{Z}$, define

$$\boldsymbol{\mathfrak{g}}_i = \{X \in \boldsymbol{\mathfrak{g}} : \mathrm{Ad}(\varphi(s))X = s^i X \text{ for all } s \in \mathbb{G}_m\}.$$

Set

$$\boldsymbol{\mathfrak{n}} := \boldsymbol{\mathfrak{n}}^+ := \bigoplus_{i>0} \boldsymbol{\mathfrak{g}}_i, \quad \boldsymbol{\mathfrak{n}}^- := \bigoplus_{i<0} \boldsymbol{\mathfrak{g}}_i, \quad \boldsymbol{\mathfrak{p}}^- := \bigoplus_{i\leq 0} \boldsymbol{\mathfrak{g}}_i.$$

The parabolic subgroup $\boldsymbol{P}^-$ of $\boldsymbol{G}$ stabilizing $\boldsymbol{\mathfrak{n}}^-$ has $\boldsymbol{\mathfrak{p}}^-$ as its Lie algebra. Let $\boldsymbol{N} = \boldsymbol{N}^+$ be the unipotent subgroup of $\boldsymbol{G}$ having the Lie algebra $\boldsymbol{\mathfrak{n}}$.

Let $G(Y)$ be the centralizer of Y in G and $Y^\#$ the centralizer of Y in $\mathfrak{g}$. The G-orbit $\mathbb{O}_Y$ of Y can be identified with $G/G(Y)$ and therefore its tangent space at Y can be identified with $\mathfrak{g}/Y^\#$. Note that

$$\begin{aligned} Y^\# &= \{X \in \mathfrak{g} : [X, Y] = 0\} \\ &= \{X \in \mathfrak{g} : B([X, Y], Z) = 0 \text{ for all } Z \in \mathfrak{g}\} \\ &= \{X \in \mathfrak{g} : B(Y, [X, Z]) = 0 \text{ for all } Z \in \mathfrak{g}\}. \end{aligned}$$

The bilinear form B induces a nondegenerate alternating form

$$B_Y : \mathfrak{g}/Y^\# \times \mathfrak{g}/Y^\# \longrightarrow E$$

defined by $B_Y(X_1, X_2) = B(Y, [X_1, X_2])$.

Let $L \subset \mathfrak{g}$ be a lattice satisfying the following conditions:

(1) $[L, L] \subset L$;

(2) $L = \bigoplus_{i\in\mathbb{Z}} L_i$, where $L_i = L \cap \mathfrak{g}_i$;

(3) The lattice L/L_Y, where $L_Y = L \cap Y^\#$, is self-dual with respect to B_Y, i.e., $(L/L_Y)^\perp = L/L_Y$. (For any vector space V with a nondegenerate bilinear form B and a lattice M in V, $M^\perp := \{X \in V : B(X, Y) \in \mathfrak{O}_E \text{ for all } Y \in V\}$.)

A lattice L satisfying these properties can be chosen by taking a suitable basis of all the $\mathfrak{g}_i$; see [Mœglin and Waldspurger 1987]. Now we summarize a few well-known properties of the exponential map, and use them to define subgroups G_n and their Iwahori decompositions.

Lemma 5. (1) *There exists a positive integer A such that* $\exp$ *is defined and injective on $\varpi^A L$, with inverse* $\log$.

(2) *The exponential map on $\varpi^n L$ is a homeomorphism onto its image $G_n := \exp(\varpi^n L)$, which is an open subgroup of G for all $n \geq A$.*

(3) *Set $P_n^- = \exp(\varpi^n L \cap \mathfrak{p}^-)$ and $N_n = \exp(\varpi^n L \cap \mathfrak{n})$. Then we have an Iwahori factorization*

$$G_n = P_n^- N_n.$$

We will be working with a certain character χ_n of G_n, which we recall in the next lemma.

Lemma 6. *For large values of n there exists a character χ_n of G_n whose restriction to $\exp((Y^\# \cap \varpi^n L) + \varpi^{n+\operatorname{val} 2} L)$ coincides with $\gamma \mapsto \psi(B(\varpi^{-2n} Y, \log \gamma))$. If P_n^- is as in Lemma 5, the character χ_n can be chosen so that*

$$\chi_n(p) = 1 \quad \textit{for all } \; p \in P_n^-.$$

For a proof of this lemma and other properties of this character χ_n see Lemma 5 in [Varma 2014].

Remark 7. If $p \neq 2$, then the map $\gamma \mapsto \psi(B(\varpi^{-2n} Y, \log \gamma))$ itself defines a character of G_n for large n and satisfies the properties stated in Lemma 6. But for $p = 2$, there is more than one such character χ_n; for more details see [Varma 2014].

3. Covering groups

Let μ_r be the group of r-th roots of unity in $\mathbb{C}$. Consider an r-fold covering $\tilde{G}$ of G. Recall that this is a central extension of locally compact groups of the group G by μ_r giving rise to the short exact sequence

$$1 \longrightarrow \mu_r \longrightarrow \tilde{G} \longrightarrow G \longrightarrow 1.$$

Lemma 8. (1) *The covering $\tilde{G} \to G$ splits uniquely over any unipotent subgroup of G.*

(2) *For large enough n the covering $\tilde{G} \to G$ splits over G_n. Moreover, there is a splitting s of $\tilde{G} \to G$ restricted to $\bigcup_{g \in G} g G_n g^{-1}$ such that $s(hth^{-1}) = hs(t)h^{-1}$ for all $h \in G$.*

Proof. (1) This is well known; see [Mœglin and Waldspurger 1995]. For a simpler proof in the case when E has characteristic zero, see Section 2.2 of [Li 2014].

(2) Recall that the subgroups G_n form a basis of neighborhoods of the identity. It is well known that the covering $\tilde{G} \to G$ splits over a neighborhood of the identity. Therefore, for large enough n, the covering splits over G_n. There is more than one possible splitting of the cover $\tilde{G} \to G$ over G_n. If a splitting is fixed, then any other splitting over G_n will differ from the above splitting by a character $G_n \to \mu_r$.

Fix some m such that the covering splits over $G_m = \exp(\varpi^m L)$. As mentioned above, any two splittings over the subgroup G_m will differ by a character $G_m \to \mu_r$ and any such character is trivial over

$$G_m^r := \{g^r : g \in G_m\}.$$

Hence all the possible splittings over G_m agree on G_m^r. The subset G_m^r is a subgroup of G_m as it equals $\exp(r \cdot \varpi^m L)$. Let $g, h \in G$. We have

$$gG_m g^{-1} \cap hG_m h^{-1} \supset gG_m^r g^{-1} \cap hG_m^r h^{-1}.$$

This implies that any two splittings of $\tilde{G} \to G$ restricted to $gG_m^r g^{-1} \cap hG_m^r h^{-1}$, one coming from the restriction of a splitting of $\tilde{G} \to G$ over $gG_m g^{-1}$ and the other coming from the restriction of a splitting over $hG_m h^{-1}$, are the same. Now choose A' so large such that $G_n \subset G_m^r$ for $n \geq A'$. We fix the splitting of G_n which comes from that of the restriction of G_m^r. This gives us a splitting over $\bigcup_{g \in G} gG_n g^{-1}$. □

Using this splitting we get that the exponential map is defined from a small enough neighborhood of $\mathfrak{g}$ to $\tilde{G}$, namely the usual exponential map composed with this splitting, which one can use to define the character expansion of an irreducible admissible genuine representation (π, W) of $\tilde{G}$, which was done in [Li 2012].

Remark 9. If r is coprime to p, then as G_n is a pro-p group and $(r, p) = 1$, there is no nontrivial character from G_n to μ_r. In that situation, the splitting in the preceding lemma is unique.

From now onwards, for large enough n, we treat G_n not only as a subgroup of G but also as one of $\tilde{G}$, with the above specified splitting. In other words, for the covering group $\tilde{G}$ (as in the linear case) we have a sequence of pairs (G_n, χ_n) using the splitting specified above which satisfies the properties described in Section 2.

Definition 10. Let $H \subset G$ be an open subgroup and $s : H \hookrightarrow \tilde{G}$ be a splitting. Then for any $\phi \in C_c^\infty(G)$ with $\operatorname{supp}(\phi) \subset H$, define $\tilde{\phi}_s \in C_c^\infty(\tilde{G})$ by

$$\tilde{\phi}_s(g) := \begin{cases} \phi(g') & \text{if } g = s(g') \in s(H), \\ 0 & \text{if } g \in \tilde{G} \setminus s(H). \end{cases}$$

Note that this definition depends upon the choice of splitting. Whenever the splitting is clear in the context or it has been fixed and there is no confusion we write just $\tilde{\phi}$

instead of $\tilde{\phi}_s$ and H for $s(H)$. Recall that the convolution $\phi * \phi'$ for $\phi, \phi' \in C_c^\infty(G)$ is defined by

$$\phi * \phi'(x) = \int_G \phi(xy^{-1})\phi'(y)\, dy.$$

Observe that

$$\operatorname{supp}(\phi * \phi') \subset \operatorname{supp}(\phi) \cdot \operatorname{supp}(\phi'),$$

which implies the lemma below.

Lemma 11. *Let H be an open subgroup of G such that the covering $\tilde{G} \to G$ has a splitting over H, say, $s : H \hookrightarrow \tilde{G}$, satisfying $s(xy) = s(x)s(y)$ whenever x and y are in H. If $\phi, \phi' \in C_c^\infty(G)$ are such that supp(ϕ) and supp(ϕ') are contained in H, then we have*

$$\widetilde{\phi * \phi'} = \tilde{\phi} * \tilde{\phi}'.$$

4. Degenerate Whittaker forms

In this section we give the definition of degenerate Whittaker forms for a smooth genuine representation π of $\tilde{G}$. This is an adaptation of Section I.7 of [Mœglin and Waldspurger 1987] and Section 5 of [Varma 2014].

Define

$$N := \exp \mathfrak{n} = \exp \bigoplus_{i \geq 1} \mathfrak{g}_i, \quad N^2 := \exp \bigoplus_{i \geq 2} \mathfrak{g}_i, \quad \text{and } N' := \exp(\mathfrak{g}_1 \cap Y^\#)N^2.$$

It is easy to see that N^2 and N' are normal subgroups of N. Let H be the Heisenberg group defined with $\mathfrak{g}_1/(\mathfrak{g}_1 \cap Y^\#) \times E$ as underlying set using the symplectic form induced by B_Y, i.e., for $X, Z \in \mathfrak{g}_1/(\mathfrak{g}_1 \cap Y^\#)$ and $a, b \in E$,

$$(X, a)(Z, b) = \big(X + Z, a + b + \tfrac{1}{2} B_Y(X, Z)\big). \tag{4}$$

Consider the map $N \to H$ given by

$$\exp X \mapsto (\bar{X}, B(Y, X)),$$

where $\bar{X}$ is the image of the $\mathfrak{g}_1$ component of X in $\mathfrak{g}_1/(\mathfrak{g}_1 \cap Y^\#)$. The Campbell–Hausdorff formula implies that this map is a homomorphism with kernel

$$N'' = \{n \in N' : B(Y, \log n) = 0\}.$$

Let $\chi : N' \to \mathbb{C}^\times$ be defined by $\gamma \mapsto \psi \circ B(Y, \log \gamma)$. Note that

$$\gamma \mapsto B(Y, \log \gamma) \in E \cong \{0\} \times E \subset H$$

induces an isomorphism $N'/N'' \cong E$.

We note that the cover $\tilde{G} \to G$ splits uniquely over the subgroups N, N' and N''. We denote the images of these splittings inside $\tilde{G}$ by the same letters. For a smooth genuine representation (π, W) of $\tilde{G}$ we define

$$N^2_\chi W = \{\pi(n)w - \chi(n)w : w \in W, n \in N^2\}$$

and

$$N'_\chi W = \{\pi(n)w - \chi(n)w : w \in W, n \in N'\}.$$

Note that N normalizes χ, therefore $H = N/N''$ acts on $W/N'_\chi W$ in a natural way. This action restricts to N'/N'' (the center of N/N'') as multiplication by the character χ. Let $\mathcal{S}$ be the unique irreducible representation of the Heisenberg group H with central character χ.

Definition 12. Define the space of degenerate Whittaker forms for (π, W) associated to (Y, φ) to be

$$\mathcal{W} := \mathrm{Hom}_H(\mathcal{S}, W/N'_\chi W).$$

Remark 13. If $\mathfrak{g}_1 = 0$, then $N = N' = N^2$. In this case, $\mathcal{W} \cong W/N_\chi W$ is the (N, χ)-twisted Jacquet functor.

Definition 14. For a smooth representation (π, W) of $\tilde{G}$, define $\mathcal{N}_{\mathrm{Wh}}(\pi)$ to be the set of nilpotent orbits $\mathbb{O}$ of $\mathfrak{g}$ such that there exist $Y \in \mathbb{O}$ and φ as in (3), such that the space of degenerate Whittaker forms for π associated to (Y, φ) is nonzero.

As $\mathfrak{g}_1/\mathfrak{g}_1 \cap Y^\#$ is a symplectic vector space and L/L_Y is self-dual, it follows that $L_H := (L \cap \mathfrak{g}_1)/(L \cap \mathfrak{g}_1 \cap Y^\#)$ is a self-dual lattice in the symplectic vector space $H/Z(H) \cong \mathfrak{g}_1/(\mathfrak{g}_1 \cap Y^\#)$.

Recall the definition (4) of the Heisenberg group H. Since ψ is trivial on $\mathfrak{O}_E$, it follows that one can extend the character ψ of $E \cong Z(H)$ to a character of the inverse image of $2L_H$ under $H \to \mathfrak{g}_1/(\mathfrak{g}_1 \cap Y^\#)$ by defining it to be trivial on $2L_H \times \{0\} \subset H$. From Lemma 4 in [Varma 2014], this character can be extended to a character $\tilde{\chi}$ on the inverse image H_0 of L_H under the natural map $H \to \mathfrak{g}_1/(\mathfrak{g}_1 \cap Y^\#)$.

Remark 15. There are one-parameter subgroups φ which do not arise from $\mathfrak{sl}_2$-triplets. If φ arises from an $\mathfrak{sl}_2$-triplet, then it is easy to see that $Y^\# \subset \bigoplus_{i \le 0} \mathfrak{g}_i$. In particular, we have $\mathfrak{g}_1 \cap Y^\# = \{0\}$ and hence the Heisenberg group H coincides with $\mathfrak{g}_1 \times E$.

Then, by Chapter 2, Section I.3 of [Mœglin et al. 1987], one knows that $\mathcal{S} = \mathrm{ind}_{H_0}^H \tilde{\chi}$ (induction with compact support). Since H_0 is an open subgroup of the locally profinite group H, we have the Frobenius reciprocity law

$$\mathrm{Hom}_H(\mathcal{S}, \tau) = \mathrm{Hom}_H(\mathrm{ind}_{H_0}^H \tilde{\chi}, \tau) = \mathrm{Hom}_{H_0}(\tilde{\chi}, \tau|_{H_0})$$

for any smooth representation τ of H. Thus, in the category of representations of N on which N' acts via the character χ, the functor $\mathrm{Hom}_H(\mathcal{S}, -)$ amounts to taking the $\tilde{\chi}|_{H_0}$-isotypic component. Since H_0 is compact modulo the center, this functor is exact. Thus, we have

$$\mathscr{W} = \mathrm{Hom}_H(\mathcal{S}, W/N'_\chi W) \cong (W/N'_\chi W)^{(H_0,\tilde{\chi})}.$$

where $(W/N'_\chi W)^{(H_0,\tilde{\chi})}$ denotes the $(H_0, \tilde{\chi})$-isotypic component of $W/N'_\chi W$.

Recall that we have defined certain characters χ_n in Section 2 and now we have a character $\tilde{\chi}$. We need to choose them in a compatible way. First we fix a character $\tilde{\chi}$ and consider it as a character of $\exp(\mathfrak{g}_1 \cap L)N'$ in the obvious way (as $\exp(\mathfrak{g}_1 \cap L)N'$ is the inverse image of H_0 under $N \to H$). Let $t := \varphi(\varpi) \in G$. Let $\tilde{t} \in \tilde{G}$ be any lift of t in $\tilde{G}$. Let

$$G'_n = \mathrm{Int}(\tilde{t}^{-n})(G_n), \quad P'_n = \mathrm{Int}(\tilde{t}^{-n})(P_n^-) \quad \text{and} \quad V'_n = \mathrm{Int}(\tilde{t}^{-n})(N_n).$$

It can be easily verified that V'_n contains $\exp(\mathfrak{g}_1 \cap L)$. We also have $V'_n \subset V'_m$ for large m, n with $n \leq m$. Moreover,

$$\exp(\mathfrak{g}_1 \cap L)N^2 = \bigcup_{n \geq 0} V'_n.$$

It can also be verified easily that $\tilde{\chi} \circ \mathrm{Int}(\tilde{t}^{-n})$ restricts to a character of N_n that extends the character on $N_{n+\mathrm{val}\,2}N'_n$ given by $\gamma \mapsto \psi(B(\varpi^{-2n}Y, \log\gamma))$. Now define

$$\chi_n(pv) = \tilde{\chi}(\tilde{t}^{-n}v\tilde{t}^n) \quad \text{for all } p \in P_n^- \text{ and all } v \in V'_n. \tag{5}$$

Lemma 16 [Varma 2014, Lemma 6]. *Let χ_n be as defined in* (5). *Then χ_n is a character of G_n and satisfies the properties stated in Lemma 6.*

Define a character χ'_n on G'_n by

$$\chi'_n := \chi_n \circ \mathrm{Int}(\tilde{t}^n).$$

Remark 17. The characters χ_n have been defined so that the χ'_n agree with χ on the intersection of their domains, namely, for large n we have

$$\chi'_n|_{V'_n} = \tilde{\chi}|_{V'_n}.$$

In particular, $\chi'_n|_{\exp(L\cap\mathfrak{g}_1)} = \tilde{\chi}|_{\exp(L\cap\mathfrak{g}_1)}$. One can also see that χ'_n and χ'_m (for large n, m) agree on $G'_n \cap G'_m$, because they agree on $V'_n \cap V'_m$ and also on $P'_n \cap P'_m$ (being trivial on it).

Set

$$W_n := \{w \in W : \pi(\gamma)w = \chi_n(\gamma)w \text{ for all } \gamma \in G_n\} \tag{6}$$

and

$$W'_n := \{w \in W : \pi(\gamma)w = \chi'_n(\gamma)w \text{ for all } \gamma \in G'_n\} = \pi(\tilde{t}^{-n})W_n. \tag{7}$$

For large m, n, define the map $I'_{n,m} : W'_n \to W'_m$ by

$$I'_{n,m}(w) = \int_{G'_m} \chi'_m(\gamma^{-1})\pi(\gamma)w \, d\gamma. \tag{8}$$

Let m, n be large with $m > n$. Since χ'_n is trivial on $P'_n \supset P'_m$ and since $G'_m = P'_m V'_m$, for a convenient choice of measures we have

$$\begin{aligned} I'_{n,m}(w) &= \int_{V'_m} \chi'_m(x^{-1})\pi(x)w \, dx \\ &= \int_{\exp(\mathfrak{g}_1 \cap L)} \tilde{\chi}^{-1}(\exp X)\pi(\exp X) \int_{N^2 \cap G'_m} \chi(x^{-1})\pi(x)w \, dx \, dX. \end{aligned}$$

Now using the fact that $\exp(\mathfrak{g}_1 \cap L)$ lies in G'_n for large n and that it normalizes the character $\chi|_{N^2}$, we get

$$I'_{n,m}(w) = \int_{N^2 \cap G'_m} \chi(x^{-1})\pi(x)w \, dx = \int_{N' \cap G'_m} \chi(x^{-1})\pi(x)w \, dx.$$

From this the following is clear for large n, m with $m > n$:

$$I'_{n,m} = I'_{n+1,m} \circ I'_{n,n+1}. \tag{9}$$

For large n, this equation gives that $\ker I'_{n,m} \subset \ker I'_{n,p}$ for $n < m \leq p$. Set

$$W'_{n,\chi} := \bigcup_{m>n} \ker I'_{n,m}.$$

Recall that for any unipotent subgroup U, character $\chi : U \to \mathbb{C}^\times$ and $w \in W$, we have that

$$\int_K \chi(x)^{-1}\pi(x)w \, dx = 0$$

for some open compact subgroup K of U if and only if $w \in U_\chi W$, where $U_\chi W$ is the span of $\{\pi(u)w - \chi(u)w : u \in U, w \in W\}$. Thus we have $W_{n,\chi} \subset N^2_\chi W$ as well as $W_{n,\chi} \subset N'_\chi W$, which gives the natural maps

$$j_n : W'_n/W'_{n,\chi} \longrightarrow W/N^2_\chi W \quad \text{and} \quad j'_n : W'_n/W'_{n,\chi} \longrightarrow W/N'_\chi W,$$

and these give the diagram

$$\begin{tikzcd} W'_n/W'_{n,\chi} \arrow[rr, "j'_n"] \arrow[dr, "j_n"'] & & W/N'_\chi W \\ & W/N^2_\chi W \arrow[ur, dashed, "\exists\,\text{natural}"'] & \end{tikzcd} \tag{10}$$

By the compatibility between χ_n' and $\tilde{\chi}$, it is easy to see that the image of j_n' is contained in $(W/N_\chi' W)^{(H_0,\tilde{\chi})}$. Let $w \in W$ such that the image $\bar{w}$ of w in $W/N_\chi' W$ belongs to $(W/N_\chi' W)^{(H_0,\tilde{\chi})}$. For large n, P_n' acts trivially on w, as (π, W) is smooth. Since $G_n' = P_n' V_n' = V_n' P_n'$, the element

$$\int_{V_n'} \chi_n'(x^{-1})\pi(x) w \, dx$$

belongs to W_n'. As χ_n' and χ are compatible, it can be seen that its image in $W/N_\chi' W$ is $\bar{w}$. This gives us the following lemma.

Lemma 18. *Let (Y, φ) be arbitrary. Then any element of $(W/N_\chi' W)^{(H_0,\chi)}$ belongs to $j_n'(W_n')$ for all sufficiently large n. In particular, if $\mathcal{W} \neq 0$, then W_n and W_n' are nonzero for large n.*

5. Main theorem

Now recall that, by the work of Li [2012], the Harish-Chandra–Howe character expansion of an irreducible admissible genuine representation of $\tilde{G}$ at the identity element has an expression of the same form as that of an irreducible admissible representation of a linear group. The proof of the following lemma for a covering group follows verbatim that of Proposition I.11 in [Mœglin and Waldspurger 1987] and Proposition 1 in [Varma 2014].

Proposition 19. *Let $\mathcal{W}$ be the space of degenerate Whittaker forms for π with respect to a given (Y, φ). If $\mathcal{W} \neq 0$, then there exists a nilpotent orbit $\mathbb{O}$ in $\mathcal{N}_{\mathrm{tr}}(\pi)$ such that $\mathbb{O}_Y \leq \mathbb{O}$, i.e., $Y \in \bar{\mathbb{O}}$.*

Let the function $\phi_n : G \to \mathbb{C}$ be defined by

$$\phi_n(\gamma) = \begin{cases} \chi_n(\gamma^{-1}) & \text{if } \gamma \in G_n, \\ 0 & \text{otherwise.} \end{cases}$$

Consider the corresponding function $\tilde{\phi_n} : \tilde{G} \to \mathbb{C}$. Write the character expansion at the identity element as

$$\Theta_\pi \circ \exp = \sum_{\mathbb{O}} c_{\mathbb{O}} \widehat{\mu_{\mathbb{O}}}.$$

Choose n large enough so that this expansion is valid over G_n, and then evaluate Θ_π at the function $\tilde{\phi_n}$. As $\pi(\tilde{\phi_n})$ is a projection from W to W_n, by definition we get $\Theta_\pi(\tilde{\phi_n}) = \text{trace } \pi(\tilde{\phi_n}) = \dim W_n$. Now assume that (Y, φ) is such that O_Y is a maximal element in $\mathcal{N}_{\mathrm{tr}}(\pi)$. On the other hand, if we evaluate $\sum_{\mathbb{O}} c_{\mathbb{O}} \widehat{\mu_{\mathbb{O}}}(\tilde{\phi_n})$, it turns out that $\widehat{\mu_{\mathbb{O}}}(\tilde{\phi_n})$ is zero unless $\mathbb{O} = \mathbb{O}_Y$. In addition, if we fix a G-invariant measure on $\mathbb{O}_Y$ as in I.8 of [Mœglin and Waldspurger 1987] (for more details about this invariant measure see Section 3 of [Varma 2014]), we get the following lemma.

Lemma 20 (Lemma I.12 in [Mœglin and Waldspurger 1987] and Lemma 7 in [Varma 2014]). *If (Y, φ) is such that $\mathbb{O}_Y$ is a maximal element of $\mathcal{N}_{\mathrm{tr}}(\pi)$. Then for large n,*

$$\dim W_n = c_{\mathbb{O}_Y}.$$

In particular, the dimension of W_n is finite and independent of n, for large n.

From Lemma 18 we know that every vector in $\mathcal{W}$ is in the image of j'_n for large n. In particular, if W_n is finite-dimensional, we get that the map j'_n is surjective. Moreover, we have the following lemma, whose proof is verbatim that of Corollary I.14 in [Mœglin and Waldspurger 1987] and Lemma 8 in [Varma 2014] in the case of a linear group.

Lemma 21. *Let (Y, φ) be such that $\mathbb{O}_Y$ is a maximal element of $\mathcal{N}_{\mathrm{tr}}(\pi)$. Then for large n, the maps j_n and j'_n are injections and the image of j'_n is $(W/N'_\chi W)^{(H_0,\tilde{\chi})}$.*

Let $\phi'_n : G \to \mathbb{C}$ be defined by

$$\phi'_n(\gamma) = \begin{cases} \chi'_n(\gamma^{-1}) & \text{if } \gamma \in G'_n, \\ 0 & \text{otherwise.} \end{cases}$$

Consider the corresponding function $\tilde{\phi}'_n : \tilde{G} \to \mathbb{C}$.

Lemma 22. *Consider a pair (Y, φ) such that $\mathbb{O} = \mathbb{O}_Y$ is a maximal in $\mathcal{N}_{\mathrm{tr}}(\pi)$. Then, for large enough n:*

(1) *Let $\mathcal{Y}_n \subset G'_{n+1} \cap G(Y)$ be a set of representatives for the G'_n double cosets in $G'_n(G_{n+1} \cap G(Y))G'_n$. Then for large enough n,*

$$\tilde{\phi}'_n * \tilde{\phi}'_{n+1} * \tilde{\phi}'_n(g) = \begin{cases} \lambda \cdot (\chi'_n)^{-1}(h_1 h_2) & \textit{if } g = h_1 y h_2 \textit{ with } y \in \mathcal{Y}_n,\, h_1, h_2 \in G'_n, \\ 0 & \textit{if } g \notin G'_n \mathcal{Y}_n G'_n, \end{cases}$$

where $\lambda = \operatorname{meas}(G'_n \cap G'_{n+1}) \operatorname{meas}(G'_n)$.

(2) *For large n, $I'_{n,n+1}$ is injective.*

Proof. From Lemma 9(a) in [Varma 2014], we have

$$\phi'_n * \phi'_{n+1} * \phi'_n(g) = \begin{cases} \lambda \cdot (\chi'_n)^{-1}(h_1 h_2) & \text{if } g = h_1 y h_2 \text{ with } y \in \mathcal{Y}_n,\, h_1, h_2 \in G'_n, \\ 0 & \text{if } g \notin G'_n \mathcal{Y}_n G'_n, \end{cases}$$

where $\lambda = \operatorname{meas}(G'_n \cap G'_{n+1}) \operatorname{meas}(G'_n)$. Now part 1 follows from Lemma 11, as we have

$$\tilde{\phi}'_n * \tilde{\phi}'_{n+1} * \tilde{\phi}'_n = (\phi'_n * \phi'_{n+1} * \phi'_n)^{\sim}. \tag{11}$$

Now we prove part 2. It is enough to say that $\pi(\tilde{\phi}'_n * \tilde{\phi}'_{n+1} * \tilde{\phi}'_n)$ acts by a nonzero multiple of the identity on W'_n. This implies that $I'_{n+1,n} \circ I'_{n,n+1}$ is a nonzero multiple of the identity on W'_n. From part 1, $\tilde{\phi}'_n * \tilde{\phi}'_{n+1} * \tilde{\phi}'_n$ is a positive linear combination of functions $\tilde{\phi}'_{n,y} : \gamma \mapsto \tilde{\phi}'_n(\gamma y^{-1})$, where $y \in G_{n+1} \cap G(Y)$ is fixed and $G(Y)$ is the centralizer of Y in G. Then the lemma follows from the fact that $\pi(y)$ acts trivially on W'_n for large n, so that

$$\pi(\tilde{\phi}'_{n,y})|_{W'_n} = \pi(\tilde{\phi}'_n)\pi(y)|_{W'_n} = \pi(\tilde{\phi}'_n)|_{W'_n}. \qquad \square$$

Theorem 23. *Let (π, W) be an irreducible admissible genuine representation of $\tilde{G}$.*

(1) *The set of maximal elements in $\mathcal{N}_{\mathrm{tr}}(\pi)$ coincides with the set of maximal elements in $\mathcal{N}_{\mathrm{Wh}}(\pi)$.*

(2) *Let $\mathbb{O}$ be a maximal element in $\mathcal{N}_{\mathrm{tr}}(\pi)$. Then the coefficient $c_{\mathbb{O}}$ equals the dimension of the space of degenerate Whittaker forms with respect to any pair (Y, φ) such that $Y \in \mathbb{O}$ is arbitrary and $\varphi : \mathbb{G}_m \to \boldsymbol{G}$ satisfies $\mathrm{Ad}(\varphi(s))Y = s^{-2}Y$ for all $s \in E^\times$.*

Proof. Let $\mathbb{O}$ be a maximal element in $\mathcal{N}_{\mathrm{tr}}(\pi)$. Choose (Y, φ) such that $Y \in \mathbb{O}$ and $\varphi : \mathbb{G}_m \to \boldsymbol{G}$ satisfies $\mathrm{Ad}(\varphi(s))Y = s^{-2}Y$. Then, from Lemma 20, for large n we have

$$\dim W_n = c_{\mathbb{O}}.$$

Therefore $W_n \neq 0$ and $W'_n \neq 0$ for large n. From Lemma 21, the map j'_n is injective and maps surjectively onto $(W/N'_\chi W)^{(H_0, \tilde{\chi})}$. But from the second part of Lemma 22 and (9), $I'_{n,m}$ is injective for large n and $m > n$, which implies that

$$W'_{n,\chi} = \bigcup_{m>n} \ker(I'_{n,m}) = 0.$$

Thus $\dim \mathcal{W} = \dim W'_n = \dim W_n = c_{\mathbb{O}}$, which proves part 2 of the theorem. In particular, $\mathcal{W} \neq 0$ and hence $\mathbb{O} \in \mathcal{N}_{\mathrm{Wh}}(\pi)$. Now we claim that $\mathbb{O}$ is maximal in $\mathcal{N}_{\mathrm{Wh}}(\pi)$. If not, there is a maximal orbit $\mathbb{O}' \in \mathcal{N}_{\mathrm{Wh}}(\pi)$ such that $\mathbb{O} \lneq \mathbb{O}'$. From Proposition 19, there is $\mathbb{O}'' \in \mathcal{N}_{\mathrm{tr}}(\pi)$ such that $\mathbb{O}' \leq \mathbb{O}''$. Therefore, $\mathbb{O} \lneq \mathbb{O}''$ and $\mathbb{O}, \mathbb{O}'' \in \mathcal{N}_{\mathrm{tr}}(\pi)$, a contradiction to the maximality of $\mathbb{O}$ in $\mathcal{N}_{\mathrm{tr}}(\pi)$.

Let $\mathbb{O}$ be a maximal element in $\mathcal{N}_{\mathrm{Wh}}(\pi)$. From Proposition 19, there is an element in $\mathbb{O}' \in \mathcal{N}_{\mathrm{tr}}(\pi)$ such that $\mathbb{O} \leq \mathbb{O}'$. Take a maximal such $\mathbb{O}'$. Then, by the result in the preceding paragraph, $\mathbb{O}'$ is a maximal element in $\mathcal{N}_{\mathrm{Wh}}(\pi)$. But $\mathbb{O}$ is also maximal in $\mathcal{N}_{\mathrm{Wh}}(\pi)$. Hence $\mathbb{O} = \mathbb{O}'$. This proves that $\mathbb{O}$ is a maximal element in $\mathcal{N}_{\mathrm{tr}}(\pi)$ too. $\square$

Acknowledgements

The author would like to express his gratitude to Professor D. Prasad and Professor Sandeep Varma for their help and numerous suggestions at various points. Without their help and continuous encouragement this paper would not have been possible.

References

[Li 2012] W.-W. Li, "La formule des traces pour les revêtements de groupes réductifs connexes, II: Analyse harmonique locale", *Ann. Sci. Éc. Norm. Supér.* (4) **45**:5 (2012), 787–859. MR 3053009 Zbl 06155586

[Li 2014] W.-W. Li, "La formule des traces pour les revêtements de groupes réductifs connexes, I: Le développement géométrique fin", *J. Reine Angew. Math.* **686** (2014), 37–109. MR 3176600 Zbl 06296344

[Mœglin and Waldspurger 1987] C. Mœglin and J.-L. Waldspurger, "Modèles de Whittaker dégénérés pour des groupes p-adiques", *Math. Z.* **196**:3 (1987), 427–452. MR 89f:22024 Zbl 0612.22008

[Mœglin and Waldspurger 1995] C. Mœglin and J.-L. Waldspurger, "Appendix I: Lifting of unipotent subgroups into a central extension", pp. 273–277 in *Spectral decomposition and Eisenstein series: a paraphrase of the Scriptures*, Cambridge Tracts in Mathematics **113**, Cambridge University Press, 1995. MR 97d:11083 Zbl 0846.11032

[Mœglin et al. 1987] C. Mœglin, M.-F. Vignéras, and J.-L. Waldspurger, *Correspondances de Howe sur un corps p-adique*, Lecture Notes in Math. **1291**, Springer, Berlin, 1987. MR 91f:11040 Zbl 0642.22002

[Prasad 1992] D. Prasad, "Invariant forms for representations of GL_2 over a local field", *Amer. J. Math.* **114**:6 (1992), 1317–1363. MR 93m:22011 Zbl 0780.22004

[Rodier 1975] F. Rodier, "Modèle de Whittaker et caractères de représentations", pp. 151–171 in *Non-commutative harmonic analysis* (Marseille-Luminy, 1974), edited by J. Carmona et al., Lecture Notes in Math. **466**, Springer, Berlin, 1975. MR 52 #14165 Zbl 0339.22014

[Varma 2014] S. Varma, "On a result of Moeglin and Waldspurger in residual characteristic 2", *Math. Z.* **277**:3-4 (2014), 1027–1048. MR 3229979 Zbl 06323349

Received February 10, 2014. Revised April 13, 2014.

SHIV PRAKASH PATEL
SCHOOL OF MATHEMATICS
TATA INSTITUTE OF FUNDAMENTAL RESEARCH
HOMI BHABHA ROAD
COLABA
MUMBAI 400005
INDIA
shiv@math.tifr.res.in

PACIFIC JOURNAL OF MATHEMATICS
Vol. 273, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.273.241

SPANNING TREES AND RANDOM WALKS ON WEIGHTED GRAPHS

XIAO CHANG, HAO XU AND SHING-TUNG YAU

Using two graph invariants arising from Chung and Yau's discrete Green's function, we derive explicit formulas and new estimates of hitting times of random walks on weighted graphs through the enumeration of spanning trees.

1. Introduction

Every reversible Markov chain can be viewed as a random walk on a weighted undirected graph $G=(V,E)$ with edge weights w_{xy}. We may assume that G has no multiedges but may have a loop of weight w_{xx} at each vertex x. The weighted degree d_x of x is the sum of all w_{xy}, $y \in V$. The *volume* of a graph is $\operatorname{vol}(G)=\sum_{v\in V} d_v$.

The *Laplacian* of G is the matrix $L=D-A$, where D is the diagonal matrix whose entries are d_x, $x\in V$ and A is the adjacency matrix of G. *Chung's normalized Laplacian*, $\mathcal{L}=D^{-1/2}LD^{-1/2}$, is

$$\mathcal{L}(x,y)=\begin{cases} 1-w_{xx}/d_x & \text{if } x=y,\\ -w_{xy}/\sqrt{d_xd_y} & \text{if } x\sim y,\\ 0 & \text{otherwise,}\end{cases}$$

where $x\sim y$ denotes that x and y are adjacent. Here we assume that $d_x\neq 0$ for all $x\in V$, since for a random walk it is natural to impose that G is connected and $w_{xy}>0$ for all $xy\in E$.

If G is connected, denote by $0=\lambda_0<\lambda_1\le\lambda_2\le\cdots\le\lambda_{n-1}$ the eigenvalues of $\mathcal{L}$ with the corresponding orthonormal basis of eigenvectors $\phi_0,\phi_1,\dots,\phi_{n-1}$, as $n\times 1$ column vectors. Obviously $\phi_0(x)=\sqrt{d_x/\operatorname{vol}(G)}$.

Chung and Yau [2000] defined the *discrete Green's function* $\mathcal{G}$ by

$$\mathcal{G}=\sum_{j=1}^{n-1}\frac{1}{\lambda_j}\phi_j\phi_j^*,$$

MSC2010: 05C81.
Keywords: hitting time, random walk, spanning tree.

which is uniquely determined by the relations $\mathscr{G}\mathscr{L} = \mathscr{L}\mathscr{G} = I - P_0$ and $\mathscr{G}P_0 = 0$, where $P_0 = \phi_0\phi_0^t$ is an $n \times n$ matrix.

A random walk on G is a Markov chain on V with transition probability matrix $(p_{xy})_{x,y\in V}$, where

$$p_{xy} = \begin{cases} w_{xy}/d_x & \text{if } x \sim y, \\ 0 & \text{otherwise.} \end{cases}$$

The *hitting time* $H(x, y)$ is the expected number of steps to reach vertex y when started from vertex x. By using a result in [Aldous and Fill 2014], Chung and Yau proved an expression for $H(x, y)$ in terms of the discrete Green's function:

Theorem 1.1 [Chung and Yau 2000]. *On a connected graph G, the hitting time $H(x, y)$ and Green's function $\mathscr{G}(x, y)$ satisfy*

$$H(x, y) = \operatorname{vol}(G)\left(\frac{\mathscr{G}(y, y)}{d_y} - \frac{\mathscr{G}(x, y)}{\sqrt{d_x d_y}}\right). \tag{1}$$

For a weighted graph G, we denote by $\Omega(G)$ the set of spanning trees of G. For $T \in \Omega(G)$, define the weight $w(T)$ of T to be $\prod_{e\in T} w_e$. Let $\tau(G)$ be the weighted counting of spanning trees:

$$\tau(G) = \sum_{T\in\Omega(G)} w(T). \tag{2}$$

Below is a typical identity expressing hitting times in terms of spanning trees (see Theorem 2.11) arising from our study of Chung and Yau's discrete Green's function.

Theorem 1.2. *Let G be a connected weighted graph and $x, y \in V(G)$. Then*

$$H(x, y) = \frac{1}{\tau(G)} \sum_{u\in V(G)} d_u \sum_{\substack{P\in\mathscr{P}_G(x,u)\\ y\notin P}} \prod_{e\in E(P)} w_e \tau(G/\{P, y\}). \tag{3}$$

The paper is organized as follows. In Section 2, we introduce two graph invariants and use them to derive explicit formulas for Chung and Yau's discrete Green's functions and hitting times of random walks on weighted graphs. In Section 3, we apply our formulas to obtain various estimates of hitting times on weighted graphs. In Section 4, we prove an explicit formula for hitting times of random walks on infinite trees. In Section 5, we apply our work to improve estimates of hitting times under different weight schemes on a given simple finite graph.

2. The hitting time of random walks on weighted graphs

Kirchhoff discovered his matrix-tree theorem in 1847 in his work on electrical networks, and this theorem gives an efficient way to calculate $\tau(G)$ using linear algebra.

Theorem 2.1 (Kirchhoff's matrix-tree theorem). *Let G be a connected weighted graph.*

(i) *If the Laplacian L of G has eigenvalues $0=\mu_0<\mu_1\le\cdots\le\mu_{n-1}$, then*

$$\prod_{k=1}^{n-1}\mu_k = n\tau(G). \tag{4}$$

(ii) *Let L_{ij} be the matrix obtained from L by deleting the i-th row and j-th column. Then all cofactors $(-1)^{i+j}\det(L_{ij})$ of L are equal and*

$$\det L_{ii} = n\tau(G) \quad \textit{for all } 1\le i\le n. \tag{5}$$

We also need the following version of Kirchhoff's matrix-tree theorem for weighted graphs. A proof, with slight changes, can be found in the cited reference.

Theorem 2.2 [Chung 2011, Theorem 1]. *For a connected weighted graph $G=(V,E)$, we have*

$$\prod_{k=1}^{n-1}\lambda_k = \frac{\mathrm{vol}(G)\tau(G)}{\prod_{v\in V} d_v}, \tag{6}$$

where $0=\lambda_0<\lambda_1\le\cdots\le\lambda_{n-1}$ are eigenvalues of $\mathcal{L}$ and $\tau(G)$ is defined in (2).

A weighted graph is a graph G equipped with a function $w: E(G)\to\mathbb{R}_+$ that assigns a positive number to each edge. In the following, we fix a weighted graph (G,w) and introduce two invariants for any induced subgraph S of G. Consider the matrix

$$B(x,y)=\begin{cases} d_x^2 s + d_x - w_{xx} & \text{if } x=y,\\ d_x d_y s - w_{xy} & \text{if } x\sim y,\\ d_x d_y s & \text{otherwise,}\end{cases} \tag{7}$$

where $d_x = w_{xx}+\sum_{y\sim x} w_{xy}$. Denote by B_S the principle submatrix of B on indices corresponding to the vertices of S; we define $R(S)$ and $Z(S)$ by

$$\det B_S = R(S) + Z(S)\cdot s.$$

In particular, $R(\varnothing)=1$, $Z(\varnothing)=0$ for the empty subgraph $\varnothing$, and $R(\{x\})=d_x-w_{xx}$, $Z(\{x\})=d_x^2$ for any $x\in V(G)$.

In the following four lemmas, S is an arbitrary induced subgraph of a fixed weighted graph (G, w). The proofs of these lemmas are similar to those in [Xu and Yau 2013a, Section 2], where we considered unweighted simple graphs with $w_e \equiv 1$ for all $e \in E(G)$.

Lemma 2.3. *If S has k connected components $S_1, \dots, S_k$, then*

$$R(S) = \prod_{i=1}^{k} R(S_i), \quad Z(S) = \sum_{i=1}^{k} Z(S_i) \prod_{\substack{j=1 \\ j \neq i}}^{k} R(S_j).$$

Lemma 2.4. *For any fixed vertex $x \in V(S)$, we have*

$$R(S) = (d_x - w_{xx}) R(S - \{x\}) - \sum_{\substack{y \in V(S) \\ y \sim x}} w_{xy} \sum_{P \in \mathcal{P}_S(x,y)} \prod_{e \in E(P)} w_e R(S - \{P\}),$$

and

$$\begin{aligned} Z(S) &= (d_x - w_{xx}) Z(S - \{x\}) \\ &\quad - \sum_{\substack{y \in V(S) \\ y \sim x}} w_{xy} \sum_{P \in \mathcal{P}_S(x,y)} \prod_{e \in E(P)} w_e Z(S - \{P\}) + (d_x - w_{xx})^2 R(S - \{x\}) \\ &\quad + \sum_{\substack{u,v \in V(S) \\ u \neq v}} d_u d_v \sum_{\substack{P_1 \in \mathcal{P}_S(x,u) \\ P_2 \in \mathcal{P}_S(x,v) \\ P_1 \cap P_2 = x}} \prod_{e \in E(P_1 \cup P_2)} w_e R(S - \{P_1, P_2\}), \end{aligned}$$

where $\mathcal{P}_S(x, y)$ is the set of all simple paths (with no repeated vertices) connecting x and y in S. We assume that $\mathcal{P}_S(x, x)$ consists of the trivial path $\{x\}$ only. Here $S - \{P\}$ means the graph obtained by removing P together with incident edges.

Lemma 2.5. *We have*

$$Z(S) = \sum_{x,y \in V(S)} d_x d_y \sum_{P \in \mathcal{P}_S(x,y)} \prod_{e \in E(P)} w_e R(S - \{P\}). \tag{8}$$

Lemma 2.6. *Regarding G as a subgraph of itself, we have $R(G) = 0$ and $Z(G) = \mathrm{vol}(G)^2 \tau(G)$. For any $x, y \in V(G)$, we have*

$$R(G - \{x\}) = \sum_{P \in \mathcal{P}_G(x,y)} \prod_{e \in E(P)} w_e R(G - \{P\}) = \tau(G). \tag{9}$$

Now we come to an explicit formula for the Green's function expressed in terms of the above two invariants.

Theorem 2.7. *For a connected graph G and $x, y \in V(G)$, the value of the Green's function $\mathcal{G}(x, y)$ is equal to*

$$\frac{\sqrt{d_x d_y}}{\operatorname{vol}(G)^2 \tau(G)} \Biggl(\sum_{P \in \mathcal{P}_G(x,y)} \prod_{e \in E(P)} w_e \bigl(R(G - \{P\}) + Z(G - \{P\})\bigr) - \sum_{\substack{u,v \in V(G) \\ u \neq v}} d_u d_v \sum_{\substack{P_1 \in \mathcal{P}_G(x,u) \\ P_2 \in \mathcal{P}_G(y,v) \\ P_1 \cap P_2 = \varnothing}} \prod_{e \in E(P_1 \cup P_2)} w_e R(G - \{P_1, P_2\}) \Biggr) - \frac{\sqrt{d_x d_y}}{\operatorname{vol}(G)^2}.$$

In particular, when $x = y$,

$$\mathcal{G}(y, y) = \frac{d_y}{\operatorname{vol}(G)^2 \tau(G)} \bigl(R(G - \{y\}) + Z(G - \{y\})\bigr) - \frac{d_y}{\operatorname{vol}(G)^2}.$$

Proof. The proof is almost the same as that of [Xu and Yau 2013a, Theorem 2.9]. □

Theorem 2.8. *Given a connected graph G and $x, y \in V(G)$, the expected hitting time $H(x, y)$ satisfies*

$$H(x, y) = \frac{1}{\operatorname{vol}(G) \tau(G)} \Biggl(Z(G - \{y\}) - \sum_{P \in \mathcal{P}_G(x,y)} \prod_{e \in E(P)} w_e Z(G - \{P\}) + \sum_{\substack{u,v \in V(G) \\ u \neq v}} d_u d_v \sum_{\substack{P_1 \in \mathcal{P}_G(x,u) \\ P_2 \in \mathcal{P}_G(y,v) \\ P_1 \cap P_2 = \varnothing}} \prod_{e \in E(P_1 \cup P_2)} w_e R(G - \{P_1, P_2\}) \Biggr). \tag{10}$$

Proof. This follows from Theorems 1.1 and 2.7 and Lemma 2.6. □

Corollary 2.9. *On a connected weighted graph G, $H(x, y) = H(y, x)$ for any $x, y \in V(G)$ if and only if $Z(G - \{x\})$ is independent of $x \in V(G)$.*

Proof. By (10), we have

$$H(x, y) - H(y, x) = \frac{1}{\operatorname{vol}(G) \tau(G)} \bigl(Z(G - \{y\}) - Z(G - \{x\})\bigr),$$

which implies the corollary. □

Corollary 2.10. *On a connected weighted graph G, $H(x, y) = H(y, x)$ for any $x, y \in V(G)$ if and only if* $\det B_{G-\{x\}}|_{s=1}$ *is independent of $x \in V(G)$.*

Proof. From $\det B_{G-\{x\}}|_{s=1} = R(G - \{x\}) + Z(G - \{x\})$, the conclusion follows from Lemma 2.6 and the previous corollary. □

Theorem 2.11. *Let G be a connected weighted graph and $x, y \in V(G)$. Then*

$$H(x, y) = \frac{1}{\tau(G)} \sum_{u \in V(G)} d_u \sum_{\substack{P \in \mathcal{P}_G(x,u) \\ y \notin P}} \prod_{e \in E(P)} w_e R(G - \{P, y\}). \tag{11}$$

In fact, $R(G - \{P, y\}) = \tau(G/\{P, y\})$, where $G/\{P, y\}$ is obtained from G by contracting $\{P, y\}$ to a point.

Proof. The proof is almost identical to that of [Xu and Yau 2013b, Theorem 2.7]. $R(G - \{P, y\}) = \tau(G/\{P, y\})$ follows from Theorem 2.1(ii). □

Corollary 2.12. *For any connected weighted graph G, we have*

$$Z(G - \{x\}) = \tau(G) \sum_{y \in V(G)} d_y H(y, x). \tag{12}$$

Proof. By (8) and (11), we have

$$\begin{aligned} Z(G - \{x\}, d_G) &= \sum_{u,y \in V(G-\{x\})} d_u d_y \sum_{P \in \mathcal{P}_{G-\{x\}}(u,y)} \prod_{e \in E(P)} w_e R(G - \{P, x\}) \\ &= \sum_{y,u \in V(G)} d_y d_u \sum_{\substack{P \in \mathcal{P}_G(y,u) \\ x \notin P}} \prod_{e \in E(P)} w_e R(G - \{P, x\}) \\ &= \tau(G) \sum_{y \in V(G)} d_y H(y, x). \end{aligned}$$
□

3. Identities and estimates of hitting times

It is natural to regard a weighted graph as an electrical network, where an edge xy has conductance w_{xy} and hence resistance $1/w_{xy}$. Chandra et al. [1989] proved that the commute time $\kappa(x, y) := H(x, y) + H(y, x)$ can be expressed in terms of the effective resistance R_{xy} between x, y,

$$\kappa(x, y) = \mathrm{vol}(G) R_{xy}. \tag{13}$$

The effective resistance R_{xy} can be expressed in terms of spanning trees (cf. Theorem 3.2):

$$R_{xy} = \frac{\tau(G/\{x, y\})}{\tau(G)}. \tag{14}$$

Tetali's formula [1991] expresses $H(x, y)$ in terms of effective resistances:

$$H(x, y) = \frac{1}{2} \sum_{z \in V(G)} d_z (R_{xy} + R_{yz} - R_{xz}). \tag{15}$$

We have the following well-known upper bound of R_{xy}.

Theorem 3.1. *Given a connected graph G and $x, y \in V(G)$, we have $R_{xy} \le d(x, y)$, where the distance $d(x, y)$ between x, y is defined by*

$$d(x, y) = \min\left\{ \sum_{e\in E(P)} \frac{1}{w_e} \;\middle|\; P \in \mathcal{P}_G(x, y) \right\}.$$

As remarked in [Lovász 1996, Corollary 4.2], the following formula could be proved by the method of electric networks. Here we give a proof by using (11).

Theorem 3.2. *Let $x, y \in V(G)$ be two distinct vertices of a connected weighted graph G. Then we have*

$$H(x, y) + H(y, x) = \mathrm{vol}(G)\frac{\tau(G/\{x, y\})}{\tau(G)}, \tag{16}$$

where $G/\{x, y\}$ is obtained from G by contracting $\{x, y\}$ to a point.

Proof. Define a graph G' by

$$G' = \begin{cases} G & \text{if } x \sim y, \\ G \cup \{xy\} & \text{otherwise.} \end{cases}$$

Namely, we modify G by adding an edge xy if x, y are not adjacent.

If $u \neq x, y$, define

$$\Omega_1 = \{T \in \Omega(G') \mid T \text{ contains } xy \text{ and a path from } u \text{ to } x \text{ not containing } y\},$$

$$\Omega_2 = \{T \in \Omega(G') \mid T \text{ contains } xy \text{ and a path from } u \text{ to } y \text{ not containing } x\},$$

$$\Omega_3 = \{T \in \Omega(G') \mid T \text{ contains } xy\}.$$

It is not difficult to see that $\Omega_1 \cup \Omega_2 = \Omega_3 = \Omega(G/\{x, y\})$. (More precisely, Ω_3 is in one-to-one correspondence with $\Omega(G/\{x, y\})$.) Then, by (11),

$$\begin{aligned}
&H(x, y) + H(y, x) \\
&\quad = \frac{1}{\tau(G)} \sum_{u\in V(G)} d_u \Biggl(\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}} \prod_{e\in E(P)} w_e \tau(G/\{P, y\}) \\
&\qquad\qquad\qquad + \sum_{\substack{P\in\mathcal{P}_G(y,u)\\ x\notin P}} \prod_{e\in E(P)} w_e \tau(G/\{P, x\}) \Biggr) \\
&\quad = \frac{1}{\tau(G)} \sum_{u\in V(G)} d_u \biggl(\sum_{T\in\Omega_1} \prod_{e\in T} w_e + \sum_{T\in\Omega_2} \prod_{e\in T} w_e \biggr) \\
&\quad = \frac{1}{\tau(G)} \sum_{u\in V(G)} d_u \tau(G/\{x, y\}).
\end{aligned}$$

The term in parenthesis on the third line is equal to $\tau(G/\{x, y\})$, independently of $u \in V(G)$. □

In the rest of this section, we apply Theorem 2.11 to prove some estimates for the hitting time on weighted graphs. It is interesting to see the role of edge weights in these estimates.

Corollary 3.3. *Let G be a connected weighted graph with n vertices and $x, y \in V(G)$. Then*

$$H(x, y) \leq (n-1)^2 \frac{d_{\max}}{w_{\min}}, \tag{17}$$

where $d_{\max} = \max\{d_v \mid v \in V(G)\}$ and $w_{\min} = \min\{w_e \mid e \in E(G)\}$.

Proof. Fix $x, y, u \in V(G)$ with $y \neq u$. Given a spanning tree $T \in \Omega(G)$ and an edge $e \in E(T)$, denote by $T(e)$ the subgraph of G' obtained from T by removing e and adding an edge uy if $uy \notin E(T)$, namely,

$$T(e) = \begin{cases} T & \text{if } uy \in T, \\ T \cup \{uy\} - \{e\} & \text{if } uy \notin T. \end{cases}$$

Define a subset S of $\Omega(G) \times E(G)$ by

$$S = \{(T, e) \mid T \in \Omega(G), e \in E(T), T(e) \in \Omega(G')\}$$

and let $S' = \{T \in \Omega(G') \mid T \text{ contains } uy\}$. Then the map $(T, e) \to T(e)$ is a surjective map from S to S'. Since

$$\bigcup_{\substack{P \in \mathcal{P}_G(x,u) \\ y \notin P}} \Omega(G/\{P, y\}) = \{T \in \Omega(G') \mid T \text{ contains } uy \text{ and a path from } u \text{ to } x \text{ not passing through } y\}$$

is a subset of S' and the left-hand side is a disjoint union over $P \in \mathcal{P}_G(x, u)$, $y \notin P$, we have

$$\sum_{\substack{P \in \mathcal{P}_G(x,u) \\ y \notin P}} \prod_{e \in E(P)} w_e \tau(G/\{P, y\}) \leq (n-1)\tau(G) \frac{1}{w_{\min}}.$$

Then (17) follows from (11). $\square$

Corollary 3.4. *Let G be a connected weighted graph and $xy \in E(G)$. Then*

$$H(x, y) \leq \frac{\mathrm{vol}(G) - d_y}{w_{xy}}.$$

Proof. Fix $x, y, u \in V(G)$ with $y \neq u$. It is not difficult to see that

$$\bigcup_{\substack{P \in \mathcal{P}_G(x,u) \\ y \notin P}} \Omega(G/\{P, y\}) = \{T \in \Omega(G) \mid T \text{ contains } xy \text{ and a path from } u \text{ to } x \text{ not passing through } y\}$$

is a subset of $\Omega(G)$ and the left-hand side is a disjoint union over $P \in \mathcal{P}_G(x, u)$, $y \notin P$. Thus (11) implies that

$$H(x, y) \le \frac{1}{\tau(G)} \sum_{\substack{u \in V(G) \\ u \neq y}} d_u \frac{\tau(G)}{w_{xy}} = \frac{\mathrm{vol}(G) - d_y}{w_{xy}},$$

as claimed. □

Corollary 3.5. *Let G be a connected weighted graph. Then for any two distinct vertices $x, y \in V(G)$, we have*

$$H(x, y) \le \max\left\{ \frac{d_u}{w_{yu}} \;\middle|\; u \in \mathcal{S} \right\}, \tag{18}$$

where $\mathcal{S} = \{u \in V(G) \mid \text{there is a path from } x \text{ to } u \text{ not passing through } y\}$.

Proof. Fix two distinct vertices $x, y \in V(G)$. We may assume that $w_{yu} > 0$ for all $u \in \mathcal{S}$, i.e., there is an edge connecting y and u. Otherwise the right-hand side of (18) is infinite. Define $\Omega_{xy} = \{T \in \Omega(G) \mid xy \in T\}$ and

$$V_T = \{u \in V(G) \mid T \text{ contains a path from } x \text{ to } u \text{ not passing through } y\}.$$

Let $S = \{(T, u) \mid T \in \Omega_{xy}, u \in V_T\}$. Define a map $f : S \to \Omega(G)$ by

$$f(T, u) = \begin{cases} T & \text{if } u = x, \\ \{T - xy\} \cup \{uy\} & \text{if } u \neq x, \end{cases}$$

where we used the fact that $d_y = n - 1$. It is not difficult to see that f is injective. Thus, we have

$$\begin{aligned} \sum_{u \in V(G)} d_u \sum_{\substack{P \in \mathcal{P}_G(x,u) \\ y \notin P}} \prod_{e \in E(P)} w_e \tau(G/\{P, y\}) &= \sum_{u \in V(G)} \frac{d_u}{w_{yu}} \sum_{(T,u) \in S} w(f(T, u)) \\ &\le \max\left\{ \frac{d_u}{w_{yu}} \;\middle|\; u \in \mathcal{S} \right\} \tau(G). \end{aligned}$$

Therefore (11) implies (19). □

There is a direct probabilistic proof of Corollary 3.5 (see [Xu and Yau 2013b, Remark 2.14]). In fact, all the above three corollaries may be obtained from the following more refined estimates (see also [Chang and Xu ≥ 2015]):

Theorem 3.6. *Let G be a connected weighted graph. Then*

$$H(x, y) \le \max\{d_u \mid u \in \Gamma(y), u \neq y\} + \sum_{\substack{u \in V(G), u \neq y \\ u \notin \Gamma(y)}} d_u \min\{d(x, y), d(u, y)\}, \tag{19}$$

where $\Gamma(y)$ is the set of vertices adjacent to y and $d(x, y)$ is defined in Theorem 3.1.

Proof. We split the right-hand side of (11) into two terms:

$$H(x,y)=\frac{1}{\tau(G)}\sum_{u\in\Gamma(y)}d_u\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}}\prod_{e\in E(P)}w_e\tau(G/\{P,y\}) + \frac{1}{\tau(G)}\sum_{\substack{u\in V(G)\\ u\notin\Gamma(y)}}d_u\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}}\prod_{e\in E(P)}w_e\tau(G/\{P,y\}). \tag{20}$$

If $u\in\Gamma(y)$, then define

$\Omega_{uy}=\{T\in\Omega(G)\mid T$ contains uy and a path from x to u not passing through $y\}$.

Since $\bigcup_{u\in\Gamma(y)}\Omega_{uy}$ is a disjoint union in $\Omega(G)$ and

$$\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}}\prod_{e\in E(P)}w_e\tau(G/\{P,y\})=\sum_{T\in\Omega_{uy}}w(T)\quad\text{for all } u\in\Gamma(G),$$

the first summand in the right-hand side of (20) satisfies

$$\frac{1}{\tau(G)}\sum_{u\in\Gamma(y)}d_u\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}}\prod_{e\in E(P)}w_e\tau(G/\{P,y\})\le\max\{d_u\mid u\in\Gamma(y),u\neq y\}. \tag{21}$$

By using (14) and the inequality

$$\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}}\prod_{e\in E(P)}w_e\tau(G/\{P,y\})\le\min\{\tau(G/\{x,y\}),\tau(G/\{u,y\})\}, \tag{22}$$

the second summand in the right-hand of (20) satisfies

$$\frac{1}{\tau(G)}\sum_{\substack{u\in V(G)\\ u\notin\Gamma(y)}}d_u\sum_{\substack{P\in\mathcal{P}_G(x,u)\\ y\notin P}}\prod_{e\in E(P)}w_e\tau(G/\{P,y\}) \le\sum_{\substack{u\in V(G)\\ u\notin\Gamma(y)}}d_u\min\{R_{xy},R_{uy}\}\le\sum_{\substack{u\in V(G)\\ u\notin\Gamma(y)}}d_u\min\{d(x,y),d(u,y)\}. \tag{23}$$

The last inequality follows from Theorem 3.1. So (19) follows from (21) and (23). □

Corollary 3.7. *Let G be a connected weighted graph. Then*

$$H(x,y)\le\sum_{\substack{u\in V(G)\\ u\neq y}}d_u\min\{R_{xy},R_{uy}\}\le\sum_{\substack{u\in V(G)\\ u\neq y}}d_u\min\{d(x,y),d(u,y)\}. \tag{24}$$

Proof. The first inequality follows from (14), (22) and Theorem 2.11. The second inequality follows from Theorem 3.1. □

4. Some examples

First we consider infinite but locally finite connected graphs. Since an infinite (locally finite) graph can be considered as a limit of a sequence of finite graphs, the hitting time formula (11) is still valid as long as the limit exists. A weighted tree T is a locally finite tree (possibly with loops) whose edges are assigned positive weights.

Theorem 4.1. *Let x, y be two distinct vertices of a weighted tree T, and denote by P_{xy} the path $[x = v_0, v_1, \dots, v_{k-1}, v_k = y]$ connecting x to y. For any $v_i \in V(P_{xy})$, we denote by T_i the component of $T - E(P_{xy})$ that contains v_i, and denote by $w_{i-1,i}$ the weight of the edge $v_{i-1}v_i$. Then the hitting time $H(x, y)$ is given by*

$$H(x, y) = \sum_{j=0}^{k-1}\Big(\sum_{u\in T_j} d_u\Big)\Big(\sum_{i=j+1}^{k} \frac{1}{w_{i-1,i}}\Big). \tag{25}$$

Proof. First we define induced subtrees $T(N)$ of T for $N \in \mathbb{N}$. The vertices of $T(N)$ are those vertices whose distances from x are within N. We may apply (11) to get $H(x, y)$ on $T(N)$, which increases as N increases. Then (25) follows easily. We omit the details. □

Corollary 4.2. *Let x, y be two distinct vertices of a weighted tree T. Then $H(x, y) < \infty$ if and only if*

$$\sum_{u\in\mathcal{S}} d_u < \infty,$$

where $\mathcal{S} = \{u \in V(T) \mid \text{there is a path from } x \text{ to } u \text{ not passing through } y\}$.

Corollary 4.3. *On the weighted one-dimensional lattice $\mathbb{Z}$,*

$$H(j, j+1) = \frac{\sum_{i\le j} d_i}{w_{j,j+1}}. \tag{26}$$

Both corollaries follow easily from (25). For unweighted trees, formula (25) was obtained in [Haiyan and Fuji 2004] (see also [Moon 1973]). Formula (26) can be found in [Palacios and Tetali 1996], where it was used to study hitting times for birth and death chains.

Now let G be a locally finite connected weighted graph and $xy \in E(G)$. Then the inequality of Corollary 3.4 still holds: $H(x, y) \le (\mathrm{vol}(G) - d_y)/w_{xy}$. Next we show that the equality essentially holds when xy is a cut edge of G.

Let $\mathcal{S} = \{u \in V(G) \mid \text{there is a path from } x \text{ to } u \text{ not passing through } y\}$. Let G' be the subgraph obtained by removing all vertices in $V(G)/\{\mathcal{S} \cup y\}$ from G. If $xy \in E(G)$ is a cut edge of G, note that $H(x, y)$ is the same for random walks on either G or G'. Moreover, for each spanning tree T of G' and $u \in \mathcal{S}$, there exists a path from x to u. Therefore,

$$H(x, y) = \frac{1}{\tau(G')} \sum_{u \in \mathcal{S}} d_u \frac{1}{w_{xy}} \tau(G') = \frac{1}{w_{xy}} \sum_{u \in \mathcal{S}} d_u = \frac{\mathrm{vol}(G') - d'_y}{w_{xy}},$$

where d'_y is the degree of y in G'.

Following [Georgakopoulos 2012], we call a weighted graph G *reversible* if $H(x, y) = H(y, x)$ holds for any $x, y \in V(G)$. For simplicity, we assume that G has no loops, i.e., $w_{xx} = 0$ for all $x \in V(G)$, and all edge weights of G are positive. It is interesting to study restrictions on edge weights for a reversible graph G.

Conjecture 4.4. *Let G be a weighted cycle on n vertices. Assume all edge weights of G are positive. Denote $w_{n,n+1} = w_{n,1}$.*

(i) *If n is odd, then G is reversible if and only if there exists some $a > 0$ such that $w_{i,i+1} = a$ for all $1 \leq i \leq n$.*

(ii) *If n is even, then G is reversible if and only if there exist $a, b > 0$ such that $w_{1,2} = w_{3,4} = \cdots = w_{n-1,n} = a$ and $w_{2,3} = w_{4,5} = \cdots = w_{n,1} = b$.*

The sufficiency in (ii) follows from Corollary 2.10.

5. Weight schemes on graphs

Given a simple, connected, undirected graph G with n vertices, we obtain a weighted graph by assigning a positive number w_e to each edge $e \in E(G)$. The hitting and cover times of a simple random walk on G (i.e., $w_e = 1$, for all $e \in E(G)$) have order $O(n^3)$. The work of [Ikeda et al. 2009; Abdullah 2011] showed that if a token knows not only the degree of the current vertex that it is on, but also the degrees of neighboring vertices, we can guarantee $O(n^2)$ hitting times.

In this section, we will denote by $d(u)$ the number of edges adjacent to a vertex u in G and assume that G has no loops.

Lemma 5.1. *Let G be connected graph with n vertices and $u_0 = x, u_1, \ldots, u_l = y$ a shortest path (achieving minimum l) connecting any two distinct vertices x and y. Then $\sum_{i=0}^{l} d(u_i) \leq 3n - 4$. More precisely,*

$$\sum_{i=0}^{l} d(u_i) \leq \begin{cases} 2n - 2 & \text{if } l = 1, \\ 3n - l - 3 & \text{if } l \geq 2. \end{cases}$$

Proof. The proof is due to [Ikeda et al. 2009, Theorem 2]. Each vertex of $V(G)$ not lying on the path can be connected to at most 3 vertices of the path, due to its minimality, which also implies that u_i, u_j are adjacent if and only if $|i - j| = 1$. The asserted inequalities follow easily. □

Next we will apply Theorem 3.6 to estimate hitting times under three different weight schemes: $w_{uv} = 1/\sqrt{d(u)d(v)}$, $1/\min\{d(u), d(v)\}$ or $1/\max\{d(u), d(v)\}$. The leading terms of the bounds in Theorems 5.2 and 5.3 were obtained in [Ikeda et al. 2009, Theorem 2] and [Abdullah 2011, Theorem 68] respectively.

Theorem 5.2. *Let G be a graph with assigned weights $w_{uv} = 1/\sqrt{d(u)d(v)}$ for each edge uv. Then the hitting time satisfies $H(x, y) \le 3n^2 - 9n + \frac{15}{2}$.*

Proof. For the two terms in the right-hand side of (19), we have the estimates

$$d_u = \sum_{v\in\Gamma(u)} \frac{1}{\sqrt{d(u)d(v)}} \le \frac{1}{2} \sum_{v\in\Gamma(u)} \left(\frac{1}{d(u)} + \frac{1}{d(v)}\right) \le \frac{1}{2} + \frac{d(u)}{2} \le \frac{n}{2}. \tag{27}$$

and, by using $\sum_{u\in V(G)} \sum_{v\in\Gamma(u)} \frac{1}{2}(1/d(u) + 1/d(v)) = n$,

$$\sum_{\substack{u\in V(G),u\neq y\\ u\notin\Gamma(y)}} d_u \le \sum_{\substack{u\in V(G),u\neq y\\ u\notin\Gamma(y)}} \sum_{v\in\Gamma(u)} \frac{1}{2}\left(\frac{1}{d(u)}+\frac{1}{d(v)}\right) \le n-1-\frac{d(y)}{2} \le n-\frac{3}{2}. \tag{28}$$

Let $u_0 = x, u_1, \ldots, u_l = y$ be a shortest path (achieving minimum l) connecting x and y. Then

$$d(x, y) \le \sum_{i=0}^{l-1} \sqrt{d(u_i)d(u_{i+1})} \le \sum_{i=0}^{l-1} \frac{d(u_i) + d(u_{i+1})}{2} \le 3n - 5. \tag{29}$$

The last inequality follows from Lemma 5.1. By (19), we have

$$H(x, y) \le (3n - 5)(n - \tfrac{3}{2}) + \tfrac{1}{2}n = 3n^2 - 9n + \tfrac{15}{2},$$

as claimed. □

Theorem 5.3. *Let G be a graph with assigned weights $w_{uv} = 1/\min\{d(u), d(v)\}$ for each edge uv. Then the hitting time satisfies $H(x, y) \le 6n^2 - 18n + 14$.*

Proof. For the two terms in the right-hand side of (19), we have

$$d_u = \sum_{v\in\Gamma(u)} \frac{1}{\min\{d(u), d(v)\}} \le d(u) \le n - 1 \tag{30}$$

and, similarly to (28),

$$\sum_{\substack{u\in V(G),u\neq y\\ u\notin\Gamma(y)}} d_u \le \sum_{\substack{u\in V(G),u\neq y\\ u\notin\Gamma(y)}} \sum_{v\in\Gamma(u)} \left(\frac{1}{d(u)} + \frac{1}{d(v)}\right) \le 2n - 3. \tag{31}$$

Similarly to (29), we have

$$d(x, y) \le \sum_{i=0}^{l-1} \min\{d(u_i), d(u_{i+1})\} \le \sum_{i=0}^{l-1} d(u_i) \le 3n - 5.$$

The desired upper bound of $H(x, y)$ follows from (19). $\square$

Theorem 5.4. *Let G be a graph with assigned weights $w_{uv} = 1/\max\{d(u), d(v)\}$ for each edge uv. Then the hitting time satisfies $H(x, y) \le 6n^2 - 23n + 23$.*

Proof. For the two terms in the right-hand side of (19), we have

$$d_u = \sum_{v\in\Gamma(u)} \frac{1}{\max\{d(u), d(v)\}} \le 1 \quad \text{and} \quad \sum_{\substack{u\in V(G), u\neq y\\ u\notin\Gamma(y)}} d_u \le n - 2.$$

Similar to (29), we have

$$d(x, y) \le \sum_{i=0}^{l-1} \max\{d(u_i), d(u_{i+1})\} \le \sum_{i=0}^{l-1} (d(u_i) + d(u_{i+1}) - 1) \le 6n - 11.$$

The desired upper bound of $H(x, y)$ follows from (19). $\square$

References

[Abdullah 2011] M. Abdullah, *The cover time of random walks on graphs*, Ph.D. thesis, King's College London, 2011. arXiv 1202.5569

[Aldous and Fill 2014] D. Aldous and J. Fill, "Reversible Markov chains and random walks on graphs", Unfinished monograph, 2014, Available at http://www.stat.berkeley.edu/~aldous/RWG/book.pdf.

[Chandra et al. 1989] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari, "The electrical resistance of a graph captures its commute and cover times", pp. 574–586 in *Proc. 21st ACM Symp. Theory of Computing* (Seattle, 1989), edited by D. S. Johnson, Association for Computing Machinery, New York, 1989.

[Chang and Xu ≥ 2015] X. Chang and H. Xu, "Chung–Yau invariants and graphs with symmetric hitting times", preprint, Available at http://pitt.edu/~haoxu/ReversibleGraph.pdf.

[Chung 2011] F. R. K. Chung, "PageRank as a discrete Green's function", pp. 285–302 in *Geometry and analysis, I* (Cambridge, MA, 2008), edited by L. Ji, Adv. Lect. Math. (ALM) **17**, Int. Press, Somerville, MA, 2011. MR 2012m:68344 Zbl 1255.68109

[Chung and Yau 2000] F. Chung and S.-T. Yau, "Discrete Green's functions", *J. Combin. Theory Ser. A* **91**:1-2 (2000), 191–214. MR 2001g:05073 Zbl 0963.65120

[Georgakopoulos 2012] A. Georgakopoulos, "On walk-regular graphs and graphs with symmetric hitting times", preprint, 2012. arXiv 1211.5689

[Haiyan and Fuji 2004] C. Haiyan and Z. Fuji, "The expected hitting times for graphs with cutpoints", *Statist. Probab. Lett.* **66**:1 (2004), 9–17. MR 2005a:60065 Zbl 1113.60046

[Ikeda et al. 2009] S. Ikeda, I. Kubo, and M. Yamashita, "The hitting and cover times of random walks on finite graphs using local degree information", *Theoret. Comput. Sci.* **410**:1 (2009), 94–100. MR 2010c:60143 Zbl 05509592

[Lovász 1996] L. Lovász, "Random walks on graphs: A survey", pp. 353–397 in *Combinatorics, Paul Erdős is eighty, II* (Keszthely, 1993), edited by D. Miklós et al., Bolyai Soc. Math. Stud. **2**, János Bolyai Math. Soc., Budapest, 1996. MR 97a:60097 Zbl 0854.60071

[Moon 1973] J. W. Moon, "Random walks on random trees", *J. Austral. Math. Soc.* **15** (1973), 42–53. MR 47 #6525 Zbl 0265.60065

[Palacios and Tetali 1996] J. L. Palacios and P. Tetali, "A note on expected hitting times for birth and death chains", *Statist. Probab. Lett.* **30**:2 (1996), 119–125. MR 97h:60069 Zbl 0883.60081

[Tetali 1991] P. Tetali, "Random walks and the effective resistance of networks", *J. Theoret. Probab.* **4**:1 (1991), 101–109. MR 92c:60097 Zbl 0722.60070

[Xu and Yau 2013a] H. Xu and S.-T. Yau, "Discrete Green's functions and random walks on graphs", *J. Combin. Theory Ser. A* **120**:2 (2013), 483–499. MR 2995053 Zbl 1256.05225

[Xu and Yau 2013b] H. Xu and S. T. Yau, "An explicit formula of hitting times for random walks on graphs", preprint, 2013. To appear in *Pure Appl. Math. Q.* arXiv 1312.0065

Received February 24, 2014.

XIAO CHANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PITTSBURGH
301 THACKERAY HALL
PITTSBURGH, PA 15260
UNITED STATES
xic58@pitt.edu

HAO XU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PITTSBURGH
301 THACKERAY HALL
PITTSBURGH, PA 15260
UNITED STATES
mathxuhao@gmail.com

SHING-TUNG YAU
DEPARTMENT OF MATHEMATICS
HARVARD UNIVERSITY
CAMBRIDGE, MA 02138
UNITED STATES
yau@math.harvard.edu

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 273 No. 1 January 2015

Maximal estimates for Schrödinger equations with inverse-square potential 1
CHANGXING MIAO, JUNYONG ZHANG and JIQIANG ZHENG

Vassiliev Invariants of Virtual Legendrian Knots 21
PATRICIA CAHN and ASA LEVI

Some results on the generic vanishing of Koszul cohomology via deformation theory 47
JIE WANG

Conformal metrics with constant curvature one and finitely many conical singularities on compact Riemann surfaces 75
QING CHEN, WEI WANG, YINGYI WU and BIN XU

$\mathbb{Q}$-bases of the Néron–Severi groups of certain elliptic surfaces 101
MASAMICHI KURODA

On a prime zeta function of a graph 123
TAKEHIRO HASEGAWA and SEIKEN SAITO

On Whittaker modules for a Lie algebra arising from the 2-dimensional torus 147
SHAOBIN TAN, QING WANG and CHENGKANG XU

Fréchet quantum supergroups 169
AXEL DE GOURSAC

Generators of the Gauss–Picard modular group in three complex dimensions 197
BAOHUA XIE, JIEYAN WANG and YUEPING JIANG

Complete characterization of isolated homogeneous hypersurface singularities 213
STEPHEN YAU and HUAIQING ZUO

A theorem of Mœglin and Waldspurger for covering groups 225
SHIV PRAKASH PATEL

Spanning trees and random walks on weighted graphs 241
XIAO CHANG, HAO XU and SHING-TUNG YAU

0030-8730(201501)273:1;1-6