Pacific Journal of Mathematics

Volume 255 No. 2 February 2012

PACIFIC JOURNAL OF MATHEMATICS

http://pacificmath.org

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

See inside back cover or pacificmath.org for submission instructions.

The subscription price for 2012 is US $420/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. 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 PACIFIC JOURNAL OF MATHEMATICS
at the University of California, Berkeley 94720-3840
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

ON THE LOCAL LANGLANDS CORRESPONDENCES OF DEBACKER–REEDER AND REEDER FOR GL(ℓ, F), WHERE ℓ IS PRIME

MOSHE ADRIAN

We prove that the conjectural depth-zero local Langlands correspondence of DeBacker and Reeder agrees with the known depth-zero local Langlands correspondence for the group GL(ℓ, F), where ℓ is prime and F is a nonarchimedean local field of characteristic 0. We also prove that if one assumes a certain compatibility condition between Adler's and Howe's constructions of supercuspidal representations, then the conjectural positive-depth local Langlands correspondence of Reeder also agrees with the known positive-depth local Langlands correspondence for GL(ℓ, F).

1. Introduction

Let F be a nonarchimedean local field of characteristic zero. Let $\mathbf{G}$ be a connected reductive group defined over F. The local Langlands correspondence asserts that there is a finite to one map from the set of admissible representations of $\mathbf{G}(F)$ to the set of Langlands parameters of $\mathbf{G}(F)$, satisfying various conditions. Until recently, this has only been proven for special cases of groups such as GL(n, F), Sp(4, F), and $U(3)$. The local Langlands correspondence for GL(n, F) was proven by Harris and Taylor, and independently by Henniart.

More recently, DeBacker and Reeder, in two papers that will be cited throughout the text, described conjectural local Langlands correspondences for a more general class of groups and certain classes of Langlands parameters. These correspondences are still conjectural, despite satisfying several requirements that the Langlands correspondence should have. One would therefore like to know whether they agree at least with the proven correspondences in the known cases.

We prove that the correspondence introduced in [DeBacker and Reeder 2009] (henceforth [DB-R]) agrees with the known correspondence for GL(ℓ, F), while the one in [Reeder 2008] (henceforth [R]) agrees with the known correspondence

MSC2010: 11S37, 22E50.
Keywords: local Langlands correspondence, p-adic groups, Langlands program, representation theory, number theory.

for $\mathrm{GL}(\ell, F)$ if one assumes a certain compatibility condition, which we describe later.

For $\mathrm{GL}(n, F)$, the constructions of Harris–Taylor, Henniart, and DeBacker–Reeder (and Reeder) use different methods. We first recall the classical construction of the tame local Langlands correspondence for $\mathrm{GL}(\ell, F)$ as in [Moy 1986]. We note that a tame local Langlands correspondence for $\mathrm{GL}(n, F)$ was conjectured there for general n. In view of [Bushnell and Henniart 2005], Moy's correspondence is indeed correct for $\mathrm{GL}(\ell, F)$, ℓ a prime.

Definition 1.1. Let E/F be an extension of degree ℓ, ℓ relatively prime to the residual characteristic of F, and let χ be a character of E^*. The pair $(E/F, \chi)$ is called *admissible* if χ does not factor through the norm from a proper subfield of E containing F.

We write $\mathbb{P}_\ell(F)$ for the set of F-isomorphism classes of admissible pairs $(E/F, \chi)$ where E/F is a degree-ℓ extension (for more information about admissible pairs, see [Moy 1986]). Let $\mathbb{A}_\ell^0(F)$ denote the set of supercuspidal representations of $\mathrm{GL}(\ell, F)$. Howe [1977] constructs a map

$$\mathbb{P}_\ell(F) \to \mathbb{A}_\ell^0(F), \quad (E/F, \chi) \mapsto \pi_\chi.$$

This map is a bijection [Moy 1986]. Let $\mathbb{G}_\ell^0(F)$ denote the set of irreducible ℓ-dimensional representations of W_F, where W_F is the Weil group of F. We then have a bijection [Moy 1986]

$$\mathbb{P}_\ell(F) \to \mathbb{G}_\ell^0(F), \quad (E/F, \chi) \mapsto \mathrm{Ind}_{W_E}^{W_F}(\chi) =: \phi(\chi).$$

The local Langlands correspondence is then given by

$$\phi(\chi) \mapsto \pi_{\chi \Delta_\chi}$$

for some subtle finite order character Δ_χ of E^* [Bushnell and Henniart 2005]. In the case of depth-zero supercuspidal representations, there is only one extension E/F to deal with, namely, the unramified extension of F of degree ℓ.

On the other hand, the constructions of [DB-R] and [R] extensively use Bruhat–Tits theory. To a certain class of Langlands parameters for an unramified connected reductive group $\mathbf{G}$, they associate a character of a torus, to which they attach a collection of supercuspidal representations on the pure inner forms of $\mathbf{G}(F)$, a conjectural L-packet. They are also able to isolate the part of this packet corresponding to a particular pure inner form, and prove that their correspondences satisfy various natural conditions, such as stability.

Specifically, we prove the following. Let E/F be the unramified degree-ℓ extension, ℓ a prime. To any tame, regular, semisimple, elliptic, Langlands parameter (TRSELP) for $\mathrm{GL}(\ell, F)$, we show that DeBacker–Reeder theory attaches the character $\chi\Delta_\chi$ of E^*, to which is attached the representation $\pi_{\chi\Delta_\chi}$. This will prove that their correspondence agrees with the correspondence of [Moy 1986] for GL(ℓ, F).

We then prove the same for Reeder's construction, if one assumes a certain compatibility condition, which we describe now. The construction in [R] begins by canonically attaching a certain admissible pair $(L/F, \Omega)$ to a Langlands parameter for GL(ℓ, F). His construction then inputs this admissible pair into the theory of [Adler 1998] in order to construct a supercuspidal representation $\pi(L, \Omega)$ of GL(ℓ, F). The compatibility condition that we will need to assume is that $\pi(L, \Omega)$ is the same supercuspidal representation that is attached to $(L/F, \Omega)$ via the construction in [Howe 1977]. We remark that this compatibility condition does not seem to be known to the experts.

Although Moy's correspondence agrees with DeBacker and Reeder's (and also with Reeder's, assuming the above compatibility), some important details are different. One interesting and subtle difference lies in the passage from a Langlands parameter to a character of a torus. To illustrate it, we rewrite both correspondences to include their factorization through characters of elliptic tori as $\{$Langlands parameters from [DB-R] or [R] for GL(ℓ, F)$\} \to \mathbb{P}_\ell(F) \to \mathbb{A}^0_\ell(F)$.

Then, the correspondence of Moy is given by

$$\phi(\chi) = \mathrm{Ind}_{W_E}^{W_F}(\chi) \mapsto (E/F, \chi) \mapsto \pi_{\chi\Delta_\chi},$$

whereas the correspondences of DeBacker–Reeder (and Reeder, assuming the compatibility) are given by

$$\phi(\chi) = \mathrm{Ind}_{W_E}^{W_F}(\chi) \mapsto (E/F, \chi\Delta_\chi) \mapsto \pi_{\chi\Delta_\chi}.$$

We now briefly present an outline of the paper. In Section 2, we introduce some notation that we will need throughout. In Section 3, we briefly recall some of the key components to the construction from [DB-R]. In Section 4, we recall the tame local Langlands correspondence for GL(ℓ, F) as explained in [Moy 1986]. In Sections 5 and 6, we work out the DeBacker–Reeder theory for GL(ℓ, F), and we show that the correspondences of DeBacker–Reeder and Moy agree for GL(ℓ, F). Finally, in Section 7, we work out the theory of [R] for GL(ℓ, F), where ℓ is prime, and we show that under the compatibility condition, the correspondences of Reeder and Moy agree for GL(ℓ, F).

2. Notation

Let F denote a nonarchimedean local field of characteristic zero. We let $\mathfrak{o}_F$ denote the ring of integers of F, $\mathfrak{p}_F$ its maximal ideal, $\mathfrak{f}$ the residue field of F, q the order of $\mathfrak{f}$, and p the characteristic of $\mathfrak{f}$. Let $\mathfrak{f}_{\mathfrak{m}}$ denote the degree-m extension of $\mathfrak{f}$. We let ϖ denote a uniformizer of F. Let F^u denote the maximal unramified extension of F. We have the canonical projection

$$\Pi : \mathfrak{o}_F^* \to \mathfrak{o}_F^*/(1+\mathfrak{p}_F) \cong \mathfrak{f}^*$$

We denote by W_F the Weil group of F, I_F the inertia subgroup of W_F, I_F^+ the wild inertia subgroup of W_F, and W_F^{ab} the abelianization of W_F. We denote by W_F' the Weil–Deligne group, we set $W_t := W_F/I_F^+$, and we set $I_t := I_F/I_F^+$. We fix an element $\Phi \in \mathrm{Gal}(\overline{F}/F)$ whose inverse induces the map $x \mapsto x^q$ on $\mathfrak{F} := \bar{\mathfrak{f}}$, and if E/F is the unramified extension of degree ℓ, we fix an element $\Phi_E \in \mathrm{Gal}(\overline{E}/E)$ whose inverse induces the map $x \mapsto x^{q^\ell}$ on $\mathfrak{F} := \bar{\mathfrak{f}}$. Let $\mathbf{G}$ be an unramified connected reductive group over F, and set $G = \mathbf{G}(F^u)$. We fix $\mathbf{T} \subset \mathbf{G}$, an F^u-split maximal torus which is defined over F and maximally F-split, and set $T = \mathbf{T}(F^u)$. We write $X := X_*(\mathbf{T})$, W_o for the finite Weyl group $N_G(T)/T$, and set $N := N_G(T)$. Recall that the extended affine Weyl group is defined by $W := X \rtimes W_o$, and that the affine Weyl group is defined by $W^o := \Psi \rtimes W_o$, where Ψ is the coroot lattice in X. We let $\mathscr{A} := \mathscr{A}(T)$ be the apartment of T. We denote by θ the automorphism of X, W induced by Φ. If E/F is a finite Galois extension, then we denote by $\aleph_{E/F}$ the local class field theory character of F^* with respect to the extension E/F. If $\chi \in \widehat{E^*}$ satisfies $\chi|_{1+\mathfrak{p}_E} \equiv 1$, then $\chi|_{\mathfrak{o}_E^*}$ factors to a character, denoted χ_o, of the multiplicative group of the residue field of E, given by $\chi_o(x) := \chi(u)$ for any $u \in \mathfrak{o}_E^*$ such that $\Pi(u) = x$. If E/F is the degree-ℓ unramified extension, where ℓ is prime, we once and for all fix a generator ξ of $\mathrm{Gal}(E/F)$. We also fix a generator of $\mathrm{Gal}(\mathfrak{f}_\ell/\mathfrak{f})$, which, abusing notation, we also denote by ξ. If χ is a character of E^* or $\mathfrak{f}_\ell^*$, we let χ^ξ denote the character given by $\chi^\xi(x) := \chi(\xi(x))$. If L/K is a Galois quadratic extension, we let the map $x \mapsto \overline{x}$ denote the nontrivial Galois automorphism of L/K. If A is a group and B is a normal subgroup of A, we denote the image of $a \in A$ in A/B by $[a]$. If $\phi : C \to D$ is a group homomorphism and ϕ is trivial on a normal subgroup $M \lhd C$, then we will abuse notation and write $\phi|_{C/M}$ for the factorization of ϕ to a map $C/M \to D$. For example, the Langlands parameters in [DB-R] are trivial on the wild inertia subgroup I_F^+ of the inertia group I_F. Therefore, if ϕ is such a Langlands parameter and $I_t := I_F/I_F^+$, we will write $\phi|_{I_t}$ to denote the factorization of $\phi|_{I_F}$ to the quotient I_t.

3. Review of construction of DeBacker and Reeder

We first review some of the basic theory from [DB-R]. We first fix a pinning $(\hat{T}, \hat{B}, \{x_\alpha\})$ for the dual group $\hat{G}$. The operator $\hat{\theta}$ dual to θ extends to an automorphism of $\hat{T}$. There is a unique extension of $\hat{\theta}$ to an automorphism of $\hat{G}$, satisfying $\hat{\theta}(x_\alpha) = x_{\theta\cdot\alpha}$ (see [DB-R, Section 3.2]). Following [DB-R], we may form the semidirect product ${}^LG := \langle\hat{\theta}\rangle \ltimes \hat{G}$.

Definition 3.1. Let W'_F denote the Weil–Deligne group. A Langlands parameter $\phi : W'_F \to {}^L G$ is called a *tame regular semisimple elliptic Langlands parameter* (abbreviated TRSELP) if

(1) ϕ is trivial on I_F^+;

(2) the centralizer of $\phi(I_F)$ in $\hat{G}$ is a torus;

(3) $C_{\hat{G}}(\phi)^o = (\hat{Z}^{\hat{\theta}})^o$, where $\hat{Z}$ denotes the center of $\hat{G}$.

Condition (2) forces ϕ to be trivial on SL(2, $\mathbb{C}$). Let $\hat{N} = N_{\hat{G}}(\hat{T})$. After conjugating by $\hat{G}$, we may assume that $\phi(I_F) \subset \hat{T}$ and $\phi(\Phi) = \hat{\theta} f$, where $f \in \hat{N}$. Let $\hat{w}$ be the image of f in $\hat{W}_o$, and let w be the element of W_o corresponding to $\hat{w}$.

Let ϕ be a TRSELP with associated w and set $\sigma = w\theta$. σ is an automorphism of X. Let $\hat{\sigma}$ be the automorphism dual to σ, and let n be the order of σ. We set $\hat{G}_{ab} := \hat{G}/\hat{G}'$, where $\hat{G}'$ denotes the derived group of $\hat{G}$. Let

$$ {}^L T_\sigma := \langle \hat{\sigma} \rangle \ltimes \hat{T}. $$

Associated to ϕ, DeBacker and Reeder [DB-R, Chapter 4] define a $\hat{T}$-conjugacy class of Langlands parameters

$$ \phi_T : W_F \to {}^L T_\sigma \tag{1} $$

as follows. Set $\phi_T := \phi$ on I_F, and $\phi_T(\Phi) := \hat{\sigma} \ltimes \tau$ where $\tau \in \hat{T}$ is any element whose class in $\hat{T}/(1-\hat{\sigma})\hat{T}$ corresponds to the image of f in $\hat{G}_{ab}/(1-\hat{\theta})\hat{G}_{ab}$ under the bijection

$$ \hat{T}/(1-\hat{\sigma})\hat{T} \xrightarrow{\sim} \hat{G}_{ab}/(1-\hat{\theta})\hat{G}_{ab} \tag{2} $$

Chapter 4 of [DB-R] gives a canonical bijection between $\hat{T}$-conjugacy classes of admissible homomorphisms $\phi : W_t \to {}^L T_\sigma$ and depth-zero characters of T^{Φ_σ} where $\Phi_\sigma := \sigma \otimes \Phi^{-1}$. We briefly summarize this construction. Let $\mathbb{T} := X \otimes \mathfrak{F}^*$. Given automorphisms α, β of abelian groups A, B, respectively, let $\mathrm{Hom}_{\alpha,\beta}(A, B)$ denote the set of homomorphisms $f : A \to B$ such that $f \circ \alpha = \beta \circ f$. The twisted norm map

$$ N_\sigma : \mathbb{T}^{\Phi_\sigma^n} \to \mathbb{T}^{\Phi_\sigma} $$

given by $N_\sigma(t) = t\,\Phi_\sigma(t)\,\Phi_\sigma^2(t)\ldots\Phi_\sigma^{n-1}(t)$ induces isomorphisms

$$ \mathrm{Hom}(\mathbb{T}^{\Phi_\sigma}, \mathbb{C}^*) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(\mathbb{T}^{\Phi_\sigma^n}, \mathbb{C}^*) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(X \otimes \mathfrak{f}_n^*, \mathbb{C}^*) $$

Moreover, the map $s \mapsto \chi_s$ gives an isomorphism

$$ \mathrm{Hom}_{\Phi,\hat{\sigma}}(\mathfrak{f}_n^*, \hat{T}) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(X \otimes \mathfrak{f}_n^*, \mathbb{C}^*), $$

where $\chi_s(\lambda \otimes a) := \lambda(s(a))$. The canonical projection $I_t \to \mathfrak{f}_m^*$ induces an isomorphism as Φ-modules

$$I_t/(1-\mathrm{Ad}(\Phi)^m)\,I_t \xrightarrow{\sim} \mathfrak{f}_m^*,$$

where Ad denotes the adjoint action. Since $\hat{\sigma}$ has order n, we have

$$\mathrm{Hom}_{\Phi,\hat{\sigma}}(\mathfrak{f}_n^*, \hat{T}) \cong \mathrm{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}(I_t, \hat{T}).$$

Therefore, the map $s \mapsto \chi_s$ is a canonical bijection

$$\mathrm{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}(I_t, \hat{T}) \xrightarrow{\sim} \mathrm{Hom}(\mathbb{T}^{\Phi_\sigma}, \mathbb{C}^*)$$

Moreover, we have an isomorphism

$${}^0T^{\Phi_\sigma} \times X^\sigma \xrightarrow{\sim} T^{\Phi_\sigma}, \quad (\gamma, \lambda) \mapsto \gamma\lambda(\varpi),$$

where ${}^0T$ is the group of $\mathfrak{o}_{F^u}$-points of $\mathbf{T}$.

Finally, note that $\hat{T}/(1-\hat{\sigma})\hat{T}$ is the character group of X^σ, whereby

$$\tau \in \hat{T}/(1-\hat{\sigma})\hat{T}$$

corresponds to $\chi_\tau \in \mathrm{Hom}(X^\sigma, \mathbb{C}^*)$, where $\chi_\tau(\lambda) := \lambda(\tau)$. Therefore, we have a canonical bijection between $\hat{T}$-conjugacy classes of admissible homomorphisms $\phi : W_t \to {}^LT_\sigma$ and depth-zero characters

$$\chi_\phi := \chi_s \otimes \chi_\tau \in \mathrm{Irr}(T^{\Phi_\sigma}), \tag{3}$$

where $s := \phi|_{I_t}$, $\phi(\Phi) = \hat{\sigma} \ltimes \tau$, and where we have inflated χ_s to ${}^0T^{\Phi_\sigma}$.

To get the depth-zero L-packet associated to ϕ, one implements the component group

$$\mathrm{Irr}(C_\phi) \cong [X/(1-w\theta)X]_{\mathrm{tor}}$$

as follows. We set X_w to be the preimage of $[X/(1-w\theta)X]_{\mathrm{tor}}$ in X. To $\lambda \in X_w$, DeBacker and Reeder associate a 1-cocycle u_λ, hence a twisted Frobenius $\Phi_\lambda = \mathrm{Ad}(u_\lambda) \circ \Phi$. Moreover, to λ, they associate a facet J_λ, and hence a parahoric subgroup G_λ associated to J_λ. Let $\mathbb{G}_\lambda := G_\lambda/G_\lambda^+$. Let W_λ be the subgroup of W^o generated by reflections in the hyperplanes containing J_λ. Then to λ, DeBacker and Reeder associate an element $w_\lambda \in W_\lambda$. Fix once and for all a lift $\dot{w}$ of w to N. Using this lift, DeBacker and Reeder also associate a lift $\dot{w}_\lambda$ of w_λ to N. By Lang's theorem, there exists $p_\lambda \in G_\lambda$ such that $p_\lambda^{-1}\Phi_\lambda(p_\lambda) = \dot{w}_\lambda$. We then define $T_\lambda := \mathrm{Ad}(p_\lambda)T$, and set $\chi_\lambda := \chi_\phi \circ \mathrm{Ad}(p_\lambda)^{-1}$. Since χ_λ is depth-zero, its restriction to ${}^0T_\lambda^{\Phi_\lambda}$ factors through a character χ_λ^0 of $\mathbb{T}_\lambda^{\Phi_\lambda}$, where $\mathbb{T}_\lambda^{\Phi_\lambda}$ is the projection of ${}^0T^{\Phi_\lambda}$ in $\mathbb{G}_\lambda$. Therefore, χ_λ^0 gives rise to an irreducible cuspidal Deligne–Lusztig representation κ_λ^0 of $\mathbb{G}_\lambda^{\Phi_\lambda}$. Inflate κ_λ^0 to a representation of $G_\lambda^{\Phi_\lambda}$, and define an extension to $Z^{\Phi_\lambda} G_\lambda^{\Phi_\lambda}$ by

$$\kappa_\lambda := \chi_\lambda \otimes \kappa_\lambda^0,$$

where Z denotes the center of G. This makes sense since $(Z \cap G_\lambda)^{\Phi_\lambda}$ acts on κ_λ^0 via the restriction of χ_λ^0. Finally, form the representation

$$\pi_\lambda := \mathrm{Ind}_{Z^{\Phi_\lambda} G_\lambda^{\Phi_\lambda}}^{G^{\Phi_\lambda}} \kappa_\lambda,$$

where Ind denotes smooth induction. Then DeBacker and Reeder construct a packet $\Pi(\phi)$ of representations on the pure inner forms of G, parametrized by $\mathrm{Irr}(C_\phi)$, using the above construction, where C_ϕ is the component group of ϕ.

4. Existing description of the tame local Langlands correspondence for GL(ℓ, F)

In this section, we describe the construction of the tame local Langlands correspondence for GL(ℓ, F) as explained in [Moy 1986], where ℓ is a prime.

4A. ***Depth-zero supercuspidal representations of* GL(ℓ, F).** Let $(E/F, \chi)$ be an admissible pair, where χ has level 0 and E/F has degree ℓ. By definition of admissible pair, this implies that E/F is unramified, and the residue field of E is $\mathfrak{f}_\ell$. We have $\chi|_{1+\mathfrak{p}_E} = 1$, so $\chi|_{\mathfrak{o}_E^*}$ is the inflation of the character χ_o of $\mathfrak{f}_\ell^*$. By the theory of finite groups of Lie type, this character gives rise to an irreducible cuspidal representation λ' of GL(ℓ, $\mathfrak{f}$), which is the irreducible cuspidal Deligne–Lusztig representation corresponding to the elliptic torus $\mathfrak{f}_\ell^* \subset \mathrm{GL}(\ell, \mathfrak{f})$ and the character χ_o of $\mathfrak{f}_\ell^*$. Let λ be the inflation of λ' to $\mathrm{GL}(\ell, \mathfrak{o}_F)$. We may extend λ to a representation Λ of $K(F) := F^*\mathrm{GL}(\ell, \mathfrak{o}_F)$ by setting $\Lambda|_{F^*} = \chi|_{F^*}$, and then induce the resulting representation to $G(F) = \mathrm{GL}(\ell, F)$. Set

$$\pi_\chi = \mathrm{cInd}_{K(F)}^{G(F)} \Lambda,$$

where cInd denotes compact induction. Let $\mathbb{P}_\ell(F)_0$ be the subset of admissible pairs $(E/F, \chi)$ such that χ has level zero and $\mathbb{A}_\ell^0(F)_0$ be the subset of depth-zero supercuspidal representations of GL(ℓ, F).

Proposition 4.1. *Suppose that $p \neq \ell$. The map $(E/F, \chi) \mapsto \pi_\chi$ induces a bijection*

$$\mathbb{P}_\ell(F)_0 \to \mathbb{A}_\ell^0(F)_0$$

Proof. See [Moy 1986]. □

4B. ***Positive depth supercuspidal representations of* GL(ℓ, F), *ℓ a prime.*** In this section we recall the parametrization of the positive depth supercuspidal representations via admissible pairs, following [Moy 1986]. Let $\mathbb{A}_\ell^0(F)^+$ denote the set of all positive depth irreducible supercuspidal representations of GL(ℓ, F), and let

$\mathbb{P}_\ell(F)^+$ denote the set of all admissible pairs $(E/F, \chi) \in \mathbb{P}_\ell(F)$ such that χ has positive level.

Proposition 4.2. *Suppose that $p \neq \ell$. There is a map $(E/F, \chi) \mapsto \pi_\chi$ that induces a bijection*

$$\mathbb{P}_\ell(F)^+ \to \mathbb{A}_\ell^0(F)^+$$

Proof. See [Moy 1986]. □

4C. *Langlands parameters.* Let $\mathbb{G}_\ell^0(F)$ be the set of equivalence classes of irreducible smooth ℓ-dimensional representations of W_F. Recall that there is a local Artin reciprocity isomorphism given by $W_E^{ab} \cong E^*$. Then, if $(E/F, \chi) \in \mathbb{P}_\ell(F)$, χ gives rise to a character of W_E^{ab}, which we can pullback to a character, also denoted χ, of W_E. We can then form the induced representation $\phi(\chi) := \mathrm{Ind}_{W_E}^{W_F} \chi$ of W_F.

Theorem 4.3. *Suppose $p \neq \ell$. If $(E/F, \chi) \in \mathbb{P}_\ell(F)$, the representation $\phi(\chi)$ of W_F is irreducible. The map $(E/F, \chi) \mapsto \phi(\chi)$ induces a bijection*

$$\mathbb{P}_\ell(F) \to \mathbb{G}_\ell^0(F)$$

Proof. See [Moy 1986]. □

For the next theorem, we will need to associate to any admissible pair $(E/F, \chi)$ in $\mathbb{P}_\ell(F)$ a specific character Δ_χ of E^*. We will not define Δ_χ in general, but only for the cases that we need in this paper. For the general definition of Δ_χ associated to any admissible pair $(E/F, \chi) \in \mathbb{P}_\ell(F)$, see [Moy 1986].

Definition 4.4. If $(E/F, \chi)$ is an admissible pair in which E/F is quadratic and unramified, define Δ_χ to be the unique quadratic unramified character of E^*. If $(E/F, \chi)$ is an admissible pair in which E/F is of degree ℓ and unramified, where ℓ is an odd prime, then define Δ_χ to be the trivial character of E^*.

Theorem 4.5 (Tame local Langlands correspondence [Moy 1986]). *Suppose $p \neq \ell$. For $\phi \in \mathbb{G}_\ell^0(F)$, define $\pi(\phi) = \pi_{\chi\Delta_\chi}$ in the notation of Propositions 4.1 and 4.2, for any $(E/F, \chi) \in \mathbb{P}_\ell(F)$ such that $\phi \cong \phi(\chi)$. The map*

$$\pi : \mathbb{G}_\ell^0(F) \to \mathbb{A}_\ell^0(F)$$

is the local Langlands correspondence for supercuspidal representations of $\mathrm{GL}(\ell, F)$.

5. The case of GL(ℓ, F)

For Sections 5 and 6, we consider the group $\mathbf{G}(F) = \mathrm{GL}(\ell, F)$, where ℓ is prime. We will show that the conjectural correspondence of [DB-R] agrees with the local Langlands correspondence for $\mathrm{GL}(\ell, F)$ given in Section 4.

Let $\phi : W_F \to {}^LG$ be a TRSELP for $\mathbf{G}(F) = \mathrm{GL}(\ell, F)$. This is equivalent to an irreducible admissible $\phi : W_F \to \mathrm{GL}(\ell, \mathbb{C})$ that is trivial on the wild inertia group.

By Section 4C, we have $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$ for some admissible pair $(E/F, \chi)$, where χ has level zero and E/F is of degree ℓ and unramified. We will need the relative Weil group [Tate 1979, Chapter 1]

$$W_{E/F} := W_F/[W_E, W_E]^c,$$

where c denotes closure and $[W_E, W_E]$ denotes the commutator subgroup of W_E. The representation $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$ factors through $W_{E/F}$, since

$$\phi|_{W_E} = \chi \oplus \chi^{\xi} \oplus \cdots \oplus \chi^{\xi^{\ell-1}}.$$

We begin by calculating the character χ_ϕ from (3). Note that

$${}^L G = \langle \hat{\theta} \rangle \times \mathrm{GL}(\ell, \mathbb{C}).$$

$\hat{T}$ is the diagonal maximal torus in $\hat{G} = \mathrm{GL}(\ell, \mathbb{C})$, and after conjugation, we may assume $\phi(I_F) \subset \hat{T}$. Moreover, $\phi(\Phi) = \hat{\theta} f$ for some $f \in \hat{N}$ such that $\hat{w}$ is a cycle of length ℓ in the Weyl group S_ℓ, the symmetric group on ℓ letters. The reason for this requirement on the Weyl group element is that ϕ is TRSELP and hence elliptic. In particular, ellipticity is equivalent to requiring that the image of ϕ is not contained in any proper Levi subgroup of ${}^L G$ [DB-R, Section 3.4]. After conjugating the TRSELP by a permutation matrix in $N_{\hat{G}}(\hat{T})$, we may assume without loss of generality that $\hat{w} = (1\ 2\ 3\ \cdots\ \ell) \in S_\ell$ since all cycles of length ℓ are conjugate in S_ℓ. Note that this choice implies that $w = (1\ 2\ 3\ \cdots\ \ell) \in S_\ell$. The arguments in the remainder of the paper are the same for all other allowable choices of $\hat{w}$.

Let us first calculate χ_s, where $s := \phi|_{I_t}$ (recall again that $\phi|_{I_F^+} \equiv 1$, so $\phi|_{I_F}$ factors to I_t).

Proposition 5.1. *Let $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$ and set $s = \phi|_{I_t}$, where $(E/F, \chi)$ is an admissible pair as above. Then, the isomorphism*

$$\mathrm{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}(I_t, \hat{T}) \xrightarrow{\sim} \mathrm{Hom}_{\Phi,\hat{\sigma}}(\mathfrak{f}_\ell^*, \hat{T})$$

sends s to $\tilde{\beta}_s$, where

$$\tilde{\beta}_s(x) = \begin{pmatrix} \chi_o(x) & 0 & 0 & \dots & 0 \\ 0 & \chi_o^{\xi}(x) & 0 & \dots & 0 \\ 0 & 0 & \chi_o^{\xi^2}(x) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \chi_o^{\xi^{\ell-1}}(x) \end{pmatrix}$$

Proof. Since $\hat{\sigma}$ has order ℓ, $s \in \mathrm{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}(I_t, \hat{T})$ is trivial on $(1-\mathrm{Ad}(\Phi)^{\ell})\, I_t$, so factors to $I_t/(1-\mathrm{Ad}(\Phi)^{\ell})\, I_t$. We first note that the isomorphisms

$$I_t \cong \varprojlim_m \mathfrak{f}_m^*, \qquad I_t/(1-\mathrm{Ad}(\Phi)^{\ell})\, I_t \xrightarrow{\sim} \mathfrak{f}_\ell^*$$

are induced by local Artin reciprocity [R, Chapter 5]. Moreover, the map

$$\mathrm{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}(I_t, \hat{T}) \to \mathrm{Hom}_{\Phi,\hat{\sigma}}(\mathfrak{f}_\ell^*, \hat{T})$$

comes from the diagram

$$\begin{tikzcd} & I_t \arrow[r, "s"] \arrow[d] & \hat{T} \\ \mathfrak{f}_\ell^* \arrow[r, "\sim"] & I_t/(1-\mathrm{Ad}(\Phi)^{\ell})\, I_t \arrow[ur] & \end{tikzcd}$$

Recall that ϕ factors through $W_{E/F}$. Hence, we also have the commutative diagram

$$\begin{tikzcd} W_F \arrow[r, "\phi"] \arrow[d] & \mathrm{GL}(\ell, \mathbb{C}) \\ W_{E/F} \arrow[ur, "\beta"] & \end{tikzcd}$$

It is a fact that $W_{E/F}$ is an extension of $\mathrm{Gal}(E/F)$ by E^*, and can be described by generators and relations as follows. The generators are $\{z \in E^*\}$ and an element j where $j \in W_{E/F}$ satisfies $j^{\ell} = \varpi$ and $jzj^{-1} = \xi(z)$. Then the map $W_F \to W_{E/F}$ sends I_F to $\mathfrak{o}_E^*$ and Φ to j.

Let us calculate the map β. Consider the canonical sequence

$$1 \to W_E/[W_E, W_E]^c \to W_F/[W_E, W_E]^c \to W_F/W_E \cong \mathrm{Gal}(E/F) \to 1$$

Recall that ϕ is trivial on $[W_E, W_E]^c$. To calculate $\beta|_{E^*}$, it suffices to calculate $\phi|_{W_E}$ since $W_E/[W_E, W_E]^c \cong E^*$ by Artin reciprocity. But

$$\phi|_{W_E} = \chi \oplus \chi^{\xi} \oplus \cdots \oplus \chi^{\xi^{\ell-1}}.$$

Therefore,

$$\beta(t) = \begin{pmatrix} \chi(t) & 0 & 0 & \dots & 0 \\ 0 & \chi^{\xi}(t) & 0 & \dots & 0 \\ 0 & 0 & \chi^{\xi^2}(t) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \chi^{\xi^{\ell-1}}(t) \end{pmatrix}$$

Moreover, since ϕ is irreducible, we have that $\beta(j) \in N_{\mathrm{GL}(\ell,\mathbb{C})}(\hat{T})$ represents $\hat{w}$.

Since $\phi|_{I_F^+} \equiv 1$, we have that $\beta|_{1+\mathfrak{p}_E} \equiv 1$, so $\beta|_{\mathfrak{o}_E^*}$ factors to a map

$$\tilde{\beta}_s : \mathfrak{f}_\ell^* \to \mathrm{GL}(\ell, \mathbb{C}),$$

given by

$$\tilde{\beta}_s(x) = \begin{pmatrix} \chi_o(x) & 0 & 0 & \dots & 0 \\ 0 & \chi_o^{\xi}(x) & 0 & \dots & 0 \\ 0 & 0 & \chi_o^{\xi^2}(x) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \chi_o^{\xi^{\ell-1}}(x) \end{pmatrix} \quad \text{for all } x \in \mathfrak{f}_\ell^*. \qquad \square$$

Proposition 5.2. *Let $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$ and set $s = \phi|_{I_t}$ as above. Then the composite isomorphism*

$$\mathrm{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}(I_t, \hat{T}) \xrightarrow{\sim} \mathrm{Hom}_{\Phi,\hat{\sigma}}(\mathfrak{f}_\ell^*, \hat{T}) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(X \otimes \mathfrak{f}_\ell^*, \mathbb{C}^*) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(\mathbb{T}^{\Phi_\sigma^\ell}, \mathbb{C}^*)$$

sends s to ${}^\ell\chi_o^E$, where $s = \phi|_{I_t}$ and ${}^\ell\chi_o^E(x_1, x_2, \dots, x_\ell) := \chi_o(x_1)\chi_o^{\xi}(x_2)\dots\chi_o^{\xi^{\ell-1}}(x_\ell)$.

Proof. The composite isomorphism

$$\mathrm{Hom}_{\Phi,\hat{\sigma}}(\mathfrak{f}_\ell^*, \hat{T}) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(X \otimes \mathfrak{f}_\ell^*, \mathbb{C}^*) \xrightarrow{\sim} \mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(\mathbb{T}^{\Phi_\sigma^\ell}, \mathbb{C}^*)$$

is given by $\tilde{\alpha} \longmapsto \{\lambda(x) \mapsto \lambda(\tilde{\alpha}(x))\}$, where $x \in \mathfrak{f}_\ell^*$ and $\lambda \in X = X_*(T) = X^*(\hat{T})$. Note that $\mathbb{T}$ splits over $\mathfrak{f}_\ell$ and $\mathbb{T}^{\Phi_\sigma^\ell} \cong \mathfrak{f}_\ell^* \times \dots \times \mathfrak{f}_\ell^*$. Then, it is easy to see that under this composite isomorphism, $\tilde{\beta}_s$ (where $\tilde{\beta}_s$ is as in Proposition 5.1) maps to the homomorphism

$$(x_1, x_2, \dots, x_\ell) \mapsto \chi_o(x_1)\, \chi_o^{\xi}(x_2) \cdots \chi_o^{\xi^{\ell-1}}(x_\ell) \quad \text{for all } x_1, x_2, \dots, x_\ell \in \mathfrak{f}_\ell^*$$

by considering the standard basis of cocharacters of X. $\square$

Proposition 5.3. *The isomorphism*

$$\mathrm{Hom}_{\Phi_\sigma,\mathrm{Id}}(\mathbb{T}^{\Phi_\sigma^\ell}, \mathbb{C}^*) \xrightarrow{\sim} \mathrm{Hom}(\mathbb{T}^{\Phi_\sigma}, \mathbb{C}^*) \tag{4}$$

is given by $\Lambda \mapsto \Lambda'$, where $\Lambda'(a) := \Lambda((x_1, x_2, \dots, x_\ell))$ whenever $a \in \mathfrak{f}_\ell^*$ and $(x_1, x_2, \dots, x_\ell) \in \mathfrak{f}_\ell^* \times \mathfrak{f}_\ell^* \times \dots \times \mathfrak{f}_\ell^*$ satisfies $x_1 x_2^{q^{\ell-1}} x_3^{q^{\ell-2}} \cdots x_\ell^q = a$.

Proof. Recall that the isomorphism (4) is abstractly given by $\Lambda \mapsto \Lambda'$, where $\Lambda'(a) := \Lambda((x_1, x_2, \dots, x_\ell))$ for any $(x_1, x_2, \dots, x_\ell) \in \mathfrak{f}_\ell^* \times \mathfrak{f}_\ell^* \times \dots \times \mathfrak{f}_\ell^*$ such that $N_\sigma((x_1, x_2, \dots, x_\ell)) = a$.

We need some preliminaries. First note that

$$\begin{aligned}\Phi_\sigma((x_1, x_2, \dots, x_\ell)) &= w\, \Phi^{-1}((x_1, x_2, \dots, x_\ell)) \\ &= w(x_1^q, x_2^q, \dots, x_\ell^q) = (x_\ell^q, x_1^q, x_2^q, \dots, x_{\ell-1}^q).\end{aligned}$$

If we make the identification of $\mathbb{T}^{\Phi_\sigma^\ell}$ with tuples $(x_1, x_2, \dots, x_\ell) \in \mathfrak{f}_\ell^* \times \mathfrak{f}_\ell^* \times \cdots \times \mathfrak{f}_\ell^*$, then we have that since we made our choice of $w = (1\ 2\ 3\ \cdots\ \ell) \in S_\ell$, we get

$$\begin{aligned}\mathbb{T}^{\Phi_\sigma} &= \{(x_1, \dots, x_\ell) \in \mathfrak{f}_\ell^* \times \mathfrak{f}_\ell^* \times \cdots \times \mathfrak{f}_\ell^* : (x_\ell^q, x_1^q, \dots, x_{\ell-1}^q) = (x_1, x_2, \dots, x_\ell)\} \\ &= \{(x_1, x_1^q, x_1^{q^2}, \dots, x_1^{q^{\ell-1}}) : x_1 \in \mathfrak{f}_\ell\}.\end{aligned}$$

If $(x_1, x_2, \dots, x_\ell) \in \mathfrak{f}_\ell^* \times \mathfrak{f}_\ell^* \times \cdots \times \mathfrak{f}_\ell^* = \mathbb{T}^{\Phi_\sigma^\ell}$, then

$$\begin{aligned}&N_\sigma((x_1, x_2, \dots, x_\ell)) \\ &\quad= (x_1, x_2, \dots, x_\ell)\, \Phi_\sigma((x_1, x_2, \dots, x_\ell)) \cdots \Phi_\sigma^{\ell-1}((x_1, x_2, \dots, x_\ell)) \\ &\quad= (x_1, x_2, \dots, x_\ell)\, (x_\ell^q, x_1^q, x_2^q, \dots, x_{\ell-1}^q) x_{\ell-2}^{q^2}) \dots (x_2^{q^{\ell-1}}, x_3^{q^{\ell-1}}, \dots, x_\ell^{q^{\ell-1}}, x_1^{q^{\ell-1}}) \\ &\quad= (x_1 x_2^{q^{\ell-1}} x_3^{q^{\ell-2}} \cdots x_\ell^q,\ x_2 x_3^{q^{\ell-1}} x_4^{q^{\ell-2}} \cdots x_1^q,\ \dots,\ x_\ell x_1^{q^{\ell-1}} x_2^{q^{\ell-2}} \cdots x_{\ell-1}^q)\end{aligned}$$

Therefore, $N_\sigma : \mathbb{T}^{\Phi_\sigma^\ell} \to \mathbb{T}^{\Phi_\sigma}$ is the map

$$(x_1, x_2, \dots, x_\ell) \mapsto x_1 x_2^{q^{\ell-1}} x_3^{q^{\ell-2}} \cdots x_\ell^q$$

for all $(x_1, x_2, \dots, x_\ell) \in \mathfrak{f}_\ell^* \times \mathfrak{f}_\ell^* \times \cdots \times \mathfrak{f}_\ell^*$. ☐

We now need to obtain a character of ${}^0T^{\Phi_\sigma}$ from a character of $\mathbb{T}^{\Phi_\sigma}$. In our case, ${}^0T^{\Phi_\sigma} = \mathfrak{o}_E^*$, which has a canonical projection map $\mathfrak{o}_E^* = {}^0T^{\Phi_\sigma} \xrightarrow{\eta} \mathbb{T}^{\Phi_\sigma} = \mathfrak{f}_\ell^*$. Then, given $\zeta \in \mathrm{Hom}(\mathbb{T}^{\Phi_\sigma}, \mathbb{C}^*)$, we obtain a character μ of ${}^0T^{\Phi_\sigma} = \mathfrak{o}_E^*$ given by $\mu(z) := \zeta(\eta(z))$, $z \in \mathfrak{o}_E^*$. Let us more explicitly calculate such a μ, given some $\Lambda' \in \mathrm{Hom}(\mathbb{T}^{\Phi_\sigma}, \mathbb{C}^*)$ that comes from $\Lambda \in \mathrm{Hom}_{\Phi_\sigma, \mathrm{Id}}(\mathbb{T}^{\Phi_\sigma^\ell}, \mathbb{C}^*)$ as in (4). Let $z \in \mathfrak{o}_E^*$. Then $\mu(z) = \Lambda'(\eta(z)) = \Lambda((\eta(z), 1, 1, \dots, 1))$, by Proposition 5.3.

We may now calculate the character χ_s that arises from ϕ, where $s = \phi|_{I_t}$ and $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$. The above analysis and Proposition 5.2 shows that

$$\chi_s(z) = {}^\ell\chi_o^E((\eta(z), 1, 1, \dots, 1)) = \chi_o(\eta(z)) = \chi(z),$$

where $z \in \mathfrak{o}_E^*$. It remains to compute χ_τ. First note that if we make the identification $X = \mathbb{Z} \times \mathbb{Z} \times \cdots \times \mathbb{Z}$, then $X^\sigma = \{(k, k, \dots, k) : k \in \mathbb{Z}\}$. Let $\lambda_{(k,k,\dots,k)} \in X^\sigma$ denote the character of $\hat{T}$ corresponding to $(k, k, \dots, k) \in \mathbb{Z} \times \mathbb{Z} \times \cdots \times \mathbb{Z}$ via this identification.

Proposition 5.4. *Let $\ell = 2$. The character $\chi_\tau : X^\sigma \to \mathbb{C}^*$ is given by*

$$\lambda_{(k,k)} \mapsto (-\chi(\varpi))^k.$$

Proof. Note that $\hat{\theta} = 1$ and $\hat{G}' = \mathrm{SL}(2, \mathbb{C})$, so τ is any element whose class in $\hat{T}/(1-\hat{\sigma})\hat{T}$ corresponds to the image of f in $\mathrm{GL}(2, \mathbb{C})/\mathrm{SL}(2, \mathbb{C})$ under the bijection

$$\hat{T}/(1-\hat{\sigma})\hat{T} \xrightarrow{\sim} \mathrm{GL}(2, \mathbb{C})/\mathrm{SL}(2, \mathbb{C})$$

as in (2). We thus need to compute f first.

Recall that $\phi(\Phi) = \beta(j)$, where β is as in the proof of Proposition 5.1. Recall that since ϕ is irreducible, then

$$\beta(j) = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

for some $a, b \in \mathbb{C}^*$. After conjugation by $\hat{G}$, we may assume that $b = 1$. But since $j^2 = \varpi$, we have

$$\begin{pmatrix} \chi(\varpi) & 0 \\ 0 & \chi^{\xi}(\varpi) \end{pmatrix} = \beta(\varpi) = \beta(j^2) = \beta(j)^2 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

Therefore, $a = \chi(\varpi)$ and so

$$\beta(j) = \begin{pmatrix} 0 & \chi(\varpi) \\ 1 & 0 \end{pmatrix},$$

and we may take

$$f = \begin{pmatrix} 0 & \chi(\varpi) \\ 1 & 0 \end{pmatrix}.$$

We now note that the bijection

$$\hat{T}/(1-\hat{\sigma})\hat{T} \xrightarrow{\sim} \hat{G}_{ab}/(1-\hat{\theta})\hat{G}_{ab}$$

is induced by the inclusion $\hat{T} \hookrightarrow \hat{G}$ [DB-R, Section 4.3]. Now, we have that

$$\left[\begin{pmatrix} -\chi(\varpi) & 0 \\ 0 & 1 \end{pmatrix}\right] = \left[\begin{pmatrix} 0 & \chi(\varpi) \\ 1 & 0 \end{pmatrix}\right] \in \mathrm{GL}(2, \mathbb{C})/\mathrm{SL}(2, \mathbb{C})$$

since

$$\begin{pmatrix} 0 & \chi(\varpi) \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -\chi(\varpi) & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C})$$

Therefore, since

$$\begin{pmatrix} -\chi(\varpi) & 0 \\ 0 & 1 \end{pmatrix} \in \hat{T},$$

we may set

$$\tau = \begin{pmatrix} -\chi(\varpi) & 0 \\ 0 & 1 \end{pmatrix}.$$

Then $\chi_\tau : X^\sigma \to \mathbb{C}^*$ is given by

$$\chi_\tau(\lambda_{(k,k)}) = \lambda_{(k,k)}(\tau) = \lambda_{(k,k)}\begin{pmatrix} -\chi(\varpi) & 0 \\ 0 & 1 \end{pmatrix} = (-\chi(\varpi))^k. \qquad \square$$

Proposition 5.5. *Let ℓ be an odd prime. The character $\chi_\tau : X^\sigma \to \mathbb{C}^*$ is given by*

$$\lambda_{(k,k,\dots,k)} \mapsto \chi(\varpi)^k.$$

Proof. Note that $\hat{\theta} = 1$ and $\hat{G}' = \mathrm{SL}(\ell, \mathbb{C})$, so τ is any element whose class in $\hat{T}/(1-\hat{\sigma})\hat{T}$ corresponds to the image of f in $\mathrm{GL}(\ell,\mathbb{C})/\mathrm{SL}(\ell,\mathbb{C})$ under the bijection

$$\hat{T}/(1-\hat{\sigma})\hat{T} \xrightarrow{\sim} \mathrm{GL}(\ell,\mathbb{C})/\mathrm{SL}(\ell,\mathbb{C})$$

as in (2). We thus need to compute f first.

Recall that ϕ factors through $W_{E/F}$, and we have the commutative diagram

$$\begin{tikzcd} W_F \arrow[r, "\phi"] \arrow[d] & \mathrm{GL}(\ell,\mathbb{C}) \\ W_{E/F} \arrow[ur, "\beta"] & \end{tikzcd}$$

From Proposition 5.1, we have $\phi(\Phi) = \beta(j)$. To compute $\beta(j)$, recall that because of our choice of $\hat{w}$, we have

$$\beta(j) = \begin{pmatrix} 0 & 0 & 0 & \dots & a_1 \\ a_2 & 0 & 0 & \dots & 0 \\ 0 & a_3 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & a_\ell & 0 \end{pmatrix}$$

for some $a_i \in \mathbb{C}^*$. After conjugation the Langlands parameter by an element in $\hat{G}$ of the form

$$\begin{pmatrix} 0 & x_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & x_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & x_4 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & x_\ell \\ x_1 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

we may assume that $a_2 = a_3 = \cdots = a_\ell = 1$. Therefore,

$$\begin{pmatrix} \chi(\varpi) & 0 & 0 & \dots & 0 \\ 0 & \chi^{\xi}(\varpi) & 0 & \dots & 0 \\ 0 & 0 & \chi^{\xi^2}(\varpi) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \chi^{\xi^{\ell-1}}(\varpi) \end{pmatrix} = \beta(\varpi) = \beta(j^{\ell}) = \beta(j)^{\ell} = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_1 & 0 & \dots & 0 \\ 0 & 0 & a_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_1 \end{pmatrix}$$

Hence, $a_1 = \chi(\varpi)$, so we may take $f = \begin{pmatrix} 0 & 0 & 0 & \dots & \chi(\varpi) \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$ Now, we have

$$\left[\begin{pmatrix} \chi(\varpi) & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}\right] = \left[\begin{pmatrix} 0 & 0 & 0 & \dots & \chi(\varpi) \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}\right] \in \mathrm{GL}(\ell, \mathbb{C})/\mathrm{SL}(\ell, \mathbb{C})$$

since

$$\begin{pmatrix} 0 & 0 & 0 & \dots & \chi(\varpi) \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \chi(\varpi) & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix},$$

which is an element of $\mathrm{SL}(\ell, \mathbb{C})$. Therefore, we may set

$$\tau = \begin{pmatrix} \chi(\varpi) & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Then $\chi_\tau : X^\sigma \to \mathbb{C}^*$ is given by

$$\chi_\tau(\lambda_{(k,k,\dots,k)}) = \lambda_{(k,k,\dots,k)}(\tau) = (\chi(\varpi))^k. \qquad \square$$

Recall that we have computed χ_ϕ on $\mathfrak{o}_E^*$. It remains to compute $\chi_\phi(\varpi)$. Because of the isomorphism

$$^0T^{\Phi_\sigma} \times X^\sigma \xrightarrow{\sim} T^{\Phi_\sigma}, \quad (\gamma, \lambda) \mapsto \gamma\lambda(\varpi),$$

we need to compute $\chi_\phi(1, \lambda_{(1,1,\ldots,1)})$.

Proposition 5.6. *Let $\ell = 2$. Then $\chi_\phi = \chi\Delta_\chi$, where $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$.*

Proof. We have that $\chi_\phi(1, \lambda_{(1,1)}) = \chi_s(1)\chi_\tau(\lambda_{(1,1)}) = -\chi(\varpi)$. Therefore, $\chi_\phi(\varpi) = -\chi(\varpi)$. Recall that we have shown that $\chi_\phi|_{\mathfrak{o}_E^*} = \chi|_{\mathfrak{o}_E^*}$. Since $\ell = 2$, Δ_χ is the unique quadratic unramified character of E^*. Therefore, we have that $\Delta_\chi(\varpi) = -1$ and $\Delta_\chi|_{\mathfrak{o}_E^*} \equiv 1$, so $\chi_\phi = \chi\Delta_\chi$. □

Proposition 5.7. *Let ℓ be an odd prime. $\chi_\phi = \chi\Delta_\chi$.*

Proof. We have that $\chi_\phi(1, \lambda_{(1,1,\ldots,1)}) = \chi_s(1)\,\chi_\tau(\lambda_{(1,1,\ldots,1)}) = \chi(\varpi)$. Therefore, $\chi_\phi(\varpi) = \chi(\varpi)$. Recall that we have shown that $\chi_\phi|_{\mathfrak{o}_E^*} = \chi|_{\mathfrak{o}_E^*}$. Therefore, $\chi_\phi = \chi$. But recall that Δ_χ is trivial since ℓ is an odd prime, so we have $\chi_\phi = \chi\Delta_\chi$. □

6. From a character of a torus to a representation for GL(ℓ, F)

In this section we determine the representation that DeBacker and Reeder assign to a TRSELP for $\mathrm{GL}(\ell, F)$, using the results from Section 5. Note that

$$[X/(1 - w\theta)X]_{\mathrm{tor}} = 0,$$

so we may let $\lambda = 0$ (recall that $\lambda \in X_w$). The proof of [DB-R, Lemma 2.7.2] implies that we may take $u_\lambda = 1$, and therefore $\Phi_\lambda = \Phi$. It is also easy to see that we may take $w_\lambda = w$ ([DB-R, Section 2.7]) and $\dot{w}_\lambda = \dot{w}$, where $\dot{w}$ is a fixed choice of lift of w. Since the theory of [DB-R] is independent of any choices, we are free to choose a specific lift $\dot{w}$, which we do now.

Let $f(x)$ be a monic irreducible polynomial of degree ℓ over $\mathfrak{f}$. Let $\tilde{f}(x)$ be a monic lift of $f(x)$ to $F[x]$. We may write $E = F(\delta)$, where δ is a root of $\tilde{f}(x)$. First set

$$\tilde{w} := \begin{pmatrix} 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}.$$

Recall that we need to find $p_\lambda \in G_\lambda$ such that $p_\lambda^{-1}\Phi(p_\lambda) = \dot{w}_\lambda$. By choosing the basis $1, \delta, \delta^2, \ldots, \delta^{\ell-1}$ for E over F, we may embed E^* into $\mathrm{GL}(\ell, F)$ in the standard way. Denote this embedding by $\varphi : E^* \hookrightarrow \mathrm{GL}(\ell, F)$.

Lemma 6.1. *There exists an $A \in G_\lambda$ such that*

$$A \begin{pmatrix} t & 0 & 0 & 0 & \dots & 0 \\ 0 & \xi(t) & 0 & 0 & \dots & 0 \\ 0 & 0 & \xi^2(t) & 0 & \dots & 0 \\ 0 & 0 & 0 & \xi^3(t) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \xi^{\ell-1}(t) \end{pmatrix} A^{-1} = \varphi(t)$$

for all $t = a_0 + a_1\delta + a_2\delta^2 + \cdots + a_{\ell-1}\delta^{\ell-1} \in E^*$, $a_i \in F$.

Proof. Suppose $R(x)$ is a polynomial of degree ℓ in $F[x]$ that splits over E. Then we get an isomorphism

$$E[x]/(R(x)) \xrightarrow{\sim} \bigoplus_{i=1}^{\ell} E,$$
$$p(x)0 \mapsto (p(a_1), p(a_2), \dots, p(a_\ell)),$$

where the a_i are the roots of $R(x)$. Setting $R(x)$ to now be the minimal polynomial of δ, and considering the basis $1, x, x^2, \dots, x^{\ell-1}$ of $E[x]/(R(x))$ over E, we get an isomorphism

$$\begin{aligned} E[x]/(R(x)) &\xrightarrow{G} E \oplus E \oplus \cdots \oplus E \\ 1 &\mapsto (1, 1, 1, \dots, 1) \\ x &\mapsto \big(\delta, \Phi(\delta), \Phi^2(\delta), \dots, \Phi^{\ell-1}(\delta)\big) \\ x^2 &\mapsto \big(\delta^2, \Phi(\delta)^2, \Phi^2(\delta)^2, \dots, \Phi^{\ell-1}(\delta)^2\big) \\ &\dots \\ x^{\ell-1} &\mapsto \big(\delta^{\ell-1}, \Phi(\delta)^{\ell-1}, \Phi^2(\delta)^{\ell-1}, \dots, \Phi^{\ell-1}(\delta)^{\ell-1}\big) \end{aligned}$$

This transformation yields the matrix

$$V := \begin{pmatrix} 1 & \delta & \delta^2 & \delta^3 & \dots & \delta^{\ell-1} \\ 1 & \Phi(\delta) & \Phi(\delta)^2 & \Phi(\delta)^3 & \dots & \Phi(\delta)^{\ell-1} \\ 1 & \Phi^2(\delta) & \Phi^2(\delta)^2 & \Phi^2(\delta)^3 & \dots & \Phi^2(\delta)^{\ell-1} \\ 1 & \Phi^3(\delta) & \Phi^3(\delta)^2 & \Phi^3(\delta)^3 & \dots & \Phi^3(\delta)^{\ell-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \Phi^{\ell-1}(\delta) & \Phi^{\ell-1}(\delta)^2 & \Phi^{\ell-1}(\delta)^3 & \dots & \Phi^{\ell-1}(\delta)^{\ell-1} \end{pmatrix}.$$

We then set $A := V^{-1}$. Note that what we have really done here is the following. We first have taken the standard E-basis $e_1', e_2', \dots, e_\ell'$ of $E \oplus E \oplus \cdots \oplus E$ and pulled it back by G to get a basis $e_1, e_2, \dots, e_\ell$ of $E[x]/(R(x))$. We have then shown that the standard embedding of an element $w \in E^*$ in GL(ℓ, F) (by considering its action on the basis $1, \delta, \delta^2, \ldots, \delta^{\ell-1}$) can be diagonalized over E with respect to the basis $e_1, e_2, \ldots, e_\ell$.

Note that V is a Vandermonde matrix. Therefore, its determinant is

$$\prod_{0\le i<j\le \ell-1} (\Phi^j(\delta) - \Phi^i(\delta)),$$

which has valuation zero. Since we also have that the entries of V are contained in $\mathfrak{o}_E^*$, we conclude that V, and hence A, is contained in G_λ. $\square$

Set $\tilde{s} = \tilde{w}^{-1}A^{-1}\Phi(A)$. We now fix our choice of lift $\dot{w}$ of w by setting $\dot{w} := \tilde{s}^{-1}\tilde{w}\,\tilde{s}\,\Phi(\tilde{s})$, which we shall show is a legitimate lift. We claim first that we may set $p_\lambda = A\tilde{s}$, and that $\tilde{s} \in G_\lambda \cap T$. Since we will show that $\tilde{s} \in G_\lambda \cap T$, this shows that $p_\lambda \in G_\lambda$, which is required. To prove all of this, consider the adjoint action of $A^{-1}\Phi(A)$ on T. First, for $s \in T^{\Phi_w}$, we have

$$\begin{aligned}(A^{-1}\Phi(A)) \cdot \Phi(s) &= A^{-1}\Phi(A)\Phi(s)\Phi(A)^{-1}A \\ &= A^{-1}\Phi(AsA^{-1})A = A^{-1}AsA^{-1}A = s\end{aligned}$$

since Lemma 6.1 implies that AsA^{-1} is fixed by Φ.

We therefore have that $(A^{-1}\Phi(A)) \cdot \Phi(s) = w \cdot \Phi(s)$ for all $s \in T^{\Phi_w}$. Since T^{Φ_w} is dense in T in the Zariski topology, we have that $(A^{-1}\Phi(A)) \cdot \Phi(s) = w \cdot \Phi(s)$ for all $s \in T$. This implies that $(\tilde{w}^{-1}A^{-1}\Phi(A)) \cdot s = s$ for all $s \in T$ since $\tilde{w}$ is clearly a lift of w, which means that

$$\tilde{w}^{-1}A^{-1}\Phi(A)\, s\, (\tilde{w}^{-1}A^{-1}\Phi(A))^{-1} = s \quad \text{for all } s \in T.$$

This means that $\tilde{w}^{-1}A^{-1}\Phi(A) \in C_G(T) = T$, so in particular $\tilde{w}^{-1}A^{-1}\Phi(A) = \tilde{s} \in T$. But $A, \tilde{w} \in G_\lambda$ implies that $\tilde{w}^{-1}A^{-1}\Phi(A) \in G_\lambda$, which implies that $\tilde{s} \in G_\lambda \cap T$. This shows that $p_\lambda \in G_\lambda$, which is required. Moreover,

$$p_\lambda^{-1}\Phi(p_\lambda) = (A\tilde{s})^{-1}\Phi(A\tilde{s}) = \tilde{s}^{-1}A^{-1}\Phi(A)\,\Phi(\tilde{s}) = \tilde{s}^{-1}\tilde{w}\,\tilde{s}\,\Phi(\tilde{s}) = \dot{w}.$$

Finally, $\dot{w}$ is a lift of w since $\tilde{w}$ is, and since $\tilde{s} \in T$, proving the claim.

Thus, we have a p_λ such that $p_\lambda^{-1}\Phi_\lambda(p_\lambda) = \dot{w}$, and $\dot{w}$ is indeed a lift of w. Then if we define $T_\lambda := \mathrm{Ad}(p_\lambda)T$, we get that $T_\lambda^{\Phi_\lambda}$ is the image of E^* under φ. This is crucial, since the depth-zero supercuspidals of $\mathrm{GL}(\ell, F)$ are constructed in Section 4A by first fixing an the embedding of E^* into $\mathrm{GL}(\ell, F)$. The overall construction does not depend on the choice of embedding. We have fixed the embedding φ. DeBacker and Reeder are attaching a depth-zero supercuspidal representation of $\mathrm{GL}(\ell, F)$ to a Langlands parameter, and we need to show that their depth-zero supercuspidal matches the depth-zero supercuspidal attached in Theorem 4.5 (the latter of which, again, uses the construction in Section 4A, which assumes a fixed embedding, which we are assuming without loss of generality is φ).

Note that we have a simple description for the map $\mathrm{Ad}(p_\lambda)^{-1} : T_\lambda^{\Phi_\lambda} \to T^{\Phi_w}$, namely,

$$\phi(t) \mapsto \mathrm{diag}\big(t, \xi(t), \xi^2(t), \ldots, \xi^{\ell-1}(t)\big),$$

where $\mathrm{diag}(d_1, d_2, \ldots, d_\ell)$ denotes the diagonal $\ell \times \ell$ matrix with $d_1, d_2, \ldots, d_\ell$ on the diagonal, and where $t = a_0 + a_1\delta + a_2\delta^2 + \cdots + a_{\ell-1}\delta^{\ell-1}$. Note that

$$T^{\Phi_w} = \big\{\mathrm{diag}\big(a_0, \xi(a_0), \xi^2(a_0), \ldots, \xi^{\ell-1}(a_0)\big) : a_0 \in E^*\big\}$$

Finally, because of Propositions 5.6 and 5.7 and the definition of χ_λ, we have:

Proposition 6.2. *χ_λ is given by*

$$\chi_\lambda(\varphi(t)) = \chi(t)\,\Delta_\chi(t)$$

for all $t = a_0 + a_1\delta + a_2\delta^2 + \cdots + a_{\ell-1}\delta^{\ell-1} \in E^*$.

Let us sum up the data that we have obtained so far. Given a TRSELP for $\mathrm{GL}(\ell, F)$, we have obtained a torus T^{Φ_w}. Given $\lambda = 0 \in X_w$, we have constructed $T_\lambda^{\Phi_\lambda}$ and p_λ. We have $T_\lambda^{\Phi_\lambda} \cong E^*$. From ϕ we have constructed a character χ_ϕ of T^{Φ_w}. Via $\mathrm{Ad}(p_\lambda)$, we transported χ_ϕ to a character χ_λ of $T_\lambda^{\Phi_\lambda}$. We have shown that $\chi_\lambda = \chi\,\Delta_\chi$. note that the restriction of χ_λ to ${}^0T_\lambda^{\Phi_\lambda}$ factors through a character χ_λ^0 of $\mathbb{T}_\lambda^{\Phi_\lambda}$. Then, the packet of representations that DeBacker–Reeder construct in [DB-R] from the data that we have obtained thus far is the single representation

$$\mathrm{Ind}_{F^*\mathrm{GL}(\ell,\mathfrak{o}_F)}^{\mathrm{GL}(\ell,F)}(\chi_\lambda \otimes \kappa_\lambda^0) = \pi_{\chi\Delta_\chi}$$

Recall that in Section 4C, the local Langlands correspondence for $\mathrm{GL}(\ell, F)$, where ℓ is prime, was given as

$$\mathrm{Ind}_{W_E}^{W_F}(\chi) \mapsto \pi_{\chi\Delta_\chi}$$

We have therefore shown that the correspondence of DeBacker–Reeder coincides with the local Langlands correspondence.

7. The positive-depth correspondence of Reeder for GL(ℓ, F)

In this section, we prove that the correspondence of [R] agrees with the local Langlands correspondence of [Moy 1986] for $\mathrm{GL}(\ell, F)$, where ℓ is an arbitrary prime, if one assumes a certain compatibility condition, which we describe now. Reeder's construction begins by canonically attaching a certain admissible pair $(L/F, \Omega)$ to a Langlands parameter for $\mathrm{GL}(\ell, F)$. His construction then inputs this admissible pair into the theory of [Adler 1998] in order to construct a supercuspidal representation $\pi(L, \Omega)$ of $\mathrm{GL}(\ell, F)$. The compatibility condition that we will need to assume is that $\pi(L, \Omega)$ is the same supercuspidal representation that is attached to $(L/F, \Omega)$ via the construction in [Howe 1977].

Most of the arguments and setup are the same as in the depth-zero case, so there is not much to prove here. We first very briefly review the construction of Reeder and refer to [R] for definitions and notions that are not explained here.

7A. *Review.* Let $\mathbf{G}$ be an F-quasisplit and F^u-split connected reductive group. Let $\mathbf{B} \subset \mathbf{G}$ be a Borel subgroup defined over F, and $\mathbf{T}$ a maximal torus of $\mathbf{B}$.

The Langlands parameters considered in [R] are the maps

$$\phi : W_F \to {}^L G = \langle \hat{\theta} \rangle \ltimes \hat{G}$$

such that:

(1) ϕ is trivial on $I^{(r+1)}$ and nontrivial on $I^{(r)}$ for some integer $r > 0$. Here, $\{I^{(k)}\}_{k \geq 0}$ is a filtration on I_F defined in [R, Section 5.2].

(2) The centralizer of $\phi(I^{(r)})$ in $\hat{G}$ is a maximal torus of $\hat{G}$. This is the *regularity* condition.

(3) $\phi(\Phi) \in \hat{\theta} \ltimes \hat{G}$, and the centralizer of $\phi(W_F)$ in $\hat{G}$ is finite, modulo $\hat{Z}^{\hat{\theta}}$. This is the *ellipticity* condition.

We may conjugate ϕ by an element of $\hat{G}$ so that $\phi(I_F) \subset \hat{T}$, and $\phi(\Phi) = \hat{\theta} f$, where $f \in \hat{N}$. Let $\hat{w}$ be the image of f in $\hat{W}_o$, and let w be the element of W_o dual to $\hat{w}$. We say that the element w is *associated* to ϕ.

Set $\sigma = w\theta$ and suppose its action on X has order n. From an above such Langlands parameter, Reeder defines a $\hat{T}$-conjugacy class of Langlands parameters

$$\phi_T : W_F \to {}^L T_\sigma$$

in the exact same way as in the depth-zero case. In particular, the element τ is defined in the same way.

As in the depth-zero case, a bijection is later given between $\hat{T}$-conjugacy classes of continuous homomorphisms

$$\phi : W_F / I^{(r+1)} \to {}^L T_\sigma$$

for which $\phi(\Phi) \in \hat{\sigma} \ltimes \hat{T}$ and characters of T^{Φ_σ} that are trivial on $T_{r+1}^{\Phi_\sigma}$, where $\{T_k\}_{k \geq 0}$ is the canonical filtration on T [R, Section 5.3]. This is done as follows. We have a composite isomorphism [R, Section 5.3]

$$\begin{aligned} \operatorname{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}\big(I_F / I^{(r+1)}, \hat{T}\big) &\cong \operatorname{Hom}_{\mathrm{Ad}(\Phi),\hat{\sigma}}\big(I_F / I_n^{(r+1)}, \hat{T}\big) \\ &= \operatorname{Hom}_{\Phi,\hat{\sigma}}\big(\mathfrak{o}_n^* / (1 + \mathfrak{p}_n^{r+1}), \hat{T}\big) \\ &= \operatorname{Hom}_{\Phi_\sigma,\mathrm{Id}}\big(X \otimes (\mathfrak{o}_n^* / (1 + \mathfrak{p}_n^{r+1})), \mathbb{C}^*\big) \\ &= \operatorname{Hom}_{\Phi_\sigma,\mathrm{Id}}\big({}^0 T^{\Phi_\sigma^n} / T_{r+1}^{\Phi_\sigma^n}, \mathbb{C}^*\big) \\ &= \operatorname{Hom}\big({}^0 T^{\Phi_\sigma} / T_{r+1}^{\Phi_\sigma}, \mathbb{C}^*\big). \end{aligned} \tag{5}$$

Under this composite isomorphism, $s := \phi|_{I_F}$ maps to a character

$$\chi_s \in \mathrm{Hom}({}^0T^{\Phi_\sigma}/T^{\Phi_\sigma}_{r+1}, \mathbb{C}^*).$$

Then, if $\phi(\Phi) = \hat{\sigma} \ltimes \tau$, we get that τ gives rise to a character of X^σ given by $\chi_\tau(\lambda) := \lambda(\tau)$ for $\lambda \in X^\sigma$, just as in the depth-zero case. Recalling that $T^{\Phi_\sigma} = {}^0T^{\Phi_\sigma} \times X^\sigma$, we define a character χ_ϕ of T^{Φ_σ} by $\chi_\phi := \chi_s \otimes \chi_\tau$, which is our desired character of T^{Φ_σ} constructed from the Langlands parameter ϕ.

As in the depth-zero case, we have the set X_w. To $\lambda \in X_w$, Reeder associates a 1-cocycle u_λ, hence a twisted Frobenius $\Phi_\lambda = \mathrm{Ad}(u_\lambda) \circ \Phi$. Moreover, to λ is associated an affine Weyl group element w_λ, a parahoric subgroup G_{x_λ}, and an element $p_\lambda \in G_{x_\lambda}$ such that $p_\lambda^{-1}\Phi_\lambda(p_\lambda)$ is a lift of w_λ. We then define $T_\lambda := \mathrm{Ad}(p_\lambda)\,T$ and set $\chi_\lambda := \chi_\phi \circ \mathrm{Ad}(p_\lambda)^{-1}$. To the torus T_λ and the character χ_λ, we apply the construction of [Adler 1998] to obtain a supercuspidal representation. Then, Reeder constructs a packet $\Pi(\phi)$ of representations on the pure inner forms of G, parametrized by $\mathrm{Irr}(C_\phi)$, using the above construction.

7B. ***The case of* GL(ℓ, F).** We now consider the group $\mathbf{G}(F) = \mathrm{GL}(\ell, F)$, for ℓ prime. Let $\phi : W_F \to {}^LG$ be one of the Langlands parameters for $\mathbf{G}(F) = \mathrm{GL}(\ell, F)$ that is considered in Section 7A.

Lemma 7.1. $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$ *for some admissible pair* $(E/F, \chi)$, *where* χ *has positive level and* E/F *is of degree* ℓ *and unramified.*

Proof. The proof is similar as in the GL(2, F) case, but we include it for completeness purposes. As in the depth-zero case in Section 5, we may conjugate ϕ by an element of $\hat{G}$ so that the Weyl group element w that is associated to ϕ is the Weyl group element $(1\ 2\ 3\ \cdots\ \ell)$ in the symmetric group on ℓ letters. We know that ϕ is an irreducible admissible $\phi : W_F \to \mathrm{GL}(\ell, \mathbb{C})$ that is trivial on $I^{(r+1)}$ and nontrivial on $I^{(r)}$ for some integer $r > 0$. Let E be the degree ℓ unramified extension of F. Again, any representation $\mathrm{Ind}_{W_E}^{W_F}(\Omega)$ where $(E/F, \Omega)$ is an admissible pair is equivalent to the representation $\kappa : W_F \to \mathrm{GL}(\ell, \mathbb{C})$ satisfying:

(1) $\kappa|_{W_E}$ is given by $\Omega \in \widehat{E^*}$ by the local Langlands correspondence for tori.

(2)
$$\kappa(\Phi) = \begin{pmatrix} 0 & 0 & 0 & \dots & \Omega(\varpi) \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

We want to show that ϕ satisfies the two conditions above, for some admissible pair $(E/F, \chi)$. Let's restrict ϕ to W_E. By the composite isomorphism (5), $\phi|_{I_E}$ gives rise to a character $\ddot{\chi}$ of $\mathfrak{o}_E^*$. Then, by following the composite isomorphism (5) backwards, one sees that

$$\phi(x) = \begin{pmatrix} \ddot{\chi}(r_\ell(x)) & 0 & 0 & \dots & 0 \\ 0 & \ddot{\chi}^{\xi}(r_\ell(x)) & 0 & \dots & 0 \\ 0 & 0 & \ddot{\chi}^{\xi^2}(r_\ell(x)) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \ddot{\chi}^{\xi^{\ell-1}}(r_\ell(x)) \end{pmatrix}$$

as in the depth-zero case. Now, as in Propositions 5.4 and 5.5, we know that

$$\phi(\Phi) = \begin{pmatrix} 0 & 0 & 0 & \dots & a \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

for some $a \in \mathbb{C}^*$, because of the ellipticity condition on ϕ. Therefore, we have that

$$\phi(\Phi_E) = \phi(\Phi^\ell) = \phi(\Phi)^\ell = \begin{pmatrix} a & 0 & 0 & \dots & 0 \\ 0 & a & 0 & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a \end{pmatrix}.$$

Then $\ddot{\chi}$ extends to a character, denoted χ, of E^*, by setting

$$\chi(\varpi) := a \quad \text{and} \quad \chi|_{\mathfrak{o}_E^*} := \ddot{\chi}|_{\mathfrak{o}_E^*}.$$

One can now see that $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$. By the regularity condition on ϕ, we get that $\ddot{\chi} \neq \ddot{\chi}^{\xi}$, and thus $(E/F, \chi)$ is an admissible pair. Finally, χ has positive level since $r > 0$. □

Proposition 7.2. *Let $\ell = 2$. Then $\chi_\phi = \chi\Delta_\chi$.*

Proof. The analogous arguments as in the depth-zero case show that $\chi_\phi|_{\mathfrak{o}_E^*} = \chi|_{\mathfrak{o}_E^*}$. In particular, let $z \in \mathfrak{o}_E^*$. Let $x \in I_F$ be any element such that $r_2(x) = z$ (where r_2 is as in [R, Section 5.1]), and let Γ be the cocharacter $t \mapsto \left(\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Then

$$N_\sigma\big(\Gamma \otimes r_2(x)\big) = \begin{pmatrix} z & 0 \\ 0 & \overline{z} \end{pmatrix}.$$

Moreover, by the same arguments as in Proposition 5.1, we get

$$\phi(x) = \begin{pmatrix} \chi\,(r_2(x)) & 0 \\ 0 & \chi\big(\overline{r_2(x)}\big) \end{pmatrix},$$

so that $\Gamma(\phi(x)) = \chi(z)$, where here we are viewing Γ as a character of $\hat{T}$. Finally, as we may take τ to be the same element as in the depth-zero case, we have that $\chi_\phi(\varpi) = -\chi(\varpi)$, so that $\chi_\phi = \chi\Delta_\chi$. □

Proposition 7.3. *Let ℓ be an odd prime. Then*

$$\chi_\phi = \chi\Delta_\chi.$$

Proof. A reasoning analogous to that of Proposition 7.2 and the depth-zero case works here. □

Note that $[X/(1-w\theta)X]_{\mathrm{tor}} = 0$, so we may let $\lambda = 0$ (recall that $\lambda \in X_w$). It is easy to see that we may again take $u_\lambda = 1$, and therefore $\Phi_\lambda = \Phi$. It is also easy to see that we may take $w_\lambda = w$ (see [R, Section 6.4]), and we may also take the same p_λ as in the depth-zero case in Section 6. So we have the same T_λ as in Section 6 and the analogous χ_λ.

We have therefore shown that if we assume the compatibility condition in the beginning of Section 7, then by Proposition 7.3, the Reeder construction attaches the representation $\pi_{\chi\Delta_\chi}$ to the Langlands parameter $\phi = \mathrm{Ind}_{W_E}^{W_F}(\chi)$. This shows that as long as we assume this compatibility condition, the correspondences of [R] and [Moy 1986] agree for GL(ℓ, F), where ℓ is an odd prime.

Acknowledgments

This paper has benefited from conversations with Jeffrey Adams, Jeffrey Adler, Gordan Savin, and Jiu-Kang Yu.

References

[Adler 1998] J. D. Adler, "Refined anisotropic K-types and supercuspidal representations", *Pacific J. Math.* **185**:1 (1998), 1–32. MR 2000f:22019 Zbl 0924.22015

[Bushnell and Henniart 2005] C. J. Bushnell and G. Henniart, "The essentially tame local Langlands correspondence. I", *J. Amer. Math. Soc.* **18**:3 (2005), 685–710. MR 2006a:22014 Zbl 1073.11070

[DeBacker and Reeder 2009] S. DeBacker and M. Reeder, "Depth-zero supercuspidal L-packets and their stability", *Ann. of Math.* (2) **169**:3 (2009), 795–901. MR 2010d:22023 Zbl 1193.11111

[Howe 1977] R. E. Howe, "Tamely ramified supercuspidal representations of Gl_n", *Pacific J. Math.* **73**:2 (1977), 437–460. MR 58 #11241 Zbl 0404.22019

[Moy 1986] A. Moy, "Local constants and the tame Langlands correspondence", *Amer. J. Math.* **108**:4 (1986), 863–930. MR 88b:11081 Zbl 0597.12019

[Reeder 2008] M. Reeder, "Supercuspidal L-packets of positive depth and twisted Coxeter elements", *J. Reine Angew. Math.* **620** (2008), 1–33. MR 2009e:22019 Zbl 1153.22021

[Tate 1979] J. Tate, "Number theoretic background", pp. 3–26 in *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), Part 2, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, R.I., 1979. MR 80m:12009 Zbl 0422.12007

Received November 10, 2010. Revised February 15, 2012.

MOSHE ADRIAN
MATHEMATICS DEPARTMENT
UNIVERSITY OF UTAH
155 S 1400 E, ROOM 233
SALT LAKE CITY, UT 84112-0090
UNITED STATES
madrian@math.utah.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

R-GROUPS AND PARAMETERS

DUBRAVKA BAN AND DAVID GOLDBERG

Let G be a p-adic group, SO_{2n+1}, Sp_{2n}, O_{2n} or U_n. Let π be an irreducible discrete series representation of a Levi subgroup of G. We prove the conjecture that the Knapp–Stein R-group of π and the Arthur R-group of π are isomorphic. Several instances of the conjecture were established earlier: for archimedean groups by Shelstad; for principal series representations by Keys; for $G = \mathrm{SO}_{2n+1}$ by Ban and Zhang; and for $G = \mathrm{SO}_n$ or Sp_{2n} in the case when π is supercuspidal, under an assumption on the parameter, by Goldberg.

1. Introduction

Central to representation theory of reductive groups over local fields is the study of parabolically induced representations. In order to classify the tempered spectrum of such a group, one must understand the structure of parabolically induced from discrete series representations, in terms of components, multiplicities, and whether or not components are elliptic. The Knapp–Stein R-group gives an explicit combinatorial method for conducting this study. On the other hand, the local Langlands conjecture predicts the parametrization of such nondiscrete tempered representations, in L-packets, by admissible homomorphisms of the Weil–Deligne group which factor through a Levi component of the Langlands dual group. Arthur [1989] gave a conjectural description of the Knapp–Stein R-group in terms of the parameter. This conjecture generalizes results of Shelstad [1982] for archimedean groups, as well as those of Keys [1987] in the case of unitary principal series of certain p-adic groups. In [Ban and Zhang 2005] this conjecture was established for odd special orthogonal groups. In [Goldberg 2011] the conjecture was established for induced from supercuspidal representations of split special orthogonal or symplectic groups, under an assumption on the parameter. In the current work, we complete the conjecture for the full tempered spectrum of all these groups.

Dubravka Ban was supported in part by NSF grant DMS-0601005 and by a Research Fellowship of the Alexander von Humboldt Foundation.
MSC2010: 11F70, 22E35, 22E50.
Keywords: R-groups, reducibility of induced representations, classical groups.

Let F be a nonarchimedean local field of characteristic zero. We denote by $\boldsymbol{G}$ a connected reductive quasi-split algebraic group defined over F. We let $G = \boldsymbol{G}(F)$, and use similar notation for other groups defined over F. Fix a maximal torus $\boldsymbol{T}$ of $\boldsymbol{G}$, and a Borel subgroup $\boldsymbol{B} = \boldsymbol{TU}$ containing $\boldsymbol{T}$. We let $\mathscr{E}(G)$ be the equivalence classes of irreducible admissible representations of G, $\mathscr{E}_t(G)$ the tempered classes, $\mathscr{E}_2(G)$ the discrete series, and ${}^\circ\mathscr{E}(G)$ the irreducible unitary supercuspidal classes. We make no distinction between a representation π and its equivalence class.

Let $\boldsymbol{P} = \boldsymbol{MN}$ be a standard, with respect to $\boldsymbol{B}$, parabolic subgroup of $\boldsymbol{G}$. Let $\boldsymbol{A} = \boldsymbol{A_M}$ be the split component of $\boldsymbol{M}$, and let $W = W(\boldsymbol{G}, \boldsymbol{A}) = N_{\boldsymbol{G}}(\boldsymbol{A})/\boldsymbol{M}$ be the Weyl group for this situation. For $\sigma \in \mathscr{E}(M)$ we let $\operatorname{Ind}_P^G(\sigma)$ be the representation unitarily induced from $\sigma \otimes \mathbf{1}_N$. Thus, if V is the space of σ, we let

$$V(\sigma) = \left\{ f \in C^\infty(G, V) \mid f(mng) = \delta_P(m)^{1/2} f(g) \text{ for all } m \in M,\ n \in N,\ g \in G \right\},$$

with δ_P the modulus character of P. The action of G is by the right regular representation, so $(\operatorname{Ind}_P^G(\sigma)(x) f)(g) = f(gx)$. Then any $\pi \in \mathscr{E}_t(G)$ is an irreducible component of $\operatorname{Ind}_P^G(\sigma)$ for some choice of M and $\sigma \in \mathscr{E}_2(M)$. In order to determine the component structure of $\operatorname{Ind}_P^G(\sigma)$, Knapp and Stein, in the archimedean case, and Harish-Chandra in the p-adic case, developed the theory of singular integral intertwining operators, leading to the theory of R-groups, due to Knapp and Stein [1971] in the archimedean case and Silberger [1978; 1979] in the p-adic case. We describe this briefly and refer the reader to the introduction of [Goldberg 1994] for more details. The poles of the intertwining operators give rise to the zeros of Plancherel measures. Let $\Phi(\boldsymbol{P}, \boldsymbol{A})$ be the reduced roots of $\boldsymbol{A}$ in $\boldsymbol{P}$. For $\alpha \in \Phi(\boldsymbol{P}, \boldsymbol{A})$ and $\sigma \in \mathscr{E}_2(M)$ we let $\mu_\alpha(\sigma)$ be the rank one Plancherel measure associated to σ and α. We let $\Delta' = \{\alpha \in \Phi(\boldsymbol{P}, \boldsymbol{A}) \mid \mu_\alpha(\sigma) = 0\}$. For $w \in W$ and $\sigma \in \mathscr{E}_2(M)$ we let $w\sigma(m) = \sigma(w^{-1} m \sigma)$. (Note, we make no distinction between $w \in W$ and its representative in $N_G(A)$.) We let

$$W(\sigma) = \{w \in W \mid w\sigma \simeq \sigma\},$$

and let W' be the subgroup of $W(\sigma)$ generated by those w_α with $\alpha \in \Delta'$. We let $R(\sigma) = \{w \in W(\sigma) \mid w\Delta' = \Delta'\} = \{w \in W(\sigma) \mid w\alpha > 0 \text{ for all } \alpha \in \Delta'\}$. Let $\mathscr{C}(\sigma) = \operatorname{End}_G(\operatorname{Ind}_P^G(\sigma))$.

Theorem 1 [Knapp and Stein 1971; Silberger 1978; 1979]. *For any $\sigma \in \mathscr{E}_2(M)$, we have $W(\sigma) = R(\sigma) \ltimes W'$, and $\mathscr{C}(\sigma) \simeq \mathbb{C}[R(\sigma)]_\eta$, the group algebra of $R(\sigma)$ twisted by a certain 2-cocycle η.*

Thus $R(\sigma)$, along with η, determines how many inequivalent components appear in $\operatorname{Ind}_P^G(\sigma)$ and the multiplicity with which each one appears. Furthermore Arthur shows $\mathbb{C}[R(\sigma)]_\eta$ also determines whether or not components of $\operatorname{Ind}_P^G(\sigma)$ are elliptic (and hence whether or not they contribute to the Plancherel formula) [Arthur 1993].

Arthur [1989] conjectured a construction of $R(\sigma)$ in terms of the local Langlands conjecture. Let W_F be the Weil group of F and $W_F' = W_F \times \mathrm{SL}_2(\mathbb{C})$ the Weil–Deligne group. Suppose $\psi : W_F' \to {}^L M$ parametrizes the L-packet, $\Pi_\psi(M)$, of M containing σ. Here ${}^L M = \hat{M} \rtimes W_F$ is the Langlands L-group, and $\hat{M}$ is the complex group whose root datum is dual to that of $\boldsymbol{M}$. Then

$$\psi : W_F' \to {}^L M \hookrightarrow {}^L G$$

must be a parameter for an L-packet $\Pi_\psi(G)$ of G. The expectation is that $\Pi_\psi(G)$ consists of all irreducible components of $\mathrm{Ind}_P^G(\sigma')$ for all $\sigma' \in \Pi_\psi(M)$. We let $S_\psi = Z_{\hat{G}}(\mathrm{Im}\,\psi)$, and take S_ψ° to be the connected component of the identity. Let T_ψ be a maximal torus in S_ψ°. Set $W_\psi = W(S_\psi, T_\psi)$, and $W_\psi^\circ = W(S_\psi^\circ, T_\psi)$. Then $R_\psi = W_\psi / W_\psi^\circ$ is called the R-group of the packet $\Pi_\psi(G)$. By duality we can identify W_ψ with a subgroup of W. With this identification, we let $W_{\psi,\sigma} = W_\psi \cap W(\sigma)$ and $W_{\psi,\sigma}^\circ = W_\psi^\circ \cap W(\sigma)$. We then set

$$R_{\psi,\sigma} = W_{\psi,\sigma} / W_{\psi,\sigma}^\circ.$$

We call $R_{\psi,\sigma}$ the Arthur R-group attached to ψ and σ.

Conjecture 2. *For any $\sigma \in \mathscr{E}_2(M)$, we have $R(\sigma) \simeq R_{\psi,\sigma}$.*

In [Ban and Zhang 2005], the first named author and Zhang proved this conjecture in the case $\boldsymbol{G} = \mathrm{SO}_{2n+1}$. In [Goldberg 2011] the second named author confirmed the conjecture when σ is supercuspidal, and $\boldsymbol{G} = \mathrm{SO}_n$ *or* Sp_{2n}, with a mild assumption on the parameter ψ. Here, we complete the proof of the conjecture for Sp_{2n}, or O_n, under assumptions given in Section 2.3.

This work is based on the classification of discrete series for classical p-adic groups of Mœglin and Tadić [2002], and on the results of Mœglin [2002; 2007b]. Subsequent to our submission, Arthur's unfinished book has become available in preprint form [Arthur 2011]. In this long awaited and impressive work, he uses the trace formula to classify the automorphic representations of special orthogonal and symplectic groups in terms of those of $\mathrm{GL}(n)$. An important ingredient in this work is a formulation of the classification at the local places. The results for irreducible tempered representations are obtained from the classification of discrete series using R-groups. Our result on isomorphism of R-groups and their dual version for $\mathrm{SO}(2n+1, F)$ and $\mathrm{Sp}(2n, F)$ (see Theorem 7) also appear in Arthur's work [2011, page 346]. Arthur's proof differs significantly from the one we use here. We work with a rather concrete description of parameters based on Jordan blocks and L-functions, while Arthur works in the general context of his theory.

We now describe the contents of the paper in more detail. In Section 2 we introduce our notation and discuss the classification of $\mathscr{E}_2(M)$ for our groups, due to Mœglin and Tadić, as well as preliminaries on Knapp–Stein and Arthur R-groups.

In Section 3 we consider the parameters ψ and compute their centralizers. In Section 4 we turn to the case of $\boldsymbol{G} = O_{2n}$. Here we show the Arthur R-group agrees with the generalization of the Knapp–Stein R-group as discussed in [Goldberg and Herb 1997]. In Section 5 we complete the proof of the theorem for the induced from discrete series representations of Sp_{2n}, SO_{2n+1}, or O_{2n}.

In Section 6, we study R-groups for unitary groups. These groups are interesting for us because they are not split and the action of the Weil group on the dual group is nontrivial. In addition, the classification of discrete series and description of L-parameters is completed [Mœglin 2007b].

The techniques used here can be used for other groups. In particular we should be able to carry out this process for similitude groups and G_2. Furthermore, the techniques of computing the Arthur R-groups will apply to $GSpin$ groups, as well, and may shed light on the Knapp–Stein R-groups in this case. We leave all of this for future work.

2. Preliminaries

2.1. ***Notation.*** Let F be a nonarchimedean local field of characteristic zero. Let G_n, $n \in \mathbb{Z}^+$, be $\mathrm{Sp}(2n, F)$, $\mathrm{SO}(2n+1, F)$ or $\mathrm{SO}(2n, F)$. We define G_0 to be the trivial group. For $G = G_n$ or $G = \mathrm{GL}(n, F)$, fix the minimal parabolic subgroup consisting of all upper triangular matrices in G and the maximal torus consisting of all diagonal matrices in G. If δ_1, δ_2 are smooth representations of $\mathrm{GL}(m, F)$, $\mathrm{GL}(n, F)$, respectively, we define

$$\delta_1 \times \delta_2 = \mathrm{Ind}_P^G(\delta_1 \otimes \delta_2)$$

where $G = \mathrm{GL}(m+n, F)$ and $P = MU$ is the standard parabolic subgroup of G with Levi factor $M \cong \mathrm{GL}(m, F) \times \mathrm{GL}(n, F)$. Similarly, if δ is a smooth representation of $\mathrm{GL}(m, F)$ and σ is a smooth representation of G_n, we define

$$\delta \rtimes \sigma = \mathrm{Ind}_P^{G_{m+n}}(\delta \otimes \sigma)$$

where $P = MU$ is the standard parabolic subgroup of G_{m+n} with Levi factor $M \cong \mathrm{GL}(m, F) \times G_n$. We denote by $\mathscr{E}_2(G)$ the set of equivalence classes of irreducible square integrable representations of G and by ${}^0\mathscr{E}(G)$ the set of equivalence classes of irreducible unitary supercuspidal representations of G.

We say that a homomorphism $h : X \to \mathrm{GL}(d, \mathbb{C})$ is symplectic (respectively, orthogonal) if h fixes an alternating form (respectively, a symmetric form) on $\mathrm{GL}(d, \mathbb{C})$. We denote by S_a the standard a-dimensional irreducible algebraic representation of $\mathrm{SL}(2, \mathbb{C})$. Then

$$S_a \text{ is } \begin{cases} \text{orthogonal} & \text{for } a \text{ odd}, \\ \text{symplectic} & \text{for } a \text{ even}. \end{cases} \tag{1}$$

Let ρ be an irreducible supercuspidal unitary representation of $\mathrm{GL}(d, F)$. According to the local Langlands correspondence for GL_d [Harris and Taylor 2001; Henniart 2000], attached to ρ is an L-parameter $\varphi : W_F \to \mathrm{GL}(d, \mathbb{C})$. Suppose $\rho \cong \tilde{\rho}$. Then $\varphi \cong \tilde{\varphi}$ and one of the Artin L-functions $L(s, \mathrm{Sym}^2\varphi)$ or $L(s, \bigwedge^2\varphi)$ has a pole. The L-function $L(s, \mathrm{Sym}^2\varphi)$ has a pole if and only if φ is orthogonal. The L-function $L(s, \bigwedge^2\varphi)$ has a pole if and only if φ is symplectic. From [Henniart 2010] we know

$$L(s, \textstyle\bigwedge^2\varphi) = L(s, \rho, \bigwedge^2), \text{ and } L(s, \mathrm{Sym}^2\varphi) = \mathrm{L}(\mathrm{s}, \rho, \mathrm{Sym}^2), \tag{2}$$

where $L(s, \rho, \bigwedge^2)$ and $L(s, \rho, \mathrm{Sym}^2)$ are the Langlands L-functions as defined in [Shahidi 1981].

Let ρ be an irreducible supercuspidal unitary representation of $\mathrm{GL}(d, F)$ and $a \in \mathbb{Z}^+$. We define $\delta(\rho, a)$ to be the unique irreducible subrepresentation of

$$\rho\|^{(a-1)/2} \times \rho\|^{(a-3)/2} \times \cdots \times \rho\|^{(-(a-1))/2};$$

see [Zelevinsky 1980].

2.2. *Jordan blocks.* We now review the definition of Jordan blocks from [Mœglin and Tadić 2002]. Let G be $\mathrm{Sp}(2n, F)$, $\mathrm{SO}(2n+1, F)$ or $O(2n, F)$. For $d \in \mathbb{N}$, let r_d denote the standard representation of $\mathrm{GL}(d, \mathbb{C})$. Define

$$R_d = \begin{cases} \bigwedge^2 r_d & \text{for } G = \mathrm{Sp}(2n, F),\ O(2n, F), \\ \mathrm{Sym}^2 r_d & \text{for } G = \mathrm{SO}(2n+1, F). \end{cases}$$

Let σ be an irreducible discrete series representation of G_n. Denote by $\mathrm{Jord}(\sigma)$ the set of pairs (ρ, a), where $\rho \in {}^0\mathscr{E}(\mathrm{GL}(d_\rho, F))$, $\rho \cong \tilde{\rho}$, and $a \in \mathbb{Z}^+$, such that

(J-1) a is even if $L(s, \rho, R_{d_\rho})$ has a pole at $s = 0$ and odd otherwise,

(J-2) $\delta(\rho, a) \rtimes \sigma$ is irreducible.

For $\rho \in {}^0\mathscr{E}(\mathrm{GL}(d_\rho, F))$, $\rho \cong \tilde{\rho}$, define

$$\mathrm{Jord}_\rho(\sigma) = \{a \mid (\rho, a) \in \mathrm{Jord}(\sigma)\}.$$

Let $\hat{G}$ denote the complex dual group of G. Then $\hat{G} = \mathrm{SO}(2n+1, \mathbb{C})$ for $G = \mathrm{Sp}(2n, F)$, $\hat{G} = \mathrm{Sp}(2n, \mathbb{C})$ for $G = \mathrm{SO}(2n+1, F)$ and $\hat{G} = O(2n, \mathbb{C})$ for $G = O(2n, F)$.

Lemma 3. *Let σ be an irreducible discrete series representation of G_n. Let ρ be an irreducible supercuspidal self-dual representation of* $\mathrm{GL}(d_\rho, F)$ *and $a \in \mathbb{Z}^+$. Then $(\rho, a) \in \mathrm{Jord}(\sigma)$ if and only if the following conditions hold*:

(J-1′) *$\rho \otimes S_a$ is of the same type as $\hat{G}$,*

(J-2) *$\delta(\rho, a) \rtimes \sigma$ is irreducible.*

Proof. We will prove that (J-1) $\Leftrightarrow$ (J-1$'$). We know from [Shahidi 1990] that one and only one of the two L-functions $L(s, \rho, \bigwedge^2)$ and $L(s, \rho, \mathrm{Sym}^2)$ has a pole at $s = 0$. Suppose $G = \mathrm{Sp}(2n, F)$ or $O(2n, F)$. We consider $L(s, \rho, \bigwedge^2)$. It has a pole at $s = 0$ if and only if the parameter $\rho : W_F \to \mathrm{GL}(d_\rho, \mathbb{C})$ is symplectic. According to (1), this is equivalent to $\rho \otimes S_a$ being orthogonal for a even. Therefore, for $(\rho, a) \in \mathrm{Jord}(\sigma)$, a is even if and only if $\rho \otimes S_a$ is orthogonal. For $G = \mathrm{SO}(2n+1, F)$, the arguments are similar. $\square$

2.3. ***Assumptions.*** In this paper, we use the classification of discrete series for classical p-adic groups of Mœglin and Tadić [Mœglin and Tadić 2002], so we have to make the same assumptions as there. Let σ be an irreducible supercuspidal representation of G_n and let ρ be an irreducible self-dual supercuspidal representation of a general linear group. We make the following assumption:

(BA) $\nu^{\pm(a+1)/2}\rho \rtimes \sigma$ *reduces for*

$$a = \begin{cases} \max \mathrm{Jord}_\rho(\sigma) & \text{if } \mathrm{Jord}_\rho(\sigma) \neq \varnothing, \\ 0 & \text{if } L(s, \rho, R_{d_\rho}) \text{ has a pole at } s = 0 \text{ and } \mathrm{Jord}_\rho(\sigma) = \varnothing, \\ -1 & \text{otherwise.} \end{cases}$$

Moreover, there are no other reducibility points in $\mathbb{R}$.

In addition, we assume that the L-parameter of σ is given by

$$\varphi_\sigma = \bigoplus_{(\rho,a)\in \mathrm{Jord}(\sigma)} \varphi_\rho \otimes S_a. \tag{3}$$

Here, φ_ρ denotes the L-parameter of ρ given in [Harris and Taylor 2001; Henniart 2000].

Mœglin [2007a], assuming certain Fundamental Lemmas, proved the validity of the assumptions for $\mathrm{SO}(2n+1, F)$ and showed how Arthur's results imply the Langlands classification of discrete series for $\mathrm{SO}(2n+1, F)$.

2.4. ***The Arthur R-group.*** Let ${}^LG = \hat{G} \rtimes W_F$ be the L-group of G, and suppose ${}^LM$ is the L-group of a Levi subgroup, M, of G. Then ${}^LM$ is a Levi subgroup of ${}^LG$ (see [Borel 1979, Section 3] for the definition of parabolic subgroups and Levi subgroups of ${}^LG$). Suppose ψ is an A-parameter of G which factors through ${}^LM$,

$$\psi : W_F \times \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \longrightarrow {}^LM \subset {}^LG.$$

Then we can regard ψ as an A-parameter of M. Suppose, in addition, the image of ψ is not contained in a smaller Levi subgroup (i.e., ψ is an elliptic parameter of M).

Let S_ψ be the centralizer in $\hat{G}$ of the image of ψ and S_ψ^0 its identity component. If T_ψ is a maximal torus of S_ψ^0, define

$$W_\psi = N_{S_\psi}(T_\psi)/Z_{S_\psi}(T_\psi), \quad W_\psi^0 = N_{S_\psi^0}(T_\psi)/Z_{S_\psi^0}(T_\psi), \quad R_\psi = W_\psi/W_\psi^0.$$

Lemma 2.3 of [Ban and Zhang 2005] and the discussion on page 326 of [Ban and Zhang 2005] imply that W_ψ can be identified with a subgroup of $W(G, A)$.

Let σ be an irreducible unitary representation of M. Assume σ belongs to the A-packet $\Pi_\psi(M)$. If $W(\sigma) = \{w \in W(G, A) \mid w\sigma \cong \sigma\}$, we let

$$W_{\psi,\sigma} = W_\psi \cap W(\sigma), \quad W_{\psi,\sigma}^0 = W_\psi^0 \cap W(\sigma),$$

and take $R_{\psi,\sigma} = W_{\psi,\sigma}/W_{\psi,\sigma}^0$ as the Arthur R-group.

3. Centralizers

Let G be $\mathrm{Sp}(2n, F)$, $\mathrm{SO}(2n+1, F)$ or $O(2n, F)$. Let $\hat{G}$ be the complex dual group of G. Let

$$\psi : W_F \times \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \longrightarrow \hat{G} \subset \mathrm{GL}(N, \mathbb{C})$$

be an A-parameter. We consider ψ as a representation. Then ψ is a direct sum of irreducible subrepresentations. Let ψ_0 be an irreducible subrepresentation. For $m \in \mathbb{N}$, set

$$m\psi_0 = \underbrace{\psi_0 \oplus \cdots \oplus \psi_0}_{m \text{ times}}.$$

If $\psi_0 \ncong \tilde{\psi}_0$, then it can be shown using the bilinear form on $\hat{G}$ that $\tilde{\psi}_0$ is also a subrepresentation of ψ. Therefore, ψ decomposes into a sum of irreducible subrepresentations

$$\psi = (m_1\psi_1 \oplus m_1\tilde{\psi}_1) \oplus \cdots \oplus (m_k\psi_k \oplus m_k\tilde{\psi}_k) \oplus m_{k+1}\psi_{k+1} \oplus \cdots \oplus m_l\psi_l,$$

where $\psi_i \ncong \psi_j$, $\psi_i \ncong \tilde{\psi}_j$ for $i \neq j$. In addition, $\psi_i \ncong \tilde{\psi}_i$ for $i = 1, \ldots, k$ and $\psi_i \cong \tilde{\psi}_i$ for $i = k+1, \ldots, l$. If $\psi_i \cong \tilde{\psi}_i$, then ψ_i factors through a symplectic or orthogonal group. In this case, if ψ_i is not of the same type as $\hat{G}$, then m_i must be even. This follows again using the bilinear form on $\hat{G}$.

We want to compute S_ψ and W_ψ. First, we consider the case $\psi = m\psi_0$ or $\psi = m\psi_0 \oplus m\tilde{\psi}_0$, where ψ_0 is irreducible. The following lemma is an extension of Proposition 6.5 of [Gross and Prasad 1992]. A part of the proof was communicated to us by Joe Hundley.

Lemma 4. *Let G be* $\mathrm{Sp}(2n, F)$, $\mathrm{SO}(2n+1, F)$ *or* $O(2n, F)$. *Let*

$$\psi_0 : W_F \times \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \to \mathrm{GL}(d_0, \mathbb{C})$$

be an irreducible parameter.

(i) *Suppose* $\psi_0 \ncong \tilde{\psi}_0$ *and* $\psi = m\psi_0 \oplus m\tilde{\psi}_0$. *Then* $S_\psi \cong \mathrm{GL}(m, \mathbb{C})$ *and* $R_\psi = 1$.

(ii) *Suppose* $\psi_0 \cong \tilde{\psi}_0$ *and* $\psi = m\psi_0$. *Suppose* ψ_0 *is of the same type as* $\hat{G}$. *Then*

$$R_\psi \cong \begin{cases} \mathbb{Z}_2 & \textit{for } m \textit{ even}, \\ 1 & \textit{for } m \textit{ odd}. \end{cases}$$

(iii) *Suppose* $\psi_0 \cong \tilde{\psi}_0$ *and* $\psi = m\psi_0$. *Suppose* ψ_0 *is not of the same type as* $\hat{G}$. *Then* m *is even*, $S_\psi \cong \mathrm{Sp}(m, \mathbb{C})$ *and* $R_\psi = 1$.

Proof. (i) The proof of the statement is the same as in [Gross and Prasad 1992].

(ii) and (iii) Suppose $G = \mathrm{Sp}(2n, F)$ or $\mathrm{SO}(2n+1, F)$. Let V and V_0 denote the spaces of the representations ψ and ψ_0, respectively. Denote by $\langle\,,\rangle$ the ψ-invariant bilinear form on V and by $\langle\,,\rangle_0$ the ψ_0-invariant bilinear form on V_0. There exists an isomorphism $V \to V_0 \oplus \cdots \oplus V_0$. Equivalently, $V \cong W \otimes V_0$, where W is a finite dimensional vector space with trivial $W_F \times \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C})$-action. The space W can be identified with $\mathrm{Hom}_{W_F \times \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C})}(V_0, V)$. Then the map $W \otimes V_0 \to V$ is

$$l \otimes v \mapsto l(v), \quad l \in \mathrm{Hom}_{W_F \times \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C})}(V_0, V),\ v \in V_0.$$

We claim there exists a nondegenerate bilinear form $\langle\,,\rangle_W$ on W such that $\langle\,,\rangle = \langle\,,\rangle_W \otimes \langle\,,\rangle_0$ in the sense that

$$\langle l_1 \otimes v_1, l_2 \otimes v_2\rangle = \langle l_1, l_2\rangle_W \langle v_1, v_2\rangle_0 \quad \text{for all } l_1, l_2 \in W,\ v_1, v_2 \in V_0.$$

The key ingredient is Schur's lemma, or rather, the variant thereof stating that every invariant bilinear form on V_0 is a scalar multiple of $\langle\,,\rangle_0$. Given any l_1, l_2 in $\mathrm{Hom}_{W_F \times \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C})}(V_0, V)$,

$$\langle l_1(v_1), l_2(v_2)\rangle$$

is an invariant bilinear form on V_0 and therefore it is equal to $c\langle\,,\rangle_0$, for some constant c. We can define $\langle l_1, l_2\rangle_W$ by

$$\langle l_1, l_2\rangle_W = \frac{\langle l_1(v_1), l_2(v_2)\rangle}{\langle v_1, v_2\rangle_0}$$

because Schur's lemma tells us that the right-hand side is independent of v_1, v_2 in V_0. This proves the claim. Observe that if ψ_0 is not of the same type as ψ, the form $\langle\,,\rangle_W$ is alternating, while in the case when ψ_0 and ψ are of the same type, the form $\langle\,,\rangle_W$ is symmetric.

Now, $\mathrm{Im}\,\psi = \{I_m \otimes g \mid g \in \mathrm{Im}\,\psi_0\}$ and

$$\begin{aligned} Z_{\mathrm{GL}(N,\mathbb{C})}(\mathrm{Im}\,\psi) &= \{g \otimes z \mid g \in \mathrm{GL}(m, \mathbb{C}), z \in \{\lambda I_{d_0} \mid \lambda \in \mathbb{C}^\times\}\} \\ &= \{g \otimes I_{d_0} \mid g \in \mathrm{GL}(m, \mathbb{C})\}. \end{aligned}$$

Let us denote by $\mathcal{W}$ the group of matrices in $\mathrm{GL}(W)$ which preserve $\langle\, , \rangle_W$, i.e., $\mathcal{W} = \mathrm{Sp}(m, \mathbb{C})$ if $\langle\, , \rangle_W$ is an alternating form and $\mathcal{W} = O(m, \mathbb{C})$ if $\langle\, , \rangle_W$ is a symmetric form. Then

$$S_\psi = Z_{\mathrm{GL}(N,\mathbb{C})}(\mathrm{Im}\,\psi) \cap \hat{G} = \{g \otimes I_{d_0} \mid g \in \mathcal{W},\ \det(g \otimes I_{d_0}) = 1\}.$$

It follows that in case (iii) we have $S_\psi \cong \mathrm{Sp}(m, \mathbb{C})$, $S_\psi^0 = S_\psi$ and $R_\psi = 1$.

In case (ii), $\mathcal{W} = O(m, \mathbb{C})$. Since $\det(g \otimes I_{d_0}) = (\det g)^{d_0}$, it follows

$$S_\psi \cong \begin{cases} O(m, \mathbb{C}), & d_0 \text{ even}, \\ \mathrm{SO}(m, \mathbb{C}), & d_0 \text{ odd}. \end{cases}$$

In the case $G = \mathrm{SO}(2n+1, F)$, ψ_0 is symplectic and d_0 is even. Then $S_\psi \cong O(m, \mathbb{C})$ and $S_\psi^0 \cong \mathrm{SO}(m, \mathbb{C})$. If m is even, this implies $R_\psi \cong \mathbb{Z}_2$. For m odd, $W_\psi = W_\psi^0$ and $R_\psi = 1$.

In the case $G = \mathrm{Sp}(2n, F)$, we have $\hat{G} = \mathrm{SO}(2n+1, \mathbb{C})$ and $md_0 = 2n+1$. It follows that m and d_0 are both odd. Then $S_\psi \cong \mathrm{SO}(m, \mathbb{C})$, $S_\psi^0 = S_\psi$ and $R_\psi = 1$.

The case $G = O(2n, F)$ is similar, but simpler, because there is no condition on determinant. It follows that $S_\psi \cong O(m, \mathbb{C})$. This implies $R_\psi \cong \mathbb{Z}_2$ for m even and $R_\psi = 1$ for m odd. $\square$

Lemma 5. *Let G be $\mathrm{Sp}(2n, F)$, $\mathrm{SO}(2n+1, F)$ or $O(2n, F)$. Let*

$$\psi : W_F \times \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \to \hat{G}$$

be an A-parameter. We can write ψ in the form

$$\psi \cong \Big(\bigoplus_{i=1}^{p} (m_i\psi_i \oplus m_i\tilde{\psi}_i)\Big) \oplus \Big(\bigoplus_{i=p+1}^{q} 2m_i\psi_i\Big) \oplus \Big(\bigoplus_{i=q+1}^{r} (2m_i+1)\psi_i\Big) \oplus \Big(\bigoplus_{i=r+1}^{s} 2m_i\psi_i\Big), \tag{4}$$

where ψ_i is irreducible for $i \in \{1, \ldots, s\}$, and

$$\begin{array}{ll} \psi_i \not\cong \psi_j,\ \psi_i \not\cong \tilde{\psi}_j & \textit{for } i, j \in \{1, \ldots, s\},\ i \neq j, \\ \psi_i \not\cong \tilde{\psi}_i & \textit{for } i \in \{1, \ldots, p\}, \\ \psi_i \cong \tilde{\psi}_i & \textit{for } i \in \{p+1, \ldots, s\}, \\ \psi_i \textit{ not of the same type as } \hat{G} & \textit{for } i \in \{p+1, \ldots, q\}, \\ \psi_i \textit{ of the same type as } \hat{G} & \textit{for } i \in \{q+1, \ldots, s\}. \end{array}$$

Let $d = s - r$. Then

$$R_\psi \cong \mathbb{Z}_2^d.$$

Proof. Set $\Psi_i = m_i\psi_i \oplus m_i\tilde{\psi}_i$ for all $i \in \{1, \ldots, p\}$, and $\Psi_i = m_i\psi_i$ for all $i \in \{p+1, \ldots, s\}$. Denote by Z_i the centralizer of the image of Ψ_i in the corresponding GL. Then

$$Z_{\mathrm{GL}(N,\mathbb{C})}(\mathrm{Im}\,\psi) = Z_1 \times \cdots \times Z_s \quad \text{and} \quad S_\psi = Z_{\mathrm{GL}(N,\mathbb{C})}(\mathrm{Im}\,\psi) \cap \hat{G}.$$

Lemma 4 tells us the factors corresponding to $i \in \{1, \ldots, q\}$ do not contribute to R_ψ. In addition, we can see from the proof of Lemma 4 that these factors do not appear in determinant considerations. Therefore, we can consider only the factors corresponding to $i \in \{q+1, \ldots, s\}$. Let $\mathcal{Z} = Z_{q+1} \times \cdots \times Z_s$ and $\mathcal{S} = \mathcal{Z} \cap \hat{G}$. In the same way as in the proof of Lemma 4, we obtain

$$\text{(5)} \quad \mathcal{S} \cong \Big\{(g_{q+1}, \ldots, g_s) \mid g_i \in O(2m_i+1, \mathbb{C}),\ i \in \{q+1, \ldots, r\},\\ g_i \in O(2m_i, \mathbb{C}), i \in \{r+1, \ldots, s\},\ \prod_{i=q+1}^{s} (\det g_i)^{\dim\psi_i} = 1\Big\},$$

for $G = \mathrm{SO}(2n+1, F)$ or $\mathrm{Sp}(2n, F)$. For $G = O(2n, F)$, we omit the condition on determinant. If $G = \mathrm{SO}(2n+1, F)$, then for $i \in \{q+1, \ldots, s\}$, ψ_i is symplectic and $\dim\psi_i$ is even. Therefore, the product in (5) is always equal to 1.

Now, for $G = \mathrm{SO}(2n+1, F)$ and $G = O(2n, F)$, we have

$$\mathcal{S} \cong \prod_{i=q+1}^{r} O(2m_i+1, \mathbb{C}) \times \prod_{i=r+1}^{s} O(2m_i, \mathbb{C}).$$

It follows that $R_\psi \cong \prod_{i=q+1}^{r} 1 \times \prod_{i=r+1}^{s} \mathbb{Z}_2 \cong \mathbb{Z}_2^d$.

It remains to consider $G = \mathrm{Sp}(2n, F)$, $\hat{G} = \mathrm{SO}(2n+1, \mathbb{C})$. We have

$$\sum_{i=1}^{q} 2m_i \dim\psi_i + \sum_{i=q+1}^{r} (2m_i+1)\dim\psi_i + \sum_{i=1}^{p} 2m_i \dim\psi_i = 2n+1.$$

Since the total sum is odd, we must have $r > q$ and $\dim\psi_i$ odd, for some $i \in \{q+1, \ldots, r\}$. Without loss of generality, we may assume $\dim\psi_{q+1}$ odd. Then

$$\mathcal{S} \cong \mathrm{SO}(2m_{q+1}+1, \mathbb{C}) \times \prod_{i=q+2}^{r} O(2m_i+1, \mathbb{C}) \times \prod_{i=r+1}^{s} O(2m_i, \mathbb{C}).$$

It follows $R_\psi \cong 1 \times \prod_{i=q+2}^{r} 1 \times \prod_{i=r+1}^{s} \mathbb{Z}_2 \cong \mathbb{Z}_2^d$. □

4. Even orthogonal groups

4.1. ***R-groups for nonconnected groups.*** In this section, we review some results of [Goldberg and Herb 1997]. Let G be a reductive F-group. Let G^0 be the connected component of the identity in G. We assume that G/G^0 is finite and abelian.

Let π be an irreducible unitary representation of G. We say that π is discrete series if the matrix coefficients of π are square integrable modulo the center of G.

We will consider the parabolic subgroups and the R-groups as defined in [Goldberg and Herb 1997]. Let $P^0 = M^0 U$ be a parabolic subgroup of G^0. Let A be the split component in the center of M^0. Define $M = C_G(A)$ and $P = MU$. Then P is called the cuspidal parabolic subgroup of G lying over P^0. The Lie algebra $\mathscr{L}(G)$ can be decomposed into root spaces with respect to the roots Φ of $\mathscr{L}(A)$,

$$\mathscr{L}(G) = \mathscr{L}(M) \oplus \sum_{\alpha \in \Phi} \mathscr{L}(G)_\alpha .$$

Let σ be an irreducible unitary representation of M. We denote by $r_{M^0,M}(\sigma)$ the restriction of σ to M^0. Then, by Lemma 2.21 of [Goldberg and Herb 1997], σ is discrete series if and only if any irreducible constituent of $r_{M^0,M}(\sigma)$ is discrete series. Now, suppose σ is discrete series. Let σ_0 be an irreducible constituent of $r_{M^0,M}(\sigma)$. Then σ_0 is discrete series and we have the Knapp–Stein R-group $R(\sigma_0)$ for $i_{G^0,M^0}(\sigma_0)$ [Knapp and Stein 1971; Silberger 1978]. We review the definition of $R(\sigma_0)$. Let $W(G^0, A) = N_{G^0}(A)/M^0$ and $W_{G^0}(\sigma_0) = \{w \in W_G(M) \mid w\sigma_0 \cong \sigma_0\}$. For $w \in W_{G^0}(\sigma_0)$, we denote by $\mathscr{A}(w, \sigma_0)$ the normalized standard intertwining operator associated to w (see [Silberger 1979]). Define

$$W^0_{G^0}(\sigma_0) = \{w \in W_{G^0}(\sigma_0) \mid \mathscr{A}(w, \sigma_0) \text{ is a scalar}\}.$$

Then $W^0_{G^0}(\sigma_0) = W(\Phi_1)$ is generated by reflections in a set Φ_1 of reduced roots of (G, A). Let Φ^+ be the positive system of reduced roots of (G, A) determined by P and let $\Phi_1^+ = \Phi_1 \cap \Phi^+$. Then

$$R(\sigma_0) = \{w \in W_{G^0}(\sigma_0) \mid w\beta \in \Phi^+ \text{ for all } \beta \in \Phi_1^+\}$$

and $W_{G^0}(\sigma_0) = R(\sigma_0) \ltimes W(\Phi_1)$.

For the definition of $R(\sigma)$, we follow [Goldberg and Herb 1997]. Define

$$\begin{aligned} N_G(\sigma) &= \{g \in N_G(M) \mid g\sigma \cong \sigma\}, \\ W_G(\sigma) &= N_G(\sigma)/M, \quad \text{and} \\ R(\sigma) &= \{w \in W_G(\sigma) \mid w\beta \in \Phi^+ \text{ for all } \beta \in \Phi_1^+\}. \end{aligned}$$

For $w \in W_G(\sigma)$, let $\mathscr{A}(w, \sigma)$ denote the intertwining operator on $i_{G,M}(\sigma)$ defined in [Goldberg and Herb 1997, page 135]. Then the $\mathscr{A}(w, \sigma)$, $w \in R(\sigma)$, form a basis for the algebra of intertwining operators on $i_{G,M}(\sigma)$, by Theorem 5.16 of [Goldberg and Herb 1997]. In addition, $W_G(\sigma) = R(\sigma) \ltimes W(\Phi_1)$. For $w \in W_G(\sigma)$, $\mathscr{A}(w, \sigma)$ is a scalar if and only if $w \in W(\Phi_1)$; see [Goldberg and Herb 1997, Lemma 5.20].

4.2. ***Even orthogonal groups.*** Let $G = O(2n, F)$ and $G^0 = \mathrm{SO}(2n, F)$. Then $G = G^0 \rtimes \{1, s\}$, where $s = \mathrm{diag}\big(I_{n-1}, \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), I_{n-1}\big)$ and it acts on G^0 by conjugation.

(a) Let

$$\begin{aligned} M^0 &= \{\operatorname{diag}(g_1, \dots, g_r, h, {}^\tau g_r^{-1}, \dots, {}^\tau g_1^{-1}) \mid g_i \in \mathrm{GL}(n_i, F), h \in \mathrm{SO}(2m, F)\} \\ &\cong \mathrm{GL}(n_1, F) \times \cdots \times \mathrm{GL}(n_r, F) \times \mathrm{SO}(2m, F), \end{aligned}$$

where $m > 1$ and $n_1 + \cdots + n_r + m = n$. Then M^0 is a Levi subgroup of G^0. The split component of M^0 is

$$A = \{\operatorname{diag}(\lambda_1 I_{n_1}, \dots, \lambda_r I_{n_r}, I_{2m}, \lambda_r^{-1} I_{n_r}, \dots, \lambda_1^{-1} I_{n_1}) \mid \lambda_i \in F^\times\}.$$

Then $M = C_G(A)$ is equal to

$$\begin{aligned} M &= \{\operatorname{diag}(g_1, \dots, g_r, h, {}^\tau g_r^{-1}, \dots, {}^\tau g_1^{-1}) \mid g_i \in \mathrm{GL}(n_i, F), h \in O(2m, F)\} \\ &\cong \mathrm{GL}(n_1, F) \times \cdots \times \mathrm{GL}(n_r, F) \times O(2m, F). \end{aligned} \tag{6}$$

Let $\pi \in \mathscr{E}_2(M)$. Then $\pi \cong \rho_1 \otimes \cdots \otimes \rho_k \otimes \sigma$, where $\rho_i \in \mathscr{E}_2(\mathrm{GL}(n_i, F))$ and $\sigma \in \mathscr{E}_2(O(2m, F))$. Let $\pi_0 \cong \rho_1 \otimes \cdots \otimes \rho_k \otimes \sigma_0$ be an irreducible component of $r_{M^0,M}(\pi)$. If $s\sigma_0 \cong \sigma_0$, then $W_G(\pi) = W_{G^0}(\pi_0)$ and $R(\pi) = R(\pi_0)$. In this case, $r_{M^0,M}(\pi) = \pi_0$, by Lemma 4.1 of [Ban and Jantzen 2003], and $\rho_i \rtimes \sigma$ is reducible if and only if $\rho_i \rtimes \sigma_0$ is reducible, by Proposition 2.2 of [Goldberg 1995]. Then Theorem 6.5 of [Goldberg 1994] tells us that $R(\pi) \cong \mathbb{Z}_2^d$, where d is the number of inequivalent ρ_i with $\rho_i \rtimes \sigma$ reducible.

Now, consider the case $s\sigma_0 \not\cong \sigma_0$. It follows from Lemma 4.1 of [Ban and Jantzen 2003] that $\pi = i_{M,M^0}(\pi_0)$. Then $i_{G,M}(\pi) = i_{G,M^0}(\pi_0)$ and we know from Theorem 3.3 of [Goldberg 1995] that $R(\pi) \cong \mathbb{Z}_2^d$, where $d = d_1 + d_2$, d_1 is the number of inequivalent ρ_i such that n_i is even and $\rho_i \rtimes \sigma$ is reducible, and d_2 is the number of inequivalent ρ_i such that n_i is odd and $\rho_i \cong \tilde{\rho}_i$. Moreover, Corollary 3.4 of [Goldberg 1995] implies if n_i is odd and $\rho_i \cong \tilde{\rho}_i$, then $\rho_i \rtimes \sigma$ is reducible. Therefore, we see that $R(\pi) \cong \mathbb{Z}_2^d$, where d is the number of inequivalent ρ_i with $\rho_i \rtimes \sigma$ reducible.

In the case $m = 1$, since

$$\mathrm{SO}(2, F) = \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \;\middle|\; a \in F^\times \right\},$$

we have

$$\begin{aligned} M^0 &= \{\operatorname{diag}(g_1, \dots, g_r, a, a^{-1}, {}^\tau g_r^{-1}, \dots, {}^\tau g_1^{-1}) \mid g_i \in \mathrm{GL}(n_i, F), a \in F^\times\} \\ &\cong \mathrm{GL}(n_1, F) \times \cdots \times \mathrm{GL}(n_r, F) \times \mathrm{GL}(1, F), \end{aligned}$$

and this case is described in (b).

(b) Let M^0 be a Levi subgroup of G^0 of the form

$$M^0 = \{\operatorname{diag}(g_1, \dots, g_r, {}^\tau g_r^{-1}, \dots, {}^\tau g_1^{-1}) \mid g_i \in \mathrm{GL}(n_i, F)\}$$

where $n_1 + \cdots + n_r = n$. The split component of M^0 is

$$A = \{\mathrm{diag}(\lambda_1 I_{n_1}, \ldots, \lambda_r I_{n_r}, \lambda_r^{-1} I_{n_r}, \ldots, \lambda_1^{-1} I_{n_1}) \mid \lambda_i \in F^\times\}$$

and $M = C_G(A) = M^0$. Therefore,

$$\begin{aligned} M &= \{\mathrm{diag}(g_1, \ldots, g_r, {}^\tau g_r^{-1}, \ldots, {}^\tau g_1^{-1}) \mid g_i \in \mathrm{GL}(n_i, F)\} \\ &\cong \mathrm{GL}(n_1, F) \times \cdots \times \mathrm{GL}(n_r, F). \end{aligned} \tag{7}$$

Let $\pi \cong \rho_1 \otimes \cdots \otimes \rho_k \otimes 1 \in \mathscr{E}_2(M)$, where 1 denotes the trivial representation of the trivial group. Since $M = M^0$, we can apply directly Theorem 3.3 of [Goldberg 1995]. It follows $R(\pi) \cong \mathbb{Z}_2^d$, where $d = d_1 + d_2$, d_1 is the number of inequivalent ρ_i such that n_i is even and $\rho_i \rtimes 1$ is reducible, and d_2 is the number of inequivalent ρ_i such that n_i is odd and $\rho_i \cong \tilde{\rho}_i$. As above, it follows from Corollary 3.4 of [Goldberg 1995] that if n_i is odd and $\rho_i \cong \tilde{\rho}_i$, then $\rho_i \rtimes \sigma$ is reducible. Again, we obtain $R(\pi) \cong \mathbb{Z}_2^d$, where d is the number of inequivalent ρ_i with $\rho_i \rtimes \sigma$ reducible.

We summarize the above considerations in the following lemma. Observe that the group $O(2, F)$ does not have square integrable representations. It also does not appear as a factor of cuspidal Levi subgroups of $O(2n, F)$. We call a subgroup M defined by (6) or (7) a standard Levi subgroup of $O(2n, F)$.

Lemma 6. *Let $G = O(2n, F)$ and consider a standard Levi subgroup of G of the form*

$$M \cong \mathrm{GL}(n_1, F) \times \cdots \times \mathrm{GL}(n_r, F) \times O(2m, F),$$

where $m \geq 0$, $m \neq 1$, $n_1 + \cdots + n_r + m = n$. Let $\pi \cong \rho_1 \otimes \cdots \otimes \rho_k \otimes \sigma \in \mathscr{E}_2(M)$. Then $R(\pi) \cong \mathbb{Z}_2^d$, where d is the number of inequivalent ρ_i with $\rho_i \rtimes \sigma$ reducible.

5. *R*-groups of discrete series

Let G be $\mathrm{Sp}(2n, F)$, $\mathrm{SO}(2n+1, F)$ or $O(2n, F)$.

Theorem 7. *Let π be an irreducible discrete series representation of a standard Levi subgroup M of G_n. Let φ be the L-parameter of π. Then $R_{\varphi,\pi} \cong R(\pi)$.*

Proof. We can write π in the form

$$\pi \cong (\otimes^{m_1} \delta_1) \otimes \cdots \otimes (\otimes^{m_r} \delta_r) \otimes \sigma \tag{8}$$

where σ is an irreducible discrete series representation of G_m and δ_i $(i = 1, \ldots, r)$ is an irreducible discrete series representation of $\mathrm{GL}(n_i, F)$ such that $\delta_i \not\cong \delta_j$ for $i \neq j$. As explained in Section 4, if $G_n = O(2n, F)$, then $m \neq 1$.

Let φ_i denote the L-parameter of δ_i and φ_σ the L-parameter of σ. Then the L-parameter φ of π is

$$\varphi \cong (m_1\varphi_1 \oplus m_1\tilde{\varphi}_1) \oplus \cdots \oplus (m_r\varphi_r \oplus m_r\tilde{\varphi}_r) \oplus \varphi_\sigma.$$

Each φ_i is irreducible. The parameter φ_σ is of the form $\varphi_\sigma = \varphi'_1 \oplus \cdots \oplus \varphi'_s$ where φ'_i are irreducible, $\varphi'_i \cong \tilde{\varphi}'_i$ and $\varphi'_i \ncong \varphi'_i$ for $i \neq j$. In addition, φ'_i factors through a group of the same type as $\hat{G}_n$. The sets $\{\varphi_i \mid i = 1, \ldots, r\}$ and $\{\varphi'_i \mid i = 1, \ldots, s\}$ can have nonempty intersection. After rearranging the indices, we can write φ as

$$\varphi \cong \Big(\bigoplus_{i=1}^{h} (m_i\varphi_i \oplus m_i\tilde{\varphi}_i) \Big) \oplus \Big(\bigoplus_{i=h+1}^{q} 2m_i\varphi_i \Big) \oplus \Big(\bigoplus_{i=q+1}^{k} 2m_i\varphi_i \Big) \oplus \Big(\bigoplus_{i=k+1}^{r} (2m_i+1)\varphi_i \Big) \oplus \Big(\bigoplus_{i=r+1}^{l} \varphi_i \Big),$$

where $\varphi_\sigma = \bigoplus_{i=k+1}^{l} \varphi_i$ and

$$\begin{aligned} &\varphi_i \ncong \varphi_j,\ \varphi_i \ncong \tilde{\varphi}_j && \text{for } i, j \in \{1, \ldots, l\},\ i \neq j,\\ &\varphi_i \ncong \tilde{\varphi}_i && \text{for } i \in \{1, \ldots, h\},\\ &\varphi_i \cong \tilde{\varphi}_i && \text{for } i \in \{h+1, \ldots, l\},\\ &\varphi_i \text{ not of the same type as } \hat{G} && \text{for } i \in \{h+1, \ldots, q\},\\ &\varphi_i \text{ of the same type as } \hat{G} && \text{for } i \in \{q+1, \ldots, k\}. \end{aligned}$$

Let $d = k - q$. Lemma 5 implies $R_\varphi \cong \mathbb{Z}_2^d$. In addition, $R_{\varphi,\pi} \cong R_\varphi$.

On the other hand, we know that $R(\pi) \cong \mathbb{Z}_2^c$, where c is cardinality of the set

$$C = \{i \in \{1, \ldots, r\} \mid \delta_i \rtimes \sigma \text{ is reducible}\}.$$

This follows from [Goldberg 1994] for $G = \mathrm{SO}(2n+1, F)$ and $G = \mathrm{Sp}(2n, F)$, and from Lemma 6 for $G = O(2n, F)$. We want to show $C = \{q+1, \ldots, k\}$. For any $i \in \{1, \ldots, l\}$, φ_i is an irreducible representation of $W_F \times \mathrm{SL}(2, \mathbb{C})$ and therefore it can be written in the form $\varphi_i = \varphi'_i \otimes S_{a_i}$, where φ'_i is an irreducible representation of W_F and S_{a_i} is the standard irreducible a_i-dimensional algebraic representation of $\mathrm{SL}(2, \mathbb{C})$. For $i \in \{1, \ldots, r\}$, this parameter corresponds to the representation $\delta(\rho_i, a_i)$. Therefore, the representation δ_i in (8) is $\delta_i = \delta(\rho_i, a_i)$. From (3), we have

$$\varphi_\sigma = \bigoplus_{i=k+1}^{l} \varphi_i = \bigoplus_{(\rho,a)\in \mathrm{Jord}(\sigma)} \varphi_\rho \otimes S_a.$$

For $i \in \{h+1, \ldots, q\}$, φ_i is not of the same type as $\hat{G}$ and $\delta(\rho_i, a_i) \rtimes \sigma$ is irreducible. For $i \in \{q+1, \ldots, k\}$, φ_i is of the same type as $\hat{G}$. Now, Lemma 3 tells us $(\rho_i, a_i) \in \mathrm{Jord}(\sigma)$ if and only if $\delta(\rho_i, a_i) \rtimes \sigma$ is irreducible. Therefore, $\delta(\rho_i, a_i) \rtimes \sigma$ is irreducible for $i \in \{k+1, \ldots, r\}$ and $\delta(\rho_i, a_i) \rtimes \sigma$ is reducible for $i \in \{q+1, \ldots, k\}$. It follows $C = \{q+1, \ldots, k\}$ and $R(\pi) \cong \mathbb{Z}_2^d \cong R_{\varphi,\pi}$, finishing the proof. $\square$

6. Unitary groups

Let E/F be a quadratic extension of p-adic fields. Fix $\theta \in W_F \setminus W_E$. Let $G = U(n)$ be a unitary group defined with respect to E/F, $U(n) \subset \mathrm{GL}(n, E)$. Let

$$J_n = \begin{pmatrix} & & & & 1 \\ & & & -1 & \\ & & 1 & & \\ & \cdot & & & \\ \cdot & & & & \end{pmatrix}.$$

We have

$${}^LG = \mathrm{GL}(n, \mathbb{C}) \rtimes W_F,$$

where W_E acts trivially on $\mathrm{GL}(n, \mathbb{C})$ and the action of $w \in W_F \setminus W_E$ on $g \in \mathrm{GL}(n, \mathbb{C})$ is given by $w(g) = J_n \, {}^t g^{-1} J_n^{-1}$.

6.1. ***L-parameters for Levi subgroups.*** Suppose we have a Levi subgroup $M \cong \mathrm{Res}_{E/F} \mathrm{GL}_k \times U(l)$. Then

$${}^LM^0 = \left\{ \begin{pmatrix} g & & \\ & m & \\ & & h \end{pmatrix} \;\middle|\; g, h \in \mathrm{GL}(k, \mathbb{C}), m \in \mathrm{GL}(l, \mathbb{C}) \right\}.$$

Direct computation shows that the action of $w \in W_F \setminus W_E$ on ${}^LM^0$ is given by

$$w\left(\begin{pmatrix} g & & \\ & m & \\ & & h \end{pmatrix}\right) = \begin{pmatrix} J_k \, {}^t h^{-1} J_k^{-1} & & \\ & J_l \, {}^t m^{-1} J_l^{-1} & \\ & & J_k \, {}^t g^{-1} J_k^{-1} \end{pmatrix}.$$

Let π be a discrete series representation of $\mathrm{GL}(k, E) = (\mathrm{Res}_{E/F} \mathrm{GL}_k)(F)$ and τ a discrete series representation of $U(l)$. Let $\varphi_\pi : W_E \times \mathrm{SL}(2, \mathbb{C}) \to \mathrm{GL}(k, \mathbb{C})$ be the L-parameter of π and $\varphi_\tau : W_F \times \mathrm{SL}(2, \mathbb{C}) \to \mathrm{GL}(l, \mathbb{C}) \rtimes W_F$ the L-parameter of τ. Write

$$\varphi_\tau(w, x) = (\varphi'_\tau(w, x), w), \quad w \in W_F, x \in \mathrm{SL}(2, \mathbb{C}).$$

According to [Borel 1979, Sections 4, 5 and 8], there exists a unique (up to equivalence) L-parameter $\varphi : W_F \times \mathrm{SL}(2, \mathbb{C}) \to {}^LM$ such that

$$(9) \quad \begin{aligned} &\varphi((w, x)) = (\varphi_\pi(w), *, *, w) && \text{for all } w \in W_E, x \in \mathrm{SL}(2, \mathbb{C}), \\ &\varphi((w, x)) = (*, \varphi'_\tau(w, x), *, w) && \text{for all } w \in W_F, x \in \mathrm{SL}(2, \mathbb{C}). \end{aligned}$$

We will define a map $\varphi : W_F \times \mathrm{SL}(2, \mathbb{C}) \to {}^LM$ satisfying (9) and show that φ is a homomorphism. Define

$$(10) \quad \varphi((w, x)) = (\varphi_\pi(w, x), \varphi'_\tau(w, x), {}^t\varphi_\pi(\theta w \theta^{-1}, x)^{-1}, w), \qquad w \in W_E, x \in \mathrm{SL}(2, \mathbb{C})$$

and

$$\varphi((\theta,1)) = (J_k^{-1}, \varphi'_\tau(\theta,1), {}^t\varphi_\pi(\theta^2,1)^{-1}J_k, \theta).$$

Note that

$$\begin{aligned}\varphi_\tau(\theta^2,1) &= (\varphi'_\tau(\theta,1),\theta)(\varphi'_\tau(\theta,1),\theta)\\ &= (\varphi'_\tau(\theta,1),1)(J_l\,{}^t\varphi'_\tau(\theta,1)^{-1}J_l^{-1},\theta^2)\\ &= (\varphi'_\tau(\theta,1)J_l\,{}^t\varphi'_\tau(\theta,1)^{-1}J_l^{-1},\theta^2).\end{aligned}$$

It follows that

$$\varphi'_\tau(\theta,1)J_l\,{}^t\varphi'_\tau(\theta,1)^{-1}J_l^{-1} = \varphi'_\tau(\theta^2,1). \tag{11}$$

Similarly, for $w \in W_E$, $x \in \mathrm{SL}(2,\mathbb{C})$,

$$\begin{aligned}\varphi_\tau(\theta w\theta^{-1},x) &= \varphi_\tau(\theta,1)\varphi_\tau(w,x)\varphi_\tau(\theta,1)^{-1}\\ &= (\varphi'_\tau(\theta,1),\theta)(\varphi'_\tau(w,x),w)(1,\theta^{-1})(\varphi'_\tau(\theta,1)^{-1},1)\\ &= (\varphi'_\tau(\theta,1),1)(J_l\,{}^t\varphi'_\tau(w,x)^{-1}J_l^{-1},\theta w\theta^{-1})(\varphi'_\tau(\theta,1)^{-1},1)\\ &= (\varphi'_\tau(\theta,1)J_l\,{}^t\varphi'_\tau(w,x)^{-1}J_l^{-1}\varphi'_\tau(\theta,1)^{-1},\theta w\theta^{-1})\end{aligned}$$

and thus

$$\varphi'_\tau(\theta,1)J_l\,{}^t\varphi'_\tau(w,x)^{-1}J_l^{-1}\varphi'_\tau(\theta,1)^{-1} = \varphi'_\tau(\theta w\theta^{-1},x). \tag{12}$$

Now,

$$\begin{aligned}&\varphi(\theta,1)\varphi(\theta,1)\\ &\quad= \big(J_k^{-1},\varphi'_\tau(\theta,1),{}^t\varphi_\pi(\theta^2,1)^{-1}J_k,\theta\big)\big(J_k^{-1},\varphi'_\tau(\theta,1),{}^t\varphi_\pi(\theta^2,1)^{-1}J_k,\theta\big)\\ &\quad= \big(J_k^{-1},\varphi'_\tau(\theta,1),{}^t\varphi_\pi(\theta^2,1)^{-1}J_k,1\big)\big(J_k\varphi_\pi(\theta^2,1),J_l\,{}^t\varphi'_\tau(\theta,1)^{-1}J_l^{-1},J_k^{-1},\theta^2\big)\\ &\quad= \big(\varphi_\pi(\theta^2,1),\varphi'_\tau(\theta^2,1),{}^t\varphi_\pi(\theta^2,1)^{-1},\theta^2\big) = \varphi(\theta^2,1),\end{aligned}$$

using (11) and (10). Further, for $w \in W_E$, $x \in \mathrm{SL}(2,\mathbb{C})$, we have

$$\begin{aligned}&\varphi(\theta,1)\varphi(w,x)\varphi(\theta,1)^{-1}\\ &\quad= \big(J_k^{-1},\varphi'_\tau(\theta,1),{}^t\varphi_\pi(\theta^2,1)^{-1}J_k,\theta\big)\big(\varphi_\pi(w,x),\varphi'_\tau(w,x),{}^t\varphi_\pi(\theta w\theta^{-1},x)^{-1},w\big)\\ &\qquad\qquad\cdot(1,1,1,\theta^{-1})\big(J_k,\varphi'_\tau(\theta,1)^{-1},J_k^{-1}\,{}^t\varphi_\pi(\theta^2,1),1\big)\\ &\quad= \big(J_k^{-1},\varphi'_\tau(\theta,1),{}^t\varphi_\pi(\theta^2,1)^{-1}J_k,1\big)\\ &\qquad\cdot\big(J_k\varphi_\pi(\theta w\theta^{-1},x)J_k^{-1},J_l\,{}^t\varphi'_\tau(w,x)^{-1}J_l^{-1},J_k\,{}^t\varphi_\pi(w,x)^{-1}J_k^{-1},\theta w\theta^{-1}\big)\\ &\qquad\qquad\cdot\big(J_k,\varphi'_\tau(\theta,1)^{-1},J_k^{-1}\,{}^t\varphi_\pi(\theta^2,1),1\big)\\ &\quad= \big(\varphi_\pi(\theta w\theta^{-1},x),\varphi'_\tau(\theta w\theta^{-1},x),{}^t\varphi_\pi(\theta^2w\theta^{-2},x)^{-1},\theta w\theta^{-1}\big)\\ &\quad= \varphi(\theta w\theta^{-1},x).\end{aligned}$$

Here, we use (12) and $J_k^2 = (J_k^{-1})^2 = (-1)^{k-1}$, so

$$ {}^t\varphi_\pi(\theta^2, 1)^{-1} J_k J_k \, {}^t\varphi_\pi(w, x)^{-1} J_k^{-1} J_k^{-1} \, {}^t\varphi_\pi(\theta^2, 1) = {}^t\varphi_\pi(\theta^2 w \theta^{-2}, x)^{-1}. $$

In conclusion, $\varphi(\theta^2, 1) = \varphi(\theta, 1)^2$ and $\varphi(\theta w \theta^{-1}, x) = \varphi(\theta, 1)\varphi(w, x)\varphi(\theta, 1)^{-1}$. Since φ is clearly multiplicative on $W_E \times \mathrm{SL}(2, \mathbb{C})$, it follows that φ is a homomorphism. Therefore, φ is the L-parameter for $\pi \otimes \tau$.

6.2. ***The coefficients λ_φ.*** Let $\varphi : W_E \times \mathrm{SL}(2, \mathbb{C}) \to \mathrm{GL}_k(\mathbb{C})$ be an irreducible L-parameter. Assume $\varphi \cong {}^t({}^\theta\varphi)^{-1}$. Let X be a nonzero matrix such that

$$ {}^t\varphi(\theta w \theta^{-1}, x)^{-1} = X^{-1}\varphi(w, x)X, $$

for all $w \in W_E$, $x \in \mathrm{SL}(2, \mathbb{C})$. We proceed similarly as in [Mœglin 2002, p. 190]. By taking transpose and inverse,

$$ \varphi(\theta w \theta^{-1}, x) = {}^tX \, {}^t\varphi(w, x)^{-1} \, {}^tX^{-1}. $$

Next, we replace w by $\theta w \theta^{-1}$. This gives

$$ \varphi(\theta^2, 1)\varphi(w, x)\varphi(\theta^{-2}, 1) = {}^tX \, {}^t\varphi(\theta w \theta^{-1}, x)^{-1} \, {}^tX^{-1} = {}^tX X^{-1}\varphi(w, x) X \, {}^tX^{-1}, $$

for all $w \in W_E$, $x \in \mathrm{SL}(2, \mathbb{C})$. Since φ is irreducible, $\varphi(\theta^{-2}, 1) \, {}^tX X^{-1}$ is a constant. Define

$$ \lambda_\varphi = \varphi(\theta^{-2}, 1) \, {}^tX X^{-1}. \tag{13} $$

As in [Mœglin 2002], we can show that $\lambda_\varphi = \pm 1$.

Lemma 8. *Let $\varphi : W_E \to \mathrm{GL}_k(\mathbb{C})$ be an irreducible L-parameter such that $\varphi \cong {}^t({}^\theta\varphi)^{-1}$. Let S_a be the standard a-dimensional irreducible algebraic representation of* $\mathrm{SL}(2, \mathbb{C})$. *Then ${}^\theta({}^t(\varphi \otimes S_a)^{-1}) \cong \varphi \otimes S_a$ and*

$$ \lambda_{\varphi \otimes S_a} = (-1)^{a+1}\lambda_\varphi. $$

Proof. We know that ${}^tS_a^{-1} \cong S_a$. Let Y be a nonzero matrix such that

$$ {}^tS_a(x)^{-1} = Y^{-1}S_a(x)Y, $$

for all $x \in \mathrm{SL}(2, \mathbb{C})$. Then ${}^tY = Y$ for a odd and ${}^tY = -Y$ for a even. Let X be a nonzero matrix such that

$$ {}^t\varphi(\theta w \theta^{-1})^{-1} = X^{-1}\varphi(w)X, $$

for all $w \in W_E$. We have

$$ \begin{aligned} {}^t(\varphi \otimes S_a(\theta w \theta^{-1}, x))^{-1} &= ({}^t\varphi(\theta w \theta^{-1})^{-1}) \otimes ({}^tS_a(x)^{-1}) \\ &= (X^{-1}\varphi(w)X) \otimes (Y^{-1}S_a(x)Y) \\ &= (X \otimes Y)^{-1}(\varphi \otimes S_a(w, x)) \otimes (X \otimes Y). \end{aligned} $$

It follows that ${}^{\theta}({}^{t}(\varphi\otimes S_a)^{-1})\cong\varphi\otimes S_a$ and

$$\begin{aligned}\lambda_{\varphi\otimes S_a}&=(\varphi\otimes S_a(\theta^{-2},1))\,{}^{t}(X\otimes Y)(X\otimes Y)^{-1}\\&=(\varphi(\theta^{-2})\,{}^{t}XX^{-1})\otimes({}^{t}YY^{-1}))=(-1)^{a+1}\lambda_{\varphi}.\end{aligned}\qquad\square$$

6.3. ***Centralizers.*** Let $\varphi: W_F\times\mathrm{SL}(2,\mathbb{C})\to{}^{L}G$ be an L-parameter. Denote by φ_E the restriction of φ to $W_E\times\mathrm{SL}(2,\mathbb{C})$. Then φ_E is a representation of $W_E\times\mathrm{SL}(2,\mathbb{C})$ on $V=\mathbb{C}^n$. Write φ_E as a sum of irreducible subrepresentations

$$\varphi_E=m_1\varphi_1\oplus\cdots\oplus m_l\varphi_l,$$

where m_i is the multiplicity of φ_i and $\varphi_i\not\cong\varphi_j$ for $i\neq j$. It follows from [Mœglin 2002] that S_φ, the centralizer in $\hat{G}$ of the image of φ, is given by

$$S_\varphi\cong\prod_{i=1}^{l}C(m_i\varphi_i),\tag{14}$$

where

$$C(m_i\varphi_i)=\begin{cases}\mathrm{GL}(m_i,\mathbb{C}) & \text{if } \varphi_i\not\cong{}^{\theta}\widetilde{\varphi}_i,\\ O(m_i,\mathbb{C}) & \text{if } \varphi_i\cong{}^{\theta}\widetilde{\varphi}_i,\ \lambda_{\varphi_i}=(-1)^{n-1},\\ \mathrm{Sp}(m_i,\mathbb{C}) & \text{if } \varphi_i\cong{}^{\theta}\widetilde{\varphi}_i,\ \lambda_{\varphi_i}=(-1)^{n}.\end{cases}$$

6.4. ***Coefficients λ_ρ.*** Let ${}^{L}M=\mathrm{GL}_k(\mathbb{C})\times\mathrm{GL}_k(\mathbb{C})\rtimes W_F$, where the action of $w\in W_F\setminus W_E$ on $\mathrm{GL}_k(\mathbb{C})\times\mathrm{GL}_k(\mathbb{C})$ is given by

$$w(g,h,1)w^{-1}=(J_n\,{}^{t}h^{-1}J_n^{-1},\ J_n\,{}^{t}g^{-1}J_n^{-1},1).$$

For $\eta=\pm1$, we denote by R_η the representation of ${}^{L}M$ on $\mathrm{End}_{\mathbb{C}}(\mathbb{C}^k)$ given by

$$\begin{aligned}R_\eta((g,h,1))\cdot X&=gXh^{-1},\\R_\eta((1,1,\theta))\cdot X&=\eta J_k\,{}^{t}XJ_k.\end{aligned}$$

Let τ denote the nontrivial element in $Gal(E/F)$. Let ρ be an irreducible unitary supercuspidal representation of $\mathrm{GL}(k,E)$. Assume $\rho\cong{}^{\tau}\tilde{\rho}$. Then precisely one of the two L-functions $L(s,\rho,R_1)$ and $L(s,\rho,R_{-1})$ has a pole at $s=0$. Denote by λ_ρ the value of η such that $L(s,\rho,R_\eta)$ has a pole at $s=0$.

Lemma 9. *Assume that ρ is an irreducible unitary supercuspidal representation of* $\mathrm{GL}(k,E)$ *such that $\rho\cong{}^{\tau}\tilde{\rho}$. Let φ_ρ be the L-parameter of ρ. Then $\lambda_{\varphi_\rho}=\lambda_\rho$.*

Proof. As shown in Section 6.1, the parameter $\varphi: W_F\to{}^{L}M$ corresponding to $\varphi_\rho: W_E\to\mathrm{GL}_k(\mathbb{C})$ is given by

$$\varphi(w)=\left(\begin{pmatrix}\varphi_\rho(w) & \\ & {}^{t}\varphi_\rho(\theta w\theta^{-1})^{-1}\end{pmatrix},w\right),\tag{15}$$

for $w \in W_E$, and

$$\varphi(\theta) = \left(\begin{pmatrix} J_k^{-1} & \\ & {}^t\varphi_\rho(\theta^2)^{-1} J_k \end{pmatrix}, \theta \right). \tag{16}$$

From [Henniart 2010], we have $L(s, \rho, R_\eta) = L(s, R_\eta \circ \varphi)$. Therefore, $L(s, R_{\lambda_\rho} \circ \varphi)$ has a pole at $s = 0$. Then $R_{\lambda_\rho} \circ \varphi$ contains the trivial representation, so there exists nonzero $X \in M_k(\mathbb{C})$ such that $(R_{\lambda_\rho} \circ \varphi)(w) \cdot X = X$ for all $w \in W_F$. In particular, (15) implies that for $w \in W_E$,

$$\varphi_\rho(w) X \, {}^t\varphi_\rho(\theta w \theta^{-1}) = X$$

so

$$\varphi_\rho(w) X = X \, {}^t\varphi_\rho(\theta w \theta^{-1})^{-1}. \tag{17}$$

Therefore, X is a nonzero intertwining operator between φ_ρ and ${}^t({}^\theta\varphi_\rho)^{-1}$. From (13), we have

$$\varphi_\rho(\theta^{-2}) \, {}^t X X^{-1} = \lambda_{\varphi_\rho}. \tag{18}$$

Now, since $(R_{\lambda_\rho} \circ \varphi)(\theta) \cdot X = X$, we have from (16)

$${}^t X \, {}^t\varphi_\rho(\theta^2) = \lambda_\rho X.$$

By transposing and multiplying by X^{-1}, we obtain

$$\varphi_\rho(\theta^2) = \lambda_\rho \, {}^t X X^{-1}.$$

We compare this to (18). It follows $\lambda_{\varphi_\rho} = \lambda_\rho$. □

6.5. ***Jordan blocks for unitary groups.*** For the unitary group $U(n)$, define

$$R_d = R_\eta, \quad \text{where} \quad \eta = (-1)^n.$$

Let σ be an irreducible discrete series representation of $U(n)$. Denote by $Jord(\sigma)$ the set of pairs (ρ, a), where $\rho \in {}^0\mathscr{E}(\mathrm{GL}(d_\rho, E))$, $\rho \cong {}^\tau\tilde{\rho}$, and $a \in \mathbb{Z}^+$, such that (ρ, a) satisfies properties (J-1) and (J-2) from Section 2.2.

Lemma 10. *Let ρ be an irreducible supercuspidal representation of* $\mathrm{GL}(d, E)$ *such that $\varphi_\rho \cong {}^\theta\widetilde{\varphi}_\rho$, where φ_ρ is the L-parameter for ρ. Then the condition* (J-1) *is equivalent to*

(J-1″) $\lambda_{\varphi_\rho \otimes S_a} = (-1)^{n+1}$.

Proof. The condition (J-1) says that a is even if $L(s, \rho, R_d)$ has a pole at $s = 0$ and odd otherwise. Observe that

$$\begin{aligned} L(s, \rho, R_d) \text{ has a pole at } s = 0 &\iff \lambda_{\varphi_\rho} = (-1)^n \\ &\iff \lambda_{\varphi_\rho \otimes S_a} = (-1)^n (-1)^{a+1} \\ &\iff \lambda_{\varphi_\rho \otimes S_a} = \begin{cases} (-1)^{n+1} & a \text{ even}, \\ (-1)^n & a \text{ odd}. \end{cases} \end{aligned}$$

From this, it is clear that (J-1) is equivalent to (J-1″). □

6.6. *R-groups for unitary groups.*

Lemma 11. *Let σ be an irreducible discrete series representation of $U(n)$ and let $\delta = \delta(\rho, a)$ be an irreducible discrete series representation of* $\mathrm{GL}(l, E)$, $l = da$, $d = \dim(\rho)$. *Let φ_ρ and φ be the L-parameters of ρ and $\pi = \delta \otimes \sigma$, respectively. Then $R_{\varphi,\pi} \cong R(\pi)$.*

Proof. Let φ_σ be the L-parameter of σ. Then

$$\varphi_E \cong \varphi_\rho \otimes S_a \oplus {}^\theta\widetilde{\varphi}_\rho \otimes S_a \oplus (\varphi_\sigma)_E.$$

This is a representation of $W_E \times \mathrm{SL}(2, \mathbb{C})$ on $V = \mathbb{C}^{n+2l}$. Write $(\varphi_\sigma)_E$ as a sum of irreducible components,

$$(\varphi_\sigma)_E = \varphi_1 \oplus \cdots \oplus \varphi_m.$$

Each component appears with multiplicity one. The centralizer S_φ is given by (14). If $\varphi_\rho \ncong {}^\theta\widetilde{\varphi}_\rho$, then

$$S_\varphi \cong \mathrm{GL}(1, \mathbb{C}) \times \mathrm{GL}(1, \mathbb{C}) \times \prod_{i=1}^{m} \mathrm{GL}(1, \mathbb{C}).$$

This implies $R_\varphi = 1$. On the other hand, $\delta \rtimes \sigma$ is irreducible, so $R(\pi) = 1$. It follows $R_{\varphi,\pi} \cong R(\pi)$.

Now, consider the case $\varphi_\rho \cong {}^\theta\widetilde{\varphi}_\rho$. If $\varphi_\rho \otimes S_a \in \{\varphi_1, \ldots, \varphi_m\}$, then

$$S_\varphi \cong O(3, \mathbb{C}) \times \prod_{i=1}^{m-1} \mathrm{GL}(1, \mathbb{C}) \quad \text{and} \quad S_\varphi^0 \cong \mathrm{SO}(3, \mathbb{C}) \times \prod_{i=1}^{m-1} \mathrm{GL}(1, \mathbb{C}).$$

This gives $W_\varphi = W_\varphi^0$ and $R_\varphi = 1$. Since $\varphi_\rho \otimes S_a \in \{\varphi_1, \ldots, \varphi_m\}$, the condition (J-2) implies that $\delta \rtimes \sigma$ is irreducible. Therefore, $R(\pi) = 1 = R_{\varphi,\pi}$.

It remains to consider the case $\varphi_\rho \cong {}^\theta\widetilde{\varphi}_\rho$ and $\varphi_\rho \otimes S_a \notin \{\varphi_1, \ldots, \varphi_m\}$. Then (ρ, a) does not satisfy (J-1″) or (J-2). Assume first that (ρ, a) does not satisfy

(J-1″). Then $\delta \rtimes \sigma$ is irreducible, so $R(\pi) = 1$. Since (ρ, a) does not satisfy (J-1″), we have $\lambda_{\varphi_\rho \otimes S_a} = (-1)^n = (-1)^{n+2l}$. Then, by (14),

$$S_\varphi \cong \mathrm{Sp}(2, \mathbb{C}) \times \prod_{i=1}^{m} \mathrm{GL}(1, \mathbb{C}).$$

It follows $R_{\varphi,\pi} = 1 = R(\pi)$.

Now, assume that (ρ, a) satisfies (J-1″), but does not satisfy (J-2). Then $\lambda_{\varphi_\rho \otimes S_a} = (-1)^{n-1} = (-1)^{n+2l-1}$, so

$$S_\varphi \cong O(2, \mathbb{C}) \times \prod_{i=1}^{m} \mathrm{GL}(1, \mathbb{C})$$

and $R_{\varphi,\pi} \cong \mathbb{Z}_2$. Since (ρ, a) does not satisfy (J-2), $\delta \rtimes \sigma$ is reducible and hence $R(\pi) \cong \mathbb{Z}_2 \cong R_{\varphi,\pi}$. □

Acknowledgments

We thank Guy Henniart, Joe Hundley and Freydoon Shahidi for valuable comments. Ban thanks Werner Müller and the Mathematical Institute of the University of Bonn for their hospitality during her three-month research stay, where a part of this work was done.

References

[Arthur 1989] J. Arthur, "Unipotent automorphic representations: conjectures", pp. 13–71 in *Orbites unipotentes et représentations, II: Groupes p-adiques et réels*, Astérisque **171-172**, Soc. Math. France, Paris, 1989. MR 91f:22030 Zbl 0728.22014

[Arthur 1993] J. Arthur, "On elliptic tempered characters", *Acta Math.* **171**:1 (1993), 73–138. MR 94i:22038 Zbl 0822.22011

[Arthur 2011] J. Arthur, "The endoscopic classification of representations: orthogonal and symplectic groups", preprint, University of Toronto, 2011, available at http://www.claymath.org/cw/arthur/pdf/Book.pdf.

[Ban and Jantzen 2003] D. Ban and C. Jantzen, "Degenerate principal series for even-orthogonal groups", *Represent. Theory* **7** (2003), 440–480. MR 2004k:22020 Zbl 1054.22015

[Ban and Zhang 2005] D. Ban and Y. Zhang, "Arthur R-groups, classical R-groups, and Aubert involutions for SO($2n+1$)", *Compos. Math.* **141**:2 (2005), 323–343. MR 2006d:22019 Zbl 1131.11033

[Borel 1979] A. Borel, "Automorphic L-functions", pp. 27–61 in *Automorphic forms, representations and L-functions, 2* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR 81m:10056 Zbl 0412.10017

[Goldberg 1994] D. Goldberg, "Reducibility of induced representations for Sp($2n$) and SO(n)", *Amer. J. Math.* **116**:5 (1994), 1101–1151. MR 95g:22016 Zbl 0851.22021

[Goldberg 1995] D. Goldberg, "Reducibility for non-connected p-adic groups, with G° of prime index", *Canad. J. Math.* **47**:2 (1995), 344–363. MR 96d:22003 Zbl 0835.22015

[Goldberg 2011] D. Goldberg, "On dual R-groups for classical groups", pp. 159–185 in *On certain L-functions*, edited by J. Arthur et al., Clay Math. Proc. **13**, Amer. Math. Soc., Providence, RI, 2011. MR 2767516 Zbl 05932913

[Goldberg and Herb 1997] D. Goldberg and R. Herb, "Some results on the admissible representations of non-connected reductive p-adic groups", *Ann. Sci. École Norm. Sup.* (4) **30**:1 (1997), 97–146. MR 98b:22033 Zbl 0874.22016

[Gross and Prasad 1992] B. H. Gross and D. Prasad, "On the decomposition of a representation of SO_n when restricted to SO_{n-1}", *Canad. J. Math.* **44**:5 (1992), 974–1002. MR 93j:22031 Zbl 0787.22018

[Harris and Taylor 2001] M. Harris and R. Taylor, *The geometry and cohomology of some simple Shimura varieties*, Annals of Mathematics Studies **151**, Princeton University Press, Princeton, NJ, 2001. MR 2002m:11050 Zbl 1036.11027

[Henniart 2000] G. Henniart, "Une preuve simple des conjectures de Langlands pour $\mathrm{GL}(n)$ sur un corps p-adique", *Invent. Math.* **139**:2 (2000), 439–455. MR 2001e:11052 Zbl 1048.11092

[Henniart 2010] G. Henniart, "Correspondance de Langlands et fonctions L des carrés extérieur et symétrique", *Int. Math. Res. Not.* **2010**:4 (2010), 633–673. MR 2011c:22028 Zbl 1184.22009

[Keys 1987] C. D. Keys, "L-indistinguishability and R-groups for quasisplit groups: unitary groups in even dimension", *Ann. Sci. École Norm. Sup.* (4) **20**:1 (1987), 31–64. MR 88m:22042 Zbl 0634.22014

[Knapp and Stein 1971] A. W. Knapp and E. M. Stein, "Intertwining operators for semisimple groups", *Ann. of Math.* (2) **93**:3 (1971), 489–578. MR 57 #536 Zbl 0257.22015

[Mœglin 2002] C. Mœglin, "Sur la classification des séries discrètes des groupes classiques p-adiques: paramètres de Langlands et exhaustivité", *J. Eur. Math. Soc.* **4**:2 (2002), 143–200. MR 2003g:22021 Zbl 1002.22009

[Mœglin 2007a] C. Mœglin, "Classification des séries discrètes pour certains groupes classiques p-adiques", pp. 209–245 in *Harmonic analysis, group representations, automorphic forms and invariant theory*, edited by J.-S. Li et al., Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **12**, World Scientific, Hackensack, NJ, 2007. MR 2009i:22023

[Mœglin 2007b] C. Mœglin, "Classification et changement de base pour les séries discrètes des groupes unitaires p-adiques", *Pacific J. Math.* **233**:1 (2007), 159–204. MR 2009d:22022 Zbl 1157.22010

[Mœglin and Tadić 2002] C. Mœglin and M. Tadić, "Construction of discrete series for classical p-adic groups", *J. Amer. Math. Soc.* **15**:3 (2002), 715–786. MR 2003g:22020 Zbl 0992.22015

[Shahidi 1981] F. Shahidi, "On certain L-functions", *Amer. J. Math.* **103**:2 (1981), 297–355. MR 82i:10030 Zbl 0467.12013

[Shahidi 1990] F. Shahidi, "A proof of Langlands' conjecture on Plancherel measures; complementary series for p-adic groups", *Ann. of Math.* (2) **132**:2 (1990), 273–330. MR 91m:11095 Zbl 0780.22005

[Shelstad 1982] D. Shelstad, "L-indistinguishability for real groups", *Math. Ann.* **259**:3 (1982), 385–430. MR 84c:22017 Zbl 0506.22014

[Silberger 1978] A. J. Silberger, "The Knapp–Stein dimension theorem for p-adic groups", *Proc. Amer. Math. Soc.* **68**:2 (1978), 243–246. MR 58 #11245 Zbl 0348.22007

[Silberger 1979] A. J. Silberger, *Introduction to harmonic analysis on reductive p-adic groups*, Math. Notes **23**, Princeton Univ. Press, Princeton, NJ, 1979. MR 81m:22025 Zbl 0458.22006

[Zelevinsky 1980] A. V. Zelevinsky, "Induced representations of reductive p-adic groups, II: On irreducible representations of $\mathrm{GL}(n)$", *Ann. Sci. École Norm. Sup.* (4) **13**:2 (1980), 165–210. MR 83g:22012 Zbl 0441.22014

Received February 25, 2011. Revised August 30, 2011.

DUBRAVKA BAN
DEPARTMENT OF MATHEMATICS
SOUTHERN ILLINOIS UNIVERSITY
CARBONDALE, IL 62901
UNITED STATES
dban@math.siu.edu

DAVID GOLDBERG
DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
WEST LAFAYETTE, IN 47907-1395
UNITED STATES
goldberg@math.purdue.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

FINITE-VOLUME COMPLEX-HYPERBOLIC SURFACES, THEIR TOROIDAL COMPACTIFICATIONS, AND GEOMETRIC APPLICATIONS

LUCA FABRIZIO DI CERBO

We study the classification of smooth toroidal compactifications of nonuniform ball quotients in the sense of Kodaira and Enriques. Several results concerning the Riemannian and complex algebraic geometry of these spaces are given. In particular we show that there are compact complex surfaces which admit Riemannian metrics of nonpositive curvature, but which do not admit Kähler metrics of nonpositive curvature. An infinite class of such examples arise as smooth toroidal compactifications of ball quotients.

1. Introduction

Let $\tilde{M}$ be a symmetric space of noncompact type, and let $\mathrm{Iso}_0(\tilde{M})$ denote the connected component of the isometry group of $\tilde{M}$ containing the identity. Recall that $\mathrm{Iso}_0(\tilde{M})$ is a semisimple Lie group. A discrete subgroup $\Gamma \subset \mathrm{Iso}_0(\tilde{M})$ is a *lattice* in $\tilde{M}$ if $\tilde{M}/\Gamma$ is of finite volume. When Γ is torsion free, then $\tilde{M}/\Gamma$ is a finite volume manifold or a locally symmetric space. A lattice Γ is *uniform* (*nonuniform*) if $\tilde{M}/\Gamma$ is compact (noncompact).

The theory of compactifications of locally symmetric spaces or varieties has been extensively studied, see for example [Borel and Ji 2006]. In fact, locally symmetric varieties of noncompact type often occur as moduli spaces in algebraic geometry and number theory, see [Ash et al. 2010]. For technical reasons this beautiful theory is mainly developed for quotients of symmetric spaces or varieties by *arithmetic* subgroups. For arithmetic subgroups of semisimple Lie groups a nice reduction theory is available [Borel and Ji 2006]. Among many other things, the aforementioned theory can be used to deduce their finite generation, the existence of finitely many conjugacy classes of maximal parabolic subgroups, and the existence of *neat* subgroups of finite index.

The celebrated work of Margulis [1984] implies that lattices in any semisimple Lie group of real rank bigger or equal than two are arithmetic subgroups. This important theorem does not cover many interesting cases such as lattices in the

MSC2010: 14J29, 53C20, 53C55.
Keywords: manifolds with nonpositive curvature, toroidal compactifications.

complex hyperbolic space $\mathbb{C}\mathcal{H}^n$, where nonarithmetic lattices are known to exist by the work of Mostow and Mostow–Deligne; see [Deligne and Mostow 1993] and the bibliography therein.

It is thus desirable to develop a theory of compactifications of locally symmetric varieties modeled on $\mathbb{C}\mathcal{H}^n$ regardless of the arithmeticity of the defining torsion free lattices. A compactification of finite-volume complex-hyperbolic manifolds as a complex spaces with isolated normal singularities was obtained by Siu and Yau [1982]. This compactification may be regarded as a generalization of the Baily–Borel compactification defined for arithmetic lattices in $\mathbb{C}\mathcal{H}^n$. A *toroidal* compactification for finite-volume complex-hyperbolic manifolds was described by Hummel and Schroeder [1996] in connection with cusps closing techniques arising from Riemannian geometry; see also [Mok 2009] and the classical reference [Ash et al. 2010] for what concerns the arithmetic case.

The constructions of both Siu–Yau and Hummel–Schroeder rely on the theory of nonpositively curved Riemannian manifolds. The key point here is that the structure theorems for finite-volume manifolds of negatively pinched curvature, or more generally for *visibility* manifolds [Eberlein 1996], can be used as a substitute of the reduction theory for arithmetic subgroups.

In this paper we study torsion-free nonuniform lattices in the complex hyperbolic plane $\mathbb{C}\mathcal{H}^2$ and their toroidal compactifications. Let Γ be a lattice as above and let $\overline{\mathbb{C}\mathcal{H}^2/\Gamma}$ denote its toroidal compactification. When $\overline{\mathbb{C}\mathcal{H}^2/\Gamma}$ is smooth, it is a compact Kähler surface [Hummel 1998]. It is then of interest to place these smooth Kähler surfaces in the framework of the *Kodaira–Enriques* classification of complex surfaces [Barth et al. 2004]. The main purpose of this paper is to prove the following:

Theorem A. *Let Γ be a nonuniform torsion-free lattice in $\mathbb{C}\mathcal{H}^2$. There exists a finite subset $\mathcal{F}' \subset \Gamma$ of parabolic isometries for which the following holds*: *for any normal subgroup $\Gamma' \lhd \Gamma$ with the property that $\mathcal{F}' \cap \Gamma'$ is empty*, *then $\overline{\mathbb{C}\mathcal{H}^2/\Gamma'}$ is a surface of general type with ample canonical line bundle. Moreover*, *$\overline{\mathbb{C}\mathcal{H}^2/\Gamma'}$ admits Riemannian metrics of nonpositive sectional curvature but it cannot support Kähler metrics of nonpositive sectional curvature.*

An outline of the paper follows. Section 2 starts with a summary of the results from [Hummel and Schroeder 1996]. Such results are then combined with the Kodaira–Enriques classification to prove that when the lattice Γ is sufficiently small then $\overline{\mathbb{C}\mathcal{H}^2/\Gamma}$ is a surface of general type with ample canonical bundle.

In Section 3 we present some examples of a surfaces of general type which do not admit any nonpositively curved Kähler metric, but whose underlying smooth manifolds admit Riemannian metrics of nonpositive curvature. Finally we prove Theorem A.

In Section 4 we show how Theorem A, combined with the theory of semistable curves on algebraic surfaces [Sakai 1980], can be used to address the problem of the projective-algebraicity of minimal compactifications (Siu–Yau) of finite-volume complex-hyperbolic surfaces. Those results are summarized in Theorem B. The result obtained is effective.

The projective-algebraicity of minimal compactifications was proved in [Mok 2009] through L^2-estimates for the $\overline{\partial}$-operator. This analytical approach works in any dimension.

2. Toroidal compactifications and the Kodaira–Enriques classification

Let PU$(1, 2)$ denote the connected component of Iso$(\mathbb{C}\mathcal{H}^2)$ containing the identity. Let Γ be a nonuniform torsion-free lattice of holomorphic isometries of the complex hyperbolic plane $\mathbb{C}\mathcal{H}^2$, that is, $\Gamma \leq \mathrm{PU}(1, 2)$. Recall that the locally symmetric space $\mathbb{C}\mathcal{H}^2/\,\Gamma$ has finitely many cusp ends $A_1, \dots, A_n$ which are in one to one correspondence with conjugacy classes of the maximal parabolic subgroups of Γ [Eberlein 1980]. The set of all parabolic elements of Γ can be written as a disjoint union of subsets Γ_x, where Γ_x is the set of all parabolic elements in Γ having x as their unique fixed point. Here x is a point in the natural point set compactification of $\mathbb{C}\mathcal{H}^2$ obtained by adjoining points at infinity corresponding to asymptotic geodesic rays. Thus, given a cusp A_i, let us consider the associated maximal parabolic subgroup $\Gamma_{x_i} \leq \Gamma$ and the horoball HB_{x_i} stabilized by Γ_{x_i}. We then have that $\mathrm{HB}_{x_i}/\,\Gamma_{x_i}$ is naturally identified with A_i.

Recall that after choosing an Iwasawa decomposition [Eberlein 1996] for PU$(1, 2)$, we get a identification of ∂HB with the three-dimensional Heisenberg Lie group N. Moreover, N comes equipped with a left invariant metric and then we may view Γ_{x_i} as a lattice in Iso(N). The cusps $A_1, \dots, A_n$ are then identified with $N/\,\Gamma_{x_i}\times[0, \infty)$, for $i = 1, \dots, n$.

The isometry group of N is isomorphic to the semi-direct product $N \rtimes U(1)$. We say that a lattice in Iso(N) is rotation free if it is a lattice in N, that is, if it is a lattice of left translations. A parabolic isometry $\phi \in \Gamma$ is called *unipotent* if it acts as a translation on its invariant horospheres.

We now briefly summarize some of the results from [Hummel 1998; Hummel and Schroeder 1996].

Theorem 2.1 (Hummel–Scroeder). *Let Γ be a nonuniform torsion-free lattice in $\mathbb{C}\mathcal{H}^2$. There exists a finite subset $\mathcal{F} \subset \Gamma$ of parabolic isometries such that, for any normal subgroup $\Gamma' \lhd \Gamma$ with the property that $\mathcal{F} \cap \Gamma'$ is empty, $\overline{\mathbb{C}\mathcal{H}^2/\,\Gamma'}$ is smooth and Kähler.*

Furthermore, using a cusp closing technique arising from Riemannian geometry they were able to prove:

Theorem 2.2 (Hummel–Schroeder). *Let Γ be a nonuniform torsion-free lattice in $\mathbb{C}\mathcal{H}^2$. Then there exists a finite subset $\mathcal{F}' \subset \Gamma$ of parabolic isometries containing $\mathcal{F}$ such that if $\Gamma' \lhd \Gamma$ is a normal subgroup with the property that $\mathcal{F}' \cap \Gamma'$ is empty, then $\overline{\mathbb{C}\mathcal{H}^2/\Gamma'}$ admits a Riemannian metric of nonpositive sectional curvature.*

A few remarks about these results. A nonuniform torsion-free lattice in $\mathbb{C}\mathcal{H}^2$ admits a smooth toroidal compactification if its parabolic isometries are all unipotent. In the arithmetic case this is achieved by choosing a neat subgroup of finite index [Ash et al. 2010]. It is also interesting to observe that we have plenty of normal subgroups satisfying the requirements of Theorems 2.1 and 2.2, in fact $\mathrm{PU}(1, 2)$ is linear and then *residually finite* by a fundamental result of Mal'tsev [1940]. Finally, it is interesting to notice that in general one expects the strict inclusion $\mathcal{F}' \supset \mathcal{F}$ to hold. Explicit examples can be derived from the construction of Hirzebruch [1984].

For simplicity, a compactification as in Theorem 2.2 will be referred to as a toroidal *Hummel–Schroeder* compactification.

Proposition 2.3. *Let M be a finite-volume complex-hyperbolic surface which admits a toroidal Hummel–Schroeder compactification. Then the Euler number of $\overline{M}$ is strictly positive.*

Proof. The idea for the proof goes back to an unpublished result of J. Milnor about the Euler number of closed four-dimensional Riemannian manifolds having sectional curvatures along perpendicular planes of the same sign; see [Chern 1955]. Let $(\overline{M}, g)$ be the Riemannian manifold obtained by closing the cusps of M under the condition of nonpositive curvature [Hummel and Schroeder 1996]. Let Ω be its curvature matrix. We can always choose [Chern 1955] a orthonormal frame $\{e_i\}_{i=1}^4$ such that $R_{1231} = R_{1241} = R_{1232} = R_{1242} = R_{1332} = R_{1341} = 0$. Hence

$$\begin{aligned}\mathrm{Pf}(\Omega) &= \Omega^1_2 \wedge \Omega^3_4 - \Omega^1_3 \wedge \Omega^2_4 + \Omega^1_4 \wedge \Omega^2_3 \\ &= \left(R_{1221}R_{3443} + R^2_{1243} + R_{1331}R_{2442} + R^2_{1342} + R_{1441}R_{2332} + R^2_{1234}\right) d\mu_g,\end{aligned}$$

where $\mathrm{Pf}(\Omega)$ is the Pfaffian of the skew symmetric matrix Ω. The statement is now a consequence of Chern–Weil theory. □

We can now use the Kodaira–Enriques classification of closed smooth surfaces [Barth et al. 2004] to derive the following theorem. The proof is in the spirit of the theory of nonpositively curved spaces.

Theorem 2.4. *Let M be a finite-volume complex-hyperbolic surface which admits a toroidal Hummel–Schroeder compactification. Then $\overline{M}$ is a surface of general type without rational curves.*

Proof. Since $\overline{M}$ admits a Riemannian metric of nonpositive sectional curvature, the Cartan–Hadamard theorem [Petersen 2006] implies that the universal cover of $\overline{M}$ is diffeomorphic to the four-dimensional euclidean space. Consequently,

$\overline{M}$ is aspherical and then it cannot contain rational curves. Moreover, the second Betti number of $\overline{M}$ is even since by construction it admits a Kähler metric. By the Kodaira–Enriques classification [Barth et al. 2004] we conclude that the Kodaira dimension of $\overline{M}$ cannot be negative.

From Proposition 2.3, we know that the Euler number of $\overline{M}$ is strictly positive. The minimal complex surfaces with Kodaira dimension equal to zero and positive Euler number are simply connected or with finite fundamental group. Since $\pi_1(\overline{M})$ is infinite, the Kodaira dimension of $\overline{M}$ is bigger or equal than one.

The fundamental group of an elliptic surface with positive Euler number is completely understood in terms of the *orbifold* fundamental group of the base of the elliptic fibration. More precisely, denoting by $\pi : S \to C$ the elliptic fibration, if S has no multiple fibers then π induces an isomorphism $\pi_1(S) \simeq \pi_1(C)$. If we allow multiple fibers we have the isomorphism $\pi_1(S) \simeq \pi_1^{\mathrm{Orb}}(C)$. For these results we refer to [Friedman and Morgan 1994]. We now show that $\overline{M}$ cannot be an elliptic surface. When S has multiple fibers, $\pi_1(S)$ always has torsion and then it cannot be the fundamental group of a nonpositively curved manifold. If we assume $\pi_1(\overline{M}) \simeq \pi_1(C)$, the fact that $\pi_1(\overline{M})$ grows exponentially [Avez 1970] forces the genus of the Riemann surface C to be bigger or equal than two. Since all closed geodesics in a manifold of nonpositive curvature are essential in π_1, we have that the fundamental group of the flats introduced in the compactification injects into $\pi_1(\overline{M})$ and then by assumption into $\pi_1(C)$. By elementary hyperbolic geometry this would imply that $\mathbb{Z} \oplus \mathbb{Z}$ acts as a discrete subgroup of $\mathbb{R}$, which is clearly impossible. □

Corollary 2.5. *A toroidal Hummel–Schroeder compactification has ample canonical line bundle.*

Proof. By Theorem 2.4 we know that $\overline{M}$ is a minimal surface of general type without rational curves. The corollary follows from Nakai's criterion for ampleness of divisors on surfaces [Barth et al. 2004]. More precisely, since for a minimal surface of general type the self-intersection of the canonical divisor is strictly positive [ibid.], it suffices to show that $K_{\overline{M}} \cdot E > 0$ for any effective divisor E. Thus, let E be an irreducible divisor and assume $K_{\overline{M}} \cdot E = 0$. By the Hodge index theorem we must have $E \cdot E < 0$. By the adjunction formula E must be isomorphic to a smooth rational curve with self-intersection -2. □

In the arithmetic case, part of the results contained in Theorem 2.4 can be derived from a theorem of Tai, see [Ash et al. 2010]. Furthermore, similar results for the so-called *Picard* modular surfaces are obtained by Holzapfel [1980].

3. Examples

In this section we present examples of surfaces of general type which do not admit nonpositively curved Kähler metrics, but such that their underlying smooth manifolds do admit Riemannian metrics with nonpositive Riemannian curvature. In order to do this one needs to understand the restrictions imposed by the nonpositive curvature assumption on the holomorphic curvature tensor.

Thus, define

$$p = 2\operatorname{Re}\xi \quad \text{and} \quad q = 2\operatorname{Re}\eta$$

where

$$\xi = \xi^\alpha \partial_\alpha \quad \text{and} \quad \eta = \eta^\alpha \partial_\alpha.$$

In real coordinates we have

$$R(p,q,q,p) = R_{hijk}p^h q^i q^j p^k$$

while in complex terms

$$\begin{aligned} &R\big(\xi+\overline{\xi},\eta+\overline{\eta},\eta+\overline{\eta},\xi+\overline{\xi}\big) \\ &\qquad = R\big(\xi,\overline{\eta},\eta,\overline{\xi}\big)+R\big(\xi,\overline{\eta},\overline{\eta},\xi\big)+R\big(\overline{\xi},\eta,\eta,\overline{\xi}\big)+R\big(\overline{\xi},\eta,\overline{\eta},\xi\big). \end{aligned}$$

We then have

$$\begin{aligned} &R_{hijk}p^h q^i q^j p^k \\ &\quad = R_{\alpha\overline{\beta}\gamma\overline{\delta}}\,\xi^\alpha\eta^{\overline{\beta}}\eta^\gamma\xi^{\overline{\delta}} + R_{\alpha\overline{\beta}\overline{\gamma}\delta}\,\xi^\alpha\eta^{\overline{\beta}}\eta^{\overline{\gamma}}\xi^\delta + R_{\overline{\alpha}\beta\gamma\overline{\delta}}\,\xi^{\overline{\alpha}}\eta^\beta\eta^\gamma\xi^{\overline{\delta}} + R_{\overline{\alpha}\beta\overline{\gamma}\delta}\xi^{\overline{\alpha}}\eta^\beta\eta^{\overline{\gamma}}\xi^\delta \\ &\quad = R_{\alpha\overline{\beta}\gamma\overline{\delta}}\,\xi^\alpha\eta^{\overline{\beta}}\eta^\gamma\xi^{\overline{\delta}} - R_{\alpha\overline{\beta}\gamma\overline{\delta}}\,\xi^\alpha\eta^{\overline{\beta}}\eta^{\overline{\delta}}\xi^\gamma - R_{\alpha\overline{\beta}\gamma\overline{\delta}}\,\xi^{\overline{\beta}}\eta^\alpha\eta^\gamma\xi^{\overline{\delta}} + R_{\alpha\overline{\beta}\gamma\overline{\delta}}\,\xi^{\overline{\beta}}\eta^\alpha\eta^{\overline{\delta}}\xi^\gamma \\ &\quad = R_{\alpha\overline{\beta}\gamma\overline{\delta}}\big(\xi^\alpha\eta^{\overline{\beta}}\eta^\gamma\xi^{\overline{\delta}} - \xi^\alpha\eta^{\overline{\beta}}\eta^{\overline{\delta}}\xi^\gamma - \xi^{\overline{\beta}}\eta^\alpha\eta^\gamma\xi^{\overline{\delta}} + \xi^{\overline{\beta}}\eta^\alpha\eta^{\overline{\delta}}\xi^\gamma\big) \\ &\quad = R_{\alpha\overline{\beta}\gamma\overline{\delta}}\big(\xi^\alpha\eta^{\overline{\beta}} - \eta^\alpha\xi^{\overline{\beta}}\big)\overline{\big(\xi^\delta\eta^{\overline{\gamma}} - \eta^\delta\xi^{\overline{\gamma}}\big)}. \end{aligned}$$

If we assume the Riemannian sectional curvature to be nonpositive we have

$$R_{hijk}p^h q^i q^j p^k = R_{\alpha\overline{\beta}\gamma\overline{\delta}}\big(\xi^\alpha\eta^{\overline{\beta}} - \eta^\alpha\xi^{\overline{\beta}}\big)\overline{\big(\xi^\delta\eta^{\overline{\gamma}} - \eta^\delta\xi^{\overline{\gamma}}\big)} \le 0.$$

In complex dimension two, the right hand side of the above equality reduces (after some manipulations) to

$$\begin{aligned} &R_{\alpha\overline{\beta}\gamma\overline{\delta}}\big(\xi^\alpha\eta^{\overline{\beta}} - \eta^\alpha\xi^{\overline{\beta}}\big)\overline{\big(\xi^\delta\eta^{\overline{\gamma}} - \eta^\delta\xi^{\overline{\gamma}}\big)} \\ &\qquad = R_{1\overline{1}1\overline{1}}\big|\xi^1\eta^{\overline{1}} - \eta^1\xi^{\overline{1}}\big|^2 + 4\operatorname{Re}\big\{R_{1\overline{1}1\overline{2}}\big(\xi^1\eta^{\overline{1}} - \eta^1\xi^{\overline{1}}\big)\overline{\big(\xi^2\eta^{\overline{1}} - \eta^2\xi^{\overline{1}}\big)}\big\} \\ &\qquad\qquad + 2R_{1\overline{1}2\overline{2}}\big\{\big|\xi^1\eta^{\overline{2}} - \eta^1\xi^{\overline{2}}\big|^2 + \operatorname{Re}\big(\xi^1\eta^{\overline{1}} - \eta^1\xi^{\overline{1}}\big)\overline{\big(\xi^2\eta^{\overline{2}} - \eta^2\xi^{\overline{2}}\big)}\big\} \\ &\qquad\qquad + 2\operatorname{Re}\big\{R_{1\overline{2}1\overline{2}}\big(\xi^1\eta^{\overline{2}} - \eta^1\xi^{\overline{2}}\big)\overline{\big(\xi^2\eta^{\overline{1}} - \eta^2\xi^{\overline{1}}\big)}\big\} \\ &\qquad\qquad + 4\operatorname{Re}\big\{R_{2\overline{2}1\overline{2}}\big(\xi^2\eta^{\overline{2}} - \eta^2\xi^{\overline{2}}\big)\overline{\big(\xi^2\eta^{\overline{1}} - \eta^2\xi^{\overline{1}}\big)}\big\} + R_{2\overline{2}2\overline{2}}\big|\xi^2\eta^{\overline{2}} - \eta^2\xi^{\overline{2}}\big|^2. \end{aligned}$$

Following Mostow and Siu [1980], we choose the ansatz

$$\xi^1 = ia, \quad \xi^2 = -i, \quad \eta^1 = a, \quad \eta^2 = 1$$

where a is a real number. We get the inequality

$$R_{1\bar{1}1\bar{1}}4a^4 - 2R_{1\bar{1}2\bar{2}}4a^2 + R_{2\bar{2}2\bar{2}}4 \leq 0.$$

Since nonpositive Riemannian sectional curvature implies nonpositive holomorphic sectional curvature, we conclude that

$$(R_{1\bar{1}2\bar{2}})^2 \leq R_{1\bar{1}1\bar{1}} R_{2\bar{2}2\bar{2}}. \tag{1}$$

Theorem 3.1. *A toroidal Hummel–Schroeder compactification does not admit any Kähler metric with nonpositive Riemannian sectional curvature.*

Proof. Let us proceed by contradiction. Consider one of the elliptic divisors added in the compactification. By the properties of submanifolds of a Kähler manifold [Kobayashi and Nomizu 1969], we have that the holomorphic sectional curvature tangent to the elliptic divisor has to be zero. Let us denote such a holomorphic sectional curvature by $R_{1\bar{1}1\bar{1}}$. By the inequality (1), we conclude that $R_{1\bar{1}2\bar{2}} = 0$. As a result, the Ricci curvature tangent to the elliptic divisor has to be zero. We conclude that

$$K_{\overline{M}} \cdot \Sigma = \int_\Sigma c_1(K_{\overline{M}}) = 0,$$

which contradicts the ampleness of $K_{\overline{M}}$, see Corollary 2.5. □

Combining Theorems 2.4 and 3.1 with Corollary 2.5, we have thus proved Theorem A.

4. Projective-algebraicity of minimal compactifications

Let $\overline{M}$ be a smooth toroidal compactification of a finite-volume complex-hyperbolic surface M and let Σ denote the compactifying divisor. The set Σ is exceptional and it can be blow down. The resulting complex surface, with isolated normal singularities, is usually referred as the minimal compactification of M [Siu and Yau 1982]. In this section we address the problem of the projective-algebraicity of minimal compactifications of finite-volume complex-hyperbolic surfaces. This is motivated by a beautiful example of Hironaka, see [Hartshorne 1977, p. 417], which shows that by contracting a smooth elliptic divisor on an algebraic surface one can obtain a nonprojective complex space. In the arithmetic case, the projective-algebraicity of minimal compactifications of finite-volume complex-hyperbolic surfaces is known by the work of Baily and Borel, see [Borel and Ji 2006].

For completeness, we recall the theory of semistable curves on algebraic surfaces and logarithmic pluricanonical maps as developed by Sakai [1980].

Let $\overline{M}$ be a smooth projective surface. Let Σ be a reduced divisor having simple normal crossings on $\overline{M}$.

Definition 4.1. The pair $(\overline{M}, \Sigma)$ is called minimal if $\overline{M}$ does not contain an exceptional curve E of the first kind such that $E \cdot \Sigma \leq 1$.

We consider the logarithmic canonical line bundle $\mathcal{L} = K_{\overline{M}} + \Sigma$ associated to Σ. Given any integer k, define $\overline{P}_m = \dim H^0(\overline{M}, \mathbb{O}(m\mathcal{L}))$. If $\overline{P}_m > 0$, we define the m-th *logarithmic canonical* map $\Phi_{m\mathcal{L}}$ of the pair $(\overline{M}, \Sigma)$ by

$$\Phi_{m\mathcal{L}}(x) = [s_1(x), \dots, s_N(x)],$$

for any $x \in \overline{M}$ and where $s_1, \dots, s_N$ is a basis for the vector space $H^0(\overline{M}, \mathbb{O}(m\mathcal{L}))$. At this point one introduces the notion of logarithmic Kodaira dimension exactly as in the closed smooth case. We denote this numerical invariant by $\overline{k}(M)$ where $M = \overline{M} \backslash \Sigma$. We refer to [Iitaka 1982] for further details.

Definition 4.2. A curve Σ is semistable if it has only normal crossings and each smooth rational component of Σ intersects the other components of Σ in more than one point.

We next give a numerical criterion for a minimal semistable pair $(\overline{M}, \Sigma)$ to be of log-general type.

Proposition 4.3 [Sakai 1980]. *Given a minimal semistable pair $(\overline{M}, \Sigma)$ we have that $\overline{k}(M) = 2$ if and only if $\mathcal{L}$ is numerically effective and $\mathcal{L}^2 > 0$.*

In what follows, we denote by $\mathscr{E}$ the set of irreducible curves E in $\overline{M}$ such that $\mathcal{L} \cdot E = 0$.

Theorem 4.4 [Sakai 1980]. *Let $(\overline{M}, \Sigma)$ be a minimal semistable pair of log-general type. The map $\Phi_{m\mathcal{L}}$ is then an embedding modulo $\mathscr{E}$ for any $m \geq 5$.*

It is then necessary to characterize the irreducible divisors in $\mathscr{E}$. In particular, we need the following proposition.

Proposition 4.5. *Let $(\overline{M}, \Sigma)$ be a minimal semistable pair with $\overline{k}(M) = 2$. Let E be an irreducible curve such that $\mathcal{L} \cdot E = 0$. If E is not contained in Σ then $E \simeq \mathbb{C}P^1$ and $E \cdot E = -2$.*

Proof. Under these assumptions we know that $\mathcal{L}^2 > 0$. By the Hodge index theorem

$$\mathcal{L}^2 > 0, \quad \mathcal{L} \cdot E = 0 \quad \Longrightarrow \quad E^2 < 0.$$

But now $\mathcal{L} \cdot E = 0$ which implies

$$K_{\overline{M}} \cdot E = -\Sigma \cdot E \leq 0.$$

We then have $K_{\overline{M}} \cdot E = 0$ if and only if E does not intersect Σ. In this case $p_a(E) = 0$ and then $E \simeq \mathbb{C}P^1$ and $E^2 = -2$. Assume now that $K_{\overline{M}} \cdot E < 0$, then $K_{\overline{M}} \cdot E = E^2 = -1$ and therefore E is an exceptional curve of the first kind such that $E \cdot \Sigma = 1$. This contradicts the minimality of the pair $(\overline{M}, \Sigma)$. □

We are now ready to prove the main results of this section. Let $\mathbb{C}\mathcal{H}^2/\Gamma$ be a finite-volume complex-hyperbolic surface that admits a smooth toroidal compactification as in Theorem 2.4. We then have that $\overline{\mathbb{C}\mathcal{H}^2/\Gamma}$ is a surface of general type with compactification divisor consisting of smooth disjoint elliptic curves.

Proposition 4.6. *Let $\overline{M}$ be a minimal surface of general type. Let Σ be a reduced divisor whose irreducible components consist of disjoint smooth elliptic curves. Then $(\overline{M}, \Sigma)$ is a minimal semistable pair with $\overline{k}(M) = 2$.*

Proof. Recall that the canonical divisor of any minimal complex surface of non-negative Kodaira dimension is numerically effective [Barth et al. 2004]. It follows that the adjoint divisor $\mathcal{L}$ is numerically effective. An elliptic curve on a minimal surface of general type has negative self intersection. Moreover, for a minimal surface of general type it is known that the self-intersection of the canonical divisor is strictly positive [ibid.]. By the adjunction formula, we have $\mathcal{L}^2 = K^2_{\overline{M}} - \Sigma^2 > 0$. By Proposition 4.3, we conclude that $\overline{k}(M) = 2$. □

Let $\mathbb{C}\mathcal{H}^2\backslash\Gamma_1$ be a finite-volume complex-hyperbolic surface which admits a smooth toroidal compactification $\overline{M}_1$. Let $(\overline{M}_1, \Sigma_1)$ be the associated minimal semistable pair. By Theorem A, we can find a normal subgroup $\Gamma_2 \lhd \Gamma_1$ of finite index such that the toroidal compactification $\overline{M}_2$ of $\mathbb{C}\mathcal{H}^2/\Gamma_2$ is a minimal surface of general type with compactification divisor Σ_2. Since

$$\pi : \mathbb{C}\mathcal{H}^2/\Gamma_2 \to \mathbb{C}\mathcal{H}^2/\Gamma_1$$

is an unramified covering we conclude that $\overline{k}(M_1) = \overline{k}(M_2)$ [Iitaka 1982]. But by Proposition 4.6 we know that $\overline{k}(M_2) = 2$, it follows that $(\overline{M}_1, \Sigma_1)$ is a minimal semistable pair of log-general type. Let us summarize this argument into a proposition.

Proposition 4.7. *Let $(\overline{M}, \Sigma)$ be a smooth pair arising as the toroidal compactification of a finite-volume complex-hyperbolic surface. The pair $(\overline{M}, \Sigma)$ is minimal and log-general.*

The following theorem is the main result of the present section.

Theorem B. *Let $(\overline{M}, \Sigma)$ be a smooth pair arising as the toroidal compactification of a finite-volume complex-hyperbolic surface. Then the associated minimal compactification is projective algebraic.*

Proof. By Proposition 4.7, the minimal pair $(\overline{M}, \Sigma)$ is log-general. By Theorem 4.4 we know that $\Phi_{m\mathcal{L}}$ is an embedding modulo $\mathcal{E}$ for any $m \geq 5$. We clearly have that Σ is contained in $\mathcal{E}$. We claim that there are no other divisors in $\mathcal{E}$. Assume the contrary. By Proposition 4.5, any other curve in $\mathcal{E}$ must be a smooth rational divisor

E with self-intersection minus two. The adjunction formula gives $K_{\overline{M}} \cdot E = 0$ which implies $\Sigma \cdot E = 0$. This is clearly impossible. By Theorem 4.4 for $m \geq 5$, the map

$$\Phi_{m\mathscr{L}} : \overline{M} \to \mathbb{C}P^{N-1}$$

gives a realization of the minimal compactification as a projective-algebraic variety. $\square$

For an approach to the projective-algebraicity problem through L^2-estimates for the $\overline{\partial}$-operator we refer to [Mok 2009].

Acknowledgments

I would like to thank Professor Claude LeBrun for his constant support and for constructive comments on the paper. I also would like to thank Professor Klaus Hulek for some useful bibliographical suggestions and the referee for pertinent comments on the manuscript.

References

[Ash et al. 2010] A. Ash, D. Mumford, M. Rapoport, and Y.-S. Tai, *Smooth compactifications of locally symmetric varieties*, 2nd ed., Cambridge Mathematical Library **10**, Cambridge University Press, 2010. MR 2010m:14067 Zbl 1209.14001

[Avez 1970] A. Avez, "Variétés riemanniennes sans points focaux", *C. R. Acad. Sci. Paris Sér. A-B* **270** (1970), A188–A191. MR 41 #961 Zbl 0188.26402

[Barth et al. 2004] W. P. Barth, K. Hulek, C. A. M. Peters, and A. Van de Ven, *Compact complex surfaces*, 2nd ed., Ergebnisse der Math. **4**, Springer, Berlin, 2004. MR 2004m:14070 Zbl 1036.14016

[Borel and Ji 2006] A. Borel and L. Ji, *Compactifications of symmetric and locally symmetric spaces*, Math. Theory Appl. **16**, Birkhäuser, Boston, 2006. MR 2007d:22030 Zbl 1100.22001

[Chern 1955] S.-s. Chern, "On curvature and characteristic classes of a Riemann manifold", *Abh. Math. Sem. Univ. Hamburg* **20** (1955), 117–126. MR 17,783e Zbl 0066.17003

[Deligne and Mostow 1993] P. Deligne and G. D. Mostow, *Commensurabilities among lattices in* PU$(1, n)$, Annals of Mathematics Studies **132**, Princeton University Press, 1993. MR 95a:22013 Zbl 0826.22011

[Eberlein 1980] P. Eberlein, "Lattices in spaces of nonpositive curvature", *Ann. of Math.* (2) **111**:3 (1980), 435–476. MR 82m:53040 Zbl 0401.53015

[Eberlein 1996] P. B. Eberlein, *Geometry of nonpositively curved manifolds*, Chicago Lectures in Mathematics **7**, University of Chicago Press, 1996. MR 98h:53002 Zbl 0883.53003

[Friedman and Morgan 1994] R. Friedman and J. W. Morgan, *Smooth four-manifolds and complex surfaces*, Ergebnisse der Math. **27**, Springer, Berlin, 1994. MR 95m:57046 Zbl 0817.14017

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Hirzebruch 1984] F. Hirzebruch, "Chern numbers of algebraic surfaces: an example", *Math. Ann.* **266**:3 (1984), 351–356. MR 85j:14069 Zbl 0504.14030

[Holzapfel 1980] R.-P. Holzapfel, "A class of minimal surfaces in the unknown region of surface geography", *Math. Nachr.* **98** (1980), 211–232. MR 83a:14031 Zbl 0474.14022

[Hummel 1998] C. Hummel, "Rank one lattices whose parabolic isometries have no rotational part", *Proc. Amer. Math. Soc.* **126**:8 (1998), 2453–2458. MR 98j:22014 Zbl 0898.22005

[Hummel and Schroeder 1996] C. Hummel and V. Schroeder, "Cusp closing in rank one symmetric spaces", *Invent. Math.* **123**:2 (1996), 283–307. MR 97e:53098 Zbl 0860.53025

[Iitaka 1982] S. Iitaka, *Algebraic geometry*, Graduate Texts in Mathematics **76**, Springer, New York, 1982. MR 84j:14001 Zbl 0491.14006

[Kobayashi and Nomizu 1969] S. Kobayashi and K. Nomizu, *Foundations of differential geometry*, vol. II, Wiley, New York, 1969. MR 97c:53001b Zbl 0175.48504

[Mal'tsev 1940] A. Mal'tsev, "On the faithful representation of infinite groups by matrices.", *Am. Math. Soc. Transl.* (2) **45** (1940), 1–18. Zbl 0158.02905

[Margulis 1984] G. A. Margulis, "Arithmeticity of the irreducible lattices in the semisimple groups of rank greater than 1", *Invent. Math.* **76**:1 (1984), 93–120. MR 85j:22021 Zbl 0551.20028

[Mok 2009] N. Mok, "Projective-algebraicity of minimal compactifications of complex-hyperbolic space forms of finite volume", preprint, University of Hong Kong, 2009.

[Mostow and Siu 1980] G. D. Mostow and Y. T. Siu, "A compact Kähler surface of negative curvature not covered by the ball", *Annals of Mathematics. Second Series* **112**:2 (1980), 321–360. MR 82f:53075 Zbl 0453.53047

[Petersen 2006] P. Petersen, *Riemannian geometry*, 2nd ed., Graduate Texts in Mathematics **171**, Springer, New York, 2006. MR 2007a:53001 Zbl 1220.53002

[Sakai 1980] F. Sakai, "Semistable curves on algebraic surfaces and logarithmic pluricanonical maps", *Math. Ann.* **254**:2 (1980), 89–120. MR 82f:14031 Zbl 0431.14011

[Siu and Yau 1982] Y. T. Siu and S. T. Yau, "Compactification of negatively curved complete Kähler manifolds of finite volume", pp. 363–380 in *Seminar on Differential Geometry*, edited by S. T. Yau, Ann. of Math. Stud. **102**, Princeton Univ. Press, 1982. MR 83g:32027 Zbl 0493.53049

Received April 1, 2011. Revised November 2, 2011.

LUCA FABRIZIO DI CERBO
MATHEMATICS DEPARTMENT
DUKE UNIVERSITY
BOX 90320
DURHAM, NC 27708
UNITED STATES
luca@math.duke.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

CHARACTER ANALOGUES OF RAMANUJAN-TYPE INTEGRALS INVOLVING THE RIEMANN Ξ-FUNCTION

ATUL DIXIT

A new class of integrals involving the product of Ξ-functions associated with primitive Dirichlet characters is considered. These integrals give rise to transformation formulas of the type

$$F(z, \alpha, \chi) = F(-z, \beta, \bar{\chi}) = F(-z, \alpha, \bar{\chi}) = F(z, \beta, \chi),$$

where $\alpha\beta = 1$. New character analogues of the Ramanujan–Guinand formula, the Koshliakov's formula, and a transformation formula of Ramanujan, as well as its recent generalization, are shown as particular examples. Finally, character analogues of a conjecture of Ramanujan, and Hardy and Littlewood involving infinite series of Möbius functions are derived.

1. Introduction

Modular transformations are ubiquitous in Ramanujan's notebooks [1957] and in his "Lost Notebook" [1988]. Ramanujan usually expressed them in a symmetric way, and they were valid under the conditions $\alpha\beta = \pi$, or $\alpha\beta = \pi^2, \ldots$. In the same spirit, on page 220 in one of the manuscripts of Ramanujan in the handwriting of Watson [Ramanujan 1988], one finds the following beautiful claim.

Theorem 1.1. *Define*

$$\lambda(x) := \psi(x) + \frac{1}{2x} - \log x,$$

where

$$\psi(x) := \frac{\Gamma'(x)}{\Gamma(x)} = -\gamma - \sum_{m=0}^{\infty}\Big(\frac{1}{m+x} - \frac{1}{m+1}\Big)$$

is the logarithmic derivative of the Gamma function. Let the Riemann ξ-function be defined by

$$\xi(s) := (s-1)\pi^{-s/2}\Gamma(1+\tfrac{1}{2}s)\zeta(s),$$

2010 Mathematics Subject Classification.
MSC2010: primary 11M06; secondary 11M35.
Keywords: Dirichlet character, Dirichlet L-function, modified Bessel function, Möbius function, Mellin transform, Ramanujan, Hardy, Littlewood, Koshliakov, Guinand.

and let

$$\Xi(t) := \xi(\tfrac{1}{2} + it)$$

be the Riemann Ξ*-function. If* α *and* β *are positive numbers such that* $\alpha\beta = 1$, *then*

$$\sqrt{\alpha}\left(\frac{\gamma - \log(2\pi\alpha)}{2\alpha} + \sum_{k=1}^{\infty} \lambda(k\alpha)\right) = \sqrt{\beta}\left(\frac{\gamma - \log(2\pi\beta)}{2\beta} + \sum_{k=1}^{\infty} \lambda(k\beta)\right) \tag{1-1}$$

$$= -\frac{1}{\pi^{3/2}} \int_0^\infty \left|\Xi(\tfrac{1}{2}t)\Gamma\left(\frac{-1+it}{4}\right)\right|^2 \frac{\cos((t/2)\log\alpha)}{1+t^2}\,dt,$$

where γ *denotes Euler's constant.*

This identity is of a special kind since it contains not only a modular transformation, but also a beautiful integral involving the Riemann Ξ-function. In fact, the invariance of the integral in (1-1) under the map $\alpha \to \beta$ establishes the equality of the first and the second expressions in (1-1). This is used in [Berndt and Dixit 2010] to prove the claim above and in [Dixit 2010; 2011a; 2011b] to obtain many transformation formulas of the type $F(\alpha) = F(\beta)$ or $F(z, \alpha) = F(z, \beta)$, where $\alpha\beta = 1$ and an integral involving the Riemann Ξ-function is always linked to them. This gives new identities involving infinite series of the Hurwitz zeta function as well as extensions of some well-known formulas like the Ramanujan–Guinand formula, discovered first by Ramanujan [1988, p. 253] and later in a different but equivalent form by Guinand [1955], and a formula of Koshliakov [1928], also in the lost notebook [Ramanujan 1988, p. 254]; see [Berndt et al. 2008; Dixit 2011b]. For example, we mention the following generalization of Theorem 1.1:

Theorem 1.2 [Dixit 2011a; 2011b]. *Let* $-1 < \operatorname{Re} z < 1$. *Define* $\varphi(z, x)$ *by*

$$\varphi(z, x) = \zeta(z+1, x) - \frac{x^{-z}}{z} - \tfrac{1}{2}x^{-z-1},$$

where $\zeta(z, x)$ *denotes the Hurwitz zeta function. If* α *and* β *are any positive numbers such that* $\alpha\beta = 1$, *then*

$$\alpha^{\frac{z+1}{2}}\left(\sum_{n=1}^{\infty} \varphi(z, n\alpha) - \frac{\zeta(z+1)}{2\alpha^{z+1}} - \frac{\zeta(z)}{\alpha z}\right) \tag{1-2}$$

$$= \beta^{(z+1)/2}\left(\sum_{n=1}^{\infty} \varphi(z, n\beta) - \frac{\zeta(z+1)}{2\beta^{z+1}} - \frac{\zeta(z)}{\beta z}\right)$$

$$= \frac{8(4\pi)^{(z-3)/2}}{\Gamma(z+1)} \int_0^\infty \Gamma\left(\frac{z-1+it}{4}\right)\Gamma\left(\frac{z-1-it}{4}\right)$$

$$\times\, \Xi\left(\frac{t+iz}{2}\right)\Xi\left(\frac{t-iz}{2}\right)\frac{\cos((t/2)\log\alpha)}{(z+1)^2+t^2}\,dt,$$

where $\Xi(t)$ *is the Riemann* Ξ*-function.*

Another example of a transformation formula of the type $F(z,\alpha)=F(z,\beta)$ along with an integral involving Riemann's Ξ-functions is the following extended version of the Ramanujan–Guinand formula just mentioned:

Theorem 1.3 [Dixit 2011b, Theorem 1.4]. *Let $K_\nu(s)$ denote the modified Bessel function of order ν, let γ denote Euler's constant and let $\sigma_k(n)=\sum_{d|n} d^k$. Let $-1<\operatorname{Re} z<1$. Then if α and β are positive numbers such that $\alpha\beta=1$, we have*

$$\begin{aligned}\sqrt{\alpha}\Bigg(&\alpha^{z/2-1}\pi^{-z/2}\Gamma\Big(\frac{z}{2}\Big)\zeta(z)\\ &+\alpha^{-z/2-1}\pi^{z/2}\Gamma\Big(\frac{-z}{2}\Big)\zeta(-z)-4\sum_{n=1}^{\infty}\sigma_{-z}(n)n^{z/2}K_{z/2}(2n\pi\alpha)\Bigg)\\ =\sqrt{\beta}\Bigg(&\beta^{z/2-1}\pi^{-z/2}\Gamma\Big(\frac{z}{2}\Big)\zeta(z)\\ &+\beta^{-z/2-1}\pi^{z/2}\Gamma\Big(\frac{-z}{2}\Big)\zeta(-z)-4\sum_{n=1}^{\infty}\sigma_{-z}(n)n^{z/2}K_{z/2}(2n\pi\beta)\Bigg)\\ =-\frac{32}{\pi}&\int_0^\infty \Xi\Big(\frac{t+iz}{2}\Big)\Xi\Big(\frac{t-iz}{2}\Big)\frac{\cos((t/2)\log\alpha)}{(t^2+(z+1)^2)(t^2+(z-1)^2)}\,dt.\end{aligned}\tag{1-3}$$

Letting $z\to 0$ in (1-3) then gives an extended version of Koshliakov's formula:

Theorem 1.4 [Dixit 2010]. *Let $d(n)$ denote the number of positive divisors of n, and let $K_0(n)$ denote the modified Bessel function of order 0. If α and β are positive numbers such that $\alpha\beta=1$, then*

$$\begin{aligned}\sqrt{\alpha}\Bigg(\frac{\gamma-\log(4\pi\alpha)}{\alpha}-4\sum_{n=1}^{\infty}d(n)K_0(2\pi n\alpha)\Bigg)&\\ =\sqrt{\beta}\Bigg(\frac{\gamma-\log(4\pi\beta)}{\beta}&-4\sum_{n=1}^{\infty}d(n)K_0(2\pi n\beta)\Bigg)\\ =-\frac{32}{\pi}\int_0^\infty&\frac{(\Xi(t/2))^2\cos(\frac{1}{2}t\log\alpha)\,dt}{(1+t^2)^2}.\end{aligned}$$

By an "extended version", we mean that the original identity known before is linked to an integral involving the Riemann Ξ-function.

N. S. Koshliakov [1934a; 1934b; 1936; 1949; 1954][1] was another mathematician who did significant research in this area after Ramanujan. Besides using contour integration, Mellin transforms, and several summation formulas that he developed, he frequently used a method similar to that developed by Ramanujan [1915; 1927, pp. 72–77] to obtain old and new transformation formulas of the form

[1]For the genesis of the monograph [Koshlyakov 1949], written under the patronymic "N. S. Sergeev", see [Bogolyubov et al. 1990, pp. 198–199].

$F(\alpha) = F(\beta)$, where $\alpha\beta = k$ for some constant k. He obtained deep generalizations of well-known formulas of Ramanujan and of Hardy (such as [Hardy 1915, (2)]), some of them being analogues in rational and number fields. Koshliakov [1934c; 1937] also used Fourier's integral theorem to obtain expressions for the Riemann Ξ-function, a method also enunciated in [Ramanujan 1915]. Around the same time, Ferrar [1936] also worked on transformation formulas of this kind.

As can be seen from (1-1), the general form of the integrals giving rise to formulas of the type $F(\alpha) = F(\beta)$, where $\alpha\beta = 1$, is

$$\int_0^\infty f\left(\frac{t}{2}\right)\Xi\left(\frac{t}{2}\right)\cos\mu t\,dt,$$

for μ real and $f(t) = \phi(it)\phi(-it)$, where ϕ is analytic in t as a function of a real variable. This integral is mentioned in [Titchmarsh 1986, p. 35]. Similarly, from (1-2) and (1-3), it is clear that the general form of the integrals giving rise to identities of the type $F(z, \alpha) = F(z, \beta)$, where $\alpha\beta = 1$, is

$$\int_0^\infty f\left(z, \frac{t}{2}\right)\Xi\left(\frac{t+iz}{2}\right)\Xi\left(\frac{t-iz}{2}\right)\cos\mu t\,dt, \tag{1-4}$$

for μ real and $f(z, t) = \phi(z, it)\phi(z, -it)$, where ϕ is both analytic in t as a function of a real variable and analytic in z in some complex domain. An integral of this kind was first introduced by Ramanujan [1915].

In this article, we find character analogues of all of the above-mentioned theorems. The character analogue of the Ramanujan–Guinand formula, and hence of Koshliakov's formula, given here differs from the ones established in [Berndt et al. 2011]. Throughout this article, we will be concerned with the principal branch of the logarithm. Since we frequently use the functional equation for L-functions (see (1-10) below), we work only with a primitive, nonprincipal Dirichlet character χ modulo q, where q is the period of the character; see [Apostol 1972, Theorem 1]. It is easy to see that its conjugate character $\overline{\chi}$ is also a primitive, nonprincipal character modulo q and $\overline{\chi}$ is even (odd) if and only if χ is even (respectively odd). Let $L(s, \chi)$ denote the Dirichlet L-function defined by $L(s, \chi) = \sum_{n=1}^\infty \chi(n)/n^s$ for $\operatorname{Re} s > 1$. This series converges conditionally for $0 < \operatorname{Re} s < 1$. Also, it can be analytically continued to an entire function of s. Let $G(\chi) := G(1, \chi)$, where $G(n, \chi)$ is the Gauss sum defined by

$$G(n, \chi) := \sum_{m=1}^{q} \chi(m)e^{2\pi imn/q}.$$

We know that [Apostol 1976, p. 168]

$$|G(\chi)|^2 = q, \tag{1-5}$$

and it is easy to see that

$$\overline{G(\chi)} = \begin{cases} G(\overline{\chi}) & \text{for } \chi \text{ even,} \\ -G(\overline{\chi}) & \text{for } \chi \text{ odd.} \end{cases} \tag{1-6}$$

Define b as follows:

$$b = \begin{cases} 0 & \text{for } \chi(-1) = 1, \\ 1 & \text{for } \chi(-1) = -1. \end{cases} \tag{1-7}$$

Then the function $\xi(s, \chi)$ is defined by

$$\xi(s, \chi) := \left(\frac{\pi}{q}\right)^{-(s+b)/2} \Gamma\left(\frac{s+b}{2}\right) L(s, \chi), \tag{1-8}$$

and the analogue of the Riemann Ξ-function for Dirichlet characters is defined as

$$\Xi(t, \chi) := \xi\left(\tfrac{1}{2} + it, \chi\right). \tag{1-9}$$

L-functions satisfy the functional equation [Apostol 1976, p. 263]

$$L(1-s, \chi) = \frac{q^{s-1}\Gamma(s)}{(2\pi)^s}\left(e^{-\pi is/2} + \chi(-1)e^{\pi is/2}\right) G(\chi) L(s, \overline{\chi}), \tag{1-10}$$

which can be rephrased in terms of $\xi(s, \chi)$ as [Davenport 2000]

$$\xi(1-s, \overline{\chi}) = \epsilon(\chi)\xi(s, \chi), \tag{1-11}$$

where $\epsilon(\chi) = i^b q^{1/2}/G(\chi)$. By (1-5), $|\epsilon(\chi)| = 1$. Next, we note Stirling's formula in a vertical strip $\alpha \le \sigma \le \beta$, $s = \sigma + it$, namely,

$$|\Gamma(s)| = (2\pi)^{\frac{1}{2}} |t|^{\sigma - \frac{1}{2}} e^{-\frac{1}{2}\pi|t|}\left(1 + O\left(\frac{1}{|t|}\right)\right) \tag{1-12}$$

uniformly as $|t| \to \infty$. Now, using (1-10) and the fact that $|L(s, \chi)| = O(q|t|)$ for $\operatorname{Re} s \ge \frac{1}{2}$ [Davenport 2000, p. 82], we easily see that for $\operatorname{Re} s \ge -\delta$, $\delta > 0$, we have

$$L(s, \chi) = O\left(q^{\frac{3}{2}+\delta}|t|^{\frac{3}{2}+\delta}\right). \tag{1-13}$$

We will subsequently use this result.

Transformation formulas involving Dirichlet characters of the form

$$\sum_{n=1}^{\infty} \chi(n) f(n) = \sum_{n=1}^{\infty} \overline{\chi}(n) g(n),$$

where

$$g(x) = \begin{cases} \dfrac{2G(\chi)}{q} \displaystyle\int_0^{\to\infty} \cos\left(\frac{2\pi xt}{q}\right) f(t)\, dt & \text{for } \chi(-1) = 1, \\ \dfrac{-2iG(\chi)}{q} \displaystyle\int_0^{\to\infty} \sin\left(\frac{2\pi xt}{q}\right) f(t)\, dt & \text{for } \chi(-1) = -1, \end{cases}$$

were considered by Guinand [1941, Theorems 4 and 5], though he did not give particular examples. Here, we derive a character analogue of the integral in (1-4). Its general form is

$$\int_0^{\infty} f\left(z, \frac{t}{2}\right) \Xi\left(\frac{t+iz}{2}, \overline{\chi}\right) \Xi\left(\frac{t-iz}{2}, \chi\right) \cos\left(\tfrac{1}{2}t \log \alpha\right) dt, \tag{1-14}$$

where f is an even function of both the variables z and t. These integrals give rise to transformation formulas of the type $F(z, \alpha, \chi) = F(-z, \beta, \overline{\chi}) = F(-z, \alpha, \overline{\chi}) = F(z, \beta, \chi)$. Then, via Fourier's integral theorem, one may be able to obtain integral representations for $\Xi((t+iz)/2, \overline{\chi})\, \Xi((t-iz)/2, \chi)$, which are of independent interest. The character analogue of Theorem 1.3 is as follows.

Theorem 1.5. *Let* $-1 < \operatorname{Re} z < 1$ *and let* χ *denote a primitive, nonprincipal character modulo* q. *Let the number* b *be defined as in* (1-7). *Let* $K_\nu(z)$, $d(n)$, *and* γ *be defined as before, and let* α *and* β *be positive numbers such that* $\alpha\beta = 1$. *If*

$$F(z, \alpha, \chi) := \alpha^{b+\frac{1}{2}} \sum_{n=1}^{\infty} \chi(n) n^{-z/2+b} \left(\sum_{d|n} \overline{\chi}^2(d) d^z \right) K_{-z/2}\left(\frac{2\pi n\alpha}{q}\right),$$

then

$$\begin{aligned} F(z, \alpha, \chi) &= F(-z, \beta, \overline{\chi}) = F(-z, \alpha, \overline{\chi}) = F(z, \beta, \chi) \\ &= \frac{1}{8\pi} \int_0^{\infty} \Xi\left(\frac{t+iz}{2}, \overline{\chi}\right) \Xi\left(\frac{t-iz}{2}, \chi\right) \cos\left(\tfrac{1}{2}t \log \alpha\right) dt. \end{aligned} \tag{1-15}$$

Define $\psi(a, \chi)$ by

$$\psi(a, \chi) = -\sum_{n=1}^{\infty} \frac{\chi(n)}{n+a}, \tag{1-16}$$

where $a \in \mathbb{C} \setminus \mathbb{Z}_{<0}$. For a real character χ, this agrees with the character analogue of the psi function obtained by the logarithmic differentiation of the following Weierstrass product form of the character analogue of the gamma function for real characters derived by Berndt [1975]:

$$\Gamma(a, \chi) = e^{-aL(1,\chi)} \prod_{n=1}^{\infty} \left(1 + \frac{a}{n}\right)^{-\chi(n)} e^{a\chi(n)/n}.$$

The character analogue of the Hurwitz zeta function $\zeta(z, a)$ is given by [Berndt 1970, Example 3.2]

$$L(z, a, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{(n+a)^z}, \tag{1-17}$$

valid for $\operatorname{Re} z > 0$, and provided $a \in \mathbb{C} \setminus \mathbb{Z}_{<0}$. The above character analogue of the Hurwitz zeta function can also be obtained as the special case when $x = 0$ of the function $L(z, x, a, \chi)$ defined by [Berndt 1975]

$$L(z, x, a, \chi) := \sum_{n=0}^{\infty}{}' e^{2\pi i n x/q} \chi(n)(n+a)^{-z},$$

where the prime indicates that the term corresponding to $n = -a$ is omitted if a is a negative integer and $\chi(a) \neq 0$. As shown in [Berndt 1975], $L(z, x, a, \chi)$ converges for $\operatorname{Re} z > 0$ if x is not an integer, or if x is an integer and $\gcd(x, q) > 1$. If x is an integer and $\gcd(x, q) = 1$, the series converges for $\operatorname{Re} z > 1$. For mean value properties of $L(z, a, \chi)$ and asymptotic formulas, see [Ma et al. 2010]. The character analogues of Theorem 1.2 are given below.

Theorem 1.6. *Let χ denote an even, primitive, nonprincipal character modulo q. Let $-1 < \operatorname{Re} z < 1$, and let $L(z, a, \chi)$ be defined as in* (1-17). *Define $T(z, \alpha, \chi)$ by*

$$T(z, \alpha, \chi) := \frac{\alpha^{z/2} q^{z/2} \Gamma(z+1)}{2^z \pi^{z/2} G(\chi)}, \tag{1-18}$$

and $\Omega(z, t)$ by

$$\Omega(z, t) := ((z+1)^2 + t^2)\Gamma\Big(\frac{-z-1+it}{4}\Big)\Gamma\Big(\frac{-z-1-it}{4}\Big) + ((z-1)^2 + t^2)\Gamma\Big(\frac{z-1+it}{4}\Big)\Gamma\Big(\frac{z-1-it}{4}\Big). \tag{1-19}$$

If α and β are positive numbers such that $\alpha\beta = 1$, then

$$\begin{aligned} \sqrt{\alpha}\Big(T(z, \alpha, \chi) \sum_{n=1}^{\infty} \chi(n) L(z+1, n\alpha, \chi) + T(-z, \alpha, \overline{\chi}) \sum_{n=1}^{\infty} \overline{\chi}(n) L(-z+1, n\alpha, \overline{\chi}) \Big) \\ = \sqrt{\beta}\Big(T(-z, \beta, \overline{\chi}) \sum_{n=1}^{\infty} \overline{\chi}(n) L(-z+1, n\beta, \overline{\chi}) + T(z, \beta, \chi) \sum_{n=1}^{\infty} \chi(n) L(z+1, n\beta, \chi) \Big) \\ = \frac{1}{64\pi^{3/2} q} \int_0^{\infty} \Omega(z, t)\, \Xi\Big(\frac{t+iz}{2}, \overline{\chi}\Big) \Xi\Big(\frac{t-iz}{2}, \chi\Big) \cos\big(\tfrac{1}{2} t \log \alpha\big)\, dt. \end{aligned} \tag{1-20}$$

Theorem 1.7. *Let χ denote an odd, primitive, nonprincipal character modulo q. Let $-1 < \operatorname{Re} z < 1$, let $L(z, a, \chi)$ be defined as in* (1-17), *and let $T(z, \alpha, \chi)$ be defined as in* (1-18). *Define $\Lambda(z, t)$ by*

$$\Lambda(z, t) := \Gamma\Big(\frac{z+1+it}{4}\Big)\Gamma\Big(\frac{z+1-it}{4}\Big) + \Gamma\Big(\frac{-z+1+it}{4}\Big)\Gamma\Big(\frac{-z+1-it}{4}\Big). \tag{1-21}$$

If α and β are positive numbers such that $\alpha\beta = 1$, then

$$\text{(1-22)}\quad \sqrt{\alpha}\Big(T(z,\alpha,\chi)\sum_{n=1}^{\infty}\chi(n)L(z+1,n\alpha,\chi) + T(-z,\alpha,\overline{\chi})\sum_{n=1}^{\infty}\overline{\chi}(n)L(-z+1,n\alpha,\overline{\chi})\Big)$$
$$= \sqrt{\beta}\Big(T(-z,\beta,\overline{\chi})\sum_{n=1}^{\infty}\overline{\chi}(n)L(-z+1,n\beta,\overline{\chi}) + T(z,\beta,\chi)\sum_{n=1}^{\infty}\chi(n)L(z+1,n\beta,\chi)\Big)$$
$$= \frac{1}{4\pi^{1/2}iq^2}\int_0^{\infty}\Lambda(z,t)\,\Xi\Big(\frac{t+iz}{2},\overline{\chi}\Big)\,\Xi\Big(\frac{t-iz}{2},\chi\Big)\cos\big(\tfrac{1}{2}t\log\alpha\big)\,dt.$$

The following interesting identity was suggested by the work of Ramanujan:

Theorem 1.8 ([Hardy and Littlewood 1916, p. 156, Section 2.5]). *Let $\mu(n)$ denote the Möbius function. Let α and β be two positive numbers such that $\alpha\beta = 1$. Assume that the series*

$$\sum_{\rho}\frac{\Gamma\left((1-\rho)/2\right)}{\zeta'(\rho)}a^{\rho}$$

converges, where ρ runs through the nontrivial zeros of $\zeta(s)$ and a denotes a positive real number, and that the nontrivial zeros of $\zeta(s)$ are simple. Then

$$\text{(1-23)}\quad \sqrt{\alpha}\sum_{n=1}^{\infty}\frac{\mu(n)}{n}e^{-\pi\alpha^2/n^2} - \frac{1}{4\sqrt{\pi}\sqrt{\alpha}}\sum_{\rho}\frac{\Gamma((1-\rho)/2)}{\zeta'(\rho)}\pi^{\rho/2}\alpha^{\rho}$$
$$= \sqrt{\beta}\sum_{n=1}^{\infty}\frac{\mu(n)}{n}e^{-\pi\beta^2/n^2} - \frac{1}{4\sqrt{\pi}\sqrt{\beta}}\sum_{\rho}\frac{\Gamma((1-\rho)/2)}{\zeta'(\rho)}\pi^{\rho/2}\beta^{\rho}.$$

Hardy and Littlewood's original formulation was slightly different from (1-23) but is readily seen to be equivalent to it. See also [Berndt 1998, p. 470; Paris and Kaminski 2001, p. 143; Titchmarsh 1986, p. 219, Section 9.8] for discussions on this identity. Based on certain assumptions, the character analogues of (1-23) for even and odd primitive Dirichlet characters, which furnish two examples of transformation formulas of the form $F(\alpha,\chi) = F(\beta,\overline{\chi})$, are derived here and are as follows.

Theorem 1.9. *Let χ be an odd, primitive character modulo q, and let α and β be two positive numbers such that $\alpha\beta = 1$. Assume that the series*

$$\sum_{\rho}\frac{\pi^{\rho/2}\alpha^{\rho}\Gamma((2-\rho)/2)}{q^{\rho/2}L'(\rho,\chi)} \quad \textit{and} \quad \sum_{\rho}\frac{\pi^{\rho/2}\beta^{\rho}\Gamma((2-\rho)/2)}{q^{\rho/2}L'(\rho,\overline{\chi})}$$

converge, where ρ *runs through the nontrivial zeros of* $L(s,\chi)$ *and* $L(s,\overline{\chi})$ *respectively, and that the nontrivial zeros of the associated Dirichlet L-functions are simple. Then*

$$\alpha\sqrt{\alpha}\sqrt{G(\chi)}\biggl(\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^2}e^{-\pi\alpha^2/(qn^2)}-\frac{q}{4\pi\alpha^2}\sum_{\rho}\frac{\Gamma\bigl(\frac{2-\rho}{2}\bigr)}{L'(\rho,\chi)}\Bigl(\frac{\pi}{q}\Bigr)^{\rho/2}\alpha^{\rho}\biggr)$$
$$=\beta\sqrt{\beta}\sqrt{G(\overline{\chi})}\biggl(\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n^2}e^{-\pi\beta^2/(qn^2)}-\frac{q}{4\pi\beta^2}\sum_{\rho}\frac{\Gamma\bigl(\frac{2-\rho}{2}\bigr)}{L'(\rho,\overline{\chi})}\Bigl(\frac{\pi}{q}\Bigr)^{\rho/2}\beta^{\rho}\biggr). \tag{1-24}$$

Theorem 1.10. *Let* χ *be an even, primitive character modulo* q, *and let* α *and* β *be two positive numbers such that* $\alpha\beta=1$. *Assume that the series*

$$\sum_{\rho}\frac{\pi^{\rho/2}\alpha^{\rho}\Gamma((2-\rho)/2)}{q^{\rho/2}L'(\rho,\chi)}\quad \textit{and}\quad \sum_{\rho}\frac{\pi^{\rho/2}\beta^{\rho}\Gamma((2-\rho)/2)}{q^{\rho/2}L'(\rho,\overline{\chi})}$$

converge, where ρ *runs through the nontrivial zeros of* $L(s,\chi)$ *and* $L(s,\overline{\chi})$ *respectively, and that the nontrivial zeros of the associated Dirichlet L-functions are simple. Then*

$$\sqrt{\alpha}\sqrt{G(\chi)}\biggl(\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n}e^{-\pi\alpha^2/(qn^2)}-\frac{\sqrt{q}}{4\sqrt{\pi}\alpha}\sum_{\rho}\frac{\Gamma\bigl(\frac{1-\rho}{2}\bigr)}{L'(\rho,\chi)}\Bigl(\frac{\pi}{q}\Bigr)^{\rho/2}\alpha^{\rho}\biggr)$$
$$=\sqrt{\beta}\sqrt{G(\overline{\chi})}\biggl(\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n}e^{-\pi\beta^2/(qn^2)}-\frac{\sqrt{q}}{4\sqrt{\pi}\beta}\sum_{\rho}\frac{\Gamma\bigl(\frac{1-\rho}{2}\bigr)}{L'(\rho,\overline{\chi})}\Bigl(\frac{\pi}{q}\Bigr)^{\rho/2}\beta^{\rho}\biggr). \tag{1-25}$$

This paper is organized as follows. In Section 2, we give a complex integral representation of (1-14) that is used in subsequent sections. In Section 3, we prove Theorem 1.5. Then in Section 4, we compute the inverse Mellin transforms and asymptotic expansions of certain functions which are subsequently used in Section 5. Section 5 is devoted to proofs of Theorems 1.6 and 1.7. Character analogues of Ramanujan's transformation formula (Theorem 1.1) are derived as special cases of these theorems. We conclude this section with a curious result on a certain double series involving characters. In Section 6, we present proofs of Theorems 1.9 and 1.10. Finally we conclude with some open problems in Section 7.

2. A complex integral representation of (1-14)

In this section, we give a formal way of transforming an integral involving a character analogue of Riemann's Ξ-function into an equivalent complex integral which allows us to use residue calculus and Mellin transform techniques for its evaluation.

Theorem 2.1. *Let*

$$f(z,t)=\frac{\phi(z,it)\phi(z,-it)+\phi(-z,it)\phi(-z,-it)}{2}, \tag{2-1}$$

where ϕ is analytic in t as a function of a real variable and analytic in z in some complex domain. Let $y=e^{\mu}$ with μ real. Then, under the assumption that the integral on the left side below converges,

$$\begin{aligned}\int_0^\infty & f(z,t)\,\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\Xi\Big(t-\frac{iz}{2},\chi\Big)\cos\mu t\,dt \\ &=\frac{1}{4i\sqrt{y}}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\big(\phi\big(z,s-\tfrac{1}{2}\big)\phi\big(z,\tfrac{1}{2}-s\big)+\phi\big(-z,s-\tfrac{1}{2}\big)\phi\big(-z,\tfrac{1}{2}-s\big)\big) \\ &\qquad\qquad\times\xi\Big(s-\frac{z}{2},\overline{\chi}\Big)\xi\Big(s+\frac{z}{2},\chi\Big)y^s\,ds.\end{aligned} \tag{2-2}$$

Proof. Let

$$I(z,\mu,\chi):=\int_0^\infty f(z,t)\,\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\Xi\Big(t-\frac{iz}{2},\chi\Big)\cos\mu t\,dt.$$

Then

$$\begin{aligned}I(z,\mu,\chi)&=\tfrac{1}{2}\bigg(\int_0^\infty f(z,t)\,\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\,\Xi\Big(t-\frac{iz}{2},\chi\Big)y^{it}\,dt \\ &\qquad+\int_0^\infty f(z,t)\,\Xi\big(t+\frac{iz}{2},\overline{\chi}\big)\Xi\Big(t-\frac{iz}{2},\chi\Big)y^{-it}\,dt\bigg) \\ &=\tfrac{1}{2}\bigg(\int_0^\infty f(z,t)\,\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\Xi\Big(t-\frac{iz}{2},\chi\Big)y^{it}\,dt \\ &\qquad+\int_{-\infty}^0 f(z,-t)\,\Xi\Big(-t+\frac{iz}{2},\overline{\chi}\Big)\Xi\Big(-t-\frac{iz}{2},\chi\Big)y^{it}\,dt\bigg).\end{aligned} \tag{2-3}$$

However, using (1-11), we readily see that

$$\begin{aligned}\Xi\Big(-t+\frac{iz}{2},\overline{\chi}\Big)&=\xi\Big(\tfrac{1}{2}-it-\frac{z}{2},\overline{\chi}\Big) \\ &=\epsilon(\chi)\xi\Big(\tfrac{1}{2}+it+\frac{z}{2},\chi\Big)=\epsilon(\chi)\,\Xi\Big(t-\frac{iz}{2},\chi\Big), \\ \Xi\Big(-t-\frac{iz}{2},\chi\Big)&=\xi\Big(\tfrac{1}{2}-it+\tfrac{z}{2},\chi\Big) \\ &=(\epsilon(\chi))^{-1}\xi\Big(\tfrac{1}{2}+it-\frac{z}{2},\overline{\chi}\Big)=(\epsilon(\chi))^{-1}\,\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\end{aligned}$$

so that

$$\Xi\Big(-t+\frac{iz}{2},\overline{\chi}\Big)\Xi\Big(-t-\frac{iz}{2},\chi\Big)=\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\Xi\Big(t-\frac{iz}{2},\chi\Big). \tag{2-4}$$

Thus from (2-3), (2-4), and the fact that f is an even function of t, we obtain

$$\begin{aligned}
&I(z,\mu,\chi)\\
&\quad=\tfrac{1}{2}\int_{-\infty}^{\infty} f(z,t)\,\Xi\Big(t+\frac{iz}{2},\overline{\chi}\Big)\,\Xi\Big(t-\frac{iz}{2},\chi\Big)y^{it}\,dt\\
&\quad=\frac{1}{4i\sqrt{y}}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\big(\phi(z,s-\tfrac{1}{2})\phi(z,\tfrac{1}{2}-s)+\phi(-z,s-\tfrac{1}{2})\phi(-z,\tfrac{1}{2}-s)\big)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\xi\Big(s-\frac{z}{2},\overline{\chi}\Big)\xi\Big(s+\frac{z}{2},\chi\Big)y^{s}\,ds,
\end{aligned}$$

where in the penultimate line, we made the change of variable $s=\frac{1}{2}+it$. □

For our purpose here, we replace μ by 2μ in (2-2) and then t by $t/2$ on the left-hand side of (2-2). Thus, with $y=e^{2\mu}$, we find that

$$\begin{aligned}
&\int_{0}^{\infty} f\Big(z,\frac{t}{2}\Big)\,\Xi\Big(\frac{t+iz}{2},\overline{\chi}\Big)\,\Xi\Big(\frac{t-iz}{2},\chi\Big)\cos\mu t\,dt\\
&\quad=\frac{1}{2i\sqrt{y}}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\big(\phi(z,s-\tfrac{1}{2})\phi(z,\tfrac{1}{2}-s)+\phi(-z,s-\tfrac{1}{2})\phi(-z,\tfrac{1}{2}-s)\big)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\xi\Big(s-\frac{z}{2},\overline{\chi}\Big)\xi\Big(s+\frac{z}{2},\chi\Big)y^{s}\,ds.
\end{aligned}\tag{2-5}$$

It is with this equation that we will be working throughout this paper.

3. Character analogues of the extended version of the Ramanujan–Guinand formula

Lemma 3.1. *For* $\operatorname{Re}s>1$ *and* $\operatorname{Re}(s-\eta)>1$,

$$L(s,\overline{\chi})L(s-\eta,\chi)=\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)}{n^{s}}\sum_{d\mid n}\chi^{2}(d)d^{\eta}.\tag{3-1}$$

Proof. Since the Dirichlet series for both the L-functions converge absolutely under the given hypotheses, using [Apostol 1976, Theorem 11.5], we see that

$$\begin{aligned}
L(s,\overline{\chi})L(s-\eta,\chi)&=\sum_{\substack{n=1\\(n,q)=1}}^{\infty}\frac{\overline{\chi}(n)}{n^{s}}\sum_{\substack{k=1\\(k,q)=1}}^{\infty}\frac{\chi(k)}{k^{s-\eta}}=\sum_{\substack{j=1\\(j,q)=1}}^{\infty}\frac{1}{j^{s}}\sum_{\substack{nk=j\\(k,q)=1}}\overline{\chi}(n)\chi(k)k^{\eta}\\
&=\sum_{\substack{j=1\\(j,q)=1}}^{\infty}\frac{\overline{\chi}(j)}{j^{s}}\sum_{\substack{nk=j\\(k,q)=1}}\frac{\chi^{2}(k)}{\chi(k)\overline{\chi}(k)}k^{\eta}=\sum_{j=1}^{\infty}\frac{\overline{\chi}(j)}{j^{s}}\sum_{nk=j}\chi^{2}(k)k^{\eta},
\end{aligned}$$

where in the last step, we make use of the fact that $\chi(k)\overline{\chi}(k)=1$ for $(k,q)=1$. □

Proof of Theorem 1.5. Assume that χ is even. Let $\phi(z,s)\equiv 1$. From (2-1) we see that $f(z,t)\equiv 1$. Using (1-9), (1-8), (1-12), and (1-13), we find that the integral

$$M(z,\mu,\chi) := \int_0^\infty \Xi\Big(\frac{t+iz}{2},\overline{\chi}\Big)\Xi\Big(\frac{t-iz}{2},\chi\Big)\cos\mu t\,dt$$

does converge. Using (2-5), we observe that

$$\begin{aligned}(3\text{-}2)\quad & M(z,\mu,\chi)\\ &= \frac{1}{i\sqrt{y}}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\xi\Big(s-\frac{z}{2},\overline{\chi}\Big)\xi\Big(s+\frac{z}{2},\chi\Big)y^s\,ds\\ &= \frac{1}{i\sqrt{y}}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\Gamma\Big(\frac{s}{2}-\frac{z}{4}\Big)\Gamma\Big(\frac{s}{2}+\frac{z}{4}\Big)L\Big(s-\frac{z}{2},\overline{\chi}\Big)L\Big(s+\frac{z}{2},\chi\Big)\Big(\frac{\pi}{qy}\Big)^{-s}ds.\end{aligned}$$

Since $\operatorname{Re} s=\frac{1}{2}$ and $-1<\operatorname{Re} z<1$, we have $0<\operatorname{Re}(s-z/2)<1$ and $0<\operatorname{Re}(s+z/2)<1$. Now replace s by $s-z/2$ and let $\eta=-z$ in Lemma 3.1. Then, for $\operatorname{Re}(s-z/2)>1$ and $\operatorname{Re}(s+z/2)>1$,

$$L\Big(s-\frac{z}{2},\overline{\chi}\Big)L\Big(s+\frac{z}{2},\chi\Big)=\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)}{n^{s-z/2}}\sum_{d|n}\chi^2(d)d^{-z}. \tag{3-3}$$

We wish to shift the line of integration from $\operatorname{Re} s=\frac{1}{2}$ to $\operatorname{Re} s=\frac{3}{2}$ in order to be able to use (3-3) in (3-2). Consider a positively oriented rectangular contour formed by $[\frac{1}{2}+iT,\frac{1}{2}-iT]$, $[\frac{1}{2}-iT,\frac{3}{2}-iT]$, $[\frac{3}{2}-iT,\frac{3}{2}+iT]$ and $[\frac{3}{2}+iT,\frac{1}{2}+iT]$, where T is any positive real number. The integrand on the extreme right side of (3-2) does not have any pole inside the contour. Also as $T\to\infty$, the integrals along the horizontal segments $[\frac{1}{2}-iT,\frac{3}{2}-iT]$ and $[\frac{3}{2}+iT,\frac{1}{2}+iT]$ tend to zero, which can be seen by using (1-12). Hence, employing Cauchy's residue theorem, letting $T\to\infty$, using (3-3) in (3-2), and interchanging the order of summation and integration, which is valid because of absolute convergence, we observe that

$$\begin{aligned}(3\text{-}4)\quad & M(z,\mu,\chi)\\ &= \frac{1}{i\sqrt{y}}\sum_{n=1}^{\infty}\overline{\chi}(n)n^{z/2}\Big(\sum_{d|n}\chi^2(d)d^{-z}\Big)\int_{\frac{3}{2}-i\infty}^{\frac{3}{2}+i\infty}\Gamma\Big(\frac{s}{2}-\frac{z}{4}\Big)\Gamma\Big(\frac{s}{2}+\frac{z}{4}\Big)\Big(\frac{n\pi}{qy}\Big)^{-s}ds.\end{aligned}$$

But from [Oberhettinger 1974, Formula 11.1, p. 115], for $c=\operatorname{Re} s>\pm\operatorname{Re}\nu$,

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}2^{s-2}w^{-s}\Gamma\Big(\frac{s}{2}-\frac{\nu}{2}\Big)\Gamma\Big(\frac{s}{2}+\frac{\nu}{2}\Big)x^{-s}\,ds=K_\nu(wx). \tag{3-5}$$

Hence, using (3-5) with $c = 3/2$, $\nu = z/2$, $w = 2$, and $x = n\pi/qy$ in (3-4), we find that

$$M(z,\mu,\chi) = \frac{8\pi}{\sqrt{y}} \sum_{n=1}^{\infty} \overline{\chi}(n) n^{z/2} \Big(\sum_{d|n} \chi^2(d) d^{-z} \Big) K_{z/2}\Big(\frac{2\pi n}{qy}\Big). \tag{3-6}$$

Now let $\mu = \frac{1}{2}\log\alpha$ in (3-6) so that $y = e^{2\mu}$ implies that $y = \alpha$. Then using the fact that $\alpha\beta = 1$, we deduce that

$$\begin{aligned} \frac{1}{8\pi} \int_0^{\infty} \Xi\Big(\frac{t+iz}{2}, \overline{\chi}\Big) \Xi\Big(\frac{t-iz}{2}, \chi\Big) \cos\big(\tfrac{1}{2} t \log\alpha\big)\, dt \\ = \sqrt{\beta} \sum_{n=1}^{\infty} \overline{\chi}(n) n^{z/2} \Big(\sum_{d|n} \chi^2(d) d^{-z} \Big) K_{\frac{z}{2}}\Big(\frac{2\pi n\beta}{q}\Big). \end{aligned} \tag{3-7}$$

Next, observing that replacing α by β and/or simultaneously replacing χ by $\overline{\chi}$ and z by $-z$ in (3-7) leaves the integral on the left side invariant, we obtain (1-15).

Now consider the case when χ is odd. Again the convergence of the integral $M(z,\mu,\chi)$ can be seen from (1-12) and (1-13). Following similar steps as in the case of even χ and using the definition of $\xi(s,\chi)$ from (1-8) for χ odd, we get

$$\begin{aligned} & M(z,\mu,\chi) \\ &= \frac{q}{i\pi\sqrt{y}} \sum_{n=1}^{\infty} \overline{\chi}(n) n^{z/2} \sum_{d|n} \chi^2(d) d^{-z} \int_{\frac{3}{2}-i\infty}^{\frac{3}{2}+i\infty} \Gamma\Big(\frac{s}{2}-\frac{z}{4}+\frac{1}{2}\Big)\Gamma\Big(\frac{s}{2}+\frac{z}{4}+\frac{1}{2}\Big)\Big(\frac{n\pi}{qy}\Big)^{-s} ds. \end{aligned} \tag{3-8}$$

Now replacing s by $s+1$ in (3-5), we find that for $c = \operatorname{Re} s > \pm \operatorname{Re} \nu - 1$,

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} 2^{s-1} w^{-s-1} \Gamma\Big(\frac{s+1}{2} - \frac{\nu}{2}\Big)\Gamma\Big(\frac{s+1}{2} + \frac{\nu}{2}\Big) x^{-s}\, ds = x K_{\nu}(wx). \tag{3-9}$$

Then using (3-9) with $c = \frac{3}{2}$, $\nu = 0$, $w = 2$ and $x = n\pi/qy$ in (3-8), we see that

$$M(z,\mu,\chi) = \frac{8\pi}{y^{3/2}} \sum_{n=1}^{\infty} \overline{\chi}(n) n^{z/2+1} \Big(\sum_{d|n} \chi^2(d) d^{-z} \Big) K_{z/2}\Big(\frac{2\pi n}{qy}\Big). \tag{3-10}$$

Now let $\mu = \frac{1}{2}\log\alpha$ in (3-10) so that $y = e^{2\mu}$ implies that $y = \alpha$. Then using the fact that $\alpha\beta = 1$, we deduce that

$$\begin{aligned} \frac{1}{8\pi} \int_0^{\infty} \Xi\Big(\frac{t+iz}{2}, \overline{\chi}\Big) \Xi\Big(\frac{t-iz}{2}, \chi\Big) \cos\big(\tfrac{1}{2} t \log\alpha\big)\, dt \\ = \beta^{3/2} \sum_{n=1}^{\infty} \overline{\chi}(n) n^{z/2+1} \Big(\sum_{d|n} \chi^2(d) d^{-z} \Big) K_{z/2}\Big(\frac{2\pi n\beta}{q}\Big). \end{aligned} \tag{3-11}$$

Next, observing that replacing α by β and/or simultaneously replacing χ by $\overline{\chi}$ and z by $-z$ in (3-11) leaves the integral on the left side invariant, we obtain (1-15). □

Remark. Letting $z \to 0$ in Theorem 1.5 gives a new character analogue of the extended version of Koshliakov's formula, that is, Theorem 1.4.

When χ is real, Theorem 1.5 reduces to the following corollary.

Corollary 3.2. *Let $-1 < \operatorname{Re} z < 1$ and let χ denote a real, primitive, nonprincipal character modulo q. Let the number b be defined as in* (1-7). *If*

$$F(z, \alpha, \chi) = \alpha^{b+\frac{1}{2}} \sum_{n=1}^{\infty} \chi(n) n^{-\frac{z}{2}+b} \sigma_z(n) K_{-z/2}\Big(\frac{2\pi n\alpha}{q}\Big),$$

then

$$\begin{aligned} F(z, \alpha, \chi) &= F(-z, \beta, \chi) = F(-z, \alpha, \chi) = F(z, \beta, \chi) \\ &= \frac{1}{8\pi} \int_0^{\infty} \Xi\Big(\frac{t+iz}{2}, \chi\Big) \Xi\Big(\frac{t-iz}{2}, \chi\Big) \cos\big(\tfrac{1}{2} t \log\alpha\big)\, dt. \end{aligned}$$

The above corollary (without the integral) is equivalent to the special cases, when χ is real, of the character analogues of the Ramanujan–Guinand formula established in [Berndt et al. 2011, Theorems 3.1 and 4.1].

4. Inverse Mellin transforms and asymptotic expansions

We will now evaluate inverse Mellin transforms of some functions and asymptotic expansions of certain other functions used in the later sections.

Lemma 4.1. *Let $z \in \mathbb{C}$ be fixed such that $-1 < \operatorname{Re} z < 1$. For a primitive, nonprincipal character χ mod q, let $L(z, a, \chi)$ be defined as in* (1-17). *Then, for*

$$-\operatorname{Re}\frac{z}{2} < c = \operatorname{Re} s < 1 + \operatorname{Re}\frac{z}{2}$$

and $x \in \mathbb{R} \setminus \mathbb{Z}_{<0}$,

$$\begin{aligned} \text{(4-1)} \quad \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big) L\Big(1-s+\frac{z}{2}, \chi\Big) x^{-s}\, ds \\ = x^{z/2}\Gamma(z+1) L(z+1, x, \chi). \end{aligned}$$

Proof. We prove the result only for even characters. The case for odd characters can be proved similarly. We first assume $|x| < 1$ and later extend it to any real $x \in \mathbb{R} \setminus \mathbb{Z}_{<0}$ by analytic continuation. Let

$$-\operatorname{Re}(z/2) < c = \operatorname{Re} s < 1 + \operatorname{Re}(z/2).$$

Consider a positively oriented rectangular contour formed by $[c - iT, c + iT]$, $[c+iT, -M+iT]$, $[-M+iT, -M-iT]$, and $[-M-iT, c-iT]$, where T is some positive real number and $M = n - \frac{1}{2}$, where n is a positive integer. Let $s = \sigma + it$. Among the poles of the function $\Gamma(s+z/2)\Gamma(1-s+z/2)L(1-s+z/2, \chi)x^{-s}$, the only ones that contribute are the poles at $s = -z/2 - m$, $m \geq 0$. Let $R_f(a)$ denote the residue of the function $f(s) := \Gamma(s+z/2)\Gamma(1-s+z/2)L(1-s+z/2,\,\chi)x^{-s}$ at a. Then for $m \geq 0$,

$$
\begin{aligned}
\text{(4-2)}\quad & R_f\Big(-\frac{z}{2}-m\Big) \\
&= \lim_{s\to -z/2-m}\Big(s+\frac{z}{2}+m\Big)\Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\,\chi\Big)x^{-s} \\
&= \frac{(-1)^m}{m!}\Gamma\Big(1+z+m\Big)L\Big(1+z+m,\,\chi\Big)x^{m+z/2}.
\end{aligned}
$$

From (4-2) and the residue theorem, we have

$$
\begin{aligned}
\text{(4-3)}\quad & \left[\int_{c-iT}^{c+iT}+\int_{c+iT}^{-M+iT}+\int_{-M+iT}^{-M-iT}+\int_{-M-iT}^{c-iT}\right] \\
&\qquad \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\,\chi\Big)x^{-s}\,ds \\
&= 2\pi i x^{z/2}\sum_{0\leq m<M}\frac{(-1)^m}{m!}\Gamma\,(1+z+m)\,L\,(1+z+m,\,\chi)\,x^m.
\end{aligned}
$$

We now estimate the integral along the upper horizontal segment. Using (1-13), we find that for $-M \leq \sigma \leq c$,

$$
L(1-\sigma\pm iT,\,\chi) = O(q^{c+1/2}T^{c+1/2}). \tag{4-4}
$$

Hence, for $-M\leq\sigma\leq c$, i.e., $-M-\operatorname{Re} z/2 \leq \sigma-\operatorname{Re} z/2 \leq c-\operatorname{Re} z/2$, we have

$$
\begin{aligned}
\text{(4-5)}\quad & L\Big(1-\Big(\sigma-\operatorname{Re}\frac{z}{2}\Big)-i\Big(T-\operatorname{Im}\frac{z}{2}\Big),\,\chi\Big) \\
&\qquad = O\bigg(q^{c-\operatorname{Re} z/2+1/2}\Big(T-\operatorname{Im}\frac{z}{2}\Big)^{c-\operatorname{Re} z/2+1/2}\bigg).
\end{aligned}
$$

By (1-12), we observe that

$$
\Big|\Gamma\Big(s+\frac{z}{2}\Big)\Big| \sim \sqrt{2\pi}\,e^{-\pi/2|T+\operatorname{Im} z/2|}\cdot\Big|T+\operatorname{Im}\frac{z}{2}\Big|^{\sigma+\operatorname{Re} z/2-1/2} \tag{4-6}
$$

and

$$
\Big|\Gamma\Big(1-s+\frac{z}{2}\Big)\Big| \sim \sqrt{2\pi}\,e^{-\frac{\pi}{2}|T-\operatorname{Im} z/2|}\cdot\Big|T-\operatorname{Im}\frac{z}{2}\Big|^{-\sigma+\operatorname{Re} z/2+1/2}. \tag{4-7}
$$

Since $|x|<1$, from (4-5), (4-6), and (4-7), we deduce that

$$
\begin{aligned}
& \left|\int_{c+iT}^{-M+iT}\Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\,\chi\Big)x^{-s}\,ds\right| \\
&\quad \leq 2\pi K_1(c+M)|x|^{-c}q^{c-\operatorname{Re} z/2+1/2}e^{-\pi/2(|T+\operatorname{Im} z/2|+|T-\operatorname{Im} z/2|)} \\
&\qquad\qquad \times\Big|T+\operatorname{Im}\frac{z}{2}\Big|^{\sigma+\operatorname{Re} z/2-1/2}\Big|T-\operatorname{Im}\frac{z}{2}\Big|^{c-\sigma+1},
\end{aligned}
$$

where K_1 is some absolute constant. Hence

$$\int_{c+i\infty}^{-M+i\infty} \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\chi\Big)x^{-s}\,ds = 0. \tag{4-8}$$

Similarly for the integral along the lower horizontal segment, using (4-6), (4-7), and the fact that

$$L\Big(1-\Big(\sigma-\operatorname{Re}\frac{z}{2}\Big)+i\Big(T+\operatorname{Im}\frac{z}{2}\Big),\chi\Big) = O\Big(q^{c-\operatorname{Re}z/2+1/2}\Big(T+\operatorname{Im}\frac{z}{2}\Big)^{c-\operatorname{Re}z/2+1/2}\Big),$$

we observe that

$$\int_{-M-i\infty}^{c-i\infty} \Gamma(s+z/2)\Gamma(1-s+z/2)L(1-s+z/2,\chi)x^{-s}\,ds = 0. \tag{4-9}$$

Hence, from (4-3), (4-8), and (4-9), it is clear that

$$\begin{aligned}\left[\int_{c-i\infty}^{c+i\infty}+\int_{-M+i\infty}^{-M-i\infty}\right]&\Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\chi\Big)x^{-s}ds\\ &= 2\pi i x^{z/2}\sum_{0\le m<M}\frac{(-1)^m}{m!}\Gamma\,(1+z+m)\,L\,(1+z+m,\chi)\,x^m.\end{aligned} \tag{4-10}$$

It remains to evaluate

$$\int_{-M+i\infty}^{-M-i\infty} \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\chi\Big)x^{-s}\,ds.$$

Using (1-12) and the reflection formula for the gamma function [Temme 1996, Equation (3.5), p. 46], we find that as $|t|\to\infty$,

$$\Gamma(-M+it) = O(|t|^{-M-\frac{1}{2}}e^{-\pi|t|/2}).$$

Hence, as $|t|\to\infty$,

$$\Gamma\Big(-M+it+\frac{z}{2}\Big) = O\Big(\Big|t+\operatorname{Im}\frac{z}{2}\Big|^{-M+\operatorname{Re}z/2-1/2}e^{-(\pi/2)|t+\operatorname{Im}z/2|}\Big). \tag{4-11}$$

Again by (1-12), as $|t|\to\infty$,

$$\begin{aligned}&\Big|\Gamma\Big(1+M-it+\frac{z}{2}\Big)\Big|\\ &\quad= \sqrt{2\pi}\,e^{-(\pi/2)|t-\operatorname{Im}z/2|}\cdot\Big|t-\operatorname{Im}\frac{z}{2}\Big|^{M+\operatorname{Re}z/2+1/2}\Big(1+O\Big(\frac{1}{|t-\operatorname{Im}z/2|}\Big)\Big).\end{aligned} \tag{4-12}$$

Also, $L(1+M-it+z/2, \chi)$ is bounded as $\mathrm{Re}(1+M-it+z/2) > 1$. Hence,

$$\begin{aligned}
&\left|\int_{-M+i\infty}^{-M-i\infty} \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2}, \chi\Big)x^{-s}\,ds\right| \\
&= \left|i\int_{-\infty}^{\infty} \Gamma\Big(-M+it+\frac{z}{2}\Big)\Gamma\Big(1+M-it+\frac{z}{2}\Big)L\Big(1+M-it+\frac{z}{2}, \chi\Big)x^{M-it}\,dt\right| \\
&= |x|^M \int_{-1}^{1} O(1)\,dt + |x|^M \int_{1}^{\pm\infty} O\bigg(\Big|t+\mathrm{Im}\frac{z}{2}\Big|^{-M+\mathrm{Re}z/2-1/2}\Big|t-\mathrm{Im}\frac{z}{2}\Big|^{M+\mathrm{Re}z/2+1/2} \\
&\qquad\qquad\qquad\qquad \times e^{-(\pi/2)(|t+\mathrm{Im}z/2|+|t-\mathrm{Im}z/2|)}\bigg)dt \\
&= O(|x|^M).
\end{aligned}$$

Since $|x| < 1$,

$$\lim_{M\to\infty} \int_{-M+i\infty}^{-M-i\infty} \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2}, \chi\Big)x^{-s}\,ds = 0. \tag{4-13}$$

From (4-10) and (4-13), we finally deduce that

$$\begin{aligned}
&\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2}, \chi\Big)x^{-s}\,ds \\
&\qquad = x^{z/2}\sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\Gamma(1+z+m)L(1+z+m, \chi)x^{m+z/2} \\
&\qquad = x^{z/2}\Gamma(z+1)\sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\frac{\Gamma(1+z+m)}{\Gamma(1+z)}\sum_{k=1}^{\infty}\frac{\chi(k)}{k^{z+m+1}}x^m \\
&\qquad = x^{z/2}\Gamma(z+1)\sum_{k=1}^{\infty}\frac{\chi(k)}{k^{z+1}}\sum_{m=0}^{\infty}\frac{\Gamma(1+z+m)}{m!\Gamma(1+z)}\Big(\frac{-x}{k}\Big)^m \\
&\qquad = x^{z/2}\Gamma(z+1)\sum_{k=1}^{\infty}\frac{\chi(k)}{k^{z+1}}\Big(1+\frac{x}{k}\Big)^{-z-1} \\
&\qquad = x^{z/2}\Gamma(z+1)\sum_{k=1}^{\infty}\frac{\chi(k)}{(k+x)^{z+1}} = x^{z/2}\Gamma(z+1)L(z+1, x, \chi),
\end{aligned}$$

where, in the fourth step, we have utilized the binomial theorem, since $|x| < 1$. Since both sides of (4-1) are analytic for any $x \in \mathbb{R}\setminus\mathbb{Z}_{<0}$, the result follows by analytic continuation. □

When $z = 0$, we get the following corollary.

Corollary 4.2. *For a primitive, nonprincipal character* χ *mod* q, *let* $\psi(a, \chi)$ *be defined as in* (1-16). *Then, for* $0 < c = \operatorname{Re} s < 1$ *and* $x \in \mathbb{R} \setminus \mathbb{Z}_{<0}$,

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{L(1-s, \chi)}{\sin \pi s} x^{-s}\, ds = -\frac{1}{\pi} \psi(x, \chi). \tag{4-14}$$

For $j \geq 1$, the generalized Bernoulli numbers $B_j(\chi)$ are given by

$$B_{2j}(\chi) = \frac{2(-1)^{j-1} G(\overline{\chi})(2j)!}{q(2\pi/q)^{2j}} L(2j, \chi)$$

for χ even and by

$$B_{2j-1}(\chi) = \frac{2(-1)^{j-1} i G(\overline{\chi})(2j-1)!}{q(2\pi/q)^{2j-1}} L(2j-1, \chi) \tag{4-15}$$

for χ odd; see [Berndt 1975, p. 426]. It is also known [Berndt 1975, Corollary 3.4, p. 423] that $B_{2j-1}(\chi) = 0$ when χ is even and $B_{2j}(\chi) = 0$ when χ is odd. The asymptotic expansion of $L(z, a, \chi)$ as $|a| \to \infty$ is given below.

Lemma 4.3. *For* $\operatorname{Re} z > 0$ *and* $-\pi < \arg a < \pi$, *as* $|a| \to \infty$,

$$L(z, a, \chi) \sim \chi(-1) \sum_{j=1}^{\infty} \frac{B_j(\overline{\chi}) \prod_{m=0}^{j-2} (z+m)}{j!\, a^{z+j-1}}.$$

Proof. One takes (4.3) and (4.4) in [Berndt 1975, p. 424], valid for χ even and odd respectively, substitutes $A = 0$, $B = N$, $r = 1$, and $f(u) = (u+a)^{-z}$, lets $N \to \infty$, and performs repeated integration by parts on the prevalent integral. □

This gives, as a special case, the following asymptotic expansion of $\psi(a, \chi)$ as $|a| \to \infty$.

Corollary 4.4. *For* $-\pi < \arg a < \pi$, *as* $|a| \to \infty$,

$$\psi(a, \chi) \sim -\frac{L(0, \chi)}{a} - \chi(-1) \sum_{j=2}^{\infty} \frac{B_j(\overline{\chi})}{j a^j}. \tag{4-16}$$

Proof. Specialize $z = 1$ in Lemma 4.3. Observe that $L(1, a, \chi) = -\psi(a, \chi)$. For χ even, we have $B_1(\overline{\chi}) = 0$. But from [Apostol 1976, p. 268], $L(0, \chi) = 0$. This yields (4-16) for χ even. For χ odd, we observe from (4-15) that

$$B_1(\overline{\chi}) = \frac{i}{\pi} G(\chi) L(1, \overline{\chi}), \tag{4-17}$$

and from (1-10), it is easy to see that

$$L(1, \overline{\chi}) = \frac{i\pi}{G(\chi)} L(0, \chi). \tag{4-18}$$

Now (4-16) follows from (4-17) and (4-18). □

5. Character analogues of Theorem 1.2

In this section, we prove analogues of Theorem 1.2 for even and odd primitive characters. Then we give character analogues of Ramanujan's transformation formula (Theorem 1.1) as special cases.

Proof of Theorem 1.6. Using Lemma 4.3, one sees that the series involving the functions $L(z, a, \chi)$ in the theorem are convergent. Let

$$\phi(z, s) = (z+1+2s)\Gamma\Big(\frac{-z-1}{4} + \frac{s}{2}\Big).$$

From (2-1) and (1-19), we find that $f(z, t/2) = \frac{1}{2}\Omega(z, t)$. From (2-5), we have

$$\int_0^\infty \Omega(z, t)\,\Xi\Big(\frac{t+iz}{2}, \overline{\chi}\Big)\Xi\Big(\frac{t-iz}{2}, \chi\Big)\cos\mu t\,dt = \frac{1}{i\sqrt{y}}(J(z, y, \chi) + J(-z, y, \overline{\chi})), \tag{5-1}$$

where

$$J(z, y, \chi) := \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} U(z, s, y, \chi)\,ds \tag{5-2}$$

with

$$U(z, s, y, \chi) := (-z+2s)(-z+2-2s)\Gamma\Big(\frac{z}{4}+\frac{s}{2}-\frac{1}{2}\Big)\Gamma\Big(\frac{z}{4}-\frac{s}{2}\Big)\xi\Big(s-\frac{z}{2}, \overline{\chi}\Big)\xi\Big(s+\frac{z}{2}, \chi\Big)y^s.$$

Using (1-12) and (1-13), one sees that indeed the integral on the left side of (5-1) converges. We first simplify the integrand in (5-2). Using (1-8) with $b = 0$, and then the duplication formula [Temme 1996, Equation (3.4), p. 46] and the reflection formula for the Gamma function in the second equality below, we have

$$\begin{aligned} U(z, s, y, \chi) &= 16\Big(\frac{\pi}{qy}\Big)^{-s}\Big\{\Gamma\Big(\frac{z}{4}+\frac{s+1}{2}\Big)\Gamma\Big(\frac{z}{4}+\frac{s}{2}\Big)\Big\}\Big\{\Gamma\Big(\frac{z}{4}-\frac{s}{2}+1\Big)\Gamma\Big(\frac{s}{2}-\frac{z}{4}\Big)\Big\} \\ &\qquad \times L\Big(s-\frac{z}{2}, \overline{\chi}\Big)L\Big(s+\frac{z}{2}, \chi\Big) \\ &= 16\Big(\frac{\pi}{qy}\Big)^{-s}\cdot\frac{\sqrt{\pi}}{2^{s+z/2-1}}\Gamma\Big(s+\frac{z}{2}\Big)\cdot\frac{\pi}{\sin(\pi(s/2-z/4))} \\ &\qquad \times L\Big(s-\frac{z}{2}, \overline{\chi}\Big)L\Big(s+\frac{z}{2}, \chi\Big). \end{aligned} \tag{5-3}$$

Substituting (1-10) in the form

$$L\Big(s-\frac{z}{2}, \overline{\chi}\Big) = \frac{(2\pi)^{s-z/2}L(1-s+z/2, \chi)}{2q^{s-z/2-1}G(\chi)\Gamma(s-z/2)\cos((\pi/2)(s-z/2))}$$

in (5-3) and then simplifying, we find that

$$U(z,s,y,\chi) = \frac{32y^s 2^{-z}\pi^{(1-z)/2}}{q^{-z/2-1}G(\chi)}\Gamma\Big(1-s+\frac{z}{2}\Big)\Gamma\Big(s+\frac{z}{2}\Big)L\Big(1-s+\frac{z}{2},\chi\Big)L\Big(s+\frac{z}{2},\chi\Big). \tag{5-4}$$

We wish to shift the line of integration from $\operatorname{Re} s = \frac{1}{2}$ to $\operatorname{Re} s = \frac{3}{2}$ in order to evaluate (5-2), since then $-1 < \operatorname{Re} z < 1$ implies that

$$\operatorname{Re}(s+z/2) > 1,$$

which allows us to use the series representation of $L(s+z/2,\chi)$. Consider a positively oriented rectangular contour formed by $[\frac{1}{2}+iT, \frac{1}{2}-iT]$, $[\frac{1}{2}-iT, \frac{3}{2}-iT]$, $[\frac{3}{2}-iT, \frac{3}{2}+iT]$, and $[\frac{3}{2}+iT, \frac{1}{2}+iT]$, where T is any positive real number. The integrand in (5-2) does not have any pole inside the contour since the pole of $\Gamma(1-s+z/2)$ at $s = 1+z/2$ is canceled by the zero of $L(1-s+z/2,\chi)$ there. Also as $T\to\infty$, the integrals along the horizontal segments $[\frac{1}{2}-iT, \frac{3}{2}-iT]$ and $[\frac{3}{2}+iT, \frac{1}{2}+iT]$ tend to zero, which can be seen using (1-12). Employing the residue theorem, letting $T\to\infty$ and using (5-4), we find that

$$\begin{aligned} J(z,y,\chi) &= \frac{32\cdot 2^{-z}\pi^{(1-z)/2}}{q^{-z/2-1}G(\chi)}\int_{\frac{3}{2}-i\infty}^{\frac{3}{2}+i\infty}\Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big) \\ &\qquad\qquad \times L\Big(1-s+\frac{z}{2},\chi\Big)L\Big(s+\frac{z}{2},\chi\Big)y^s\,ds \\ &= \frac{32\cdot 2^{-z}\pi^{(1-z)/2}}{q^{-z/2-1}G(\chi)}\sum_{n=1}^{\infty}\frac{\chi(n)}{n^{z/2}}\int_{\frac{3}{2}-i\infty}^{\frac{3}{2}+i\infty}\Gamma\Big(s+\frac{z}{2}\Big)\Gamma\Big(1-s+\frac{z}{2}\Big) \\ &\qquad\qquad \times L\Big(1-s+\frac{z}{2},\chi\Big)\Big(\frac{n}{y}\Big)^{-s}ds. \end{aligned} \tag{5-5}$$

Now, in order to use Lemma 4.1 to evaluate the integral in (5-5), we again have to shift the line of integration from $\operatorname{Re} s > \frac{3}{2}$ to $\operatorname{Re} s = d$, where

$$-\operatorname{Re} z/2 < d < 1+\operatorname{Re} z/2.$$

Again, we do not encounter any pole in this process. Hence

$$J(z,y,\chi) = \frac{64i2^{-z}y^{-z/2}\pi^{(3-z)/2}\Gamma(z+1)}{q^{-z/2-1}G(\chi)}\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\chi(n)\chi(k)}{(k+n/y)^{z+1}}. \tag{5-6}$$

Since $-1 < \operatorname{Re}(z) < 1$, the other integral, namely $J(-z,y,\overline{\chi})$, can be evaluated by simply replacing z by $-z$ and χ by $\overline{\chi}$ in (5-6). Now (5-1), (5-6), (1-18), and the discussion in the previous line give

$$\int_0^\infty \Omega(z,t)\,\Xi\Big(\frac{t+iz}{2},\overline{\chi}\Big)\Xi\Big(\frac{t-iz}{2},\chi\Big)\cos\mu t\,dt \tag{5-7}$$
$$= \frac{64\pi^{3/2}q}{\sqrt{y}}\bigg(T(z,y^{-1},\chi)\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\chi(n)\chi(k)}{(k+n/y)^{z+1}} + T(-z,y^{-1},\overline{\chi})\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\overline{\chi}(n)\overline{\chi}(k)}{(k+n/y)^{-z+1}}\bigg),$$

where it is easy to see from the fact that $-1 < \operatorname{Re} z < 1$, from the discussion just preceding the statement of Theorem 1.6, and from Lemma 4.3, that both the double series on the right side of (5-7) converge.

Now let $\mu = \frac{1}{2}\log\alpha$ in (5-7) so that $y = e^{2\mu}$ implies that $y = \alpha$. Then using the fact that $\alpha\beta = 1$ and using (1-17) in the second equality below, we deduce that

$$\int_0^\infty \Omega(z,t)\,\Xi\Big(\frac{t+iz}{2},\overline{\chi}\Big)\Xi\Big(\frac{t-iz}{2},\chi\Big)\cos\Big(\tfrac{1}{2}t\log\alpha\Big)\,dt$$
$$= 64\pi^{3/2}q\sqrt{\beta}\bigg(T(z,\beta,\chi)\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\chi(n)\chi(k)}{(k+n\beta)^{z+1}} + T(-z,\beta,\overline{\chi})\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\overline{\chi}(n)\overline{\chi}(k)}{(k+n\beta)^{-z+1}}\bigg)$$
$$= 64\pi^{3/2}q\sqrt{\beta}\bigg(T(z,\beta,\chi)\sum_{n=1}^{\infty}\chi(n)L(z+1,n\beta,\chi) + T(-z,\beta,\overline{\chi})\sum_{n=1}^{\infty}\overline{\chi}(n)L(-z+1,n\beta,\overline{\chi})\bigg).$$

The integral on the extreme left side above is invariant under the transformation $\alpha \to \beta$ or under the simultaneous application of the transformations

$$\alpha \to \beta, \quad \chi \to \overline{\chi}, \quad z \to -z.$$

Thus we obtain (1-20). $\square$

Next we give an analogue of Ramanujan's transformation formula (Theorem 1.1) for even characters.

Corollary 5.1. *For an even, primitive, and nonprincipal character χ modulo q, define $P(\alpha,\chi)$ by*

$$P(\alpha,\chi) := \sqrt{\alpha}\operatorname{Re}\bigg(G(\chi)\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha}\bigg) = -\sqrt{\alpha}\operatorname{Re}\bigg(G(\chi)\sum_{n=1}^{\infty}\overline{\chi}(n)\psi(n\alpha,\overline{\chi})\bigg),$$

where $\psi(a,\chi)$ *is defined in* (1-16). *Then we have*

$$\begin{aligned}\text{(5-8)}\qquad P(\alpha,\chi)&=P(\beta,\overline{\chi})=P(\alpha,\overline{\chi})=P(\beta,\chi)\\&=\frac{1}{64\pi^{3/2}}\int_0^\infty(1+t^2)\Gamma\Big(\frac{-1+it}{4}\Big)\Gamma\Big(\frac{-1-it}{4}\Big)\\&\qquad\times\Xi\Big(\frac{t}{2},\overline{\chi}\Big)\Xi\Big(\frac{t}{2},\chi\Big)\cos\Big(\frac{1}{2}t\log\alpha\Big)\,dt.\end{aligned}$$

Proof. Using Corollary 4.4, we readily see that the double series in the definition of $P(\alpha,\chi)$ converges. Let $z\to 0$ in (1-20). Then, multiplying both sides by q and using (1-6), we have

$$\begin{aligned}\text{(5-9)}\quad \sqrt{\alpha}\bigg(G(\overline{\chi})\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\chi(n)\chi(k)}{k+n\alpha}+G(\chi)\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha}\bigg)\\=\sqrt{\beta}\bigg(G(\chi)\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\beta}+G(\overline{\chi})\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\chi(n)\chi(k)}{k+n\beta}\bigg)\\=\frac{1}{32\pi^{3/2}}\int_0^\infty(1+t^2)\Gamma\Big(\frac{-1+it}{4}\Big)\Gamma\Big(\frac{-1-it}{4}\Big)\\\times\Xi\Big(\frac{t}{2},\overline{\chi}\Big)\Xi\Big(\frac{t}{2},\chi\Big)\cos\big(\tfrac{1}{2}t\log\alpha\big)\,dt.\end{aligned}$$

Each of the first two expressions in (5-9) can be written in two different ways as real parts of a double series. Thus,

$$\begin{aligned}\sqrt{\alpha}\,\mathrm{Re}\bigg(G(\overline{\chi})\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\chi(n)\chi(k)}{k+n\alpha}\bigg)\\&=\sqrt{\alpha}\,\mathrm{Re}\bigg(G(\chi)\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha}\bigg)\\&=\sqrt{\beta}\,\mathrm{Re}\bigg(G(\chi)\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\beta}\bigg)\\&=\sqrt{\beta}\,\mathrm{Re}\bigg(G(\overline{\chi})\sum_{n=1}^\infty\sum_{k=1}^\infty\frac{\chi(n)\chi(k)}{k+n\beta}\bigg)\\&=\frac{1}{64\pi^{3/2}}\int_0^\infty(1+t^2)\Gamma\Big(\frac{-1+it}{4}\Big)\Gamma\Big(\frac{-1-it}{4}\Big)\\&\qquad\times\Xi\Big(\frac{t}{2},\overline{\chi}\Big)\Xi\Big(\frac{t}{2},\chi\Big)\cos\big(\tfrac{1}{2}t\log\alpha\big)\,dt.\end{aligned}$$

This implies (5-8). □

Moreover, if we start with the integral in Corollary 5.1, evaluate it using (2-5) with $z=0$, and make use of Corollary 4.2 when χ is even, we obtain the same result as in Corollary 5.1, except that the function $P(\alpha, \chi)$ is replaced by the function $F(\alpha, \chi)$ defined by

$$F(\alpha, \chi) := \sqrt{\alpha} G(\chi) \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha} = -\sqrt{\alpha} G(\chi) \sum_{n=1}^{\infty} \overline{\chi}(n) \psi(n\alpha, \overline{\chi}). \tag{5-10}$$

It is then trivial to see that $F(\alpha, \chi) = P(\alpha, \chi)$.

Theorem 1.7 can be analogously proved using Lemma 4.1 for χ odd. We just note that there we have to take care of the pole of

$$\Gamma\Big(1 - s + \frac{1}{2}z\Big)$$

in the integrands of two separate integrals. However, in the calculations that follow later, the two residues turn out to be additive inverses of each other and hence do not contribute anything.

The following is an analogue of Theorem 1.1 (Ramanujan's transformation formula) for odd characters.

Corollary 5.2. *For an odd, primitive and, nonprincipal character χ modulo q, define $Q(\alpha, \chi)$ by*

$$Q(\alpha, \chi) := \sqrt{\alpha} \operatorname{Im}\Big(G(\chi) \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha}\Big) = -\sqrt{\alpha} \operatorname{Im}\Big(G(\chi) \sum_{n=1}^{\infty} \overline{\chi}(n) \psi(n\alpha, \overline{\chi})\Big),$$

where $\psi(a, \chi)$ is defined as in (1-16). *Then we have*

$$\begin{aligned} Q(\alpha, \chi) &= Q(\beta, \overline{\chi}) = Q(\alpha, \overline{\chi}) = Q(\beta, \chi) \\ &= \frac{1}{4\pi^{1/2} q} \int_0^{\infty} \Gamma\Big(\frac{1+it}{4}\Big) \Gamma\Big(\frac{1-it}{4}\Big) \\ &\qquad \times \Xi\Big(\frac{t}{2}, \overline{\chi}\Big) \Xi\Big(\frac{t}{2}, \chi\Big) \cos\big(\tfrac{1}{2} t \log \alpha\big)\, dt. \end{aligned} \tag{5-11}$$

Proof. Using Corollary 4.4, we find that the double series in the definition of $Q(\alpha, \chi)$ converges. Let $z \to 0$ in Theorem 1.7. Multiplying both sides by $-q$ and using (1-5) and (1-6), we observe that

$$\begin{aligned} &\sqrt{\alpha}\Big(G(\overline{\chi}) \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\chi(n)\chi(k)}{k+n\alpha} + G(\chi) \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha}\Big) \\ &= \sqrt{\beta}\Big(G(\chi) \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\beta} + G(\overline{\chi}) \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\chi(n)\chi(k)}{k+n\beta}\Big) \\ &= \frac{i}{2\pi^{1/2} q} \int_0^{\infty} \Gamma\Big(\frac{1+it}{4}\Big) \Gamma\Big(\frac{1-it}{4}\Big) \Xi\Big(\frac{t}{2}, \overline{\chi}\Big) \Xi\Big(\frac{t}{2}, \chi\Big) \cos\big(\tfrac{1}{2} t \log \alpha\big)\, dt. \end{aligned} \tag{5-12}$$

Now, using (1-6) for odd characters to simplify (5-12), we see that

$$2i\sqrt{\alpha}\,\mathrm{Im}\biggl(G(\overline{\chi})\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\chi(n)\chi(k)}{k+n\alpha}\biggr)$$
$$=2i\sqrt{\alpha}\,\mathrm{Im}\biggl(G(\chi)\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\alpha}\biggr)$$
$$=2i\sqrt{\beta}\,\mathrm{Im}\biggl(G(\chi)\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\overline{\chi}(n)\overline{\chi}(k)}{k+n\beta}\biggr)$$
$$=2i\sqrt{\beta}\,\mathrm{Im}\biggl(G(\overline{\chi})\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\frac{\chi(n)\chi(k)}{k+n\beta}\biggr)$$
$$=\frac{i}{2\pi^{1/2}q}\int_0^{\infty}\Gamma\Bigl(\frac{1+it}{4}\Bigr)\Gamma\Bigl(\frac{1-it}{4}\Bigr)\times\Xi\Bigl(\frac{t}{2},\overline{\chi}\Bigr)\Xi\Bigl(\frac{t}{2},\chi\Bigr)\cos\bigl(\tfrac{1}{2}t\log\alpha\bigr)\,dt. \qquad \square$$

If we now start with the integral in Corollary 5.2, evaluate it using (2-5) with $z=0$, and make use of Corollary 4.2 when χ is odd, we obtain the same result as in Corollary 5.2, except that the function $Q(\alpha,\chi)$ is replaced by $-iF(\alpha,\chi)$, where $F(\alpha,\chi)$ is defined in (5-10). It is then trivial to see that $F(\alpha,\chi)=iQ(\alpha,\chi)$.

We separately record the following corollary resulting from the discussion on the previous line and the one succeeding Corollary 5.1.

Corollary 5.3. *The sum $F(\alpha,\chi)$ defined in* (5-10) *is real if χ is even and purely imaginary if χ is odd.*

6. Character analogues of the Ramanujan–Hardy–Littlewood conjecture

In this section, we prove Theorems 1.9 and 1.10. We require [Ahlgren et al. 2002, Lemma 3.1] which states that if χ is a primitive character of conductor N and $k\geq 2$ is an integer such that $\chi(-1)=(-1)^k$,

$$\frac{(k-2)!\,N^{k-2}G(\chi)}{2^{k-1}\pi^{k-2}i^{k-2}}L(k-1,\overline{\chi})=L'(2-k,\chi). \tag{6-1}$$

Proof of Theorem 1.9. From [Landau 1905], we have for $\mathrm{Re}\,s>1$,

$$\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^s}=\frac{1}{L(s,\chi)}. \tag{6-2}$$

Also, since for $-1<c=\mathrm{Re}\,s<0$,

$$(1-e^{-x})=-\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Gamma(s)x^{-s}\,ds, \tag{6-3}$$

replacing s by $s+1$, we find that for $-2 < c < -1$,

$$(1 - e^{-x}) = -\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \Gamma(s+1)x^{-s-1}\,ds. \tag{6-4}$$

Using (6-2) and (6-4), we observe that

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^2}e^{-\pi\alpha^2/(n^2q)} &= \frac{1}{L(2,\chi)} - \sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^2}(1-e^{-\pi\alpha^2/(n^2q)})\\ &= \frac{1}{L(2,\chi)} + \frac{q}{2\pi^2 i\alpha^2}\int_{c-i\infty}^{c+i\infty}\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^{-2s}}\Gamma(s+1)\Big(\frac{\pi\alpha^2}{q}\Big)^{-s}ds\\ &= \frac{1}{L(2,\chi)} + \frac{q}{2\pi^2 i\alpha^2}\int_{c-i\infty}^{c+i\infty}\frac{\Gamma(s+1)}{L(-2s,\chi)}\Big(\frac{\pi\alpha^2}{q}\Big)^{-s}ds,\end{aligned} \tag{6-5}$$

where in the second step above, we interchanged the order of summation and integration, which is valid because of absolute convergence. For χ odd, (1-10) can be put in the form

$$\Big(\frac{\pi}{q}\Big)^{-(2-s)/2}\Gamma\Big(\frac{2-s}{2}\Big)L(1-s,\overline{\chi}) = \frac{iq^{1/2}}{G(\chi)}\Big(\frac{\pi}{q}\Big)^{-(s+1)/2}\Gamma\Big(\frac{s+1}{2}\Big)L(s,\chi).$$

Hence

$$\frac{\Gamma(s+1)}{L(-2s,\chi)} = \frac{G(\overline{\chi})}{iq^{1/2}}\Big(\frac{\pi}{q}\Big)^{2s+\frac{1}{2}}\frac{\Gamma\big(\frac{1}{2}-s\big)}{L(2s+1,\overline{\chi})}. \tag{6-6}$$

Substituting (6-6) in (6-5), we observe that

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^2}e^{-\pi\alpha^2/(n^2q)} &= \frac{1}{L(2,\chi)} - \frac{G(\overline{\chi})}{2\pi^{3/2}\alpha^2}\int_{c-i\infty}^{c+i\infty}\frac{\Gamma\big(\frac{1}{2}-s\big)}{L(2s+1,\overline{\chi})}\Big(\frac{q\alpha^2}{\pi}\Big)^{-s}ds.\end{aligned} \tag{6-7}$$

We wish to shift the line of integration from Re $s = c$, $-2 < c < -1$, to Re $s = \lambda$, where $\frac{1}{2} < \lambda < \frac{3}{2}$. Consider a positively oriented rectangular contour formed by $[c - iT, \lambda - iT]$, $[\lambda - iT, \lambda + iT]$, $[\lambda + iT, c + iT]$, and $[c + iT, c - iT]$, where T is any positive real number. Let $\rho = \delta + i\gamma$ denote a nontrivial zero of $L(s,\overline{\chi})$. Let $T \to \infty$ through values such that $|T - \gamma| > \exp(-A_1\gamma/\log\gamma)$ for every ordinate γ of a zero of $L(s,\overline{\chi})$. It is known [Davenport 2000, p. 102] that for t not coinciding with the ordinate γ of a zero, and $-1 \le \sigma \le 2$,

$$\frac{L^{'}(s,\overline{\chi})}{L(s,\overline{\chi})} = \sum_{|t-\gamma|\le 1} \frac{1}{s-\rho} + O\big(\log(q(|t|+2))\big),$$

from which we can conclude that

$$\log L(s,\overline{\chi}) = \sum_{|t-\gamma|\le 1} \log(s-\rho) + O\big(\log(q(|t|+2))\big). \tag{6-8}$$

Taking real parts in (6-8) gives

$$\begin{aligned}\log|L(s,\overline{\chi})| &= \sum_{|t-\gamma|\le 1} \log|s-\rho| + O\big(\log(q(|t|+2))\big)\\ &\ge \sum_{|t-\gamma|\le 1} \log|t-\gamma| + O\big(\log(q(|t|+2))\big).\end{aligned} \tag{6-9}$$

Hence, from (6-9), we have

$$\begin{aligned}\log|L(\sigma+iT,\overline{\chi})| &\ge -\sum_{|T-\gamma|\le 1} A_1\gamma/\log\gamma + O\big(\log(q(|T|+2))\big)\\ &> -A_2T,\end{aligned} \tag{6-10}$$

where $A_2 < \pi/4$ if A_1 is small enough and $T > T_0$ for some fixed T_0. From (6-10), we see that

$$\left|\frac{1}{L(2s+1,\overline{\chi})}\right| < e^{A_3T}, \tag{6-11}$$

where $A_3 < \pi/2$. Using (1-12) and (6-11), we observe that as $T \to \infty$ through the above values, the integrals along the horizontal segments tend to zero. Now let $(\rho-1)/2 := \delta + i\gamma$ denote a nontrivial zero of $L(2s+1,\overline{\chi})$. Let $R_f(a)$ denote the residue at a of the function

$$f(s) := \frac{\Gamma\left(\frac{1}{2}-s\right)}{L(2s+1,\overline{\chi})}\left(\frac{q\alpha^2}{\pi}\right)^{-s}.$$

The nontrivial zeros of $L(2s+1,\overline{\chi})$ lie in the critical strip $-\frac{1}{2} < \operatorname{Re} s < 0$, whereas the trivial zeros are at $-1, -2, -3, \dots$. Also, $\Gamma\big(\frac{1}{2}-s\big)$ has poles at $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$. Then the residue theorem yields

$$\begin{aligned}&\int_{c-i\infty}^{c+i\infty} \frac{\Gamma\big(\frac{1}{2}-s\big)}{L(2s+1,\overline{\chi})}\Big(\frac{q\alpha^2}{\pi}\Big)^{-s} ds \\ &= \int_{\lambda-i\infty}^{\lambda+i\infty} \frac{\Gamma\big(\frac{1}{2}-s\big)}{L(2s+1,\overline{\chi})}\Big(\frac{q\alpha^2}{\pi}\Big)^{-s} ds - 2\pi i\bigg(R_f(-1) + \sum_{\rho} R_f\Big(\frac{\rho-1}{2}\Big) + R_f\big(\tfrac{1}{2}\big)\bigg),\end{aligned} \tag{6-12}$$

where

$$R_f(-1) = \lim_{s\to -1}(s+1)\frac{\Gamma\left(\frac{1}{2}-s\right)}{L(2s+1,\overline{\chi})}\left(\frac{q\alpha^2}{\pi}\right)^{-s} = \frac{\alpha^2 q}{4\sqrt{\pi}L'(-1,\overline{\chi})}, \tag{6-13}$$

$$\begin{aligned} R_f\left(\frac{\rho-1}{2}\right) &= \lim_{s\to(\rho-1)/2}\left(s-\frac{\rho-1}{2}\right)\frac{\Gamma\left(\frac{1}{2}-s\right)}{L(2s+1,\overline{\chi})}\left(\frac{q\alpha^2}{\pi}\right)^{-s} \\ &= \frac{\Gamma((2-\rho)/2)}{2L'(\rho,\overline{\chi})}\left(\frac{\pi}{q\alpha^2}\right)^{(\rho-1)/2}, \end{aligned} \tag{6-14}$$

$$R_f(1/2) = -\frac{\sqrt{\pi}}{\alpha\sqrt{q}L(2,\overline{\chi})}. \tag{6-15}$$

Of course, here we have assumed that the nontrivial zeros of $L(2s+1,\overline{\chi})$ are all simple and that $\sum_\rho R_f((\rho-1)/2)$ converges, since the aforementioned discussion regarding the integrals along the horizontal segments tending to zero as $T\to\infty$ through the chosen sequence does not imply the convergence of $\sum_\rho R_f((\rho-1)/2)$ in the ordinary sense. It only means that the series converges only when we bracket the terms in such a way that the two terms for which

$$|\gamma-\gamma'| < \exp\left(\frac{-A_1|\gamma|}{\log(|\gamma|+2)}\right)+\exp\left(\frac{-A_1|\gamma'|}{\log(|\gamma'|+2)}\right)$$

are included in the same bracket. Using (6-2) and interchanging the order of summation and integration, which is valid because of absolute convergence, we obtain

$$\begin{aligned} &\int_{\lambda-i\infty}^{\lambda+i\infty}\frac{\Gamma\left(\frac{1}{2}-s\right)}{L(2s+1,\overline{\chi})}\left(\frac{q\alpha^2}{\pi}\right)^{-s}ds \\ &\qquad = \sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n}\int_{\lambda-i\infty}^{\lambda+i\infty}\Gamma\left(\tfrac{1}{2}-s\right)\left(\frac{q\alpha^2n^2}{\pi}\right)^{-s}ds \\ &\qquad = \frac{\sqrt{\pi}}{\alpha\sqrt{q}}\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n^2}\int_{d-i\infty}^{d+i\infty}\Gamma(s)\left(\frac{\pi}{\alpha^2n^2q}\right)^{-s}ds, \end{aligned} \tag{6-16}$$

where in the penultimate line, we have made the change of variable $s\to\frac{1}{2}-s$ so that $-1<d<0$. Thus, Equations (6-3) and (6-12)–(6-16) imply

$$\begin{aligned} &\int_{c-i\infty}^{c+i\infty}\frac{\Gamma\left(\frac{1}{2}-s\right)}{L(2s+1,\overline{\chi})}\left(\frac{q\alpha^2}{\pi}\right)^{-s}ds = -\frac{2\pi^{3/2}i}{\alpha\sqrt{q}}\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n^2}\left(1-e^{-\pi/(\alpha^2n^2q)}\right) \\ &\quad -2\pi i\left(\frac{\alpha^2 q}{4\sqrt{\pi}L'(-1,\overline{\chi})}+\sum_\rho\frac{\Gamma((2-\rho)/2)}{2L'(\rho,\overline{\chi})}\left(\frac{\pi}{q\alpha^2}\right)^{(\rho-1)/2}-\frac{\sqrt{\pi}}{\alpha\sqrt{q}L(2,\overline{\chi})}\right). \end{aligned} \tag{6-17}$$

From (6-7), (6-17), and the fact that $\alpha\beta = 1$ and $\sqrt{G(\chi)G(\overline{\chi})} = i\sqrt{q}$, we find that

$$\begin{aligned}(6\text{-}18)\quad & \alpha\sqrt{\alpha}\sqrt{G(\chi)}\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^2}e^{-\pi\alpha^2/(n^2q)} \\ &= \frac{\alpha\sqrt{\alpha}\sqrt{G(\chi)}}{L(2,\chi)} - \frac{\beta\sqrt{\beta}\sqrt{G(\overline{\chi})}}{L(2,\overline{\chi})} + \beta\sqrt{\beta}\sqrt{G(\overline{\chi})}\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n^2}e^{-\pi\beta^2/(n^2q)} \\ &\quad - \frac{\alpha\sqrt{\alpha}q^{3/2}\sqrt{G(\overline{\chi})}}{4\pi L'(-1,\overline{\chi})} - \frac{q\sqrt{G(\overline{\chi})}}{2\pi\sqrt{\beta}}\sum_{\rho}\frac{\Gamma((2-\rho)/2)}{L'(\rho,\overline{\chi})}\left(\frac{\pi}{q}\right)^{\rho/2}\beta^{\rho} + \frac{\beta\sqrt{\beta}\sqrt{G(\overline{\chi})}}{L(2,\overline{\chi})}.\end{aligned}$$

Applying (6-1) with $N = q$ and $k = 3$, and replacing χ by $\overline{\chi}$ gives

$$(6\text{-}19)\qquad \frac{1}{L'(-1,\overline{\chi})} = \frac{4\pi i}{qG(\overline{\chi})L(2,\chi)}.$$

Thus (6-18) and (6-19) yield

$$\begin{aligned}(6\text{-}20)\quad & \alpha\sqrt{\alpha}\sqrt{G(\chi)}\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n^2}e^{-\pi\alpha^2/(n^2q)} \\ &\qquad - \beta\sqrt{\beta}\sqrt{G(\overline{\chi})}\sum_{n=1}^{\infty}\frac{\overline{\chi}(n)\mu(n)}{n^2}e^{-\pi\beta^2/(n^2q)} \\ &\qquad\qquad = -\frac{q\sqrt{G(\overline{\chi})}}{2\pi\sqrt{\beta}}\sum_{\rho}\frac{\Gamma((2-\rho)/2)}{L'(\rho,\overline{\chi})}\left(\frac{\pi}{q}\right)^{\rho/2}\beta^{\rho}.\end{aligned}$$

Switching the roles of α and β and those of χ and $\overline{\chi}$ gives

$$\begin{aligned}(6\text{-}21)\quad & \frac{q\sqrt{G(\chi)}}{2\pi\sqrt{\alpha}}\sum_{\rho}\frac{\Gamma((2-\rho)/2)}{L'(\rho,\chi)}\left(\frac{\pi}{q}\right)^{\rho/2}\alpha^{\rho} \\ &\qquad + \frac{q\sqrt{G(\overline{\chi})}}{2\pi\sqrt{\beta}}\sum_{\rho}\frac{\Gamma((2-\rho)/2)}{L'(\rho,\overline{\chi})}\left(\frac{\pi}{q}\right)^{\rho/2}\beta^{\rho} = 0.\end{aligned}$$

Finally (6-20) and (6-21) give (1-24) upon simplification. □

Remark. The approach used above for proving that the integrals along the horizontal segments tend to zero as $T \to \infty$ through the chosen sequence is adapted from [Titchmarsh 1986, p. 219].

To prove Theorem 1.10, we require the following lemma.

Lemma 6.1. $$\sum_{n=1}^{\infty}\frac{\chi(n)\mu(n)}{n} = \frac{1}{L(1,\chi)}.$$

Proof. Dividing n into its residue classes mod q by letting $n = qr + b$, $0 \leq r < \infty$, $0 \leq b \leq q-1$, we find that since χ has period q,

$$\sum_{n=1}^{\infty} \frac{\chi(n)\mu(n)}{n} = \sum_{r=0}^{\infty}\sum_{b=0}^{q-1} \frac{\chi(b)\mu(qr+b)}{qr+b} = \sum_{b=0}^{q-1} \chi(b) \sum_{r=0}^{\infty} \frac{\mu(qr+b)}{qr+b}. \tag{6-22}$$

The series $\sum_{r=0}^{\infty} \mu(qr+b)/(qr+b)$ was first studied by Kluyver [1904] and its convergence was proved by Landau [1905]. In fact, Landau gave an explicit representation for this series in terms of a finite sum consisting of L-functions. Thus (6-22) implies convergence of $\sum_{n=1}^{\infty} \chi(n)\mu(n)/n$. Then using (6-2) and an analogue of Abel's theorem for power series, we see that

$$\sum_{n=1}^{\infty} \frac{\chi(n)\mu(n)}{n} = \lim_{s\to 1} \sum_{n=1}^{\infty} \frac{\chi(n)\mu(n)}{n^s} = \lim_{s\to 1} \frac{1}{L(s,\chi)} = \frac{1}{L(1,\chi)}. \qquad \square$$

Proof of Theorem 1.10. The proof is very similar to that of Theorem 1.9 and hence we omit the details. However we note that Lemma 6.1, (1-10) in the form [Davenport 2000, p. 69]

$$\pi^{-(1-s)/2} q^{(1-s)/2} \Gamma\Big(\frac{1-s}{2}\Big) L(1-s,\overline{\chi}) = \frac{q^{1/2}}{G(\chi)} \pi^{-s/2} q^{s/2} \Gamma\Big(\frac{s}{2}\Big) L(s,\chi),$$

and (6-1) with $N = q$ and $k = 2$ are used in the proof. $\square$

7. Open problems

Following are some open problems with which we will conclude.

(1) We have indirectly given the proof of the fact that function $F(\alpha,\chi)$ defined in (5-10) is real (respectively purely imaginary) when χ is even (respectively odd). Prove this directly; that is, without using Corollaries 5.1 and 5.2 and the integrals in those corollaries.

(2) Since (1-23) is of the form $F(\alpha) = F(\beta)$, where $\alpha\beta = 1$, it is natural to ask if there exists an integral representation involving the Riemann Ξ-function equal to the two expressions in (1-23). Finding an integral representation for either side of (1-23) may shed light on the convergence of

$$\sum_{\rho} \frac{\Gamma\left((1-\rho)/2\right) a^{\rho}}{\zeta'(\rho)},$$

provided, of course, that the integral converges in the first place. It should be remarked here that Hardy and Littlewood [1916, p. 161] have shown that the relation

$$P(y) = O(y^{-\frac{1}{4}+\delta}), \tag{7-1}$$

where $P(y)=\sum_{n=1}^{\infty}(-y)^n/(n!\zeta(2n+1))$ can be derived from (1-23) if we assume the Riemann hypothesis and the absolute convergence of

$$\sum_{\rho}\frac{\Gamma((1-\rho)/2)}{\zeta'(\rho)}.$$

They have further shown that (7-1) is a necessary and sufficient condition for the Riemann hypothesis to be true.

Similarly, it is natural to ask if the expressions in (1-24) and (1-25) have integral representations involving $\Xi(t/2,\chi)$.

Acknowledgements

The author thanks Professor Bruce C. Berndt for his constant support, careful reading of the manuscript and for several helpful comments, and to Professors Alexandru Zaharescu, Kevin Ford, and Adolf Hildebrand for their help. He also thanks Becky S. Burner, Timothy W. Cole, and Margaret A. Lewis from the Mathematics library at University of Illinois at Urbana-Champaign, and Anton Lukyanenko for their great help in locating the Russian references and citations. He thanks Reetesh Ranjan for numerically verifying some of the identities by programming in Fortran. The author also thanks the referee for his valuable suggestions which improved the quality of this paper.

References

[Ahlgren et al. 2002] S. Ahlgren, B. C. Berndt, A. J. Yee, and A. Zaharescu, "Integrals of Eisenstein series and derivatives of L-functions", *Int. Math. Res. Not.* **2002**:32 (2002), 1723–1738. MR 2003f:11056 Zbl 1086.11023

[Apostol 1972] T. M. Apostol, "Dirichlet L-functions and primitive characters", *Proc. Amer. Math. Soc.* **31**:2 (1972), 384–386. MR 44 #2717 Zbl 0239.10023

[Apostol 1976] T. M. Apostol, *Introduction to analytic number theory*, Springer, New York, 1976. MR 55 #7892 Zbl 0335.10001

[Berndt 1970] B. C. Berndt, "Identities involving the coefficients of a class of Dirichlet series, IV", *Trans. Amer. Math. Soc.* **149**:1 (1970), 179–185. MR 41 #5309 Zbl 0207.05504

[Berndt 1975] B. C. Berndt, "Character analogues of the Poisson and Euler–MacLaurin summation formulas with applications", *J. Number Theory* **7**:4 (1975), 413–445. MR 52 #3075 Zbl 0316.10023

[Berndt 1998] B. C. Berndt, *Ramanujan's notebooks, V*, Springer, New York, 1998. MR 99f:11024 Zbl 0886.11001

[Berndt and Dixit 2010] B. C. Berndt and A. Dixit, "A transformation formula involving the gamma and Riemann zeta functions in Ramanujan's lost notebook", pp. 199–210 in *The legacy of Alladi Ramakrishnan in the mathematical sciences*, edited by K. Alladi et al., Springer, New York, 2010. MR 2011m:11171

[Berndt et al. 2008] B. C. Berndt, Y. Lee, and J. Sohn, "Koshliakov's formula and Guinand's formula in Ramanujan's lost notebook", pp. 21–42 in *Surveys in number theory*, edited by K. Alladi, Dev. Math. **17**, Springer, New York, 2008. MR 2010a:33011 Zbl 1183.33007

[Berndt et al. 2011] B. C. Berndt, A. Dixit, and J. Sohn, "Character analogues of theorems of Ramanujan, Koshliakov and Guinand", *Adv. in Appl. Math.* **46**:1-4 (2011), 54–70. MR 2012c:33017 Zbl 1216.33012

[Bogolyubov et al. 1990] N. N. Bogolyubov, L. D. Faddeev, A. Y. Ishlinskiĭ, V. N. Koshlyakov, and Y. A. Mitropol'skiĭ, "Nikolai Sergeevich Koshlyakov (on the centenary of his birth)", *Uspekhi Mat. Nauk* **45**:4 (1990), 173–176. In Russian; translated in *Russian Math. Surveys* **45**:4 (1990), 197–202. MR 91i:01064

[Davenport 2000] H. Davenport, *Multiplicative number theory*, 3rd ed., Grad. Texts in Math. **74**, Springer, New York, 2000. MR 2001f:11001 Zbl 1002.11001

[Dixit 2010] A. Dixit, "Series transformations and integrals involving the Riemann Ξ-function", *J. Math. Anal. Appl.* **368**:1 (2010), 358–373. MR 2011h:11091 Zbl 1206.11104

[Dixit 2011a] A. Dixit, "Analogues of a transformation formula of Ramanujan", *Int. J. Number Theory* **7**:5 (2011), 1151–1172. MR 2825965 Zbl 1229.11113

[Dixit 2011b] A. Dixit, "Transformation formulas associated with integrals involving the Riemann Ξ-function", *Monatsh. Math.* **164**:2 (2011), 133–156. Zbl 05973790

[Ferrar 1936] W. L. Ferrar, "Some solutions of the equation $F(t) = F(t^{-1})$", *J. London Math. Soc.* **11**:2 (1936), 99–103. Zbl 0013.39204

[Guinand 1941] A. P. Guinand, "On Poisson's summation formula", *Ann. of Math.* (2) **42**:3 (1941), 591–603. MR 3,109e Zbl 0025.33802

[Guinand 1955] A. P. Guinand, "Some rapidly convergent series for the Riemann ξ-function", *Quart. J. Math. Oxford Ser.* (2) **6**:1 (1955), 156–160. MR 17,16a Zbl 0065.27703

[Hardy 1915] G. H. Hardy, "Note by Mr. G. H. Hardy on the preceding paper", *Quart. J. Math.* **46** (1915), 260–261.

[Hardy and Littlewood 1916] G. H. Hardy and J. E. Littlewood, "Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes", *Acta Math.* **41**:1 (1916), 119–196. MR 1555148 Zbl 46.0498.01

[Kluyver 1904] J. C. Kluyver, "Series derived from the series $\sum \frac{\mu(m)}{m}$", *Koningl. Akad. Wetensch. Amsterdam Proc. Sect. Sci.* **6** (1904), 305–312.

[Koshlyakov 1928] N. S. Koshlyakov, "On Voronoï's sum-formula", *Mess. Math.* **58** (1928), 30–32. Zbl 54.0197.01

[Koshlyakov 1934a] N. S. Koshlyakov, "On a general summation formula and its applications", *C. R. (Dokl.) Acad. Sci. URSS* **4** (1934), 187–191. In Russian. Zbl 0010.30502

[Koshlyakov 1934b] N. S. Koshlyakov, "Some identities in quadratic fields", *C. R. (Dokl.) Acad. Sci. URSS* **2**:9 (1934), 527–531. In Russian. Zbl 0009.34004

[Koshlyakov 1934c] N. S. Koshlyakov, "Some integral representations of the square of Riemann's function $\Xi(t)$", *C. R. (Dokl.) Acad. Sci. URSS* **2** (1934), 401–405. In Russian. Zbl 0009.19801

[Koshlyakov 1936] N. S. Koshlyakov, "On an extension of some formulae of Ramanujan", *Proc. London Math. Soc.* **41**:1 (1936), 26–32. Zbl 0013.35402

[Koshlyakov 1937] N. S. Koshlyakov, "On a transformation of definite integrals and its application to the theory of Riemann's function $\zeta(s)$", *C. R. (Dokl.) Acad. Sci. URSS* **15** (1937), 3–8. Zbl 0017.01403

[Koshlyakov 1949] N. S. Sergeev, *A study of a class of transcendental functions defined by the generalized Riemann equation*, Izdat. Akad. Nauk SSSR, Moscow, 1949. In Russian. MR 13,24d Zbl 0045.18401

[Koshlyakov 1954] N. S. Koshlyakov, "Investigation of some questions of the analytic theory of a rational and quadratic field, I", *Izvestiya Akad. Nauk SSSR. Ser. Mat.* **18**:2 (1954), 113–144. In Russian. MR 16,15d

[Landau 1905] E. Landau, "Remarks on the paper of Mr. Kluyver on page 305 of Vol. VI: Series derived from the series $\sum \frac{\mu(m)}{m}$", *Koningl. Akad. Wetensch. Amsterdam Proc. Sect. Sci.* **7** (1905), 66–77.

[Ma et al. 2010] R. Ma, Y. Yi, and Y. Zhang, "On the mean value of the generalized Dirichlet L-functions", *Czechoslovak Math. J.* **60**:3 (2010), 597–620. MR 2011j:11162 Zbl 1224.11077

[Oberhettinger 1974] F. Oberhettinger, *Tables of Mellin transforms*, Springer, New York, 1974. MR 50 #5376 Zbl 0289.44003

[Paris and Kaminski 2001] R. B. Paris and D. Kaminski, *Asymptotics and Mellin–Barnes integrals*, Encyclopedia Math. Appl. **85**, Cambridge University Press, Cambridge, 2001. MR 2002h:33001 Zbl 0983.41019

[Ramanujan 1915] S. Ramanujan, "New expressions for Riemann's functions $\xi(s)$ and $\Xi(t)$", *Quart. J. Math.* **46** (1915), 253–261. MR 2280857 Zbl 45.1343.05

[Ramanujan 1927] S. Ramanujan, *Collected papers of Srinivasa Ramanujan*, edited by G. H. Hardy et al., Cambridge University Press, Cambridge, 1927. Reprinted by the Amer. Math. Soc., Providence, RI, 2000. MR 2008b:11002 Zbl 53.0030.02

[Ramanujan 1957] S. Ramanujan, *Notebooks, 1, 2*, Tata Institute of Fundamental Research, Bombay, 1957. MR 20 #6340 Zbl 0138.24201

[Ramanujan 1988] S. Ramanujan, *The lost notebook and other unpublished papers*, Springer, Berlin, 1988. MR 89j:01078 Zbl 0639.01023

[Temme 1996] N. M. Temme, *Special functions: an introduction to the classical functions of mathematical physics*, Wiley, New York, 1996. MR 97e:33002 Zbl 0856.33001

[Titchmarsh 1986] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., edited by D. R. Heath-Brown, Clarendon, Oxford, 1986. MR 88c:11049 Zbl 0601.10026

Received February 13, 2011.

ATUL DIXIT
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
1409 WEST GREEN STREET
URBANA, ILLINOIS 61801
UNITED STATES

aadixit2@illinois.edu
http://www.math.uiuc.edu/~aadixit2/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

SPECTRAL THEORY FOR LINEAR RELATIONS VIA LINEAR OPERATORS

DANA GHEORGHE AND FLORIAN-HORIA VASILESCU

We develop a spectral theory for closed linear operators of the form $T : D(T) \subset X \mapsto X/X_0$, where X is a complex Banach space and X_0 a closed vector subspace of it. This approach, essentially expressed in terms of linear operators, provides a better understanding of the spectral theory for closed linear relations.

1. Introduction

As in the case of linear operators, the spectral theory of linear relations, including the associated analytic functional calculus, is an important tool for studying various properties of these objects and for deriving some of their applications. Results related to the spectral theory of linear relations and its applications can be found in [Baskakov and Chernyshov 2002; Baskakov and Zagorskiĭ 2007; Cross 1998; Favini and Yagi 1993; Favini and Yagi 1999] and elsewhere.

In this paper we emphasize the strong connection between the spectral theory of closed linear relations and that of some closed linear operators. As a matter of fact, we develop a spectral theory for a certain class of linear operators, obtaining as consequences most of the main spectral properties of linear relations.

Our concept of spectrum is equivalent to that of *extended spectrum* of a linear relation, as given by [Baskakov and Chernyshov 2002, Definition 1.5]; see also [Cross 1998, Section VI.4], where it is called *augmented spectrum*. In particular, the point ∞ is in the spectrum unless the quotient range operator is an ordinary everywhere-defined bounded operator (see Proposition 11).

Let us introduce some notation and definitions.

Let X be a complex Banach space and let $\mathcal{B}(X)$ denote the Banach algebra of all bounded linear operators from X into X. Let also $X_0 \subset X$ be a closed vector subspace, and let $J_0 : X \mapsto X/X_0$ be the canonical projection. The identity operator on X will be usually denoted by I.

D. Gheorghe gratefully acknowledges support from the grant PN-II-RU-PD-2011-3-0052 (CNCS-Romania).
MSC2010: primary 47A06, 47A10, 47A60; secondary 47A56.
Keywords: linear relations, spectrum, analytic functional calculus.

We are interested in linear operators of the form $T : D(T) \subset X \mapsto X/X_0$, where $D(T)$ is, of course, the domain of T. (The use of such operators is inspired by the works [Albrecht and Vasilescu 1986; Waelbroeck 1982]; see also [Gheorghe and Vasilescu 2009].) Such an operator is said to be a *quotient range operator*. Although X/X_0 is itself a Banach space, its quotient space form plays an important role in what follows. As a matter of fact, the class of closed quotient range operators is in one-to-one correspondence with the class of closed linear relations (see the definition below), and they have important similar properties. Note that the formula $T : D(T) \subset X \mapsto X/X_0$ implies that T is a quotient range operator, and in such situations the expression "quotient range" will be often omitted.

If $T : D(T) \subset X \mapsto X/X_0$, we denote, as usual, by $N(T)$, $R(T)$ and $G(T)$ the null-space, the range and the graph of T. Let $R_0(T)$ be given by $R(T) = R_0(T)/X_0$, and $G_0(T) = \{(x, y) \in X \times X; x \in D(T), J_0(y) = T(x)\}$, which are called, with the terminology of [Albrecht and Vasilescu 1986], the *lifted range* and *lifted graph*, respectively.

Following Arens [1961], any linear subspace Z of $X \times X$ is called a *linear relation* in X. Given a linear relation $Z \subset X \times X$, we associate it, as usual (see [Arens 1961; Cross 1998]), with the following subspaces:

$$D(Z) = \{u \in X; (u, v) \in Z \text{ for some } v \in X\}, \quad N(Z) = \{u \in D(Z); (u, 0) \in Z\},$$
$$R(Z) = \{v \in X; (u, v) \in Z \text{ for some } u \in X\}, \quad M(Z) = \{v \in R(Z); (0, v) \in Z\}.$$

The left two are called the domain of Z, the range of Z; the right are the kernel of Z and the multivalued part of Z. When $M(Z) = \{0\}$, then Z is the graph of a linear operator. We often identify the relation given by the graph of an operator with the operator itself.

Given an arbitrary relation $Z \subset X \times X$, to avoid any confusion with the inverse of an operator, we will denote the *reverse relation* $\{(y, x) \in X \times X; (x, y) \in Z\}$ by $Z^\dagger$.

The strong connection between linear relations and quotient range operators is well known and easily explained; see [Cross 1998; Gheorghe and Vasilescu 2009] for example. Namely, given an operator $T : D(T) \subset X \mapsto X/X_0$, the space $Z_T = G_0(T) \subset X \times X$ is a linear relation. Conversely, given a linear relation $Z \subset X \times X$, with $M(T)$ closed (which is automatic in the framework which will be used in the sequel), the linear operator $Q_Z : D(Z) \mapsto X/M(Z)$, given by $Q_Z(x) = y + M(Z)$ whenever $(x, y) \in Z$, is a quotient range operator. Moreover, this correspondence is one-to-one. This connection will be exploited to develop a spectral theory for linear relations. The simple but crucial remark leading to this development is that for a closed relation $Z \subset X \times X$, the reverse relation $Z^\dagger$ is (the graph of) a bounded operator if and only if the operator $Q_Z : D(Z) \mapsto X/M(Z)$ has a bounded inverse.

Given a linear relation Z and a complex number $\lambda \in \mathbb{C}$, we consider the linear relations $\lambda I - Z = \{(u, \lambda u - v);\ (u, v) \in Z\}$ and $(\lambda I - Z)^\dagger = \{(\lambda u - v, u);\ (u, v) \in Z\}$ (see Section 5). If we assume that Z is closed, $N(\lambda I - Z) = \{0\}$ and $R(\lambda I - Z) = X$, then we have that $(\lambda I - Z)^\dagger$ is (the graph of) a closed everywhere-defined linear operator (which is, in general, neither surjective nor injective; see Example 32), and hence $(\lambda I - Z)^\dagger \in \mathcal{B}(X)$. Because the bounded operator $(\lambda I - Z)^\dagger$ exists if and only if the operator $\lambda J_Z - Q_Z : D(Z) \subset X \mapsto X/M(Z)$ has a bounded inverse (see Remark 4(ii)), where $J_Z : X \mapsto X/M(Z)$ is the canonical projection, the spectral theory of these objects can be simultaneously developed. However, in our opinion, the spectral theory of quotient range operators is easier to handle.

Our main tool is an analytic functional calculus for quotient range operators, defined in Section 2 by using the classical Riesz–Dunford–Waelbroeck integral formula; see [Dunford and Schwartz 1958; Waelbroeck 1954]. A similar formula, valid for linear relations, is also used in [Baskakov and Chernyshov 2002]. Nevertheless, an analytic functional calculus in its full generality seems to appear only in the present work.

The analytic functional calculus allows us to recapture, in terms of operators, most of the main spectral properties known for linear relations; see especially [Cross 1998; Baskakov and Chernyshov 2002]. Among some simplifications, we mention that our approach avoids the use of the concept of *pseudoresolvent*, as well as that of *invariant subspace*, as done in [Baskakov and Chernyshov 2002]. Other differences between our approach and that of the quoted works will be discussed in due course. We should also mention that a calculus with the exponential function and with fractional powers has been already used in [Favini and Yagi 1993] to obtain a Hille–Yoshida–Phillips-type theorem for linear relations.

The paper is organized as follows. In Section 2, we introduce a notion of spectrum for quotient range operators (equivalent to that for linear relations) in the Riemann sphere $\mathbb{C}_\infty$, and construct a functional calculus with analytic functions in neighborhoods of this spectrum. As mentioned above, our Theorem 16, asserting in particular the multiplicativity of the analytic functional calculus, seems to be new in this context (as well as in that of linear relations). In Section 3, we study quotient range operators with unbounded spectrum and nonempty resolvent set. The existence of a spectral decomposition corresponding to separate parts of the spectrum as well as a spectral mapping theorem are presented herein. In Section 4, we study the class of quotient range operators for which the point ∞ is an isolated point of the spectrum. In Section 5, we investigate some connections between the analytic functional calculus and the Arens polynomial calculus [Arens 1961].

2. Spectrum and analytic functional calculus for closed quotient range operators

As in the introduction, X denotes a complex Banach space, X_0 a closed vector subspace of it, and $J_0 : X \mapsto X/X_0$ the canonical projection. The symbol $\mathbb{C}_\infty$ denotes the one-point compactification of $\mathbb{C}$. We designate by $\mathcal{B}(X, Y)$ the Banach space of all bounded linear operators from X into another Banach space Y. As usually, $\mathcal{B}(X, X)$ is denoted by $\mathcal{B}(X)$.

Let $T : D(T) \subset X \mapsto X/X_0$ be a closed linear operator. We denote by $\rho_A(T)$ the *Arens resolvent set* of T, that is, the set of those $\lambda \in \mathbb{C}$ such that $(\lambda J_0 - T)^{-1} \in \mathcal{B}(X/X_0, X)$. The *Arens spectrum* of T is the set $\sigma_A(T) := \mathbb{C} \setminus \rho_A(T)$. Because $\lambda J_0 - T : D(T) \subset X \mapsto X/X_0$ is closed, we have $\lambda \in \rho_A(T)$ if and only if $\lambda J_0 - T$ is bijective.

Remark 1. Given two complex Banach spaces X_1, X_2, we denote by $X_1 \oplus X_2$ their direct sum, endowed with a convenient norm, compatible with the norms of X_1, X_2.

Let $T_j : D(T_j) \subset X_j \mapsto X_j/X_{0j}$ for $j = 1, 2$ be quotient range operators. Then the map

$$T_1 \oplus T_2 : D(T_1) \oplus D(T_2) \subset X_1 \oplus X_2 \mapsto (X_1/X_{01}) \oplus (X_2/X_{02})$$

may be regarded as a quotient range operator, provided we identify the Banach space $(X_1/X_{01}) \oplus (X_2/X_{02})$ with the Banach space $(X_1 \oplus X_2)/(X_{01} \oplus X_{02})$, using the natural isomorphism

$$V : (X_1/X_{01}) \oplus (X_2/X_{02}) \mapsto (X_1 \oplus X_2)/(X_{01} \oplus X_{02}) \tag{1}$$

given by the assignment

$$(X_1/X_{01}) \oplus (X_2/X_{02}) \ni (x_1 + X_{01}) \oplus (x_2 + X_{02}) \mapsto$$

$$x_1 \oplus x_2 + X_{01} \oplus X_{02} \in (X_1 \oplus X_2)/(X_{01} \oplus X_{02}).$$

We write

$$T_1 \oplus_q T_2 := V(T_1 \oplus T_2).$$

In particular, given $T : D(T) \subset X \mapsto X/X_0$ closed such that there are closed vector subspaces X_1, X_2 of X and X_{01}, X_{02} of X_0 with $X = X_1 \oplus X_2$, $X_0 = X_{01} \oplus X_{02}$, $D(T) = (D(T) \cap X_1) \oplus (D(T) \cap X_2)$, and closed operators $T_j : D(T_j) \subset X_j \mapsto X_j/X_{0j}$ with $D(T_j) = D(T) \cap X_j$ for $j = 1, 2$ and $T(x_1 \oplus x_2) = V(T_1 x_1 \oplus T_2 x_2)$ for all $x_1 \oplus x_2 \in D(T_1) \oplus D(T_2)$, we have $T = T_1 \oplus_q T_2$.

Definition 2. Let $T : D(T) \subset X \mapsto X/X_0$ be closed.

(1) Assume $\sigma_A(T)$ bounded, and let $m \geq 0$ be an integer. The point ∞ is said to be *m-regular* for T if the set $\{\lambda^{1-m}(\lambda J_0 - T)^{-1}J_0; |\lambda| \geq r\}$ is bounded in $\mathcal{B}(X)$ for some $r > \sup_{\lambda\in\sigma_A(T)} |\lambda|$.

(2) If ∞ is not 0-regular we put $\sigma(T) = \sigma_A(T) \cup \{\infty\}$.

(3) Assume ∞ to be 0-regular and $X_0 \neq \{0\}$. If $T = T_0 \oplus_q T_1$, $T_0 : \{0\} \subset X_0 \mapsto X_0/X_0 = \{0\}$, we put $\sigma(T) = \sigma_A(T) \cup \{\infty\}$; otherwise, $\sigma(T) = \sigma_A(T)$.

(4) If ∞ is 0-regular and $X_0 = \{0\}$, we put $\sigma(T) = \sigma_A(T)$.

The set $\sigma(T)$ is called the *spectrum* of T, and the set $\rho(T) = \mathbb{C}_\infty \setminus \sigma(T)$ is called the *resolvent set* of T.

The set $\sigma(T)$ is nonempty except for $X_0 = X = \{0\}$ (see Proposition 7), but it may be equal to $\mathbb{C}_\infty$. For practical reasons, in this paper we work only with (quotient range) operators with nonempty resolvent set.

Example 3. The well-known fact that any continuous linear operator on a Banach space X has a bounded spectrum is no longer true in the case of quotient range operators, as we can see in the following example.

Let X be the Hilbert space of all square-summable complex sequences, let $A \in \mathcal{B}(X)$ be the shift

$$A((x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots)$$

and let

$$X_0 = \{(x_1, x_2, 0, 0, \dots) : x_1, x_2 \in \mathbb{C}\}.$$

Consider the operator T defined by $Tx = Ax + X_0$ for $x \in X$. Clearly T is continuous and thus closed. We will show that $\sigma(T)$ is unbounded. Let $\lambda \in \mathbb{C}$. We have

$$\begin{aligned} x \in N(\lambda J_0 - T) &\iff \lambda x - Ax + y = 0 \qquad \text{for some } y \in X_0,\\ &\iff \left.\begin{aligned} -\lambda x_1 &= y_1,\\ x_1 - \lambda x_2 &= y_2,\\ x_k - \lambda x_{k+1} &= 0 \quad \text{for } k \geq 2 \end{aligned}\right\} \quad \text{for some } y_1, y_2 \in \mathbb{C}. \end{aligned}$$

For $|\lambda| > 1$ and $x_1, x_2 \in \mathbb{C}$ consider $x_3 = -x_1/\lambda^2 - x_2/\lambda$. Then

$$(-x_1/\lambda, x_3, x_3/\lambda, \dots, x_3/\lambda^k, \dots) \in N(\lambda J_0 - T),$$

which implies that $N(\lambda J_0 - T) \neq \{0\}$ for $|\lambda| > 1$. Consequently,

$$\{\lambda \in \mathbb{C} : |\lambda| > 1\} \subset \sigma(T),$$

so $\sigma(T)$ is unbounded.

Remark 4. (i) For a closed operator $T : D(T) \subset X \mapsto X$ — in particular for an everywhere defined bounded operator on X — Definition 2 provides the usual definition of the spectrum. Note that the density of $D(T)$ in X is not required. For instance, if A is the operator from Example 3, which is injective, and $T = A^{-1}$, then T is not densely defined but $0 \notin \sigma(T)$, so $\rho(T) \neq \varnothing$.

Note also that if $0 : \{0\} \subset X \mapsto X$, we have $\sigma(0) = \mathbb{C}_\infty$ if $X \neq \{0\}$ and $\sigma(0) = \varnothing$ if $X = \{0\}$, by Definition 2.

If $X_0 = X \neq \{0\}$ and $0 : \{0\} \subset X \mapsto X/X_0 = \{0\}$, then 0 is a quotient range operator, whose Arens spectrum is empty, and $\sigma(T) = \{\infty\}$, by Definition 2.

(ii) Let $Z \subset X \times X$ be a closed relation. It is clear that the subspace $M(Z) \subset X$ is closed. As in the introduction, we consider the (quotient range) operator $Q_Z : D(Z) \mapsto X/M(Z)$ given by $Q_Z(x) = y + M(Z)$ whenever $(x, y) \in Z$, which is closed.

Let $J_Z : X \mapsto X/M(Z)$ be the canonical projection. Given $\lambda \in \mathbb{C}$, the operator $\lambda J_Z - Q_Z$ is again closed. If $(\lambda I - Z)^\dagger \in \mathcal{B}(X)$, then $\lambda J_Z - Q_Z$ has an everywhere-defined, and hence bounded, inverse. Indeed, $(\lambda I - Z)^\dagger$ exists if and only if for every $u \in X$ we can find a unique $x \in X$ such that $(x, y) \in Z$ and $\lambda x - y = u$ for some $y \in X$. Moreover, $x = 0$ if and only if $u \in M(Z)$. Hence $\lambda J_Z x - Q_Z x = J_Z u$, showing that $\lambda J_Z - Q_Z$ is bijective.

Conversely, if Q_Z is closed, then Z is closed. In addition, if $\lambda J_Z - Q_Z$ is bijective, for every $u \in X$ we put $x = (\lambda J_Z - Q_Z)^{-1} J_Z u$. Then we have $\lambda x - y = u$ for some $y \in X$ with $(x, y) \in Z$, and so $(\lambda I - Z)^\dagger$ does exist. Evidently, $(\lambda I - Z)^\dagger = (\lambda J_Z - Q_Z)^{-1} J_Z$.

From this discussion it clearly follows that we may define the *Arens resolvent set* and *Arens spectrum* of a closed relation $Z \subset X \times X$ by the equalities $\rho_A(Z) = \rho_A(Q_Z)$ and $\sigma_A(Z) = \sigma_A(Q_Z)$, respectively. Similarly, we may define the *resolvent set* and *spectrum* of a closed relation $Z \subset X \times X$ via the equalities $\rho(Z) = \rho(Q_Z)$ and $\sigma(Z) = \sigma(Q_Z)$. Consequently, most of the spectral properties obtained for a quotient range operator can be translated into properties for linear relations. This definition of the spectrum of a linear relation coincides with the corresponding definition [Cross 1998, Definition VI.4.1] or [Baskakov and Chernyshov 2002, Definition 1.5], because the condition $\lim_{|\lambda| \to \infty} (\lambda I - Z)^\dagger = 0$ is equivalent to the fact that ∞ is a 0-regular point for Q_Z.

In fact, given an integer $m \geq 0$, we may say that the point ∞ is *m-regular* for the closed linear relation Z if ∞ is m-regular for the operator Q_Z.

As an example, if $Z = \{0\} \times X$ ($X \neq \{0\}$), then $D(Z) = \{0\}$, $M(Z) = X$ and $Q_Z : \{0\} \subset X \mapsto X/X = \{0\}$. Therefore, $\sigma(Z) = \sigma(Q_Z) = \{\infty\}$, as in (i).

Definition 5. Let $T : D(T) \subset X \mapsto X/X_0$ be closed, with $\rho_A(T) \neq \varnothing$. The function

$$\rho_A(T) \ni \lambda \mapsto (\lambda J_0 - T)^{-1} J_0 \in \mathcal{B}(X)$$

is called the *resolvent (function)* of T. We also put $R(\lambda, T) = (\lambda J_0 - T)^{-1} J_0$.

As in the case of linear relations (see [Cross 1998; Favini and Yagi 1993]), we have a *resolvent equation*, which is very useful for the construction of the analytic functional calculus.

Lemma 6. *If $\lambda, \mu \in \rho_A(T)$, then*

$$R(\mu, T) - R(\lambda - T) = (\lambda - \mu) R(\mu, T) R(\lambda, T).$$

Proof. Indeed, for all $\lambda, \mu \in \rho_A(T)$, we have the identity

$$(\mu J_0 - T)^{-1} J_0 - (\lambda J_0 - T)^{-1} J_0 = (\lambda - \mu)(\mu J_0 - T)^{-1} J_0 (\lambda J_0 - T)^{-1} J_0,$$

which is easily checked. □

As in the case of ordinary operators, the resolvent set is open and the resolvent function is holomorphic on it.

Proposition 7. *The resolvent sets $\rho_A(T)$ and $\rho(T)$ are open subsets of $\mathbb{C}$ and $\mathbb{C}_\infty$ respectively, and the resolvent function $\lambda \mapsto R(\lambda, T)$ is holomorphic on $\rho_A(T)$, with values in $\mathcal{B}(X)$, having an analytic extension to $\rho(T)$, whenever $\infty \in \rho(T)$. In particular, the spectrum $\sigma(T)$ is a closed subset of $\mathbb{C}_\infty$, which is nonempty provided $X_0 \neq X$ or $X = X_0 \neq \{0\}$.*

Proof. We may assume $\rho(T) \neq \varnothing$. The proof is similar to the corresponding one for linear relations; see for instance [Cross 1998, Section VI.1]. Because of some differences, we shall sketch an appropriate proof.

Let $\lambda_0 \in \rho(T)$. We show that there exists a neighborhood $V \subset \mathbb{C}_\infty$ of λ_0 such that $V \subset \rho(T)$. We have the following situations.

First, if $\lambda_0 = \infty$, it follows from Definition 2 that there exists $r > 0$ such that $\{|\lambda| > r\} \subset \rho(T)$.

Second, assume $\lambda_0 \in \rho_A(T)$ and that $R(\lambda_0, T) \neq 0$. Then, if $|\lambda - \lambda_0| < \|R(\lambda_0, T)\|^{-1}$, then $\lambda \in \rho(T)$ and

$$R(\lambda, T) = R(\lambda_0, T)(I + (\lambda - \lambda_0) R(\lambda_0, T))^{-1},$$

implying, in particular, the holomorphy of $R(\lambda, T)$ in this open disc.

Third, next assume $R(\lambda_0, T) = 0$. Then $J_0 = 0$, and so $X = X_0$. Moreover, $R(\lambda, T) = 0$ for all $\lambda \in \mathbb{C}$.

If $X = X_0 = \{0\}$, then $\rho_A(T) = \mathbb{C}$, $\rho(T) = \mathbb{C}_\infty$ by Definition 2.

If $X = X_0 \neq \{0\}$, then $D(T) = \{0\}$ (otherwise $\rho(T) = \varnothing$) and $\rho_A(T) = \rho(T) = \mathbb{C}$, again by Definition 2.

Note that the assumption $X = X_0 \neq \{0\}$ implies $\sigma(T) \ni \{\infty\}$. Finally, suppose that $X_0 \neq X$ and $\sigma(T) = \varnothing$. Then $R(\lambda, T)$ is analytic in $\mathbb{C}$ and has an analytic extension at ∞. By Liouville's theorem, it follows that $R(\lambda, T)$ is a constant operator, say C_0. Since ∞ is a 0-regular point of T, we must have $C_0 = 0$. Therefore, as above, $X = X_0$, which is not possible. □

Remark 8. If $\sigma_A(T)$ is bounded, according to Proposition 7 we have a development in $\mathcal{B}(X)$ of the form

$$R(\lambda, T) = \sum_{k=-\infty}^{\infty} \lambda^k C_k,$$

where the series is uniformly convergent when $r_1 \le |\lambda| \le r_2$ for fixed $r_2 \ge r_1 > \sup_{\lambda \in \sigma_A(T)} |\lambda|$. This representation shows that ∞ is m-regular for some integer $m \ge 0$ if and only if $C_k = 0$ for all $k \ge m$. In particular, if $m \ge 2$, the point ∞ is m-regular for T if and only if ∞ is a pole of $R(\lambda, T)$ of order $\le m-1$. As already noted, ∞ is a 0-regular point of T if and only if $\lim_{\lambda\to\infty} R(\lambda, T) = 0$, while ∞ is a 1-regular point if and only if $\lim_{\lambda\to\infty} R(\lambda, T)$ exists in $\mathcal{B}(X)$.

Henceforth, to avoid quotient range operators with empty spectrum, we assume that either $X_0 \ne X$ or $X = X_0 \ne \{0\}$, if not otherwise specified.

Definition 9. Let $T : D(T) \subset X \mapsto X/X_0$ be closed, with $\varnothing \ne \rho(T)$.

(i) We denote by $\mathcal{O}(T)$ the set of all complex-valued functions f, each of them defined and analytic in an open set containing $\sigma(T)$ and depending on f. By identifying any two functions equal in a neighborhood of $\sigma(T)$ (that is, considering $\mathcal{O}(T)$ as the set of germs of analytic functions in neighborhoods of $\sigma(T)$), we may and will regard $\mathcal{O}(T)$ as an algebra.

(ii) Let $F \subset \mathbb{C}_\infty$ be closed and let U be an open neighborhood of F. An *admissible contour* surrounding F in U is a finite system of rectifiable Jordan curves Γ, positively oriented, which is the boundary of an open set $\Delta \subset \overline{\Delta} \subset U$, with $\Delta \supset F$. Note that $\Gamma \cap F = \varnothing$ and that Γ is a compact set in $\mathbb{C}$.

(iii) We define the *analytic functional calculus* for the quotient range operator T as follows. Let $f \in \mathcal{O}(T)$. We set

$$f(T) := \begin{cases} (2\pi i)^{-1} \int_\Gamma f(\lambda) R(\lambda, T)\, d\lambda & \text{if } \infty \notin \sigma(T), \\ f(\infty) I + (2\pi i)^{-1} \int_\Gamma f(\lambda) R(\lambda, T)\, d\lambda & \text{if } \infty \in \sigma(T), \end{cases}$$

where Γ is an admissible contour surrounding $\sigma(T)$ in the domain of definition of f.

Remark. Via Proposition 7, $f(T)$ is a continuous linear operator on X that does not depend on Γ.

The next result seems to be new even in the context of linear relations.

Proposition 10. *For every quotient range closed operator T with $\emptyset \neq \rho(T)$, the map $f \mapsto f(T)$ of $\mathcal{O}(T)$ into $\mathcal{B}(X)$ is an algebra morphism. If $\sigma(T) \ni \infty$, this morphism is unital.*

Proof. Clearly the map $f \mapsto f(T)$ is linear. To prove the multiplicativity of the application $f \mapsto f(T)$, we follow the lines of [Vasilescu 1982, Proposition III.3.4], via Lemma 6.

Consider first the case $\infty \in \sigma(T)$.

Let $f, g \in \mathcal{O}(T)$ and let $U \subset \mathbb{C}_\infty$ be open in the domain of definition of both f, g, with $\sigma(T) \subset U$. Let Δ and Δ_1 be open sets such that their boundaries Γ and Γ_1, respectively, are admissible contours surrounding $\sigma(T)$ in U, and such that $\sigma(T) \subset \Delta \subset \overline{\Delta} \subset \Delta_1 \subset \overline{\Delta}_1 \subset U$. Then we have

$$
\begin{aligned}
&f(T)g(T)\\
&= f(\infty)g(\infty)I + f(\infty)\frac{1}{2\pi i}\int_{\Gamma_1} g(\mu)R(\mu,T)d\mu + g(\infty)\frac{1}{2\pi i}\int_{\Gamma} f(\lambda)R(\lambda,T)d\lambda\\
&\qquad\qquad + \frac{1}{2\pi i}\int_{\Gamma} f(\lambda)R(\lambda,T)d\lambda \frac{1}{2\pi i}\int_{\Gamma_1} g(\mu)R(\mu,T)d\mu\\
&= f(\infty)g(\infty)I + f(\infty)\frac{1}{2\pi i}\int_{\Gamma_1} g(\mu)R(\mu,T)d\mu + g(\infty)\frac{1}{2\pi i}\int_{\Gamma} f(\lambda)R(\lambda,T)d\lambda\\
&\qquad\qquad + \frac{1}{2\pi i}\int_{\Gamma} f(\lambda)\Big(\frac{1}{2\pi i}\int_{\Gamma_1} (\mu-\lambda)^{-1}g(\mu)(R(\lambda,T) - R(\mu,T))d\mu\Big)\\
&= f(\infty)g(\infty)I + \frac{1}{2\pi i}\int_{\Gamma} f(\lambda)g(\lambda)R(\lambda,T)d\lambda = (fg)(T),
\end{aligned}
$$

via Lemma 6 and the Cauchy formula at infinity for analytic functions.

If $\infty \notin \sigma(T)$, the proof is similar and will be omitted.

If $\sigma(T) \ni \infty$, by letting p_0 be the constant polynomial equal to 1, we may take as Γ the boundary of a closed disc in $\rho(T)$ (negatively oriented). Since $R(\lambda, T)$ is analytic in $\rho(T)$, it follows that $\int_\Gamma R(\lambda, T)d\lambda = 0$, so $p_0(T) = I$. □

The next result corresponds to [Baskakov and Chernyshov 2002, Lemma 2.2], whose proof uses an ergodic theorem from [Hille and Phillips 1957]. We give a direct proof based on Proposition 10.

Proposition 11. *Given a closed operator $T : D(T) \subset X \mapsto X/X_0$, the spectrum $\sigma(T)$ is a bounded subset of $\mathbb{C}$ if and only if $X_0 = 0$ and $T \in \mathcal{B}(X)$.*

Proof. We use some ideas from [Vasilescu 1982, Lemma III.3.5]; see also [Hille and Phillips 1957].

Assume $\sigma(T)$ bounded, and fix an $r > 0$ such that $\sigma(T) \subset \{\lambda \in \mathbb{C}; |\lambda| < r\}$. From the analyticity of the resolvent function (Proposition 7), it follows that there exists a sequence $(C_n)_{n\geq 0} \subset \mathcal{B}(X)$ such that

$$R(\lambda, T) = \sum_{n=0}^{\infty} \lambda^{-n} C_n \quad \text{uniformly with respect to} \quad |\lambda| \geq r.$$

The operator C_0, given by the equality $C_0 = \lim_{\lambda\to\infty} R(\lambda, T)$, is necessarily 0 because ∞ is 0-regular

We define the bounded linear operators

$$E = \frac{1}{2\pi i} \int_{|\lambda|=r} R(\lambda, T)\,d\lambda \quad \text{and} \quad A = \frac{1}{2\pi i} \int_{|\lambda|=r} \lambda R(\lambda, T)\,d\lambda.$$

Because

$$\frac{1}{2\pi i} \int_{|\lambda|=r} \lambda^n d\lambda = \begin{cases} 0 & \text{if } n \neq -1, \\ 1 & \text{if } n = -1, \end{cases}$$

we have

$$\begin{aligned} \frac{1}{2\pi i} \int_{|\lambda|=r} \lambda^k R(\lambda, T)\,d\lambda &= \frac{1}{2\pi i} \int_{|\lambda|=r} \lambda^k \sum_{n=0}^{\infty} \lambda^{-n} C_n d\lambda \\ &= \sum_{n=0}^{\infty} \Big(\frac{1}{2\pi i} \int_{|\lambda|=r} \lambda^{k-n} d\lambda \Big) C_n r = C_{k+1} \end{aligned}$$

for all integers $k \geq 0$. Consequently $C_1 = E$ and $A^n = C_{n+1}$. The same proposition shows that $E^2 = E$ and $A^n E = E A^n = A^n$. On the other hand, if $|\lambda| \geq r$, then

$$\lambda^{-1} I + \lambda^{-2} A + \cdots = (\lambda I - A)^{-1},$$

which implies that

$$R(\lambda, T) = E(\lambda I - A)^{-1}. \tag{2}$$

Let

$$X_1 = (I - E)(X) \quad \text{and} \quad X_2 = E(X).$$

Hence $X = X_1 \oplus X_2$ because E is a projection. Setting $A_2 = A|_{X_2}$ and using the fact that $AE = EA$, we have

$$(\lambda I - A)^{-1}|_{X_2} = (\lambda I_2 - A_2)^{-1},$$

whenever $|\lambda| \geq r$, where I_2 is the identity on X_2. This together with (2) implies

$$\begin{aligned} &R(\lambda, T)(X_1) \subset X_1, \quad R(\lambda, T)|_{X_1} = 0|_{X_1}, \\ &R(\lambda, T)(X_2) \subset X_2, \quad R(\lambda, T)|_{X_2} = (\lambda I_2 - A_2)^{-1}|_{X_2}, \end{aligned} \tag{3}$$

whenever $|\lambda| \geq r$.

Set $0_1 = 0|_{X_1}$. Let $u \in D(T)$ and let $v \in X$ with $J_0 v = Tu$. Then we have $R(\lambda, T)(\lambda u - v) = u$ for a fixed λ with $|\lambda| \geq r$. Write $u = u_1 + u_2$, $v = v_1 + v_2$, with $u_j, v_j \in X_j$ for $j = 1, 2$. Using (3), we have in fact that

$$R(\lambda, T)(\lambda u_1 - v_1) = 0 = u_1 \quad \text{and} \quad R(\lambda, T)(\lambda u_2 - v_2) = u_2.$$

These relations imply that

$$v_1 \in N(0_1) = X_1, \tag{4}$$

$$(\lambda I_2 - A_2)^{-1}(\lambda u_2 - v_2) = u_2, \tag{5}$$

From (5) we obtain that $A_2 u_2 = v_2$. This calculation shows that $D(T) \subset \{0\} \oplus X_2$, and that $T(0 \oplus u_2) = v_1 + A_2 u_2 + X_0$ whenever $0 \oplus u_2 \in D(T)$.

If $u = 0$, then we may take as $v \in X$ with $J_0 v = Tu = 0$ any vector $v \in X_0$. The decomposition $0 = u_1 + u_2$ shows that $u_1 = u_2 = 0$. Then, from (4) and (5) we derive $v_1 \in X_1$ and $v_2 = 0$. Therefore, $X_0 \subset X_1$. As $T(0_1 v_1 \oplus 0) = 0 = v_1 + X_0$ for every $v_1 \in X_1$, we must have $X_0 = X_1$.

In fact, $D(T) = \{0\} \oplus X_2$. Indeed, if $A_2 u_2 = v_2$ for some $u_2 \in X_2$, taking into account (5), we have

$$(\lambda I_2 - A_2)^{-1}(\lambda u_2 - v_2) = u_2 = R(\lambda, T)(\lambda u_2 - v_2) \in D(T).$$

In summary, we have now two closed vector subspaces X_1 and X_2 of X with $X = X_1 \oplus X_2$, the operator $0_1 \in \mathcal{B}(X_1)$, an operator $A_2 \in \mathcal{B}(X_2)$, $X_0 = X_1$, $D(T) = \{0\} \oplus X_2$, and $T : \{0\} \oplus X_2 \mapsto (X_1 \oplus X_2)/X_1$ is given by $T(0 \oplus x_2) = 0 \oplus A_2 x_2 + X_1$ for all $0 \oplus x_2 \in \{0\} \oplus X_2$. Setting $T_1 : \{0\} \subset X_1 \mapsto X_1/X_1 = \{0\}$ and $T_2 = A_2 : X_2 \mapsto X_2$, we obtain $T = T_1 \oplus_q T_2$. Assuming $X_1 \neq \{0\}$, we must have $\sigma(T) \ni \{\infty\}$ via Definition 2, which is not possible. Therefore, which is not possible. Therefore, $X_1 = \{0\}$, and so $T = A_2 \in \mathcal{B}(X_2) = \mathcal{B}(X)$.

Conversely, the conditions in the statement from above are obviously sufficient to insure the boundedness of the spectrum of T. □

Remark 12. From the previous proof it follows that if ∞ is 0-regular for T, then $T = T_1 \oplus_q T_2$, where $T_1 : \{0\} \subset X_1 \mapsto +X_1/X_1 = \{0\}$, and $T_2 : X_2 \mapsto X_2$ is bounded.

Corollary 13. *Let $T : D(T) \subset X \mapsto X/X_0$ be closed. Then $\sigma(T) = \sigma_A(T)$ if and only if $T \in \mathcal{B}(X)$, and $\sigma(T) = \sigma_A(T) \cup \{\infty\}$ otherwise.*

In particular, if $T : D(T) \subset X \mapsto X$ is a closed operator, the spectrum of T is a bounded subset of $\mathbb{C}$ if and only if $T \in \mathcal{B}(X)$.

The next result is related to [Baskakov and Chernyshov 2002, Lemma 2.2].

Corollary 14. *Let $Z \subset X \times X$ be a closed relation. The spectrum of Z is a bounded subset of $\mathbb{C}$ if and only if Z is the graph of an operator in $\mathcal{B}(X)$.*

The spectrum of a direct sum of two quotient range operators behaves as one expects (see also Lemma 2.1 from [Baskakov and Chernyshov 2002], in the context of linear relations):

Corollary 15. *If $T : D(T) \subset X \mapsto X/X_0$ is closed and has the form $T = T_1 \oplus_q T_2$, then $\sigma(T) = \sigma(T_1) \cup \sigma(T_2)$.*

Proof. Note that $J_0 = J_{01} \oplus_q J_{02}$, where $J_{0j} : X_j \mapsto X_j/X_{0j}$ are the canonical projections for $j = 1, 2$. We have to show that $\rho(T) = \rho(T_1) \cap \rho(T_2)$. We have the following cases.

First, fix $\lambda \in \rho(T) \cap \mathbb{C}$. Setting $S = \lambda J_0 - T$, $S_j = \lambda J_{0j} - T_j$, $j = 1, 2$, we have to show that $S = S_1 \oplus_q S_2$ is bijective if and only if both S_1, S_2 are bijective, which is routine and is left to the reader. In fact, we obtain that

$$(\lambda J_0 - T)^{-1} = ((\lambda J_{01} - T_1)^{-1} \oplus (\lambda J_{02} - T_2)^{-1})V^{-1}, \tag{6}$$

where V is given by (1). Therefore,

$$R(\lambda, T) = R(\lambda, T_1) \oplus R(\lambda, T_2). \tag{7}$$

This clearly shows that $\sigma_A(T) = \sigma_A(T_1) \cup \sigma_A(T_2)$.

Second, we have only to note that $T \in \mathscr{B}(X)$ if and only if $T_j \in \mathscr{B}(X_j)$ $(j = 1, 2)$, which easily leads to the equality $\sigma(T) = \sigma(T_1) \cup \sigma(T_2)$, via Corollary 13. □

A general result concerning the existence of an analytic functional calculus for quotient range closed operators is the following.

Theorem 16. *For every quotient range closed operator T with $\varnothing \neq \rho(T)$, the map $f \mapsto f(T)$ of $\mathbb{O}(T)$ into $\mathscr{B}(X)$ is a unital algebra morphism. If $\sigma(T)$ is bounded, then $T \in \mathscr{B}(X)$ and $p_1(T) = T$, where $p_1(\lambda) = \lambda$ for all $\lambda \in \mathbb{C}$.*

Proof. If $\sigma(T)$ is unbounded, the assertion follows from Proposition 10. If $\sigma(T)$ is bounded, then $T \in \mathscr{B}(X)$ by Proposition 11, and the assertion is classical. □

Remark 17. (i) For every $f \in \mathbb{O}(T)$, we have

$$f(T)|_{X_0} = \begin{cases} f(T)|_{\{0\}} = 0 & \text{if } \infty \notin \sigma(T), \\ f(\infty)I & \text{if } \infty \in \sigma(T). \end{cases}$$

Indeed, we clearly have

$$\int_\Gamma f(\lambda)R(\lambda, T)d\lambda = \Big(\int_\Gamma f(\lambda)(\lambda J_0 - T)^{-1}d\lambda\Big)J_0,$$

for each admissible contour Γ surrounding $\sigma(T)$, which in turn implies the desired equalities.

This remark also shows that $f(T)(X_0) \subset X_0$ for every $f \in \mathcal{O}(T)$. This allows us to define an operator $f^\circ(T) \in \mathcal{B}(X/X_0)$, given by

$$f^\circ(T)(x + X_0) := f(T)x + X_0 \quad \text{for } x \in X,$$

for all $f \in \mathcal{O}(T)$. In other words, $f^\circ(T)J_0 = J_0 f(T)$ for all $f \in \mathcal{O}(T)$.

(ii) If ∞ is an isolated point of $\sigma(T)$, then $E = (2\pi i)^{-1}\int_\Gamma f(\lambda)R(\lambda, T)d\lambda$ is a projection, where Γ is a contour surrounding $\sigma_A(T)$.

(iii) If $Z \subset X \times X$ is a closed relation with nonempty resolvent set, we may define the operator $f(Z) := f(Q_Z)$ for every analytic function from $\mathcal{O}(Z) := \mathcal{O}(Q_Z)$ (see Remark 4(ii)). This provides an analytic functional calculus for Z, whose properties are easily derived from those valid for Q_Z (see also [Baskakov and Chernyshov 2002, formula (2.8)] for a similar but partial approach.)

3. Quotient range operators with unbounded spectrum

As before, let X be a complex Banach space, let X_0 be a closed vector subspace of X, and let $J_0 : X \mapsto X/X_0$ be the canonical projection. Let also $T : D(T) \subset X \mapsto X/X_0$ be closed. We may consider on $D(T)$ the *graph norm* given by

$$\|x\|_T := \|x\| + \inf_{J_0 y = Tx} \|y\| \quad \text{for } x \in D(T).$$

It is well known that when endowed with this norm, the vector space $D(T)$ becomes a Banach space; see for instance [Cross 1998, Section IV.3]. With the terminology from [Waelbroeck 1982], $(D(T), \|\cdot\|_T)$ becomes a *Banach subspace* of X, which will be occasionally denoted by D_T.

It is obvious that the maps $T : D_T \mapsto X/X_0$ and $J_T : D_T \mapsto X/X_0$, with $J_T = J_0|_{D_T}$, are continuous.

Throughout this section, $T : D(T) \subset X \mapsto X/X_0$ will be a closed (quotient range) operator, with $\infty \in \sigma(T)$ and a nonempty resolvent set.

Lemma 18. *For every function $f \in \mathcal{O}(T)$ and each admissible contour Γ surrounding $\sigma(T)$, the map*

$$X \ni x \mapsto \int_\Gamma f(\lambda)R(\lambda, T)x d\lambda$$

has values into the Banach space D_T and is continuous.

In particular, if $f(\infty) = 0$, then $f(T)$ is a continuous operator from X into D_T.

Proof. Indeed, $R(\lambda, T) = (\lambda J_0 - T)^{-1} J_0 : X \mapsto D_T$ is in $\mathcal{B}(X, D_T)$ for all $\lambda \in \rho(T)$, and hence

$$2\pi i(f(T) - f(\infty)) = \int_\Gamma f(\lambda)(\lambda J_0 - T)^{-1} J_0 d\lambda \in \mathcal{B}(X, D_T),$$

which implies the assertions. □

We recall that for any quotient range operator T and each function $f \in \mathcal{O}(T)$, we denote by $f^\circ(T)$ the operator induced by $f(T)$ in X/X_0; see Remark 17(i).

Lemma 19. *Let $f \in \mathcal{O}(T)$ be such that $f_1(\lambda) = \lambda f(\lambda) \in \mathcal{O}(T)$. Then $Tf(T) = J_0 f_1(T) = f_1^\circ(T) J_0$.*

Proof. It is clear that $f(\infty) = 0$. Let Γ be an admissible contour surrounding $\sigma(T)$ in the domain of definition of f. We have

$$\begin{aligned} T(f(T)x) &= \frac{1}{2\pi i}\int_\Gamma f(\lambda)T(\lambda J_0 - T)^{-1}J_0 x\,d\lambda \\ &= J_0\Big(-\frac{1}{2\pi i}\int_\Gamma f(\lambda)x\,d\lambda + \frac{1}{2\pi i}\int_\Gamma \lambda f(\lambda)(\lambda J_0 - T)^{-1}J_0 x\,d\lambda\Big) \\ &= J_0(f_1(T)x) = f_1^\circ(T)J_0 x, \end{aligned}$$

because $-(1/2\pi i)\int_\Gamma f(\lambda)d\lambda = f_1(\infty)$. □

Remark 20. With the notation from the previous lemma, if $x \in D(T)$ and $y \in X$ satisfy $J_0 y = Tx$, then $f_1(T)x = f(T)y$. Indeed,

$$\begin{aligned} f_1(T)x &= f_1(\infty)x + \frac{1}{2\pi i}\int_\Gamma \lambda f(\lambda)(\lambda J_0 - T)^{-1}J_0 x\,d\lambda \\ &= \frac{1}{2\pi i}\int_\Gamma f(\lambda)(\lambda J_0 - T)^{-1}J_0 y\,d\lambda = f(T)y, \end{aligned}$$

because $(1/2\pi i)\int_\Gamma f(\lambda)x\,d\lambda = -f_1(\infty)x$, as noticed before.

Lemma 21. *For all $f \in \mathcal{O}(T)$ and $x \in D(T)$, we have $Tf(T)x = f^\circ(T)Tx$.*

Proof. Because the function $\lambda f(\lambda)$ is not necessarily in $\mathcal{O}(T)$, we need an argument different from that in the proof of Lemma 19.

If $(x, y) \in G_0(T)$, then $J_0 y = Tx$. Therefore, for a fixed $\lambda \in \rho(T)$,

$$TR(\lambda, T)x = -J_0 x + \lambda J_0(\lambda J_0 - T)^{-1}J_0 x = J_0 R(\lambda, T)y. \tag{8}$$

Let Γ be an admissible contour surrounding $\sigma(T)$ in the domain of $f \in \mathcal{O}(T)$, positively oriented. We have, via (8), that

$$T\int_\Gamma f(\lambda)R(\lambda, T)x\,d\lambda = J_0\int_\Gamma f(\lambda)R(\lambda, T)y\,d\lambda,$$

implying $Tf(T)x = J_0 f(T)y$. Consequently,

$$Tf(T)x = J_0 f(T)y = f^\circ(T)J_0 y = f^\circ(T)Tx \quad \text{for all } x \in D(T). \qquad \square$$

The next result is a version of the idempotent theorem in the context of quotient range operators. For a similar result in the context of linear relations, see [Baskakov and Chernyshov 2002, Theorem 2.3]. Unlike the result there, our proof uses essentially Theorem 16.

Theorem 22. *Let $T : D(T) \subset X \mapsto X/X_0$ be a quotient range operator with $\sigma(T) \ni \infty$ and assume that there are two nonempty disjoint closed sets $F, H \subset \mathbb{C}_\infty$ such that $\sigma(T) = F \cup H$. Then there exist closed vector subspaces X_F and X_H with $X = X_F \oplus X_H$, and operators $T_F : D(T_F) \subset X_F \mapsto X_F/X_{0F}$ and $T_H : D(T_H) \subset X_H \mapsto X_H/X_{0H}$, where $X_{0F} \subset X_F$, $X_{0H} \subset X_H$ and $X_0 = X_{0F} \oplus X_{0H}$, such that $D(T) = D(T_F) \oplus D(T_H)$ and $T = T_F \oplus_q T_H$.*

In addition, $\sigma(T_F) = F$ and $\sigma(T_H) = H$.

Proof. To fix the ideas, assume that $\infty \in F$. We choose open sets U and V in $\mathbb{C}_\infty$ such that $U \supset F$, $V \supset H$ and $U \cap V = \varnothing$. Then the characteristic functions χ_U and χ_V of the sets U and V respectively, restricted to $U \cup V$, are analytic. We put $P_F = \chi_U(T)$ and $P_H = \chi_V(T)$. Since $\chi_U^2 = \chi_U$, and by a similar relation for χ_V, the operators P_F and P_H are projections via Proposition 10. Moreover, $P_F P_H = P_H P_F = 0$ and $P_F + P_H = I$.

In fact, since $\infty \in F$, we have

$$P_F = I + \frac{1}{2\pi i}\int_{\Gamma_F} R(\lambda, T)\,d\lambda, \quad \text{and} \quad P_H = \frac{1}{2\pi i}\int_{\Gamma_H} R(\lambda, T)\,d\lambda,$$

where Γ_F and Γ_H are admissible contours surrounding F and H in U and V, respectively.

Note that $P_H|_{X_0} = 0$ and $P_F|_{X_0}$ is the identity on X_0; see Remark 17(i).

Lemma 21 shows that if $x \in D(T)$, then $P_F x \in D(T)$, and $T P_F x = P_F^\circ T x$, where $P_F^\circ = \chi_U^\circ(T)$. Similarly, $P_H x \in D(T)$ and $T P_H x = P_H^\circ T x$. This also shows that $D(T) = (D(T) \cap P_F(X)) \oplus (D(T) \cap P_F(H))$.

Let $X_F = P_F(X)$ and $X_H = P_H(X)$. Obviously, $X = X_F \oplus X_H$. We have $X_0 \subset X_F$, and we put $X_{0F} = X_0$ and $X_{0H} = \{0\}$.

Let $T_F = T|_{(D(T)\cap X_F)}$. For each $x \in D(T_F) := D(T) \cap X_F$, Lemma 21 gives $T_F x \in X_F/X_{0F}$. Similarly, if $T_H = T|_{(D(T)\cap X_H)}$ for each $x \in D(T_H) := D(T) \cap X_H$, we have $T_H x \in X_H/X_{0H} = X_H$. Consequently,

$$T(x_F \oplus x_H) = T_F(x_F) \oplus T_H(x_H) \in (X_F/(X_{0F}) \oplus (X_H/(X_{0H})$$

for all $x_F \in D(T_F)$ and $x_H \in D(T_H)$, and so $T = T_F \oplus_q T_H$.

Let us show that $\sigma(T_F) \subset F$.

Let $\mu \in \mathbb{C} \setminus F$. With no loss of generality we may suppose that $\mu \notin U$. Then the function $f_\mu(\lambda) = (\mu - \lambda)^{-1}\chi_U(\lambda)$ is analytic in $U \cup V$, null at infinity, and we may define the operator

$$f_\mu(T) = \frac{1}{2\pi i}\int_{\Gamma_F} f_\mu(\lambda)R(\lambda, T)d\lambda = P_F f_\mu(T).$$

Because we have $(\mu-\lambda)f_\mu(\lambda) = \chi_U(\lambda)$, it follows that $\mu f_\mu(T) - f_{1,\mu}(T) = P_F$, where $f_{1,\mu}(\lambda) = \lambda f_\mu(\lambda) \in \mathcal{O}(T)$.

Let us show that $\mu J_F - T_F$ is injective, where $J_F : X_F \mapsto X_F/X_0$ is $J_0|_{X_F}$. Assuming that for an $x \in D(T_F)$ one has $\mu J_F x = T_F x$, and fixing an $y \in X_F$ with $J_F y = T_F x$, we have $\mu x - y \in X_0$. Because $f_\mu(\infty) = 0$, we infer that

$$\begin{aligned}
0 &= f_\mu(T)(\mu x - y) = \frac{1}{2\pi i}\int_{\Gamma_F} f_\mu(\lambda)(\lambda J_0 - T)^{-1}J_0(\mu x - y)d\lambda \\
&= \frac{1}{2\pi i}\int_{\Gamma_F}((\chi_U(\lambda)R(\lambda, T)x + f_{1,\mu}(\lambda)R(\lambda, T)x - f_\mu(\lambda)(\lambda J_0 - T)^{-1}J_0 y)d\lambda \\
&= P_F x = x,
\end{aligned}$$

where we have used the equality $f_{1,\mu}(T)x = f_\mu(T)y$, via Remark 28.

Let us show that $\mu J_F - T_F$ is surjective. Let $y = P_F y \in X_F$. Note that $y = \mu f_\mu(T)y - f_{1,\mu}(T)y$, as we have seen above. Moreover, by Lemma 19 $J_0 y = \mu J_0 f_\mu(T)y - T f_\mu(T)y$. Therefore, $(\mu J_F - T_F)^{-1}$ exists for all $\mu \notin U$. Since U is an arbitrary open neighborhood of F, it follows that $(\mu J_F - T_F)^{-1}J_F = f_\mu(T)|_{X_F}$ for all $\mu \notin F$.

We show now that $\sigma(T_H) \subset H$. First of all, we identify the space $(X_H + X_0)/X_0$ with X_H, and so $J_0|_{X_H} = I_H$, where I_H is the identity on X_H. Note also that $T_H : D(T_H) \mapsto X_H$ is a simply closed operator.

Fixing $\mu \in \mathbb{C} \setminus H$, we may suppose that $\mu \notin V$. Then the function $g_\mu(\lambda) = (\mu - \lambda)^{-1}\chi_V(\lambda)$ is analytic in $U \cup V$, null at infinity, and we can consider the operator $g_\mu(T) = P_H g_\mu(T)$.

Because we have $(\mu-\lambda)g_\mu(\lambda) = \chi_V(\lambda)$, it follows that $\mu g_\mu(T) - g_{1,\mu}(T) = P_H$, where $g_{1,\mu}(\lambda) = \lambda g_\mu(\lambda) \in \mathcal{O}(T)$.

Proceeding as in the previous case, we derive that $\mu I_H - T_H : D(T_H) \mapsto X_H$ is bijective. In fact, $(\mu I_H - T_H)^{-1} = g_\mu(T)|_{X_H}$ for all $\mu \notin H$. We omit the details.

We have only to note that

$$\|g_\mu(T)|X_H\| \le \frac{1}{2\pi\,\mathrm{dist}(\mu, \Gamma_H)}\int_{\Gamma_H}\|R(\lambda, T)\||d\lambda|,$$

implying that ∞ is 0-regular for T_H. In other words, $\sigma(T_H) \subset H$.

Since we already have $\sigma(T_F) \subset F$ and $\sigma(T_H) \subset H$, it suffices to prove that $\sigma(T_F) \cup \sigma(T_H) = \sigma(T)$. Indeed, this follows from Corollary 15, showing that we must have $\sigma(T_F) = F$ and $\sigma(T_H) = H$. $\square$

A result similar to [Baskakov and Chernyshov 2002, Theorem 2.3] follows directly from the previous theorem:

Corollary 23. *Let $Z \subset X \times X$ be a closed relation with $\sigma(Z) \ni \infty$. Assume that there are two nonempty disjoint closed sets $F, H \subset \mathbb{C}_\infty$ such that $\sigma(Z) = F \cup H$. Then we have a decomposition $Z = Z_F \oplus Z_H$ with Z_F and Z_H closed relations and $\sigma(Z_F) = F$ and $\sigma(Z_H) = H$.*

We end this section with a version of the *spectral mapping theorem*. A similar result valid for linear relations can be found in [Baskakov and Chernyshov 2002, Theorem 2.5], whose proof uses Gelfand's theory (see also Corollary 10 there). Our proof is different and is based on Theorems 16 and 22.

Theorem 24. *For every $f \in \mathbb{O}(T)$, we have $\sigma(f(T)) = f(\sigma(T))$.*

Proof. Fix an $f \in \mathbb{O}(T)$. Let $\mu \notin f(\sigma(T))$ with $\mu \neq \infty$. Then the function $g_\mu(\lambda) = (\mu - f(\lambda))^{-1}$ is in $\mathbb{O}(T)$. It is plain that $(\mu I - f(T))g_\mu(T) = I$, showing that $g_\mu(T) = (\mu I - f(T))^{-1}$, and so $\sigma(f(T)) \subset f(\sigma(T))$ (that it is 0-regular for $f(T)$ is obvious).

Conversely, let $\mu_0 \in f(\sigma(T))$, so $\mu_0 = f(\lambda_0)$ for some $\lambda_0 \in \sigma(T)$. Assume that $\mu_0 \notin \sigma(f(T))$.

In the case $\lambda_0 \neq \infty$, we consider the function $h(\lambda) = (\lambda_0 - \lambda)^{-1}(\mu_0 - f(\lambda))$, which can be clearly extended at $\lambda = \lambda_0$, and this extension belongs to $\mathbb{O}(T)$. Note that $\lambda_0 h(T) - h_1(T) = \mu_0 I - f(T)$, where $h_1(\lambda) = \lambda h(\lambda) \in \mathbb{O}(T)$. Therefore,

$$\lambda_0 h(T)(\mu_0 I - f(T))^{-1} - h_1(T)(\mu_0 I - f(T))^{-1} = I. \tag{9}$$

This shows that for each $v \in X$ we have

$$(\lambda_0 J_0 - T)h(T)(\mu_0 I - f(T))^{-1} v = J_0 v$$

via Lemma 19. Therefore, $\lambda_0 J_0 - T$ is surjective.

Further, let $x \in X$ be such that $(\lambda_0 J_0 - T)x = 0$, and let $y \in X$ with $J_0 y = Tx$. Using (9), we have

$$x = (\mu_0 I - f(T))^{-1}(\lambda_0 h(T) - h_1(T)x) = (\mu_0 I - f(T))^{-1} h(T)(\lambda_0 x - y) = 0$$

via Remark 20, and that $J_0(\lambda_0 x - y) = 0$ and $h(\infty) = 0$; see also Remark 17(i). This shows that $\lambda_0 J_0 - T$ is injective too. Consequently, $\lambda_0 J_0 - T$ is invertible, which is not possible.

In the case that $\lambda_0 = \infty$, and there exists a sequence $(\lambda_m)_{m \geq 1}$ in $\sigma_A(T)$ such that $\lim_{m \to \infty} \lambda_m = \lambda_0$, then $f(\lambda_m) \in \sigma(f(T))$ for all $m \geq 1$ by the first part of the proof, implying $f(\infty) \in \sigma(f(T))$.

Finally, if ∞ is an isolated point of $\sigma(T)$, then, according to Theorem 22, there is a decomposition $X = X_1 \oplus X_\infty$, and setting $T_\infty = T|_{D(T) \cap X_\infty}$, we have $\sigma(T_\infty) = \{\infty\}$. Because we have $\sigma(f(T_\infty)) \subset f(\sigma(T_\infty)) = \{f(\infty)\}$ by the first part of the proof, we must actually have $\sigma(f(T_\infty)) = \{f(\infty)\}$ since $\sigma(f(T_\infty))$ is nonempty. Consequently, $f(\infty) \in \sigma(f(T))$, as a consequence of Corollary 15 and of the equality $f(T) = f(T_1) \oplus_q f(T_\infty)$, where $T_1 = T|_{X_1} \in \mathcal{B}(X_1)$. □

Theorem 2.5 from [Baskakov and Chernyshov 2002] is then a consequence of the preceding theorem:

Corollary 25. *If Z is a closed relation with nonempty resolvent set and unbounded spectrum, we have $\sigma(f(Z)) = f(\sigma(Z))$ for all $f \in \mathbb{O}(Z)$.*

Using Theorem 24, we get the superposition of the analytic functional calculus:

Proposition 26. *Let $f \in \mathbb{O}(T)$ and let $g \in \mathbb{O}(f(T))$. Then we have $g \circ f \in \mathbb{O}(T)$ and $(g \circ f)(T) = g(f(T))$.*

Proof. The property $g \circ f \in \mathbb{O}(T)$ follows easily from Theorem 24. The proof of the equality $(g \circ f)(T) = g(f(T))$ follows the lines of the similar assertion in [Vasilescu 1982, Theorem III.3.10(4)]. Specifically, we may choose an admissible contour Γ surrounding $\sigma(T)$ such that $\Gamma_1 = f(\Gamma)$ surrounds $\sigma(f(T))$. Then

$$\begin{aligned} g(f(T)) &= \frac{1}{2\pi i}\int_{\Gamma_1} g(\mu)R(\mu, f(T))d\mu \\ &= \frac{1}{2\pi i}\int_{\Gamma_1} g(\mu)\Big((\mu - f(\infty))^{-1}I + \frac{1}{2\pi i}\int_{\Gamma}(\mu - f(\lambda)^{-1})R(\lambda, T)d\lambda\Big)d\mu \\ &= g(f(\infty))I + \frac{1}{2\pi i}\int_{\Gamma}\Big(\frac{1}{2\pi i}\int_{\Gamma_1} g(\mu)(\mu - f(\lambda)^{-1})d\mu\Big)R(\lambda, T)d\lambda \\ &= g(f(\infty))I + \frac{1}{2\pi i}\int_{\Gamma} g(f(\lambda))R(\lambda, T)d\lambda = (g \circ f)(T), \end{aligned}$$

which proves the result. □

A result similar to [Baskakov and Chernyshov 2002, Corollary 2.4] can be also obtained with our techniques:

Proposition 27. *We have $\sigma(T) = \{\infty\}$ if and only if there is a quasinilpotent operator $Q \in \mathcal{B}(X)$ such that $T : R(Q) \mapsto X/N(Q)$, $T(Qx) = x + N(Q)$ for $x \in X$.*

Proof. Assume $\sigma(T) = \{\infty\}$. If $h(\lambda) = \lambda^{-1}$ ($\lambda \neq 0$), we have $h \in \mathbb{O}(T)$ and $h(\infty) = 0$. Therefore, by Lemma 19, $h(T)x \in D(T)$ for all $x \in X$, and $Th(T)x = J_0 h_1(T)x = J_0 x$, where $h_1(\lambda) = 1$ for all λ. Hence $h(T) = T^{-1}J_0$, showing that $D(T) = R(h(T))$ and $N(h(T)) = X_0$. We have only to remark that $\sigma(h(T)) = h(\{\infty\}) = \{0\}$, showing that $Q = h(T)$ is quasinilpotent.

Conversely, if there is a quasinilpotent operator $Q \in \mathcal{B}(X)$ such that $T : R(Q) \mapsto X/N(Q)$, $T(Qx) = x + N(Q)$ for $x \in X$, then one has $(\lambda J_0 - T)^{-1}(y + N(Q)) = (\lambda Q - I)^{-1}Qy$ for all $y \in X$ and $\lambda \in \mathbb{C}$. Hence, $\sigma(T) = \{\infty\}$.

Note also that $R(\lambda, T) = (\lambda Q - I)^{-1}Q$, $\lambda \in \mathbb{C}$. □

Remark 28. The spectrum of the relation $Z \subset X \times X$ is equal to $\{\infty\}$ if and only if Z is the reverse of the graph of a quasinilpotent operator $Q \in \mathcal{B}(X)$. This can be deduced either from the previous result or directly, from the fact that $Z^{\dagger}$ is a bounded operator and the equality

$$(\lambda^{-1}I - Z^{\dagger})^{\dagger} = \lambda I + \lambda^2(\lambda I - Z)^{\dagger} \quad \text{for } \lambda \neq 0;$$

see for instance [Sandovici 2006, (2.1.2)] or [Baskakov and Chernyshov 2002, Corollary 2.4].

4. Quotient range operators with bounded Arens spectrum

In this section we study those quotient range operators for which the point ∞ is isolated and m-regular, for some integer $m \geq 1$. We discuss the case $m = 0$ in Remark 12. Similar results for linear relations can be also found in [Baskakov and Chernyshov 2002, Section 3]. We start with a version of Proposition 27.

Proposition 29. *Let $T : D(T) \subset X \mapsto X/X_0$ be closed with $\sigma(T) = \{\infty\}$. The point ∞ is m-regular for T for some integer $m \geq 1$ if and only if there exists $Q \in \mathcal{B}(X)$ such that $Q^{m+1} = 0$, and $T : R(Q) \mapsto X/N(Q)$ is given by $T(Qx) = x + N(Q)$ for all $x \in X$.*

Proof. The condition is sufficient by Proposition 27. Let us prove its necessity.

With the notation from Remark 8, because $\sigma(T) = \{\infty\}$ and so $R(\lambda, T)$ should be of the form $-\sum_{k=0}^{\infty} \lambda^k C_k$ for all $\lambda \in \mathbb{C}$, we must have $C_k = 0$ for all $k \geq m$. Therefore, $R(\lambda, T) = -\sum_{k=0}^{m-1} \lambda^k C_k$. For the rest of the proof, we sketch an algebraic argument.

For any two distinct points λ and μ in $\mathbb{C}$, the resolvent equation shows that

$$(\mu - \lambda) \sum_{k=0}^{m-1} \sum_{p+q=k} \lambda^p \mu^q C_p C_q = -\sum_{k=0}^{m-1} (\lambda^k - \mu^k) C_k.$$

Hence

$$\sum_{p+q=k-1} (\lambda - \mu) \lambda^p \mu^q C_p C_q = (\lambda^k - \mu^k) C_k$$

whenever $1 \leq k \leq m-1$, implying by recurrence $C_0 C_{k-1} = C_k$, and so $C_k = C_0^{k+1}$. Therefore, taking $Q = C_0$, we must have $Q^{m+1} = C_m = 0$.

Finally, since $R(\lambda, T) = Q(\lambda Q - I)^{-1}$, we infer the equality, $T^{-1} J_0 = Q$, showing that $X_0 = N(Q)$, $D(T) = R(Q)$, and $TQx = x + N(Q)$ for all $x \in X$. □

The next result is related to [Baskakov and Chernyshov 2002, Theorem 3.1].

Theorem 30. *Let $T : D(T) \subset X \mapsto X/X_0$ be closed, with $\sigma_A(T)$ bounded and $\infty \in \sigma(T)$. The point ∞ is m-regular for some integer $m \geq 1$ if and only if there* *are closed vector subspaces* X_1 *and* X_2 *of* X *with* $X = X_1 \oplus X_2$, *an operator* $A_1 \in \mathcal{B}(X_1)$ *with* $A_1^{m+1} = 0$, *another operator* $A_2 \in \mathcal{B}(X_2)$, *with* $X_0 = N(A_1) \oplus \{0\}$, $D(T) = R(A_1) \oplus X_2$, *and* $T = T_1 \oplus_q T_2$, *where* $T_1(A_1 x_1) = x_1 + N(A_1)$ *for all* $x_1 \in X_1$, *and* $T_2 = A_2$.

In addition, $\sigma_A(T) = \sigma(A_2)$.

Proof. Assume that T is closed, with $\sigma_A(T)$ bounded, such that the point ∞ is m-regular for some integer $m \geq 1$. Then $\sigma(T) = F \cup \{\infty\}$, where $F := \sigma_A(T)$. Since F is bounded, according to Theorem 22 and Proposition 27, there exist closed vector subspaces X_F and X_∞ with $X = X_F \oplus X_\infty$, and operators $T_F : X_F \mapsto X_F$ and $T_\infty : D(T_\infty) \subset X_\infty \mapsto X_\infty / X_{0\infty}$, with $\sigma(T_F) = F$ and $\sigma(T_\infty) = \{\infty\}$, where $X_{0\infty} = N(Q_\infty) = X_0$, $D(T_\infty) = R(Q_\infty) \oplus X_F$, and $Q_\infty \in \mathcal{B}(X_\infty)$ is quasinilpotent. Moreover, $T_\infty(Q_\infty x) = x + N(Q_\infty)$ for all $x \in X_\infty$, and $T = T_\infty \oplus_q T_F$. In fact, since ∞ is m-regular for T, it is also m-regular for T_∞. Therefore, $Q_\infty^{m+1} = 0$ by Proposition 29. The assertion from the statement is obtained for $A_1 = Q_\infty$ and $A_2 = T_F$.

Conversely, if $T = T_1 \oplus_q T_2$ with the stated properties, then $\sigma(T_1) = \{\infty\}$ and ∞ is m-regular for T_2 by Proposition 29, and so $\sigma_A(T) = \sigma(T_2)$ is bounded and ∞ is m-regular also for T, by (7). □

A part of [Baskakov and Chernyshov 2002, Theorem 3.1] is now obtained as a consequence of the previous theorem.

Corollary 31. *Given a closed linear relation* $Z \subset X \times X$ *with* $\sigma_A(Z)$ *a bounded subset of* $\mathbb{C}$ *and* ∞ *not* 0*-regular, the set* $\{|\lambda|^{1-m} \|(\lambda - Z)^\dagger\|; |\lambda| \geq r\}$ *is bounded for an integer* $m \geq 1$ *and some* $r > \sup\{|\lambda|; \lambda \in \sigma_A(Z)\}$ *if and only if there exist closed linear subspaces* X_1 *and* X_2 *with* $X_1 \oplus X_2 = X$, *and operators* $A_1 \in \mathcal{B}(X_1)$ *with* $A_1^{m+1} = 0$, *and* $A_2 \in \mathcal{B}(X_2)$, *such that*

$$Z = G(A_1)^\dagger \oplus G(A_2).$$

In this case, one has $\sigma_A(Z) = \sigma(A_2)$.

Example 32. Let $P \in \mathcal{B}(X)$ be a proper projection, and let $Z = G(P)^\dagger$. Clearly $Z^\dagger = P$ and thus $0 \in \rho(Z)$, and so $Z^\dagger$ is neither injective nor surjective. In fact, we can now easily compute the spectrum of Z. Setting $X_1 = N(P)$ and $X_2 = R(P)$, we have that $X = X_1 \oplus X_2$. Therefore $Z = G(0_1)^\dagger \oplus G(I_2)$, where 0_1 is the null operator on X_1 and I_2 is the identity on X_2. Using Corollary 31, it follows that $\sigma(Z) = \sigma(\{0\}) \cup \sigma(I_2) = \{\infty\} \cup \{1\}$.

Remark 33. Let Z be a densely defined closed linear relation such that, for some $r > 0$, we have $\{\lambda; |\lambda| > r\} \subset \rho_A(Z)$ and $\mathcal{R} = \{(\lambda I - Z)^\dagger; |\lambda| > r\}$ is a bounded subset of $\mathcal{B}(X)$. Then $\sigma_A(Z)$ is bounded, possibly empty. Let us show that $\sigma_A(Z)$ is nonempty. If ∞ is 0-regular, the assertion follows via Corollary 14 (see also

Remark 12). Assuming that $\sigma_A(Z)$ is empty and ∞ is not 0-regular, Corollary 31 shows that $\sigma(A_2)$ is empty, leading to $X_2=\{0\}$, and $A_1^2=0$. Since $D(Z)=R(A_1)$ is dense, the closure of $R(A_1)$ should be equal to X. Therefore $A_1=0$ implying $R(A_1)=\{0\}$, and so $X=\{0\}$, which is not possible. Consequently, $\sigma_A(Z)$ is nonempty.

One can see that the conditions from above on Z are more general than those from [Cross 1998, Theorem VI.3.3], leading to the same conclusion.

5. Applications to Arens polynomial calculus

Given the linear relations $Z,\ Z_1,\ Z_2$ in $X \times X$, and $\alpha \in \mathbb{C}$, we may consider, as usual (see e.g., [Arens 1961; Cross 1998]), the following linear relations in X. The *composition* of Z_1 and Z_2:

$$Z_1 \circ Z_2 = \{(u, w) \in X \times X;\ (u, v) \in Z_2,\ (v, w) \in Z_1 \text{ for some } v \in X\},$$

which will be also denoted by $Z_1 Z_2$. The *sum* of Z_1 and Z_2:

$$Z_1 + Z_2 = \{(u, v+w);\ u \in D(Z_1) \cap D(Z_2),\ (u, v) \in Z_1,\ (u, w) \in Z_2\}.$$

The *product* of Z by a number $\alpha \in \mathbb{C}$:

$$\alpha Z = \{(u, \alpha v);\ (u, v) \in Z\} = \alpha I \circ Z,$$

where we identify the operator αI with its graph. Note that $Z_1 + Z_2$ is not an algebraic sum and that $0Z$ is the null operator on $D(Z)$.

For a linear relation $Z \subset X \times X$ we write

$$Z^n := \underbrace{Z \circ Z \circ \cdots \circ Z}_{\text{n}} \quad \text{for } n \in \mathbb{N}^*.$$

If $p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n$ for $z \in \mathbb{C}$, following Arens [1961] we define the relation

$$p_A(Z) := \alpha_0 I + \alpha_1 Z + \cdots + \alpha_n Z^n.$$

Remark 34. Let Z, Z_1 and Z_2 be linear relations defined on a linear space X. The following assertions, which are well known, follow by a simple calculation.

(i) For any $\xi, \eta \in \mathbb{C}$, one has that $(\xi I - Z)(\eta I - Z) = (\eta I - Z)(\xi I - Z)$.

(ii) $(Z_1 Z_2)^\dagger = Z_2^\dagger Z_1^\dagger$.

We recall that the symbol $\sigma_A(Z)$ denotes the Arens spectrum of the linear relation Z; see Remark 4(ii). We also define $\rho_A(Z) := \mathbb{C} \setminus \sigma_A(Z)$.

The next proposition enables us to apply the results from the previous sections to linear relations of the form $p_A(Z)$; see also [Kascic 1968, Theorem 3.16].

Proposition 35. *If Z is a closed linear relation on the Banach space X such that $\rho_A(Z) \neq \varnothing$ and p is a polynomial, then $p_A(Z)$ is a closed linear relation on X.*

Proof. Fix a $\lambda \in \rho_A(Z)$, so $(\lambda I - Z)^\dagger \in \mathcal{B}(X)$. Using [Brezis 1983, Theorem III.9], we obtain that $(\lambda I - Z)^\dagger$ is continuous from $(X, \sigma(X, X'))$ to $(X, \sigma(X, X'))$. Therefore we can finish by applying [Kascic 1968, Theorem 3.16]. □

The next results show that the functional calculus introduced in Theorem 16 agrees, in some sense, with the Arens polynomial calculus.

Remark 36. Let Z be a closed linear relation in X such that $\sigma(Z) = \{\infty\}$ and the point ∞ is m-regular for some integer $m \geq 1$. Let us compute $p_A(Z)$, where $p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_n z^n$ for $z \in \mathbb{C}$. According to Corollary 31 (see also Remark 12), there exists $Q \in \mathcal{B}(X)$ such that $Q^{m+1} = 0$ and $Z = G(Q)^\dagger$. Hence $Z^k = G(Q^k)^\dagger$ for all integers $k \geq 0$. In particular, $Z^k = G(0)^\dagger$ if $k \geq m+1$. In other words, $p_A(Z) = p_A(G(Q)^\dagger)$. Therefore, if $n = 0$ we have $p_A(Z) = \alpha_0 G(I)$; if $1 \leq n \leq m$ we have

$$\begin{aligned} p_A(Z) &= \alpha_0 G(I) + \alpha_1 G(Q)^\dagger + \cdots + \alpha_n G(Q^n)^\dagger \\ &= \{(x_0, \alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_n x_n);\, x_0 = Q x_1 = \cdots = Q^n x_n\}; \end{aligned}$$

and if $n \geq m+1$,

$$p_A(Z) = \{(0, \alpha_1 x_1 + \cdots + \alpha_m x_m + y_m);\, Q x_1 = \cdots = Q^n x_n = 0,\, y_m \in X\},$$

Proposition 37. *Let Z be a closed linear relation in X such that the point ∞ is m-regular for Z for some integer $m \geq 1$. Then there exist closed linear subspaces X_1 and X_2 with $X_1 \oplus X_2 = X$, and operators $A_1 \in \mathcal{B}(X_1)$ with $A_1^{m+1} = 0$, and $A_2 \in \mathcal{B}(X_2)$, such that*

$$p_A(Z) = p_A(G(A_1)^\dagger) \oplus G(p_A(A_2)),$$

with $p_A(G(A_1)^\dagger)$ computed as in Remark 36.

Proof. If $Z = Z_1 \oplus Z_2$, then $p_A(Z) = p_A(Z_1) \oplus p_A(Z_2)$. In particular, using $Z_1 = G(A_1)^\dagger$ and $Z_2 = G(A_2)$ obtained by Corollary 31 (see also Remark 12), we deduce the formula from the statement. Clearly, the computation of $p_A(G(A_1)^\dagger)$ is given by Remark 36 for $Q = A_1$. □

Proposition 38. *Let Z be a closed linear relation with $\sigma(Z) \ni \infty$, and let $f \in \mathcal{O}(Z)$. Assume that $f_n(\lambda) = \lambda^n f(\lambda) \in \mathcal{O}(Z)$, where $n \geq 1$ is an integer. Then we have $(f(Z)x, (pf)(Z)x) \in p_A(Z)$ for all polynomials p of degree n and all vectors $x \in X$.*

Proof. Set $f_k(\lambda) = \lambda^k f(\lambda) \in \mathcal{O}(Z)$ for $1 \leq k \leq n$. It follows, as in Lemma 19, that $(f(Z)x, f_1(Z)x) \in Z$. Similarly, $(f_{k-1}(Z)x, f_k(Z)x) \in Z$ for all $k = 2, \ldots, n$. Consequently, $(f(Z)x, f_k(Z)x) \in Z^k$ for all $k = 1, \ldots, n$.

If $p(z)=a_0+a_1z+\cdots+a_nz^n$ ($z\in\mathbb{C}$), then

$$(f(Z)x,(pf)(Z)x)=(f(Z)x,a_0f(Z)x+a_1f_1(Z)x+\cdots+a_nf_n(Z)x\in p_A(Z).$$

□

We have the following *spectral mapping theorem* for polynomials.

Proposition 39. *Let Z be a closed linear relation on the Banach space X such that $\rho(Z)\neq\varnothing$ and let p be a nonconstant polynomial.*

(i) $\sigma_A(p_A(Z))=p(\sigma_A(Z))$.

(ii) *If $\infty\in\sigma(p_A(Z))$, then $\infty\in\sigma(Z)$. Conversely, if $\infty\in\sigma(Z)$ and ∞ is not isolated in $\sigma(Z)$, then $\infty\in\sigma(p_A(Z))$.*

Proof. (i) This part follows with minor changes as [Arens 1961, Theorem 2.5]. For this reason, we omit the details.

(ii) Assume that $\infty\in\sigma(p_A(Z))$. Assuming $\infty\notin\sigma(Z)$, we deduce that $Z=G(T)$, with $T\in\mathcal{B}(X)$, via Corollary 14. In this case, as we have $p_A(Z)=G(p_A(T))$ and $p_A(T)\in\mathcal{B}(X)$, we infer that $\infty\notin\sigma(p_A(Z))$, which is not possible.

Conversely, assume that $\infty\in\sigma(Z)$ and that ∞ is not isolated in $\sigma(Z)$. Then we can find a sequence $(\lambda_n)_n$ in $\sigma(Z)$ such that $\lim_{n\to\infty}\lambda_n=\infty$. Since $\mu_n=p_A(\lambda_n)\in\sigma(p(Z))$ for all n by (i), it follows that $\infty=\lim_{n\to\infty}\mu_n\in\sigma(p_A(Z))$. □

Remark. If $Z=\{0\}\times X$ and $p(z)=\alpha_0$, then $\sigma(Z)=\{\infty\}$, while $\sigma(p_A(Z))=\sigma(\alpha_0I)=\{\alpha_0\}$. In other words, there is a linear relation Z with ∞ isolated in $\sigma(Z)$ such that $\infty\notin\sigma(p_A(Z))$ for some polynomial p_A.

References

[Albrecht and Vasilescu 1986] E. Albrecht and F.-H. Vasilescu, "Stability of the index of a semi-Fredholm complex of Banach spaces", *J. Funct. Anal.* **66**:2 (1986), 141–172. MR 87g:58011 Zbl 0592.47008

[Arens 1961] R. Arens, "Operational calculus of linear relations", *Pacific J. Math.* **11** (1961), 9–23. MR 23 #A517 Zbl 0102.10201

[Baskakov and Chernyshov 2002] A. G. Baskakov and K. I. Chernyshov, "Spectral analysis of linear relations, and degenerate semigroups of operators", *Mat. Sb.* **193**:11 (2002), 3–42. In Russian; translated in *Sb. Math.* **193**:11–12 (2002), 1573–1610. MR 2004k:47001 Zbl 1085.47002

[Baskakov and Zagorskiĭ 2007] A. G. Baskakov and A. S. Zagorskiĭ, "On the spectral theory of linear relations on real Banach spaces", *Mat. Zametki* **81**:1 (2007), 17–31. In Russian; translated in *Math. Notes* **81**:1–2 (2007), 15–27. MR 2009c:47002 Zbl 1180.47005

[Brezis 1983] H. Brezis, *Analyse fonctionnelle: Théorie et applications*, Masson, Paris, 1983. MR 85a:46001 Zbl 0511.46001

[Cross 1998] R. Cross, *Multivalued linear operators*, Monographs and Textbooks in Pure and Applied Mathematics **213**, Marcel Dekker, New York, 1998. MR 99j:47003 Zbl 0911.47002

[Dunford and Schwartz 1958] N. Dunford and J. T. Schwartz, *Linear Operators, I: General Theory*, Pure and Applied Mathematics **7**, Interscience Publishers, New York, 1958. MR 22 #8302

[Favini and Yagi 1993] A. Favini and A. Yagi, "Multivalued linear operators and degenerate evolution equations", *Ann. Mat. Pura Appl.* (4) **163** (1993), 353–384. MR 94f:34116 Zbl 0786.47037

[Favini and Yagi 1999] A. Favini and A. Yagi, *Degenerate differential equations in Banach spaces*, Monographs and Textbooks in Pure and Applied Mathematics **215**, Marcel Dekker, New York, 1999. MR 99i:34079 Zbl 0913.34001

[Gheorghe and Vasilescu 2009] D. Gheorghe and F.-H. Vasilescu, "Quotient morphisms, compositions, and Fredholm index", *Linear Algebra Appl.* **431**:11 (2009), 2049–2061. MR 2010i:47033 Zbl 1178.47007

[Hille and Phillips 1957] E. Hille and R. S. Phillips, *Functional analysis and semi-groups*, American Mathematical Society Colloquium Publications **31**, American Mathematical Society, Providence, R. I., 1957. MR 19,664d Zbl 0392.46001

[Kascic 1968] M. J. Kascic, Jr., "Polynomials in linear relations", *Pacific J. Math.* **24** (1968), 291–295. MR 36 #5720 Zbl 0155.19004

[Sandovici 2006] A. Sandovici, *Contributions to the extension theory in Hilbert spaces*, thesis, University of Groningen, 2006.

[Vasilescu 1982] F.-H. Vasilescu, *Analytic functional calculus and spectral decompositions*, Mathematics and its Applications (East European Series) **1**, D. Reidel Publishing Co., Dordrecht, 1982. MR 85b:47016 Zbl 0495.47013

[Waelbroeck 1954] L. Waelbroeck, "Le calcul symbolique dans les algèbres commutatives", *C. R. Acad. Sci. Paris* **238** (1954), 556–558. MR 17,512g Zbl 0055.10704

[Waelbroeck 1982] L. Waelbroeck, "Quotient Banach spaces", pp. 553–562 in *Spectral theory* (Warsaw, 1977), edited by W. Żelazko, Banach Center Publ. **8**, PWN, Warsaw, 1982. MR 85m:46072a Zbl 0492.46012

Received March 8, 2011.

DANA GHEORGHE
MILITARY TECHNICAL ACADEMY
BLD - GEORGE COSBUC, NO. 81-83
RO-050141 BUCHAREST
ROMANIA

gheorghedana@yahoo.com

FLORIAN-HORIA VASILESCU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LILLE 1
59655 VILLENEUVE D'ASCQ
FRANCE

fhvasil@math.univ-lille1.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

HOMOGENEOUS LINKS AND THE SEIFERT MATRIX

PEDRO M. GONZÁLEZ MANCHÓN

Homogeneous links were introduced by Peter Cromwell, who proved that the projection surface of these links, given by the Seifert algorithm, has minimal genus. Here we provide a different proof, with a geometric rather than combinatorial flavor. To do this, we first show a direct relation between the Seifert matrix and the decomposition into blocks of the Seifert graph. Precisely, we prove that the Seifert matrix can be arranged in a block triangular form, with small boxes in the diagonal corresponding to the blocks of the Seifert graph. Then we prove that the boxes in the diagonal have nonzero determinant, by looking at an explicit matrix of degrees given by the planar structure of the Seifert graph. The paper also contains a complete classification of homogeneous knots of genus one.

1. Introduction

Throughout this paper, we assume that all links and diagrams are oriented. Let F be a spanning surface for an oriented link L, and let $b : F \times [0, 1] \to \mathbb{R}^3$ be a regular neighborhood. Identify F with $F \times \{0\}$. The associated Seifert matrix $M = (a_{ij})_{1\le i,j\le n}$ of order n is defined by the linking numbers $a_{ij} = \mathrm{lk}(a_i, a_j^+)$, where the a_i are simple closed oriented curves in F whose homology classes form a basis $\mathcal{B}$ of $H_1(F)$, and $a_i^+ = b(a_i \times 1)$ is the lifting of a_i out of F, in $F \times \{1\}$. Then

$$n = \mathrm{rk}\, H_1(F) = 2g(F) + \mu - 1 = 1 - \chi(F),$$

where $g(F)$ and $\chi(F)$ are the genus and Euler characteristic of F, and μ is the number of components of the link. Homology with coefficients in $\mathbb{Z}$ is assumed throughout the paper.

Let $\nabla_L(z)$ and $\Delta_L(x)$ be the Conway and Alexander polynomials of L, in the variables z and x respectively, as defined in [Cromwell 2004]. Upon the substitution $z = x^{-1} - x$, we have $\nabla_L(z) = \Delta_L(x) = \det(xM - x^{-1}M^t)$. Therefore the coefficient c of the highest degree term in $\nabla_L(z)$ is $(-1)^n \det M$ and the degree

Partially supported by Spanish Project MTM2010-19355 and FEDER.
MSC2010: primary 57M25, 57M27; secondary 57M15, 05C50.
Keywords: homogeneous link, projection surface, Seifert graph, Seifert matrix, Conway polynomial, knot genus, blocks of a graph.

of $\nabla_L(z)$ is n, whenever $\det M$ does not vanish. In general $\deg \nabla_L(z) \le n$, which provides the famous lower bound on the genus, $\deg \nabla_L(z) - \mu + 1 \le 2g(F)$, and in particular it allows us to deduce that F is a minimal genus spanning surface for L if $\det M \neq 0$.

Now suppose that the spanning surface F has been constructed by applying the Seifert algorithm to a diagram D of the link L. We briefly summarize the main features of this construction: start with a diagram D in the xy-plane. For each Seifert circle α a Seifert disc a is built in the plane $z = k$, if there are exactly k Seifert circles that contain α; we say that the height of a is k and write $h(a) = k$. This collection of discs lives in the upper half-space $\mathbb{R}^3_+$ and they are stacked in such a way that when viewed from above, the boundary of each disc is visible. To complete the projection surface, insert small twisted rectangles (called bands from now on) at the site of each crossing, choosing the half-twist according to the corresponding crossing. Following [Cromwell 2004], we call F a projection surface.

We can now define a graph G contained in F as follows: take a vertex in each Seifert disc of F and, if two discs are joined by a band, join the corresponding vertices by an edge contained in the band. Label the edge with the sign of the associated crossing in the diagram D. This graph, called the Seifert graph of D, is in fact a planar graph. The rank $\operatorname{rk} G$ of G, as defined in graph theory, is one minus the number of vertices plus the number of edges. Since $\chi(F) = s(D) - c(D)$, where $s(D)$ is the number of Seifert circles and $c(D)$ is the number of crossings of D, it follows that $\operatorname{rk} G = \operatorname{rk} H_1(F)$.

In general, we can consider the decomposition $G = B_1 \cup \cdots \cup B_k$ of the graph G into its blocks, which are the maximal connected subgraphs without cut vertices. The part of the projection surface (bands and Seifert discs) that corresponds to a block B_i is a submanifold of F and will be denoted by F_{B_i}, or simply F_i. The graph G is a deformation retract of the surface F, taking F_i onto B_i; in particular $H_1(F) \cong H_1(G)$ taking $H_1(F_i)$ onto $H_1(B_i)$ and $\operatorname{rk} G = \operatorname{rk} H_1(G)$, an equality sometimes taken as a definition. Now, a basis of $H_1(G)$, hence a basis $\mathcal{B}$ of $H_1(F)$, can be obtained by juxtaposing basis $\mathcal{B}_i$ of $H_1(B_i)$, since the cycles in G are precisely the cycles of its blocks [Diestel 2005, Lemma 3.1.1]. In particular, the rank of G is the sum of the ranks of its blocks.

Let M_i, where $i = 1, \dots, k$, be the Seifert matrix defined by any basis $\mathcal{B}_i$ of $H_1(B_i)$ (hence of $H_1(F_i)$). Our main result is this:

Theorem 6. *Let D be a connected diagram of an oriented link L. Let G be the corresponding Seifert graph and $G = B_1 \cup \cdots \cup B_k$ its decomposition into blocks. Then there is an order in the set of blocks of G for which the Seifert matrix for the projection surface is upper block triangular. More precisely, if M_i is the Seifert matrix that corresponds to any basis $\mathcal{B}_i$ of $H_1(B_i)$, $i = 1, \dots, k$, there exists a* *permutation $\sigma \in S_k$ such that the Seifert matrix takes on the form*

$$\begin{array}{c} \\ \mathcal{B}_{\sigma(1)} \\ \mathcal{B}_{\sigma(2)} \\ \vdots \\ \mathcal{B}_{\sigma(k)} \end{array}\begin{array}{c} \begin{array}{cccc} \mathcal{B}^+_{\sigma(1)} & \mathcal{B}^+_{\sigma(2)} & \dots & \mathcal{B}^+_{\sigma(k)} \end{array} \\ \begin{pmatrix} M_{\sigma(1)} & 0 & \dots & 0 \\ * & M_{\sigma(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \dots & * & M_{\sigma(k)} \end{pmatrix} \end{array}.$$

A link is homogeneous if it has a homogeneous diagram, which is a diagram in which all the edges of each block of its Seifert graph have the same sign. Alternating and positive diagrams (links) are homogeneous diagrams (links). The knot 9_{43} is an example of a homogeneous link that is neither positive nor alternating. Homogeneous links were introduced in [Cromwell 1989]. In knot theory the adjective homogeneous was first applied to a certain class of braids in [Stallings 1978]. Certainly, the closure of a homogeneous braid is a homogeneous diagram, although there are homogeneous links that cannot be presented as the closure of a homogeneous braid, just as there are alternating links that cannot be presented as the closure of alternating braids. In Cromwell proved the following basic result on homogeneous links:

Theorem [Cromwell 1989; 2004]. *Let D be a connected homogeneous diagram of an oriented homogeneous link L and let G be the corresponding Seifert graph. Then the highest degree of $\nabla_L(z)$ is the rank of G. Let $G = B_1 \cup \dots \cup B_k$ be the decomposition of G into blocks and M_i, $i = 1, \dots, k$, the corresponding Seifert matrices. Then* $\det M_i \neq 0$ *for $i = 1, \dots, k$, and the leading coefficient of $\nabla_L(z)$ is*

$$\prod_{i=1}^{k} \epsilon_i^{r_i} |\det M_i|$$

where ϵ_i is the sign of the edges in B_i and $r_i = \operatorname{rk} B_i$.

Corollary. *A projection surface constructed from a connected homogeneous diagram of an oriented link is a minimal-genus spanning surface for the link.*

Cromwell's proof is based on a previous construction of a specific resolving tree for calculating the Conway polynomial [Cromwell 2004, Lemma 7.5.1]. This means that no crossing is switched more than once on any path from the root of the tree to one of its leaves. The skein relation is then considered, at both the level of the diagram and the corresponding Seifert graph, having in mind that to obtain terms involving powers of z when resolving the resolution tree, a crossing must be smoothed in the diagram D, or equivalently, an edge must be deleted from the graph G. A direct proof of the corollary has been recent and independently suggested by M. Hirasawa. The proof, outlined in [Abe 2011] (see also [Ozawa 2011]), is strongly based on a difficult result from [Gabai 1983], which states that the sum of Murasugi of minimal genus surfaces is a minimal-genus surface. Hirasawa applies this result to the portions F_i above defined.

In this paper we give a different proof of Cromwell's theorem, based on the close relation between the Seifert matrix and the decomposition into blocks of the Seifert graph stated in Theorem 6. The key point is the understanding of how the parts of the projection surface corresponding to the blocks are geometrically positioned among them. We remark that Theorem 6 can be useful even when the diagram is not homogeneous. A special case, involving fibered knots of genus two formed by plumbing Hopf bands, was already considered in [Melvin and Morton 1986]. We deal with this topic in Section 2.

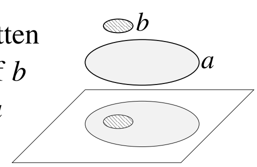
Since a homogeneous block of the Seifert graph corresponds to an alternating diagram, each little box in the diagonal of the Seifert matrix has nonzero determinant, according to the work by K. Murasugi [1958a; 1958b; 1960] and independently Crowell [1959]. Murasugi's proof was accomplished by working on the Alexander matrix of the Dehn presentation of the link, while Crowell worked with the Wirtinger presentation of the fundamental group. In this paper we will prove this result, the second ingredient of our argument, by looking at an explicit matrix of degrees that uses the planar structure of the Seifert graph (Theorem 9). This will be done in Section 3.

Section 4 contains a complete classification of genus-one homogeneous knots.

2. An order for the blocks and the Seifert matrix

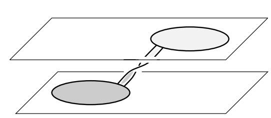
The main achievement of this paper is to prove that there is a certain ordered basis of the first homology group of the projection surface for which the Seifert matrix has a block triangular form. We need first to prove that, in a certain sense, there are only two types of blocks, or more precisely, there are only two possible configurations for the portions F_B associated to a block B.

Let a, b be two Seifert discs. We say that a contains b (written $a \supset b$) if the projection onto the xy-plane of a contains that of b (see figure). Equivalently, the Seifert circle associated to a contains that associated to b, in the xy-plane.



Remark 1. If we project the projection surface onto the xy-plane, the only self-intersections of its boundary are given by the crossings of the original diagram D, and they are produced by the half-twists of the bands.

In particular the arrangement on the right is not possible. As a result we obtain the next lemma.



Lemma 2. *Let a, b be two Seifert discs connected by a band. Then exactly one of the following three statements holds*:

(1) $a \supset b$ *and* $h(b) = h(a) + 1$.

(2) $b \supset a$ *and* $h(a) = h(b) + 1$.

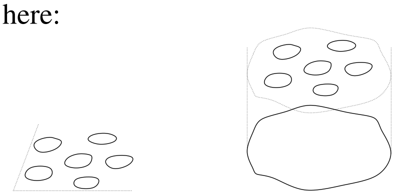
(3) $h(a) = h(b)$.

The proof is easy and left to the reader. Now, we can prove that there are basically two types of blocks. Precisely:

Theorem 3. *Let D be a diagram, F its projection surface and G the corresponding Seifert graph. Then all the Seifert discs associated to a block of G have the same height, except possibly one of them which contains all the other, being its height one less.*

Proof. Suppose that a and b are two Seifert discs with different height connected by a band, both associated to the same block. By Lemma 2 we may assume that one contains the other; say $a \supset b$. It turns out that there is no other Seifert disc associated to the block with height lower than b, since that would make the vertex corresponding to a a cut vertex, according to Remark 1. Analogously, any other disc above b would make (the vertex corresponding to) b a cut vertex. □

Hence we have two possible arrangements for (the Seifert discs that correspond to) a block: type I (*fried eggs* type) and type II (*fried eggs with a pan* type). In a type II block, the pan is the Seifert disc with lowest height. The two types of blocks are illustrated here:



Fried eggs Fried eggs with a pan

Following Cromwell [1989] or Murasugi [1958a; 1958b] we say that a (Seifert) circle is of type I if it does not contain any other circle; otherwise it is of type II. When a type II circle has other circles outside, it is called a decomposing circle. By definition, a special diagram does not contain any decomposing circle. Note that a type II circle is the boundary of the *pan* of a type II block, assuming that the diagram is connected.

Now, recall from the introduction that the part of the projection surface that corresponds to a block B_i is denoted by F_i, which is a submanifold of F. Recall also that, since the cycles of a graph are the cycles of its blocks, we have that a basis of $H_1(F)$ can be obtained by juxtaposing a basis for each block.

Remark 4. Two different F_i's can have at most one common Seifert disc; hence F is the Murasugi sum of the portions F_i's. The proof by Hirasawa mentioned in the introduction follows from this fact.

In order to prove the main theorem, we need the following result of graph theory:

Lemma 5. *Let G be a connected finite graph with at least one cut vertex. Then there is a block of G which has exactly one cut vertex of G.*

Proof. It can be deduced from Proposition 3.1.2 of [Diestel 2005]. It follows a direct argument: delete any cut vertex v_0 of $C_0 = G$ and consider $C_1 = C_1' \cup \{v_0\}$ where C_1' is any connected component of $C_0 - \{v_0\}$. We remark that, under these assumptions, the cut vertices of C_1 are exactly the cut vertices of G that lie in C_1, except for v_0, and that any block of C_1 is a block of G. If C_1 has no cut vertices, then it is the wanted block. Otherwise we select a cut vertex v_1 of C_1 and consider $C_2 = C_2' \cup \{v_1\}$ where C_2' is a connected component of $C_1 - \{v_1\}$ with $v_0 \notin C_2'$. Repeating this process, we finally get a $k \in \mathbb{N}$ such that C_k has no cut vertices, hence being the wanted block. Otherwise we would obtain an infinite sequence of distinct vertices $\{v_0, v_1, v_2, \dots\}$ in the finite graph G, a contradiction. $\square$

Theorem 6. *Let D be a connected diagram of an oriented link L. Let G be the corresponding Seifert graph and $G = B_1 \cup \cdots \cup B_k$ its decomposition into blocks. Then there is an order in the set of blocks of G for which the Seifert matrix for the projection surface is upper block triangular. More precisely, if M_i is the Seifert matrix that corresponds to any basis $\mathcal{B}_i$ of $H_1(B_i)$, $i = 1, \dots, k$, there exists a permutation $\sigma \in S_k$ such that the Seifert matrix takes on the form*

$$\begin{array}{c} \\ \mathcal{B}_{\sigma(1)} \\ \mathcal{B}_{\sigma(2)} \\ \vdots \\ \mathcal{B}_{\sigma(k)} \end{array} \begin{array}{c} \begin{array}{cccc} \mathcal{B}^+_{\sigma(1)} & \mathcal{B}^+_{\sigma(2)} & \cdots & \mathcal{B}^+_{\sigma(k)} \end{array} \\ \begin{pmatrix} M_{\sigma(1)} & 0 & \cdots & 0 \\ * & M_{\sigma(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & M_{\sigma(k)} \end{pmatrix} \end{array}$$

Proof. By Lemma 5, there exists a block B which has exactly one cut vertex. Let D be the Seifert disc associated to the unique cut vertex in B. Translated to the surface, this means that the geometric block F_B is separated from the rest of the surface F, with D as the unique intersection.

We may assume, by induction on the number of blocks, that the Seifert matrix for $F \backslash (F_B \backslash D)$ is upper triangular for a suitable order of the rest of blocks B_i's. Suppose now that the positive orientation of the disc D, that looking at $F \times \{1\}$, is upwards. Then, the basis that corresponds to the block B must be added

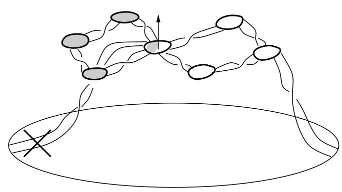
– at the beginning if B is of type I, or D is an egg of the type II block B,

– at the end if D is the pan of the type II block B.

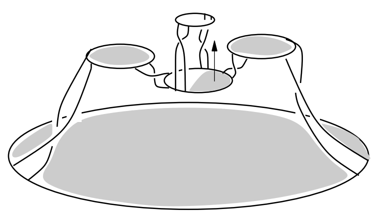
Indeed, if the positive orientation of D were downward, these statements must be interchanged.

In the following displayed figures, the shadowed discs correspond all to the block B; the disc D, partially shadowed, is part of the two considered blocks, B and any other block B_i previously ordered. On D there is an oriented arrow looking upwards, indicating the positive orientation. Suppose now that $g, g_i \in H_1(F)$ correspond to the blocks B and B_i respectively. We have to analyze the three possible cases:

(1) Suppose that B is of type I. We have to see that $\mathrm{lk}(g, g_i^+) = 0$. This can be easily checked if B_i is of type I, or B_i is of type II and the disc D is its pan. And it is also true if B_i is of type II being D an egg of B_i, since in this case the eggs would be on different half parts of the pan. To see this, project both blocks B and B_i onto the plane $z = h(D) - 1$, hence the Seifert discs at height $h(D)$ are now nested inside the pan of the block B_i (*all the eggs in the same pan*). By Remark 1 there is no intersections other than those given by the half-twists of the bands, which means that the two blocks are basically in separated half parts of the pan of B_i. In particular, a band crossed like this is not possible:



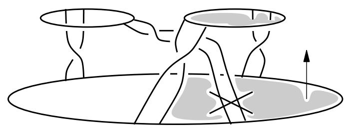
(2) Suppose that B is of type II, and the disc D is an egg of B. As in the previous case, we have to see that $\mathrm{lk}(g, g_i^+) = 0$. This can be easily checked if any other block B_i is of type I, or (see figure below) B_i is of type II and the disc D is the pan of B_i.



Note that D cannot be an egg of another type II block B_i. Indeed, if it were, again by Remark 1, the pan would be the same for B and B_i, hence the blocks B and B_i would share at least two vertices. But, by their maximality, different blocks of G overlap in at most one vertex.

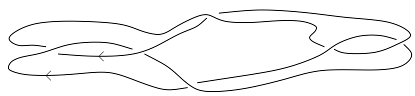
(3) Suppose that B is of type II, and the disc D is its pan. In this case we have to see that $\mathrm{lk}(g_i, g^+) = 0$. This can be easily checked if the block B_i is of type I, or

the disc D is an egg of a type II block B_i. And it is also true if the disc D is the pan of another type II block B_i, since in this case, by a similar argument to that used in the first case, the eggs would be on different half parts of the pan, the crossed band shown here not being possible:

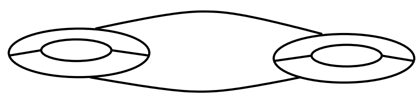


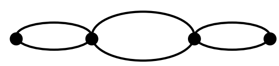
□

Example 7. Suppose that we wish to find the block triangular form for the Seifert matrix of the link shown here:

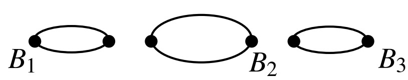


We draw the corresponding Seifert circles and Seifert graph:

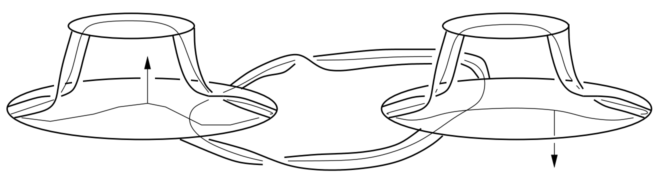




We decompose the Seifert graph into blocks B_1, B_2 and B_3, from left to right:



Here is projection surface:



We can consider $B = B_1$ as the block with only one cut vertex. Then, if the positive orientation of D is upwards, for the other two blocks the suitable basis is given by the order of blocks $\{B_3, B_2\}$, which gives the matrix

$$\begin{array}{ccc} & \mathcal{B}_3^+ & \mathcal{B}_2^+ \\ \mathcal{B}_3 & * & 0 \\ \mathcal{B}_2 & * & * \end{array}$$

Since $B = B_1$ is a block of type II and the disc D that corresponds to the cut vertex is a pan of B, according to the proof of Theorem 6 we must add the basis for B at the end, obtaining the order $\{B_3, B_2, B_1\}$ and the matrix

$$\begin{array}{cccc} & \mathcal{B}_3^+ & \mathcal{B}_2^+ & \mathcal{B}_1^+ \\ \mathcal{B}_3 & * & 0 & 0 \\ \mathcal{B}_2 & * & * & 0 \\ \mathcal{B}_1 & * & * & * \end{array}$$

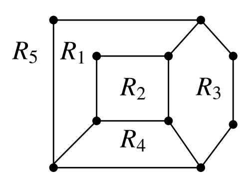
3. The box matrix associated to a block

Recall from the introduction that the coefficient c of the highest degree term in $\nabla_L(z)$ is equal to $(-1)^n \det M$ and the degree of $\nabla_L(z)$ is $n = \operatorname{rk} H_1(F)$, whenever $\det M$ does not vanish. By Theorem 6, $\det M = \prod_{i=1}^{k} \det M_i$ where M_i is the Seifert matrix that corresponds to the surface F_i associated to the block B_i of G. Then, in order to prove the theorem stated in the introduction, it is enough to show that, if B_i is a block with rank r_i and all its edges have sign ϵ_i, then the determinant of its Seifert matrix does not vanish and has sign $(-\epsilon_i)^{r_i}$. Indeed, since $n = \operatorname{rk} G$ is the sum of the ranks r_i of its blocks, we would have

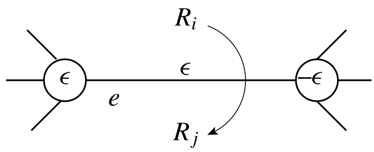
$$\begin{aligned} c = (-1)^n \det M &= (-1)^n \prod_{i=1}^{k} \det M_i \\ &= (-1)^n \prod_{i=1}^{k} (-\epsilon_i)^{r_i} |\det M_i| = \prod_{i=1}^{k} \epsilon_i^{r_i} |\det M_i| \neq 0. \end{aligned}$$

Now, the part of the diagram that corresponds to a homogeneous block is alternating (in fact, it is a special alternating diagram), and the result for these links follows from [Murasugi 1960] and [Crowell 1959]. Murasugi's proof was accomplished by working on the Alexander matrix of the Dehn presentation, while Crowell worked with the Wirtinger presentation of the fundamental group of the link. In fact, Crowell's paper rests on a striking application of a graph theoretical result, the Bott–Mayberry matrix tree theorem, an approach also explained in [Burde and Zieschang 2003, Proposition 13.24]. In this section we will prove it (Theorems 9 and 10) by looking at an explicit matrix of degrees defined using the planar structure of the Seifert graph.

Let D be an oriented diagram, F its projection surface and G the corresponding Seifert graph. Let B be a block of G. A basis $\{g_1, \ldots, g_r\}$ of $H_1(B)$ (hence of $H_1(F_B)$) can be obtained collecting the counterclockwise oriented cycles defined by the boundaries of the bounded regions R_i defined by B. Let R_{r+1} be the unbounded region defined by this planar graph B (like R_5 in the figure on the right).



The Seifert graph is a bipartite graph because the projection surface is orientable, hence every circuit in the graph must have an even length. In particular, we can choose a sign for an arbitrary vertex, and extend this labelling to the other vertices in an alternating fashion, when moving along the edges. We also have, for each edge e in B, its corresponding sign $\epsilon(e)$ (if the original diagram is homogeneous, this sign is constant in the block). We define E_{ij} as the set of edges in $\partial R_i \cap \partial R_j$ with the sign arrangement exemplified by the figure. (The edge e belongs to E_{ij} with this arrangement of signs.)



It turns out that $\mathrm{lk}(g_i, g_i^+) = \frac{1}{2}\sum_{e\in\partial R_i} -\epsilon(e)$ and $\mathrm{lk}(g_i, g_j^+) = \sum_{e\in E_{ij}} \epsilon(e)$.

In particular, if the block is homogeneous, let say with sign ϵ, then

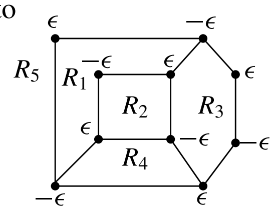
$$\mathrm{lk}(g_i, g_i^+) = -\epsilon k_i,$$

where $2k_i$ is the number of edges in ∂R_i, and

$$\mathrm{lk}(g_i, g_j^+) = \epsilon|E_{ij}|.$$

In other words, $\mathrm{lk}(g_i, g_j^+)$ is the number (with sign ϵ) of the edges e in the frontier of the regions R_i and R_j, such that one leaves the $-\epsilon$ signed vertex on the left when going from R_i to R_j through the edge e (see figure above).

As an example, we display the Seifert matrix associated to the graph of the previous page, assuming that the top left vertex is labelled with sign ϵ; the figure on the right shows the other vertex labels (note that this constitutes a homogeneous block; all the edges have sign ϵ):

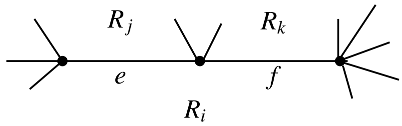


$$\begin{array}{c} \\ g_1 \\ g_2 \\ g_3 \\ g_4 \end{array} \begin{array}{c} \begin{array}{cccc} g_1^+ & g_2^+ & g_3^+ & g_4^+ \end{array} \\ \begin{pmatrix} -3\epsilon & \epsilon & \epsilon & 0 \\ \epsilon & -2\epsilon & 0 & \epsilon \\ 0 & \epsilon & -3\epsilon & 0 \\ \epsilon & 0 & \epsilon & -2\epsilon \end{pmatrix} \end{array}$$

The sets E_{ij} satisfy two properties, which will play later a central role, especially in Theorem 9:

(1) If $e \in \partial R_i \cap \partial R_j$, then $e \in E_{ij} \Longleftrightarrow e \notin E_{ji}$, and in particular $|E_{ij}| + |E_{ji}|$ is the cardinal of the edges in $\partial R_i \cap \partial R_j$.

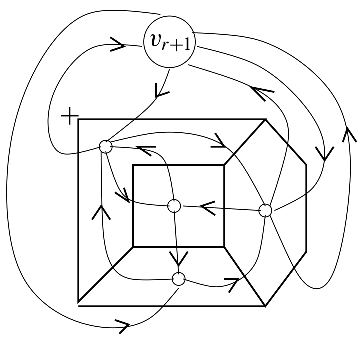
(2) Consider two consecutive edges e and f in the boundary of a certain region R_i, which separate R_i from R_j and R_k respectively (see figure), with possibly $j = k$. Suppose that both edges have the same sign, which is the case if we have a homogeneous graph. Then $e \in E_{ij} \iff f \in E_{ki}$.



Remark 8. For a homogeneous block with sign ϵ, the sum of two transposed elements in the corresponding Seifert matrix gives

$$\mathrm{lk}(g_i, g_j^+) + \mathrm{lk}(g_j, g_i^+) = \epsilon|E_{ij}| + \epsilon|E_{ji}| = \epsilon\ |\partial R_i \cap \partial R_j|.$$

The directed dual graph. A description of the Seifert matrix corresponding to a homogeneous block can be better understood as a certain matrix of degrees for the oriented dual graph. To construct the directed dual graph we draw a vertex v_i in the region R_i, including a vertex v_{r+1} for the unbounded region R_{r+1}, and for each edge e in $\partial R_i \cap \partial R_j$ we draw an edge $\bar{e}$ joining v_i and v_j, the edge $\bar{e}$ intersecting the original graph only in e. Moreover, the edge $\bar{e}$ is oriented from v_i to v_j if (and only if) $e \in E_{ij}$. The directed dual graph in the case of our running example is exhibited in the figure, assuming the sign $\epsilon = +1$ for all the edges and for the top left vertex.



Note that the edges incident at any vertex have alternative orientations, which is equivalent to the second property of the sets E_{ij}'s. In particular the degrees of the vertices are even numbers.

We now define $m_{ii} = -\epsilon\ \deg_i$ and $m_{ij} = \epsilon \deg_{ij}$, where $\deg_i$ is the number of edges leaving (or going to) v_i and $\deg_{ij}$ is the number of edges from v_i to v_j. It turns out that the matrix $(m_{ij})_{1 \le i, j \le r+1}$ has determinant zero, and we obtain the Seifert matrix of the block by just deleting its last row and column. In our running example, for $\epsilon = +1$, we would have

$$\begin{pmatrix} -3 & 1 & 1 & 0 & 1 \\ 1 & -2 & 0 & 1 & 0 \\ 0 & 1 & -3 & 0 & 2 \\ 1 & 0 & 1 & -2 & 0 \\ 1 & 0 & 1 & 1 & -3 \end{pmatrix}.$$

One should note that this is essentially what Proposition 13.21 in [Burde and Zieschang 2003] states, where the adjective special is applied to a diagram if the union of the black regions (assuming a chessboard coloring in which the unbounded region is white) is the image of a Seifert surface under the projection that defines the diagram.

Properties of the matrix for a homogeneous block. Sard matrices. Let ϵ be a sign, $+1$ or -1. A square matrix A is said to be ϵ-signed if its diagonal elements have sign $-\epsilon$ (in particular they do not vanish) and the elements out of the diagonal are zero or have sign ϵ. The matrix A is said to be row-dominant if for any row i we have $|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|$. The matrix A is said to be strictly ascending row-dominant (abbreviated, *sard*) if A is row-dominant and, in addition, there is an order of its rows $i_1 < \cdots < i_r$ such that $|a_{i_r i_r}| > 0$ and for any $k \in \{1, \ldots, r-1\}$ we have that $|a_{i_k i_k}| > \sum_{j\neq i_1, i_2, \ldots, i_k} |a_{i_k j}|$.

The following matrix can be seen to be $(+)$-signed and sard choosing the order 3, 1, 2 for its rows (note that the condition $|a_{i_r i_r}| > 0$ is for sure if A is ϵ-signed):

$$\begin{pmatrix} -3 & 0 & 3 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{pmatrix}$$

Theorem 9. *Let B be a homogeneous block with sign ϵ. Then there exists a basis of $H_1(B)$ such that the associated Seifert matrix M is ϵ-signed and sard.*

Proof. Consider the basis of $H_1(B)$ given by the counterclockwise oriented cycles $\{g_1, \ldots, g_r\}$, boundaries of the bounded regions R_i of B. Then the Seifert matrix $M = (a_{ij})_{1\leq i,j\leq r}$ is obviously ϵ-signed since $a_{ii} = \operatorname{lk}(g_i, g_i^+) = -\epsilon k_i$ where k_i is half the number of edges in the boundary of R_i, and $a_{ij} = \operatorname{lk}(g_i, g_j^+) = \epsilon |E_{ij}|$ if $i \neq j$. To see that A is row-dominant note that $|a_{ii}| = k_i$, and on the other hand $\sum_{j\neq i} |a_{ij}| = \sum_{j\neq i} |E_{ij}| \leq k_i$, the inequality by the second property of the sets E_{ij}.

We finally check that the matrix M is sard, by finding an order $i_1 < \cdots < i_r$ for its rows such that $|a_{i_k i_k}| > \sum_{j\neq i_1, i_2, \ldots, i_k} |a_{i_k j}|$ for any $k \in \{1, \ldots, r-1\}$. By the second property of the sets E_{ij} there is always a bounded region R_i such that $E_{i,r+1} \neq \varnothing$. The corresponding row is chosen to be the first one in this order, that is, $i_1 = i$. Note that, since $|a_{ii}| \geq \sum_{1\leq j\leq r+1, j\neq i} |E_{ij}|$ and $E_{i,r+1} \neq \varnothing$, it follows that $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$. Now, when we delete the i-th row and column, the remaining matrix corresponds to the graph that remains after deleting the region R_i (precisely, deleting the intersection between R_i and R_{r+1}). The region can be also taken in such a way that the remaining graph is still a homogeneous block, hence the repetition of this process provides the wanted order for the rows of M. □

The determinant for a homogeneous block. In this section we will prove that, given a block with sign ϵ and rank r, the determinant of the corresponding submatrix is nonzero, and its sign is equal to $(-\epsilon)^r$. To see this we just need a purely algebraic result due to Murasugi. For the convenience of the reader, we reproduce here its proof in a slightly different way:

Theorem 10 [Murasugi 1960, Section 2]. *Let A be a square matrix of order r, ϵ-signed and sard. Then* $\det A < 0$ *if* $\epsilon = +1$ *and r is odd*, *and* $\det A > 0$ *otherwise. In other words*, $\det A$ *does not vanish and has sign* $(-\epsilon)^r$.

Proof. By induction on r. The case $r = 1$ (odd) is trivial; for $A = (a)$ we have $\det A = a$, and the result follows from the fact that A is ϵ-signed.

Now assume the statement for cases 1 to $r-1$, and consider the case r. Since A is sard, there is an order $i_1 < \cdots < i_r$ of the rows such that for any $k \in \{1, \ldots, r-1\}$ we have $|a_{i_k i_k}| > \sum_{j\neq i_1, i_2, \ldots, i_k} |a_{i_k j}|$. In particular, we have

$$a_{i_1 i_1} = -\sum_{j\neq i_1} a_{i_1 j} - \lambda$$

with $\lambda \neq 0$ and sign ϵ. We now develop the determinant by the i_1-row, obtaining

$$\det A = \det \begin{pmatrix} & \cdots & \\ a_{i_1 1} \ \ldots \ a_{i_1 i_1} \ \ldots \ a_{i_1 r} \\ & \cdots & \end{pmatrix} = x - \lambda y,$$

where

$$x = \det \begin{pmatrix} & \cdots & \\ a_{i_1 1} \ \ldots \ -\sum_{j\neq i_1} a_{i_1 j} \ \ldots \ a_{i_1 r} \\ & \cdots & \end{pmatrix}$$

and y is the determinant of the square matrix of order $r-1$, obtained by deleting the i_1-th row and column. Since this matrix is also ϵ-signed and sard, by induction $y = (-\epsilon)^{r-1}|y| \neq 0$. Moreover, if each $a_{i_1 j} = 0$ for $j \neq i_1$ then $x = 0$ obviously; otherwise it is a square matrix of order r, ϵ-signed and row-dominant, and by Lemma 11, either $x = 0$ or x has sign $(-\epsilon)^r$. Then

$$\det A = x - \lambda y = (-\epsilon)^r|x| - \epsilon|\lambda|(-\epsilon)^{r-1}|y| = (-\epsilon)^r(|x| + |\lambda||y|)$$

and the result follows since $|\lambda| > 0$, $|y| > 0$ and $|x| \geq 0$. □

Lemma 11. *Let A be a square matrix of order r, ϵ-signed and row-dominant. Then* $\det A \leq 0$ *if* $\epsilon = +1$ *and r is odd, and* $\det A \geq 0$ *otherwise.*

Proof. For technical reasons in the induction argument, we will prove this result for a slightly wider category of matrices, the weak ϵ-signed and row-dominant matrices. For this matrices the condition of being ϵ-signed is relaxed for allowing zeros in the diagonal.

We proceed by induction on r. The case $r = 1$ is trivial. Assume now the statement for cases 1 to $r-1$, and consider the case r. Since A is weak ϵ-signed and row-dominant, each diagonal element of A can be written as $a_{ii} = -\sum_{j\neq i} a_{ij} - \lambda_i$ with $\lambda_i = 0$ or with sign ϵ.

Let A_i be the same matrix as A except for possibly the first i elements of its diagonal, where a_{ii} is replaced by $a_{ii} + \lambda_i$. Let $A_0 = A$. It turns out that

$$\det A = \det A_r - \sum_{i=1}^{r} \lambda_i \ \det\big((A_{i-1})_i^i\big),$$

where the notation B_i^i is used to denote the matrix obtained from B by deleting its i-th row and i-th column. This follows from the fact that, for $k = 1, \ldots, r$,

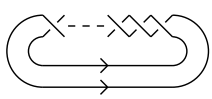
$$\det A_{k-1} = \det A_k - \lambda_k \det\big((A_{k-1})_k^k\big).$$

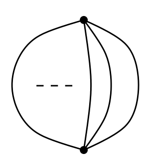
Note that the determinant of A_r is equal to zero, since the sum of all the elements of each row is zero. Moreover, each matrix $(A_{i-1})_i^i$ is also weak ϵ-signed and row-dominant, and has order $r - 1$. By induction, its determinant is zero or has sign $(-\epsilon)^{r-1}$. Since each λ_i is zero or has sign $-\epsilon$, the result follows. □

Here is an application of the argument developed in this section:

Claim. *Let L be an oriented link which has a special alternating diagram. Then the leading coefficient of $\nabla_L(z)$ is ± 1 if and only if L is the connected sum of $(2, q)$-torus links.*

Proof. Assume that L is the connected sum of $(2, q)$-torus links. Since $\nabla_{L\sharp L'}(z) = \nabla_L(z)\nabla_{L'}(z)$, it is enough to show that the leading coefficient of $\nabla_L(z)$ is ± 1 if L is a $(2, q)$-torus link. The diagram of L is then of this form, or its mirror image:



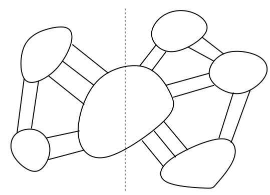


It has q crossings, all with the same sign ϵ. The corresponding Seifert graph, shown on the right, is a homogeneous block B with two vertices and q edges, all of them with sign ϵ.

Following the process explained at the beginning of this section, we obtain the Seifert matrix $M = (m_{i,j})_{i,j=1,\ldots,q-1}$ where $m_{i,i} = -\epsilon$ for $i = 1, \ldots, q - 1$,

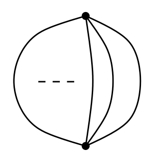
$m_{i,i+1} = \epsilon$ for $i = 1, \ldots, q-2$, and $m_{i,j} = 0$ otherwise. Then the leading coefficient of $\nabla_L(z)$ is $\epsilon^{\mathrm{rk}\, B} |\det M| = \epsilon^{q-1}$ since $\mathrm{rk}\, B = 1 - v + e = 1 - 2 + q = q - 1$.

Suppose now that the leading coefficient of $\nabla_L(z)$ is ± 1, and L has a special alternating diagram D. Then D is the connected sum of diagrams $D_1, \ldots, D_r$, where each D_i is a diagram (of a link L_i) such that its Seifert graph has only one (homogeneous) block:



Clearly, $L = \sharp_{i=1}^r L_i$. Since $\nabla_L(z) = \prod_{i=1}^r \nabla_{L_i}(z)$ and $\nabla_L(z) \in Z[z^{\pm 1}]$, the leading coefficient of each $\nabla_{L_i}(z)$ is ± 1. Hence it is enough to prove that L is a $(2, q)$-torus link assuming that the leading coefficient of $\nabla_L(z)$ is ± 1, and L has a diagram D whose associated Seifert graph is a homogeneous block B, let's say with sign ϵ.

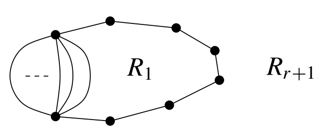
We will prove that B has the desired form (reproduced on the right for convenience). We do this by induction on the number of edges of B. With this aim, we order the r bounded regions of B as in the proof of Theorem 9. The corresponding Seifert matrix A is then ϵ-signed and sard, and by the proof of Theorem 10, we have



$$\det A = (-\epsilon)^r (|x| + |\lambda|\,|y|),$$

where $y = \det A_1^1 \neq 0$. Since the leading coefficient of $\nabla_L(z)$ is ± 1, we have $\det A = \pm 1$; since λ and y are nonzero integers, we have $y = \det A_1^1 = \pm 1$.

Now, according to the proof of Theorem 9, A_1^1 is the Seifert matrix associated to the diagram D' whose Seifert graph is $B' = B \setminus (R_1 \cap R_{r+1})$, where R_{r+1} is the unbounded region of B. Since B' is still a homogeneous block, by induction we have that B' has the form shown above and to the right, and B adds a path connecting the two vertices of B' in the unbounded region of B':



Let $2k$ be the number of edges bounding R_1 in B. Then the original Seifert matrix is

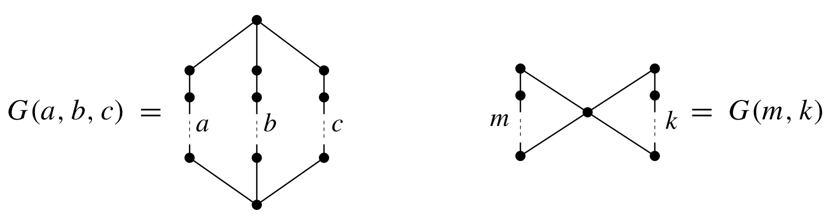
$$A = \left(\begin{array}{c|c} k & \pm 1\ 0\ \cdots\ 0 \\ \hline 0 & \\ \vdots & A_1^1 \\ 0 & \end{array}\right)$$

or its transpose, in any case with determinant $\pm k$. Hence $k = 1$ and the result follows. □

Corollary 12. *Let L be an oriented homogeneous link. Then the leading coefficient of $\nabla_L(z)$ equals ± 1 if and only if L is the Murasugi sum of connected sums of $(2, q)$-torus links.*

4. Homogeneous knots of genus one

We finish the paper with a complete classification of the family of homogeneous knots of genus one. Let D be a homogeneous diagram of a homogeneous knot K of genus one. Let F and G be respectively the projection surface and the Seifert graph associated to the diagram D. We already know that the genus of F is exactly the genus of the knot. Since $2g(F)+\mu-1 = \operatorname{rk} G$ and K is a link with one component, we deduce that G has rank two. Here are the two types of graphs:



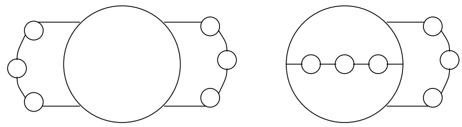
Homogeneous graphs with rank two: one and two blocks

As indicated, we name them $G(a, b, c)$ and $G(m, k)$, respectively; the absolute values of the integers a, b, c, m, k are the numbers of corresponding edges, and their signs are the signs of these edges.

Note that these graphs could have some tails, but this would not affect to the knot type. Since $G(a, b, c)$ is homogeneous and has only one block, a, b, c must have all the same sign; since F is orientable, they have also the same parity. On the contrary, $G(m, k)$ has two blocks, hence m and k can have different signs, but both must be even because of the orientability. Note also that the second graph can be considered a degenerated form of the first one, with $b = 0$.

In general, the Seifert graph does not determine the link where it comes from, although in the first case it does. In $G(a, b, c)$ there are exactly two trivalent vertices; the corresponding Seifert circles can be one inside the other, or separated. When viewed this in the sphere S^2 there is no difference, and the corresponding knot is the pretzel knot with diagram $P(a, b, c)$. Moreover, since $P(a, b, c)$ must be a knot, the numbers a, b, c should be all odd, or exactly one of them should be even. It follows that all of them are odd.

Now consider the graph with two blocks, $G(m, k)$. There is only one vertex with valence four, given the two possible configurations for the Seifert circles shown in the figure (which illustrates the case $|m| = |k| = 4$):



The first configuration corresponds to a link with three components, and the second corresponds to a knot K (see figure below). Moreover, the knot K obtained is also a pretzel knot, given by the pretzel diagram $D(m, k) = P(m, \epsilon, \overset{|k|}{\ldots}, \epsilon)$, where m and k are even integers and ϵ is the sign of k. For example, $D(4, -2) = P(4, -1, -1)$ is the example in the right diagram:

What we have done is to prove the following result:

Theorem 13. *A genus-one knot is homogeneous if and only if it belongs to one of the two following classes of knots*:

(1) *Pretzel knots with diagram $P(a, b, c)$, where a, b, c are odd integers with the same sign.*

(2) *Pretzel knots with diagram $D(m, k) = P(m, \epsilon, \overset{|k|}{\ldots}, \epsilon)$, where m and k are nonzero even integers and $\epsilon = k/|k|$ is the sign of k.*

This classification and some partial information from the Jones polynomial allow us to give another proof of the following result:

Corollary 14 [Cromwell 1989]. *Pretzel knots $P(p, -q, -r)$ with $3 \le p \le q \le r$, all of them odd, are not homogeneous.*

In the original proof, Cromwell calculated the Homfly polynomial $P(v, z) = \sum_{i=0}^{r} \alpha_i(v) z^i$ and checked that $\alpha_r(v)$ contains terms of both signs [Cromwell 1989, Theorem 10]. But, for homogeneous links, these coefficients are all nonnegative or all nonpositive, according to a result [ibid, Corollary 4.3] due to Traczyk.

Proof. We want to prove that the knot K defined by a pretzel diagram $P(p, -q, -r)$ is not homogeneous. First note that K has genus one, since the projection surface defined by the diagram $P(p, -q, -r)$ has Euler characteristic -1, hence genus one, and K is not the trivial knot; for example, according to [Manchón 2003, Theorem 2, case (iv)(a)], the span of its Jones polynomial (with normalization

$-t^{-1/2} - t^{1/2})$ is $p + q + r - \min\{p, q - 1\}$, which is different from one since $3 \le p \le q, r$.

Now, the lowest degree and the coefficient of the highest degree term of the Jones polynomial tell us that K does not belong to any of the two classes of homogeneous knots of genus one given by Theorem 13, as the following table shows:

Knot diagram		Lowest degree	Coefficient of the highest degree term
$P(p, -q, -r)$	$3 \le p < q \le r$	$1/2$	-1
	$3 \le p = q \le r$	$-1/2$	
$P(a, b, c)$	$0 \le a, b, c$	$-3/2 - a - b - c$	
	$a, b, c \le 0$	$1/2$	1
$D(m, k)$	$m, k > 0$	$-m - 1/2$	
	$m < 0, k > 0$	$1/2$	1
	$m > 0, k < 0$	$k - m - 1/2$	
	$m, k < 0$	$k - 1/2$	

□

Note that the Conway polynomial together with the span of the Jones polynomial are not enough in order to prove Corollary 14. According to the values displayed in the following table, we have for example that the knots defined by the diagrams $P(3, -45, -91)$ and $P(11, 23, 101)$ share Conway polynomial and the span of their Jones polynomials, and the same happens to the pair of knots defined by the diagrams $P(11, -15, -15)$ and $D(-4, 26)$.

Knot diagram		Conway polynomial $1 + \lambda z^2$, where λ is	Jones polynomial span
$P(p, -q, -r)$	$3 \le p < q \le r$	$(qr - pq - pr + 1)/4$	$q + r$
	$3 \le p = q \le r$		$q + r + 1$
$P(a, b, c)$	$0 \le a, b, c$	$(ab + ac + bc + 1)/4$	$1 + a + b + c$
	$a, b, c \le 0$		$1 - a - b - c$
$D(m, k)$	$m, k > 0$	$mk/4$	$1 + m + k$
	$m < 0, k > 0$		$k - m$
	$m > 0, k < 0$		$m - k$
	$m, k < 0$		$1 - m - k$

We also have the following result (as above, the Jones polynomial of the pretzel links and their spans have been calculated following [Manchón 2003]):

Corollary 15. *At least one of the extreme coefficients of the Jones polynomial of a homogeneous knot of genus one is* -1.

Finally, we remark that Stoimenow [2008] has showed that a genus-two homogeneous knot is alternating or positive. Jong and Kishimoto [2009] have studied genus-two positive knots extensively.

Acknowledgments

I am grateful to Prof. Hugh R. Morton for several helpful comments on a previous version of this paper. I am also grateful to the referee for a number of helpful suggestions for improving the article. An overview of this work was presented at the Conference *Knots in Poland III*; I am grateful to its organizers for their hospitality.

References

[Abe 2011] T. Abe, "The Rasmussen invariant of a homogeneous knot", *Proc. Amer. Math. Soc.* **139**:7 (2011), 2647–2656. MR 2012c:57010 Zbl 1226.57004

[Burde and Zieschang 2003] G. Burde and H. Zieschang, *Knots*, 2nd ed., De Gruyter Studies in Mathematics **5**, De Gruyter, Berlin, 2003. MR 2003m:57005 Zbl 1009.57003

[Cromwell 1989] P. R. Cromwell, "Homogeneous links", *J. London Math. Soc.* (2) **39**:3 (1989), 535–552. MR 90f:57001 Zbl 0685.57004

[Cromwell 2004] P. R. Cromwell, *Knots and links*, Cambridge University Press, Cambridge, 2004. MR 2005k:57011 Zbl 1066.57007

[Crowell 1959] R. H. Crowell, "Genus of alternating link types", *Ann. of Math.* (2) **69**:2 (1959), 258–275. MR 20 #6103b Zbl 0111.35803

[Diestel 2005] R. Diestel, *Graph theory*, 3rd ed., Grad. Texts in Math. **173**, Springer, Berlin, 2005. MR 2006e:05001 Zbl 1093.05001

[Gabai 1983] D. Gabai, "The Murasugi sum is a natural geometric operation", pp. 131–143 in *Low dimensional topology* (San Francisco, 1981), edited by S. J. Lomonaco, Jr., Contemp. Math. **20**, Amer. Math. Soc., Providence, RI, 1983. MR 85d:57003 Zbl 0524.57004

[Jong and Kishimoto 2009] I. D. Jong and K. Kishimoto, "On positive knots of genus two", preprint, Osaka City University, 2009, available at http://jong.web.fc2.com/posi.pdf.

[Manchón 2003] P. M. G. Manchón, "On the Kauffman bracket of pretzel links", *Marie Curie Fellowships Annals* **2** (2003), 118–122.

[Melvin and Morton 1986] P. M. Melvin and H. R. Morton, "Fibred knots of genus 2 formed by plumbing Hopf bands", *J. London Math. Soc.* (2) **34**:1 (1986), 159–168. MR 87m:57006 Zbl 0576.57008

[Murasugi 1958a] K. Murasugi, "On the genus of the alternating knot, I", *J. Math. Soc. Japan* **10**:1 (1958), 94–105. MR 20 #6103a Zbl 0084.19301

[Murasugi 1958b] K. Murasugi, "On the genus of the alternating knot, II", *J. Math. Soc. Japan* **10**:3 (1958), 235–248. MR 20 #6103a Zbl 0106.16701

[Murasugi 1960] K. Murasugi, “On alternating knots”, *Osaka Math. J.* **12**:2 (1960), 277–303. MR 25 #563 Zbl 0113.38603

[Ozawa 2011] M. Ozawa, “Essential state surfaces for knots and links”, *J. Austral. Math. Soc.* **391** (2011), 365–390.

[Stallings 1978] J. R. Stallings, “Constructions of fibred knots and links”, pp. 55–60 in *Algebraic and geometric topology, II* (Stanford, CA, 1976), edited by R. J. Milgram, Proc. Sympos. Pure Math. **32**, Amer. Math. Soc., Providence, RI, 1978. MR 80e:57004 Zbl 0394.57007

[Stoimenow 2008] A. Stoimenow, “Knots of (canonical) genus two”, *Fund. Math.* **200**:1 (2008), 1–67. MR 2009m:57015 Zbl 1190.57008

Received February 16, 2011. Revised August 29, 2011.

PEDRO M. GONZÁLEZ MANCHÓN
DEPARTMENT OF APPLIED MATHEMATICS
EUITI – UNIVERSIDAD POLITÉCNICA DE MADRID
RONDA DE VALENCIA 3
28012 MADRID
SPAIN

pedro.gmanchon@upm.es
http://gestion.euiti.upm.es/index/departamentos/matematicas/manchon/index.htm

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

QUANTUM AFFINE ALGEBRAS, CANONICAL BASES, AND q-DEFORMATION OF ARITHMETICAL FUNCTIONS

HENRY H. KIM AND KYU-HWAN LEE

We obtain affine analogs of the Gindikin–Karpelevich and Casselman–Shalika formulas as sums over Kashiwara and Lusztig's canonical bases. As suggested by these formulas, we define natural q-deformation of arithmetical functions such as (multi)partition functions and Ramanujan τ-functions, and prove various identities among them. In some examples we recover classical identities by taking limits. Additionally, we consider q-deformation of the Kostant function and study certain q-polynomials whose special values are weight multiplicities.

Introduction

This paper is a continuation of [Kim and Lee 2011]. The classical Gindikin–Karpelevich formula and the Casselman–Shalika formula express certain integrals of spherical functions over maximal unipotent subgroups of p-adic groups as products over all positive roots. In the previous paper, we expressed the products over positive roots as sums over Kashiwara and Lusztig's canonical bases. This idea first appeared in [Bump and Nakasuji 2010]. Let G be a split reductive p-adic group, χ an unramified character of T, the maximal torus, and f^0 the standard spherical vector corresponding to χ. Let z be the element of ${}^L T \subset {}^L G$, the L-group of G, corresponding to χ by the Satake isomorphism. Then

$$\int_{N_-(F)} f^0(n)\,dn = \prod_{\alpha\in\Delta^+} \frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}} = \sum_{b\in\boldsymbol{B}} (1-q^{-1})^{d(\phi_i(b))} z^{\mathrm{wt}(b)}, \tag{0-1}$$

$$\begin{aligned}\int_{N_-(F)} f^0(n)\psi_\lambda(n)\,dn &= \chi(V(\lambda)) \prod_{\alpha\in\Delta^+} (1-q^{-1}z^{\alpha}) \\ &= (-t)^M z^{2\rho}\, \chi(V(\lambda)) \prod_{\alpha\in\Delta^+} (1-t^{-1}z^{-\alpha}) \\ &= (-t)^M z^{\rho} \sum_{b'\otimes b\in\mathfrak{B}_\lambda\otimes\mathfrak{B}_\rho} G_\rho(b;q) z^{\mathrm{wt}(b'\otimes b)},\end{aligned} \tag{0-2}$$

Henry Kim was partially supported by an NSERC grant.
MSC2010: primary 17B37; secondary 05E10.
Keywords: quantum affine algebras, canonical bases, q-deformation of arithmetic functions.

where Δ^+ is the set of positive roots, $\boldsymbol{B}$ is the canonical basis, $\mathfrak{B}_\lambda$ is the crystal basis with highest weight λ, and we set $M = |\Delta^+|$ and $t = q^{-1}$. Notice that in the Casselman–Shalika formula, we used crystal bases because they behave well with respect to the tensor product.

In the affine Kac–Moody groups, A. Braverman, D. Kazhdan, and M. Patnaik [Braverman et al. $\geq$ 2012] calculated the integral (0-1) and obtained a formula of the form

$$\int_{N_-(F)} f^0(n)\,dn = A \prod_{\alpha\in\Delta^+}\left(\frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}}\right)^{\operatorname{mult}\alpha}, \tag{0-3}$$

where A is a certain correction factor. When the underlying finite simple Lie algebra $\mathfrak{g}_{\mathrm{cl}}$ is simply laced of rank n, A is given by

$$\prod_{i=1}^{n}\prod_{j=1}^{\infty}\frac{1-q^{-d_i}z^{j\delta}}{1-q^{-d_i-1}z^{j\delta}},$$

where d_i's are the exponents of $\mathfrak{g}_{\mathrm{cl}}$, and δ is the minimal positive imaginary root.

In this paper, we use the explicit description of the canonical basis introduced by Beck, Chari, Pressley, and Nakajima [Beck et al. 1999; Beck and Nakajima 2004] to write the right-hand side of (0-3) as a sum over the canonical basis. Moreover, we obtain the generalization of (0-2). Namely, we prove the following (Theorem 1-16 and Corollary 2-12, respectively).

$$\prod_{\alpha\in\Delta^+}\left(\frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}}\right)^{\operatorname{mult}\alpha} = \sum_{b\in\boldsymbol{B}}(1-q^{-1})^{d(\phi(b))}z^{\operatorname{wt}(b)}, \tag{0-4}$$

$$\chi(V(\lambda))z^{\rho}\prod_{\alpha\in\Delta^+}(1-q^{-1}z^{-\alpha})^{\operatorname{mult}\alpha} = \sum_{b'\otimes b\in\mathfrak{B}_\lambda\otimes\mathfrak{B}_\rho} G_\rho(b;q)\,z^{\operatorname{wt}(b'\otimes b)}, \tag{0-5}$$

where $\boldsymbol{B}$ is the canonical basis of $\boldsymbol{U}^+$ (the positive part of the quantum affine algebra), and $\mathfrak{B}_\lambda$ is the crystal basis with highest weight λ. Here z is a formal variable. We also write the correction factor A as a sum over a canonical basis in the case when $\mathfrak{g}_{\mathrm{cl}}$ is simply laced.

We first prove (0-4) by induction, and deduce (0-5) from (0-4) and the Weyl–Kac character formula. In the course of the proof, we see that (0-5) can be considered as a q-deformation of the Weyl–Kac character formula. We also introduce $H_{\lambda+\rho}(\mu;q)\in\mathbb{Z}[q^{-1}]$ (Definition 2-2). It has many remarkable properties; its constant term is the multiplicity of the weight $\lambda-\mu$ in $V(\lambda)$, and the value at $q=-1$ is the multiplicity of the weight $\lambda+\rho-\mu$ in the tensor product $V(\lambda)\otimes V(\rho)$. It is also related to Kazhdan–Lusztig polynomials when $\mathfrak{g}$ is of finite type (Corollary 3-30).

When $q=-1$ and λ is a strictly dominant weight, the Casselman–Shalika formula (0-5) gives a formula for multiplicity of the weight ν in the tensor product

$V(\lambda - \rho) \otimes V(\rho)$ in terms of q-deformation of the Kostant partition function, generalizing the result of [Guillemin and Rassart 2004, Theorem 1] to affine Kac–Moody algebras; see (3-24). More precisely, we define $K_q^\infty(\mu)$ in a similar way as in [Guillemin and Rassart 2004], by

$$\sum_{\mu \in Q_+} K_q^\infty(\mu) z^\mu = \prod_{\alpha \in \Delta^+} \left(\frac{1 - q^{-1} z^\alpha}{1 - z^\alpha} \right)^{\operatorname{mult} \alpha}.$$

Note that when $q = \infty$, $K_q^\infty(\mu)$ is the classical Kostant partition function. Then we have

$$\dim(V(\lambda - \rho) \otimes V(\rho))_\nu = \sum_{w \in W} (-1)^{l(w)} K_{-1}^\infty(w\lambda - \nu).$$

Since the set of positive roots is infinite, the left-hand sides of (0-4) and (0-5) become infinite products. This leads to very interesting q-deformation of arithmetical functions such as multipartition functions and Fourier coefficients of modular forms. We indicate one example here.

We define $\epsilon_{q,n}(k)$ as

$$\prod_{k=1}^{\infty} (1 - q^{-1} t^k)^n = \sum_{k=0}^{\infty} \epsilon_{q,n}(k) t^k.$$

Note that $\epsilon_{1,n}(k)$ is a classical arithmetic function related to modular forms. For example, we have $\epsilon_{1,24}(k) = \tau(k+1)$, where $\tau(k)$ is the Ramanujan τ-function. Thus the function $\epsilon_{q,n}(k)$ should be considered as a q-deformation of the function $\epsilon_{1,n}(k)$.

For a multipartition $\boldsymbol{p} = (\rho^{(1)}, \dots, \rho^{(n)}) \in \mathcal{P}(n)$, we define

$$p_{q,n}(k) = \sum_{\substack{\boldsymbol{p} \in \mathcal{P}(n) \\ |\boldsymbol{p}| = k}} (1 - q^{-1})^{d(\boldsymbol{p})}, \qquad k \geq 1,$$

and set $p_{q,n}(0) = 1$. Here $|\boldsymbol{p}|$ is the weight of the multipartition and the number $d(\boldsymbol{p})$ is defined in Section 1. Notice that if $q \to \infty$ and $k > 0$, the function $p_{\infty,n}(k)$ is just the multipartition function with n-components. In particular, $p_{\infty,1}(k) = p(k)$, the usual partition function. Hence we can think of $p_{q,n}(k)$ as a q-deformation of the multipartition function.

It turns out that there are remarkable relations among these q-deformations. We prove (Theorem 3-8)

$$\epsilon_{q,n}(k) = \sum_{r=0}^{k} \epsilon_{1,n}(r) p_{q,n}(k - r),$$

which yields an infinite family of q-polynomial identities. We also obtain “classical” identities by taking limits.

When $n = 24$ and $q \to \infty$, the identity becomes a well-known recurrence formula for the Ramanujan τ-function:

$$0 = \sum_{r=0}^{k} \tau(r+1) p_{\infty,24}(k-r).$$

In fact, we prove another family of identities (Proposition 3-13) and obtain an intriguing characterization of the function $\epsilon_{q,n}(k)$. In Example 3-14, by taking $q = 1$, we write $\tau(k+1)$ as a sum of certain integers arising from the structure of the affine Lie algebra of type $A_4^{(1)}$.

These q-deformations of arithmetic functions essentially come from the observation that the Casselman–Shalika formula may be interpreted as a q-deformation of the Weyl–Kac character formula. In a forthcoming paper, we intend to study q-deformation of other arithmetical functions such as the divisor function, and obtain identities which become classical identities when $q = 1$ or $q \to \infty$.

1. Gindikin–Karpelevich formula

Let $\mathfrak{g}$ be an untwisted affine Kac–Moody algebra over $\mathbb{C}$. We denote by $I = \{0, 1, \dots, n\}$ the set of indices for simple roots. Let W be the Weyl group. We keep almost all the notations from [Beck and Nakajima 2004, Sections 2 and 3]. However, we use v for the parameter of a quantum group and reserve q for another parameter. Whenever there is a discrepancy in notations, we will make it clear.

We fix $\boldsymbol{h} = (\dots, i_{-1}, i_0, i_1, \dots)$ as in [Beck and Nakajima 2004, Section 3.1]. Then for any integers $m < k$, the product $s_{i_m} s_{i_{m+1}} \cdots s_{i_k} \in W$ is a reduced expression, as is the product $s_{i_k} s_{i_{k-1}} \cdots s_{i_m} \in W$. We set

$$\beta_k = \begin{cases} s_{i_0} s_{i_{-1}} \cdots s_{i_{k+1}}(\alpha_{i_k}) & \text{if } k \le 0, \\ s_{i_1} s_{i_2} \cdots s_{i_{k-1}}(\alpha_{i_k}) & \text{if } k > 0, \end{cases}$$

and define

$$\mathcal{R}(k) = \{\beta_0, \beta_{-1}, \dots, \beta_k\} \text{ for } k \le 0 \quad \text{and} \quad \mathcal{R}(k) = \{\beta_1, \beta_2, \dots, \beta_k\} \text{ for } k > 0.$$

Let $T_i = T''_{i,1}$ be the automorphism of $\boldsymbol{U}$ as in [Lusztig 1993, Section 37.1.3], and let

$$\boldsymbol{c}_+ = (c_0, c_{-1}, c_{-2}, \dots) \in \mathbb{N}^{\mathbb{Z}_{\le 0}} \quad \text{and} \quad \boldsymbol{c}_- = (c_1, c_2, \dots) \in \mathbb{N}^{\mathbb{Z}_{>0}}$$

be functions (or sequences) that are zero almost everywhere. We denote by $\mathscr{C}_>$ (respectively $\mathscr{C}_<$) the set of such functions $\boldsymbol{c}_+$ (respectively $\boldsymbol{c}_-$). Then we define

$$E_{\boldsymbol{c}_+} = E_{i_0}^{(c_0)} T_{i_0}^{-1}(E_{i_{-1}}^{(c_{-1})}) T_{i_0}^{-1} T_{i_{-1}}^{-1}(E_{i_{-2}}^{(c_{-2})}) \cdots$$

and

$$E_{\boldsymbol{c}_-} = \cdots T_{i_1} T_{i_2}(E_{i_3}^{(c_3)}) T_{i_1}(E_{i_2}^{(c_2)}) E_{i_1}^{(c_1)}.$$

We set

$$B(k)=\begin{cases}\{E_{\boldsymbol{c}_+} : c_m=0 \text{ for } m<k\} & \text{for } k\le 0,\\ \{E_{\boldsymbol{c}_-} : c_m=0 \text{ for } m>k\} & \text{for } k>0.\end{cases}$$

We denote by $\boldsymbol{B}$ the Kashiwara-Lusztig canonical basis for $\boldsymbol{U}^+$, the positive part of the quantum affine algebra.

Proposition 1-1 [Beck et al. 1999; Beck and Nakajima 2004]. *For each $E_{\boldsymbol{c}_+}\in B(k)$, $k\le 0$ (respectively $E_{\boldsymbol{c}_-}\in B(k)$, $k>0$), there exists a unique $b\in\boldsymbol{B}$ such that*

$$b\equiv E_{\boldsymbol{c}_+} \text{ (respectively } E_{\boldsymbol{c}_-}) \mod v^{-1}\mathbb{Z}[v^{-1}]. \tag{1-2}$$

We denote by $\boldsymbol{B}(k)$ the subset of $\boldsymbol{B}$ corresponding to $B(k)$ as in the above theorem. Then we define the map $\phi:\boldsymbol{B}(k)\to\mathscr{C}_>$ for $k\le 0$ (respectively $\mathscr{C}_<$ for $k>0$) to be $b\mapsto\boldsymbol{c}_+$ (respectively $\boldsymbol{c}_-$) such that the condition (1-2) holds. For an element $\boldsymbol{c}_+=(c_0,c_{-1},\dots)\in\mathscr{C}_>$ (respectively $\boldsymbol{c}_-=(c_1,c_2,\dots)\in\mathscr{C}_>$), we define $d(\boldsymbol{c}_+)$ (respectively $d(\boldsymbol{c}_-)$) to be the number of nonzero c_i's.

Proposition 1-3. *For each $k\in\mathbb{Z}$, we have*

$$\prod_{\alpha\in\mathscr{R}(k)}\frac{1-q^{-1}\boldsymbol{z}^{\alpha}}{1-\boldsymbol{z}^{\alpha}}=\sum_{b\in\boldsymbol{B}(k)}(1-q^{-1})^{d(\phi(b))}\boldsymbol{z}^{\mathrm{wt}(b)}. \tag{1-4}$$

Proof. First we assume $k>0$ and use induction on k. If $k=1$, then the identity (1-4) is easily verified. Now, using an induction argument, we obtain

$$\begin{aligned}
&\prod_{\alpha\in\mathscr{R}(k)}\frac{1-q^{-1}\boldsymbol{z}^{\alpha}}{1-\boldsymbol{z}^{\alpha}}\\
&\quad=\Bigg(\prod_{\alpha\in\mathscr{R}(k-1)}\frac{1-q^{-1}\boldsymbol{z}^{\alpha}}{1-\boldsymbol{z}^{\alpha}}\Bigg)\frac{1-q^{-1}\boldsymbol{z}^{\beta_k}}{1-\boldsymbol{z}^{\beta_k}}\\
&\quad=\Bigg(\sum_{b\in\boldsymbol{B}(k-1)}(1-q^{-1})^{d(\phi(b))}\boldsymbol{z}^{\mathrm{wt}(b)}\Bigg)\Bigg(1+\sum_{j\ge1}(1-q^{-1})\boldsymbol{z}^{j\beta_k}\Bigg)\\
&\quad=\sum_{b\in\boldsymbol{B}(k-1)}(1-q^{-1})^{d(\phi(b))}\boldsymbol{z}^{\mathrm{wt}(b)}+\sum_{j\ge1}\sum_{b\in\boldsymbol{B}(k-1)}(1-q^{-1})^{d(\phi(b))+1}\boldsymbol{z}^{\mathrm{wt}(b)+j\beta_k}.
\end{aligned}$$

On the other hand, since $b'\in\boldsymbol{B}(k)$ satisfies

$$b'\equiv bT_{i_1}T_{i_2}\cdots T_{i_k}(E_k^{(j)}) \mod v^{-1}\mathbb{Z}[v^{-1}]$$

for unique $b\in\boldsymbol{B}(k-1)$ and $j\ge 0$, we can write $\boldsymbol{B}(k)$ as a disjoint union

$$\boldsymbol{B}(k)=\bigcup_{j\ge0}\{b'\in\boldsymbol{B}(k):\phi(b')=(c_1,\dots,c_{k-1},j,0,0,\dots),\ c_i\in\mathbb{N}\}.$$

Now it is clear that

$$\sum_{b\in\boldsymbol{B}(k)} (1-q^{-1})^{d(\phi(b))} z^{\mathrm{wt}(b)} = \sum_{b\in\boldsymbol{B}(k-1)} (1-q^{-1})^{d(\phi(b))} z^{\mathrm{wt}(b)} + \sum_{j\geq 1}\sum_{b\in\boldsymbol{B}(k-1)} (1-q^{-1})^{d(\phi(b))+1} z^{\mathrm{wt}(b)+j\beta_k}.$$

This completes the proof of the case $k>0$. The case $k\leq 0$ can be proved in a similar way through a downward induction. □

We set

$$\mathcal{R}_> = \bigcup_{k\leq 0}\mathcal{R}(k) \quad\text{and}\quad \mathcal{R}_< = \bigcup_{k>0}\mathcal{R}(k).$$

Similarly, we set

$$\boldsymbol{B}_> = \bigcup_{k\leq 0}\boldsymbol{B}(k) \quad\text{and}\quad \boldsymbol{B}_< = \bigcup_{k>0}\boldsymbol{B}(k).$$

Corollary 1-5. *We have*

$$\prod_{\alpha\in\mathcal{R}_>} \frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}} = \sum_{b\in\boldsymbol{B}_>} (1-q^{-1})^{d(\phi(b))} z^{\mathrm{wt}(b)}. \tag{1-6}$$

The same identity is true if $\mathcal{R}_>$ and $\boldsymbol{B}_>$ are replaced with $\mathcal{R}_<$ and $\boldsymbol{B}_<$, respectively.

Let $\boldsymbol{c}_0 = (\rho^{(1)}, \rho^{(2)}, \dots, \rho^{(n)})$ be a multipartition with n components, that is, each component $\rho^{(i)}$ is a partition. We denote by $\mathcal{P}(n)$ the set of all multipartitions with n components. Let $S_{\boldsymbol{c}_0}$ be defined as in [Beck and Nakajima 2004, p. 352] and set

$$B_0 = \{S_{\boldsymbol{c}_0} : \boldsymbol{c}_0 \in \mathcal{P}(n)\}.$$

Proposition 1-7 [Beck et al. 1999; Beck and Nakajima 2004]. *For each $S_{\boldsymbol{c}_0}\in B_0$, there exists a unique $b\in\boldsymbol{B}$ such that*

$$b \equiv S_{\boldsymbol{c}_0} \mod v^{-1}\mathbb{Z}[v^{-1}]. \tag{1-8}$$

We denote by $\boldsymbol{B}_0$ the subset of $\boldsymbol{B}$ corresponding to B_0. Using the same notation ϕ as we used for $\boldsymbol{B}(k)$, we define a function $\phi : \boldsymbol{B}_0 \to \mathcal{P}(n)$, $b\mapsto \boldsymbol{c}_0$, such that the condition (1-8) is satisfied.

For a partition $\boldsymbol{p} = (1^{m_1}2^{m_2}\cdots r^{m_r}\cdots)$, we define

$$d(\boldsymbol{p}) = \#\{r : m_r \neq 0\} \quad\text{and}\quad |\boldsymbol{p}| = m_1 + 2m_2 + 3m_3 + \cdots .$$

Then for a multipartition $\boldsymbol{c}_0 = (\rho^{(1)}, \rho^{(2)}, \dots, \rho^{(n)}) \in \mathcal{P}(n)$, we set

$$d(\boldsymbol{c}_0) = d(\rho^{(1)}) + d(\rho^{(2)}) + \cdots + d(\rho^{(n)}).$$

We obtain from the definition of $S_{\boldsymbol{c}_0}$ that if $\phi(b) = \boldsymbol{c}_0$ then

$$\mathrm{wt}(b) = |\boldsymbol{c}_0|\delta,$$

where $|\boldsymbol{c}_0| = |\rho^{(1)}| + \cdots + |\rho^{(n)}|$ is the weight of the multipartition $\boldsymbol{c}_0$.

Proposition 1-9. *We have*

$$\prod_{\alpha\in\Delta_{\mathrm{im}}^{+}} \left(\frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}}\right)^{\mathrm{mult}\alpha} = \prod_{k=1}^{\infty}\left(\frac{1-q^{-1}z^{k\delta}}{1-z^{k\delta}}\right)^{n} = \sum_{b\in\boldsymbol{B}_0}(1-q^{-1})^{d(\phi(b))}z^{\mathrm{wt}(b)}, \tag{1-10}$$

where Δ_{im}^{+} is the set of positive imaginary roots of $\mathfrak{g}$.

Proof. The first equality follows from the fact that $\Delta_{\mathrm{im}}^{+} = \{\delta, 2\delta, 3\delta, \ldots\}$ and $\mathrm{mult}(k\delta) = n$ for all $k = 1, 2, \ldots$. Now we consider the second equality and assume $n = 1$. Then we have

$$\prod_{k=1}^{\infty}\left(\frac{1-q^{-1}z^{k\delta}}{1-z^{k\delta}}\right) = \prod_{k=1}^{\infty}\left(1+\sum_{j=1}^{\infty}(1-q^{-1})z^{jk\delta}\right). \tag{1-11}$$

We consider the generating function of the partition function $p(m)$:

$$\sum_{m=0}^{\infty} p(m)z^{m\delta} = \prod_{k=1}^{\infty}\left(1+\sum_{j=1}^{\infty} z^{jk\delta}\right) = \sum_{\rho^{(1)}\in\mathcal{P}(1)} z^{|\rho^{(1)}|\delta} = \sum_{b\in\boldsymbol{B}_0} z^{\mathrm{wt}(b)}. \tag{1-12}$$

Comparing (1-11) and (1-12), we see that if we expand the product in the right-hand side of (1-11) into a sum, the coefficient of $z^{|\rho^{(1)}|\delta}$ will be a power of $(1-q^{-1})$ and the exponent of $(1-q^{-1})$ is exactly the number $d(\rho^{(1)})$. Therefore, we obtain

$$\prod_{k=1}^{\infty}\left(\frac{1-q^{-1}z^{k\delta}}{1-z^{k\delta}}\right) = \sum_{\rho^{(1)}\in\mathcal{P}(1)} (1-q^{-1})^{d(\rho^{(1)})}z^{|\rho^{(1)}|\delta} = \sum_{b\in\boldsymbol{B}_0}(1-q^{-1})^{d(b)}z^{\mathrm{wt}(b)}.$$

Next we assume that $n = 2$. Then we have

$$\begin{aligned}
&\prod_{k=1}^{\infty}\left(\frac{1-q^{-1}z^{k\delta}}{1-z^{k\delta}}\right)^{2} \\
&\qquad = \left(\sum_{\rho^{(1)}\in\mathcal{P}(1)} (1-q^{-1})^{d(\rho^{(1)})}z^{|\rho^{(1)}|\delta}\right)\left(\sum_{\rho^{(2)}\in\mathcal{P}(1)} (1-q^{-1})^{d(\rho^{(2)})}z^{|\rho^{(2)}|\delta}\right) \\
&\qquad = \sum_{(\rho^{(1)},\rho^{(2)})\in\mathcal{P}(2)} (1-q^{-1})^{d(\rho^{(1)})+d(\rho^{(2)})}z^{(|\rho^{(1)}|+|\rho^{(2)}|)\delta} \\
&\qquad = \sum_{b\in\boldsymbol{B}_0}(1-q^{-1})^{d(b)}z^{\mathrm{wt}(b)}.
\end{aligned}$$

It is now clear that this argument naturally generalizes to the case $n > 2$. $\square$

Let us consider the correction factor A in (0-3). We will make a modification of the formula (1-10) to write A as a sum over $\boldsymbol{B}_0$ in the case when the underlying classical Lie algebra $\mathfrak{g}_{\mathrm{cl}}$ is simply laced. For a partition $\boldsymbol{p} = (1^{m_1}2^{m_2}\cdots)$ and $d_i \in \mathbb{N}$, we define

$$Q_{d_i}(\boldsymbol{p}, j) = \begin{cases} (1-q)q^{-(d_i+1)m_j} & \text{if } m_j \neq 0, \\ 1 & \text{if } m_j = 0, \end{cases} \quad \text{and} \quad Q_{d_i}(\boldsymbol{p}) = \prod_{j=1}^{\infty} Q_{d_i}(\boldsymbol{p}, j).$$

For a multipartition $\boldsymbol{p} = (\rho^{(1)}, \ldots, \rho^{(n)})$ and $d_i \in \mathbb{N}$, $i = 1, \ldots, n$, we define

$$Q_{d_1,\ldots,d_n}(\boldsymbol{p}) = \prod_{i=1}^{n} Q_{d_i}(\rho^{(i)}).$$

Then we obtain:

Corollary 1-13. *Assume that $\mathfrak{g}_{\mathrm{cl}}$ is simply laced. Then we have*

$$A = \prod_{i=1}^{n}\prod_{j=1}^{\infty} \frac{1-q^{-d_i}z^{j\delta}}{1-q^{-d_i-1}z^{j\delta}} = \sum_{b\in \boldsymbol{B}_0} Q(\phi(b))z^{\mathrm{wt}(b)},$$

where the d_i's are the exponents of $\mathfrak{g}_{\mathrm{cl}}$ and we write $Q(\boldsymbol{p}) = Q_{d_1,\ldots,d_n}(\boldsymbol{p})$.

Proof. The first equality is a result in [Braverman et al. $\geq$ 2012] and the second can be obtained using a similar argument as in the proof of Proposition 1-9. $\square$

Let $\mathscr{C} = \mathscr{C}_{>} \times \mathscr{P}(n) \times \mathscr{C}_{<}$ as in [Beck and Nakajima 2004].

Theorem 1-14 [Beck et al. 1999; Beck and Nakajima 2004]. *There is a bijection between the sets $\boldsymbol{B}$ and $\mathscr{C}$ such that for each $\boldsymbol{c} = (\boldsymbol{c}_+, \boldsymbol{c}_0, \boldsymbol{c}_-) \in \mathscr{C}$, there exists a unique $b \in \boldsymbol{B}$ such that*

$$b \equiv E_{\boldsymbol{c}_+}S_{\boldsymbol{c}_0}E_{\boldsymbol{c}_-} \mod v^{-1}\mathbb{Z}[v^{-1}]. \tag{1-15}$$

Then we naturally extend the function ϕ to a bijection of $\boldsymbol{B}$ onto $\mathscr{C}$ and the number $d(\boldsymbol{c})$ is also defined by $d(\boldsymbol{c}) = d(\boldsymbol{c}_+) + d(\boldsymbol{c}_0) + d(\boldsymbol{c}_-)$ for each $\boldsymbol{c} \in \mathscr{C}$.

Theorem 1-16. *We have*

$$\prod_{\alpha\in\Delta^+}\left(\frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}}\right)^{\mathrm{mult}\,\alpha} = \sum_{b\in\boldsymbol{B}}(1-q^{-1})^{d(\phi(b))}z^{\mathrm{wt}(b)}. \tag{1-17}$$

Proof. Recall that $\Delta^+ = \Delta^+_{\mathrm{re}} \cup \Delta^+_{\mathrm{im}}$, $\Delta^+_{\mathrm{re}} = \mathscr{R}_{>} \cup \mathscr{R}_{<}$, and $\mathrm{mult}\,\alpha = 1$ for $\alpha \in \Delta^+_{\mathrm{re}}$. Then the identity of the theorem follows from Corollary 1-5, Proposition 1-9, and Theorem 1-14. $\square$

2. Casselman–Shalika formula

For the functions $\boldsymbol{c}_+ = (c_0, c_{-1}, c_{-2}, \dots) \in \mathscr{C}_>$ and $\boldsymbol{c}_- = (c_1, c_2, \dots) \in \mathscr{C}_<$, we define

$$|\boldsymbol{c}_+| = c_0 + c_{-1} + c_{-2} + \cdots \quad \text{and} \quad |\boldsymbol{c}_-| = c_1 + c_2 + \cdots .$$

For a multipartition $\boldsymbol{c}_0 = (\rho^{(1)}, \rho^{(2)}, \dots, \rho^{(n)}) \in \mathscr{P}(n)$, set $|\boldsymbol{c}_0| = |\rho^{(1)}| + \cdots + |\rho^{(n)}|$ as in Section 1.

Using similar arguments as in Section 1, we obtain the following identities.

Proposition 2-1. (1) *For each $k \in \mathbb{Z}$,*

$$\prod_{\alpha \in \mathscr{R}(k)} (1 - q^{-1} z^{\alpha})^{-1} = \sum_{b \in \boldsymbol{B}(k)} q^{-|\phi(b)|} z^{\mathrm{wt}(b)}.$$

(2) *The following identity is still true if $\mathscr{R}_>$ and $\boldsymbol{B}_>$ are replaced with $\mathscr{R}_<$ and $\boldsymbol{B}_<$, respectively.*

$$\prod_{\alpha \in \mathscr{R}_>} (1 - q^{-1} z^{\alpha})^{-1} = \sum_{b \in \boldsymbol{B}_>} q^{-|\phi(b)|} z^{\mathrm{wt}(b)}.$$

(3) $$\prod_{\alpha \in \Delta^+_{\mathrm{im}}} \left(1 - q^{-1} z^{\alpha}\right)^{-\mathrm{mult}\,\alpha} = \prod_{k=1}^{\infty} \left(1 - q^{-1} z^{k\delta}\right)^{-n} = \sum_{b \in \boldsymbol{B}_0} q^{-|\phi(b)|} z^{\mathrm{wt}(b)}.$$

(4) $$\prod_{\alpha \in \Delta^+} (1 - q^{-1} z^{\alpha})^{-\mathrm{mult}\,\alpha} = \sum_{b \in \boldsymbol{B}} q^{-|\phi(b)|} z^{\mathrm{wt}(b)}.$$

Let $P_+ = \{\lambda \in P : \langle h_i, \lambda\rangle \geq 0 \text{ for all } i \in I\}$. Recall that the irreducible $\mathfrak{g}$-module $V(\lambda)$ is integrable if and only if $\lambda \in P_+$ [Kac 1990, Lemma 10.1].

Definition 2-2. Let $\lambda \in P_+$. We define $H_\lambda(\,\cdot\,; q) : Q_+ \to \mathbb{Z}[q^{-1}]$ using the generating series

$$\begin{aligned}\sum_{\mu \in Q_+} H_\lambda(\mu; q) z^{\lambda - \mu} &= \sum_{w \in W} (-1)^{\ell(w)} \sum_{b \in \boldsymbol{B}} (1 - q^{-1})^{d(\phi(b))} z^{w\lambda - \mathrm{wt}(b)} \\ &= \Bigl(\sum_{w \in W} (-1)^{\ell(w)} z^{w\lambda}\Bigr)\Bigl(\sum_{b \in \boldsymbol{B}} (1 - q^{-1})^{d(\phi(b))} z^{-\mathrm{wt}(b)}\Bigr),\end{aligned}$$

and we write

$$\chi_q(V(\lambda)) = \sum_{\mu \in Q_+} H_\lambda(\mu; q) z^{\lambda - \mu}.$$

We denote by $\chi(V(\lambda))$ the usual character of $V(\lambda)$. We have the element $d \in \mathfrak{h}$ such that $\alpha_0(d) = 1$ and $\alpha_j(d) = 0$, $j \in I \setminus \{0\}$. We define $\rho \in \mathfrak{h}^*$ as in [Kac 1990, Chapter 6] by $\rho(h_j)=1$, $j\in I$ and $\rho(d)=0$. By the Weyl–Kac character formula,

$$\frac{\displaystyle\sum_{w\in W}(-1)^{\ell(w)}z^{w(\lambda+\rho)-\rho}}{\displaystyle\prod_{\alpha\in\Delta^+}(1-z^{-\alpha})^{\operatorname{mult}\alpha}}=\chi(V(\lambda)).$$

In particular, if $\lambda=0$, then

$$\sum_{w\in W}(-1)^{\ell(w)}z^{w\rho}=z^{\rho}\prod_{\alpha\in\Delta^+}(1-z^{-\alpha})^{\operatorname{mult}\alpha}.$$

By Theorem 1-16,

$$\sum_{b\in\boldsymbol{B}}(1-q^{-1})^{d(\phi(b))}z^{-\mathrm{wt}(b)}=\prod_{\alpha\in\Delta^+}\left(\frac{1-q^{-1}z^{-\alpha}}{1-z^{-\alpha}}\right)^{\operatorname{mult}\alpha}.$$

Thus we obtain

$$\begin{aligned}\chi_q(V(\rho))&=\Big(\sum_{w\in W}(-1)^{\ell(w)}z^{w\rho}\Big)\Big(\sum_{b\in\boldsymbol{B}}(1-q^{-1})^{d(\phi(b))}z^{-\mathrm{wt}(b)}\Big)\\&=z^{\rho}\prod_{\alpha\in\Delta^+}(1-z^{-\alpha})^{\operatorname{mult}\alpha}\prod_{\alpha\in\Delta^+}\left(\frac{1-q^{-1}z^{-\alpha}}{1-z^{-\alpha}}\right)^{\operatorname{mult}\alpha}\\&=z^{\rho}\prod_{\alpha\in\Delta^+}(1-q^{-1}z^{-\alpha})^{\operatorname{mult}\alpha}.\end{aligned}$$

Therefore we have proved that

$$\chi_q(V(\rho))=z^{\rho}\prod_{\alpha\in\Delta^+}(1-q^{-1}z^{-\alpha})^{\operatorname{mult}\alpha}. \tag{2-3}$$

When $q=-1$ in (2-3), we have the following identity from [Kac 1990, Exercise 10.1].

Lemma 2-4. $\chi_{-1}(V(\rho))=z^{\rho}\prod_{\alpha\in\Delta^+}(1+z^{-\alpha})^{\operatorname{mult}\alpha}=\chi(V(\rho)).$

Remark 2-5. By Definition 2-2,

$$\chi_{-1}(V(\rho))=\sum_{\mu\in Q_+}H_\rho(\mu;-1)z^{\rho-\mu}=z^{\rho}\prod_{\alpha\in\Delta^+}(1+z^{-\alpha})^{\operatorname{mult}\alpha}.$$

Therefore, if $H_\rho(\mu;-1)\neq 0$, $\rho-\mu$ must be a weight of $V(\rho)$ and $H_\rho(\mu;-1)$ is the multiplicity of $\rho-\mu$ in $V(\rho)$.

Now we have the following affine analog of the Casselman–Shalika formula.

Corollary 2-6.

$$\chi_q(V(\lambda+\rho))=\chi(V(\lambda))\chi_q(V(\rho)). \tag{2-7}$$

Proof. By Definition 2-2 and Theorem 1-16,

$$\chi_q(V(\lambda+\rho)) = \Big(\sum_{w\in W}(-1)^{\ell(w)} z^{w(\lambda+\rho)}\Big)\prod_{\alpha\in\Delta^+}\Big(\frac{1-q^{-1}z^{-\alpha}}{1-z^{-\alpha}}\Big)^{\operatorname{mult}\alpha}.$$

By the Weyl–Kac character formula and (2-3), the right-hand side is

$$\chi(V(\lambda))\chi_q(V(\rho)).$$

□

Remark 2-8. When $q=1$, we see that $\chi_1(V(\lambda+\rho))z^{-\rho}$ is the numerator of the Weyl–Kac character formula. Hence we can think of (2-7) as a q-deformation of Weyl–Kac character formula. Since $\chi_\infty(V(\rho)) = z^\rho$, by setting $q=\infty$, we have

$$\chi_\infty(V(\lambda+\rho)) = z^\rho\chi(V(\lambda)).$$

Hence we may consider $\chi_q(V(\lambda+\rho))z^{-\rho}$ as a q-deformation of $\chi(V(\lambda))$. Moreover, by Definition 2-2,

$$\sum_{\mu\in Q_+} H_{\lambda+\rho}(\mu;\infty)z^{\lambda-\mu} = \chi(V(\lambda)).$$

Therefore, $H_{\lambda+\rho}(\mu;\infty)$ is the multiplicity of the weight $\lambda-\mu$ in $V(\lambda)$.

By setting $q=-1$ in (2-7), and by Lemma 2-4 we get the following.

Lemma 2-9.
$$\begin{aligned}\chi_{-1}(V(\lambda+\rho)) &= \sum_{\mu\in Q_+} H_{\lambda+\rho}(\mu;-1)z^{\lambda+\rho-\mu}\\ &= \chi(V(\lambda))\chi(V(\rho)) = \chi(V(\lambda)\otimes V(\rho)).\end{aligned}$$

Hence, $H_{\lambda+\rho}(\mu;-1)$ is the multiplicity of the weight $\lambda+\rho-\mu$ in the tensor product $V(\lambda)\otimes V(\rho)$.

Before we investigate further the implication of the Casselman–Shalika formula (2-7), we need the following lemma.

Lemma 2-10. *Assume that $\lambda_1, \lambda_2 \in P_+$. Then the set of weights of $V(\lambda_1)\otimes V(\lambda_2)$ is the same as that of $V(\lambda_1+\lambda_2)$.*

Proof. Suppose that $\lambda_1, \lambda_2 \in P_+$. Let $V(\lambda_1)$ and $V(\lambda_2)$ be the integrable highest weight modules with highest weights λ_1 and λ_2, respectively. By [Kac 1990, p. 211], $V(\lambda_1+\lambda_2)$ occurs in $V(\lambda_1)\otimes V(\lambda_2)$ with multiplicity one. Hence it is enough to prove that any weight of $V(\lambda_1)\otimes V(\lambda_2)$ is a weight of $V(\lambda_1+\lambda_2)$.

If V_1 and V_2 are modules in the category $\mathcal{O}$, the weight space of $(V_1\otimes V_2)_\mu$ for $\mu\in\mathfrak{h}^*$, is given by

$$(V_1\otimes V_2)_\mu = \sum_{\nu\in\mathfrak{h}^*}(V_1)_\nu\otimes(V_2)_{\mu-\nu}.$$

Hence weights of $V(\lambda_1) \otimes V(\lambda_2)$ are of the form $\mu_1 + \mu_2$, where μ_1 and μ_2 are weights of $V(\lambda_1)$ and $V(\lambda_2)$, respectively. Furthermore, since $V(\lambda_1) \otimes V(\lambda_2)$ is completely reducible, a weight $\mu_1+\mu_2$ of $V(\lambda_1)\otimes V(\lambda_2)$ is a weight of the module $V(\lambda)$ for some $\lambda \in P_+$ that appears in the decomposition of $V(\lambda_1) \otimes V(\lambda_2)$.

It follows from [Kac 1990, Corollary 10.1] that we can choose $w \in W$ such that $w(\mu_1+\mu_2) \in P_+$. Then, by [Kac 1990, Proposition 11.2], we need only show that $w(\mu_1+\mu_2)$ is nondegenerate with respect to $\lambda_1+\lambda_2$. By [Kac 1990, Lemma 11.2], $w\mu_1$ and $w\mu_2$ are nondegenerate with respect to λ_1 and λ_2, respectively. Now, from the definition of nondegeneracy [Kac 1990, p. 190], we see that $w\mu_1 + w\mu_2$ is nondegenerate with respect to $\lambda_1 + \lambda_2$. □

Now we use crystal bases, namely, bases at $v = 0$, since they behave nicely under tensor products. Let $\mathfrak{B}_\lambda$ be the crystal basis associated to a dominant integral weight $\lambda \in P_+$. We choose $G_\rho(\,\cdot\,; q) : \mathfrak{B}_\rho \to \mathbb{Z}[q^{-1}]$ by assigning any element of $\mathbb{Z}[q^{-1}]$ to each $b \in \mathfrak{B}_\rho$ so that

$$H_\rho(\mu; q) = \sum_{\substack{b\in\mathfrak{B}_\rho \\ \mathrm{wt}(b)=\rho-\mu}} G_\rho(b; q). \tag{2-11}$$

By Remark 2-5, it is enough to consider $\mu \in Q_+$ such that $\rho - \mu$ is a weight of $b \in \mathfrak{B}_\rho$.

Using the function $G_\rho(\,\cdot\,; q)$, we can rewrite the Casselman–Shalika formula in Corollary 2-6 in a familiar form:

Corollary 2-12.

$$\begin{aligned} \sum_{\mu\in Q_+} H_{\lambda+\rho}(\mu; q) z^{\lambda+\rho-\mu} &= \chi(V(\lambda)) z^\rho \prod_{\alpha\in\Delta^+} (1 - q^{-1} z^{-\alpha})^{\operatorname{mult}\alpha} \\ &= \sum_{b'\otimes b\in\mathfrak{B}_\lambda\otimes\mathfrak{B}_\rho} G_\rho(b; q) z^{\mathrm{wt}(b'\otimes b)}. \end{aligned} \tag{2-13}$$

Proof. The first equality is obvious from (2-3) and Corollary 2-6. For the second equality, we obtain

$$\begin{aligned} &\chi(V(\lambda)) z^\rho \prod_{\alpha\in\Delta^+} (1 - q^{-1} z^{-\alpha})^{\operatorname{mult}\alpha} \\ &\quad= \chi(V(\lambda))\chi_q(V(\rho)) = \Bigl(\sum_{b'\in\mathfrak{B}_\lambda} z^{\mathrm{wt}(b')}\Bigr)\Bigl(\sum_{\mu\in Q_+} H_\rho(\mu; q)\, z^{\rho-\mu}\Bigr) \\ &\quad= \Bigl(\sum_{b'\in\mathfrak{B}_\lambda} z^{\mathrm{wt}(b')}\Bigr)\Bigl(\sum_{b\in\mathfrak{B}_\rho} G_\rho(b; q) z^{\mathrm{wt}(b)}\Bigr) = \sum_{b'\otimes b\in\mathfrak{B}_\lambda\otimes\mathfrak{B}_\rho} G_\rho(b; q) z^{\mathrm{wt}(b'\otimes b)}. \end{aligned}$$
□

The following proposition provides useful information on $H_{\lambda+\rho}(\mu; q) \in \mathbb{Z}[q^{-1}]$.

Proposition 2-14. *Assume that $\lambda \in P_+$. We then have that $H_{\lambda+\rho}(\mu; q)$ is a nonzero polynomial if and only if $\lambda+\rho-\mu$ is a weight of $V(\lambda+\rho)$.*

Proof. We obtain from (2-13) that if $H_{\lambda+\rho}(\mu; q) \neq 0$, then $\lambda+\rho-\mu$ is a weight of $V(\lambda)\otimes V(\rho)$. Then $\lambda+\rho-\mu$ is a weight of $V(\lambda+\rho)$ by Lemma 2-10. Conversely, assuming that $\lambda+\rho-\mu$ is a weight of $V(\lambda+\rho)$, it is also a weight of $V(\lambda)\otimes V(\rho)$. By Lemma 2-9,

$$\sum_{\mu' \in Q_+} H_{\lambda+\rho}(\mu'; -1) z^{\lambda+\rho-\mu'} = \chi(V(\lambda)\otimes V(\rho)).$$

Since $\lambda+\rho-\mu$ is a weight of $V(\lambda)\otimes V(\rho)$, the coefficient $H_{\lambda+\rho}(\mu; -1) \neq 0$. Then $H_{\lambda+\rho}(\mu; q)$ is a nonzero polynomial. □

3. Applications

We give several applications of our formulas to q-deformation of (multi)partition functions and modular forms, and the Kostant function and the multiplicity formula. We also obtain formulas for $H_\lambda(\mu; q)$.

3.1. ***Multipartition functions and modular forms.*** We will write $\mathcal{P} = \mathcal{P}(1)$. For a partition $\boldsymbol{p} = (1^{m_1} 2^{m_2} \cdots r^{m_r} \cdots) \in \mathcal{P}$, we define

$$\kappa_q(\boldsymbol{p}) = \begin{cases} (-q^{-1})^{\sum m_r} & \text{if } m_r = 0 \text{ or } 1 \text{ for all } r, \\ 0 & \text{otherwise.} \end{cases}$$

We define for $k \geq 1$

$$\epsilon_q(k) = \sum_{\substack{\boldsymbol{p}\in\mathcal{P} \\ |\boldsymbol{p}|=k}} \kappa_q(\boldsymbol{p})$$

and set $\epsilon_q(0) = 1$. For example, $\epsilon_q(5) = 2q^{-2} - q^{-1}$ and $\epsilon_q(6) = -q^{-3} + 2q^{-2} - q^{-1}$.

From the definitions, we have

$$\prod_{k=1}^{\infty}(1-q^{-1}t^k) = 1 + \sum_{\boldsymbol{p}\in\mathcal{P}} \kappa_q(\boldsymbol{p}) t^{|\boldsymbol{p}|} = 1 + \sum_{k=1}^{\infty} \epsilon_q(k) t^k.$$

Then it follows from Euler's pentagonal number theorem that when $q = 1$, we have

$$\epsilon_1(k) = \begin{cases} (-1)^m & \text{if } k = \frac{1}{2}m(3m\pm 1), \\ 0 & \text{otherwise.} \end{cases} \tag{3-1}$$

We also define for $k \geq 1$

$$p_q(k) = \sum_{\substack{\boldsymbol{p}\in\mathcal{P} \\ |\boldsymbol{p}|=k}} (1-q^{-1})^{d(\boldsymbol{p})},$$

where $d(\boldsymbol{p})$ is the same as in the previous sections, and we set $p_q(0) = 1$. Note that if $k > 0$, $p_\infty(k) = p(k)$. Hence we can think of $p_q(k)$ as a q-deformation of the partition function.

Proposition 3-2. *If* $k > 0$, *then*

$$\epsilon_q(k) - p_q(k) = \sum_{m=1}^{\infty}(-1)^m\{p_q(k - \tfrac{1}{2}m(3m-1)) + p_q(k - \tfrac{1}{2}m(3m+1))\}, \tag{3-3}$$

where we define $p_q(M) = 0$ *for all negative integers* M.

Proof. We put $n = 1$ in Proposition 1-9 and obtain

$$\prod_{k=1}^{\infty}(1 - q^{-1}z^{k\delta}) = \left(\sum_{\boldsymbol{p}\in\mathcal{P}}(1 - q^{-1})^{d(\boldsymbol{p})}z^{|\boldsymbol{p}|\delta}\right)\prod_{k=1}^{\infty}(1 - z^{k\delta}).$$

After the change of variables $z^\delta = t$, we obtain

$$\begin{aligned}
1 + \sum_{k=1}^{\infty}\epsilon_q(k)t^k &= \prod_{k=1}^{\infty}(1 - q^{-1}t^k) \\
&= \left(\sum_{\boldsymbol{p}\in\mathcal{P}}(1 - q^{-1})^{d(\boldsymbol{p})}t^{|\boldsymbol{p}|}\right)\prod_{k=1}^{\infty}(1 - t^k) \\
&= \left(1 + \sum_{k=1}^{\infty}p_q(k)t^k\right)\left(1 + \sum_{m=1}^{\infty}(-1)^m\left\{t^{\frac{1}{2}m(3m-1)} + t^{\frac{1}{2}m(3m+1)}\right\}\right),
\end{aligned}$$

where we use the definition of $p_q(k)$ and (3-1) in the last equality. We obtain the identity (3-3) by expanding the product and equating the coefficient of t^k with $\epsilon_q(k)$. □

As a corollary of the proof of Proposition 3-2, we obtain:

Corollary 3-4. *Let* $(a; q)_n = \prod_{k=0}^{n-1}(1 - aq^k)$. *Then*

$$\sum_{n=0}^{\infty}\frac{(q^{-1}; t)_n}{(t; t)_n}\,t^n = \sum_{k=0}^{\infty}p_q(k)t^k.$$

Proof. By the q-binomial theorem,

$$\prod_{k=1}^{\infty}(1 - q^{-1}t^k) = \left(\sum_{n=0}^{\infty}\frac{(q^{-1}; t)_n}{(t; t)_n}\,t^n\right)\prod_{k=1}^{\infty}(1 - t^k).$$

Comparing this with the identity in the proof of Proposition 3-2, we obtain the result. □

Remark 3-5. When $q \to \infty$, we have

$$\sum_{n=0}^{\infty} \frac{t^n}{(t;t)_n} = \sum_{\boldsymbol{p} \in \mathcal{P}} t^{|\boldsymbol{p}|} = \sum_{n=0}^{\infty} p(n) t^n.$$

This is a special case of [Andrews 1976, Corollary 2.2].

We generalize Proposition 3-2 to the case of multipartitions. For a multipartition $\boldsymbol{p} = (\rho^{(1)}, \dots, \rho^{(n)}) \in \mathcal{P}(n)$, we define

$$\kappa_q(\boldsymbol{p}) = \prod_{i=1}^{n} \kappa_q(\rho^{(i)}),$$

and for $k \geq 1$,

$$\epsilon_{q,n}(k) = \sum_{\substack{\boldsymbol{p} \in \mathcal{P}(n) \\ |\boldsymbol{p}|=k}} \kappa_q(\boldsymbol{p}), \tag{3-6}$$

and set $\epsilon_{q,n}(0) = 1$. From the definitions, we have

$$\prod_{k=1}^{\infty} (1 - q^{-1} t^k)^n = 1 + \sum_{\boldsymbol{p} \in \mathcal{P}(n)} \kappa_q(\boldsymbol{p}) t^{|\boldsymbol{p}|} = \sum_{k=0}^{\infty} \epsilon_{q,n}(k) t^k.$$

One can see that if $k > 0$, we have $\epsilon_{\infty,n}(k) = 0$.

Remark 3-7. Note that $\epsilon_{1,n}(k)$ is a classical arithmetic function related to modular forms. For example, we have $\epsilon_{1,24}(k) = \tau(k+1)$, where $\tau(k)$ is the Ramanujan τ-function. Thus the function $\epsilon_{q,n}(k)$ should be considered as a q-deformation of the function $\epsilon_{1,n}(k)$.

We also define for $k \geq 1$

$$p_{q,n}(k) = \sum_{\substack{\boldsymbol{p} \in \mathcal{P}(n) \\ |\boldsymbol{p}|=k}} (1 - q^{-1})^{d(\boldsymbol{p})},$$

and set $p_{q,n}(0) = 1$. Notice that if $k > 0$, the function $p_{\infty,n}(k)$ is nothing but the multipartition function with n-components. Hence we can think of $p_{q,n}(k)$ as a q-deformation of the multipartition function.

Theorem 3-8. *If $k > 0$, then*

$$\epsilon_{q,n}(k) = \sum_{r=0}^{k} \epsilon_{1,n}(r) p_{q,n}(k-r). \tag{3-9}$$

Proof. From Proposition 1-9 we obtain

$$\prod_{k=1}^{\infty} (1 - q^{-1} z^{k\delta})^n = \left(\sum_{\boldsymbol{p} \in \mathcal{P}(n)} (1 - q^{-1})^{d(\boldsymbol{p})} z^{|\boldsymbol{p}|\delta} \right) \prod_{k=1}^{\infty} (1 - z^{k\delta})^n.$$

After the change of variables $z^\delta = t$, we obtain from the definitions

$$\sum_{k=0}^{\infty} \epsilon_{q,n}(k)t^k = \left(\sum_{\boldsymbol{p}\in\mathcal{P}(n)} (1-q^{-1})^{d(\boldsymbol{p})} t^{|\boldsymbol{p}|}\right) \prod_{k=1}^{\infty}(1-t^k)^n$$
$$= \left(\sum_{r=0}^{\infty} p_{q,n}(r)t^r\right)\left(\sum_{s=0}^{\infty} \epsilon_{1,n}(s)t^s\right). \qquad \square$$

By taking $q \to \infty$, we obtain the identity

$$0 = \sum_{r=0}^{k} \epsilon_{1,n}(r) p_{\infty,n}(k-r),$$

where $p_{\infty,n}(k)$ is the multipartition function with n-components. This is an easy consequence of the identities

$$\prod_{k=1}^{\infty}(1-t^k)^n = \sum_{k=0}^{\infty} \epsilon_{1,n}(k)t^k \quad \text{and} \quad \prod_{k=1}^{\infty}(1-t^k)^{-n} = \sum_{k=0}^{\infty} p_{\infty,n}(k)t^k.$$

Example 3-10. When the affine Kac–Moody algebra $\mathfrak{g}$ is of type $X_{24}^{(1)}$, with $X = A, B, C$, or D, we have

$$\epsilon_{q,24}(k) = \sum_{r=0}^{k} \tau(r+1) p_{q,24}(k-r) \quad \text{and} \quad 0 = \sum_{r=0}^{k} \tau(r+1) p_{\infty,24}(k-r),$$

where $\tau(k)$ is the Ramanujan τ-function. If $k=2$, the first identity becomes

$$\epsilon_{q,24}(2) = \tau(1)p_{q,24}(2) + \tau(2)p_{q,24}(1) + \tau(3)p_{q,24}(0).$$

Through some computations, we obtain

$$\epsilon_{q,24}(2) = 276q^{-2} - 24q^{-1}.$$

On the other hand, we have

$$\begin{aligned}\tau(1)p_{q,24}(2) + \tau(2)p_{q,24}(1) + \tau(3)p_{q,24}(0)&\\ &= p_{q,24}(2) - 24p_{q,24}(1) + 252\\ &= \{276(1-q^{-1})^2 + 48(1-q^{-1})\} - 24\cdot 24(1-q^{-1}) + 252\\ &= 276(1-q^{-1})^2 - 528(1-q^{-1}) + 252\\ &= 276q^{-2} - 24q^{-1} = \epsilon_{q,24}(2).\end{aligned}$$

We also see that

$$\tau(1)p_{\infty,24}(2) + \tau(2)p_{\infty,24}(1) + \tau(3)p_{\infty,24}(0) = 324 - 24\cdot 24 + 252 = 0.$$

Now we consider the whole set of positive roots, not just the set of imaginary positive roots, and obtain interesting identities. We begin with the identity (2-3). Recalling the description of the set of positive roots, we obtain

$$\begin{aligned}
(3\text{-}11)\qquad & \sum_{\mu\in Q_+} H_\rho(\mu;q)z^{-\mu} \\
&= z^{-\rho}\chi_q(V(\rho)) = \prod_{\alpha\in\Delta_+}(1-q^{-1}z^{-\alpha})^{\operatorname{mult}\alpha} \\
&= \left(\prod_{k=1}^{\infty}(1-q^{-1}z^{-k\delta})^n \prod_{\alpha\in\Delta_{\mathrm{cl}}}(1-q^{-1}z^{\alpha-k\delta})\right)\prod_{\alpha\in\Delta_{\mathrm{cl}}^+}(1-q^{-1}z^{-\alpha}),
\end{aligned}$$

where Δ_{cl} is the set of classical roots.

Let

$$\mathscr{X} = \left\{\sum_{\alpha\in Q_+} c_\alpha z^{-\alpha} : c_\alpha\in\mathbb{C}\right\}$$

be the set of (infinite) formal sums. Recall that we have the element $d\in\mathfrak{h}$ such that $\alpha_0(d)=1$ and $\alpha_j(d)=0$, $j\in I\setminus\{0\}$. Let $\mathfrak{h}_{\mathbb{Z}}$ be the $\mathbb{Z}$-span of $\{h_0,h_1,\dots,h_n,d\}$. We then define the evaluation map $\mathrm{EV}_t:\mathscr{X}\times\mathfrak{h}_{\mathbb{Z}}\to\mathbb{C}[[t]]$ by

$$\mathrm{EV}_t\left(\sum_\alpha c_\alpha z^{-\alpha}, s\right) = \sum_\alpha c_\alpha t^{\alpha(s)},\qquad s\in\mathfrak{h}_{\mathbb{Z}}.$$

Then we see that $\mathrm{EV}_t(\,\cdot\,,d)$ is the same as the *basic specialization* in [Kac 1990, p. 219] with q replaced by t. We apply $\mathrm{EV}_t(\,\cdot\,,d)$ to (3-11) and obtain

$$(3\text{-}12)\qquad (1-q^{-1})^{|\Delta_{\mathrm{cl}}^+|}\prod_{k=1}^{\infty}(1-q^{-1}t^k)^{\dim\mathfrak{g}_{\mathrm{cl}}} = \sum_{k=0}^{\infty}\left(\sum_{\mu\in Q_{+,\mathrm{cl}}} H_\rho(k\alpha_0+\mu;q)\right)t^k,$$

where $\mathfrak{g}_{\mathrm{cl}}$ is the finite-dimensional simple Lie algebra corresponding to $\mathfrak{g}$, and $Q_{+,\mathrm{cl}}$ is the $\mathbb{Z}_{\ge 0}$-span of $\{\alpha_1,\dots,\alpha_n\}$. We write $|\Delta_{\mathrm{cl}}^+|=r$ and $\dim\mathfrak{g}_{\mathrm{cl}}=N$ so that $N=2r+n$. By comparing (3-12) with the identity

$$\prod_{k=1}^{\infty}(1-q^{-1}t^k)^n = \sum_{k=0}^{\infty}\epsilon_{q,n}(k)t^k,$$

we obtain:

Proposition 3-13. $$\epsilon_{q,N}(k) = \sum_{\mu\in Q_{+,\mathrm{cl}}}\frac{H_\rho(k\alpha_0+\mu;q)}{(1-q^{-1})^r}.$$

By Definition 2-2, $\epsilon_{q,N}(k)$ is a power series in q^{-1} in the above formula. However, one can see from (3-6) that $\epsilon_{q,N}(k)$ is actually a polynomial in q^{-1}.

Example 3-14. We take $\mathfrak{g}$ to be of type $A_4^{(1)}$. Then the classical Lie algebra $\mathfrak{g}_{\mathrm{cl}}$ is of type A_4, and $r = |\Delta_{\mathrm{cl}}^+| = 10$ and $N = \dim \mathfrak{g}_{\mathrm{cl}} = 24$. Taking the limit $q \to 1$, we obtain

$$\tau(k+1) = \lim_{q\to 1} \sum_{\mu\in Q_{+,\mathrm{cl}}} \frac{H_\rho(k\alpha_0+\mu;q)}{(1-q^{-1})^{10}}.$$

Therefore the sum $\sum_{\mu\in Q_{+,\mathrm{cl}}} H_\rho(k\alpha_0+\mu;q)$ is always divisible by $(1-q^{-1})^{10}$. But Lehmer's conjecture predicts that the sum is never divisible by $(1-q^{-1})^{11}$.

3.2. ***The Kostant function and $H_\lambda(\mu;q)$.*** In this section, let $\mathfrak{g}$ be an untwisted affine Kac–Moody algebra (affine type) or a finite-dimensional simple Lie algebra (finite type).

Definition 3-15. We define the functions $K_q^\infty(\mu)$ and $K_q^1(\mu)$ by

$$\sum_{\mu\in Q_+} K_q^\infty(\mu) z^\mu = \prod_{\alpha\in\Delta_+}\left(\frac{1-q^{-1}z^\alpha}{1-z^\alpha}\right)^{\operatorname{mult}\alpha} = \sum_{b\in\boldsymbol{G}} (1-q^{-1})^{d(\phi(b))} z^{\mathrm{wt}(b)}$$

and

$$\sum_{\mu\in Q_+} K_q^1(\mu) z^\mu = \prod_{\alpha\in\Delta_+} (1-q^{-1}z^\alpha)^{-\operatorname{mult}\alpha} = \sum_{b\in\boldsymbol{G}} q^{-|\phi(b)|} z^{\mathrm{wt}(b)}.$$

We set $K_q^\infty(\mu) = K_q^1(\mu) = 0$ if $\mu \notin Q_+$.

Remark 3-16. (1) Note that both $K_\infty^\infty(\mu)$ with $q=\infty$ and $K_1^1(\mu)$ with $q=1$ are equal to the classical Kostant partition function $K(\mu)$. Hence both of them can be considered as q-deformations of the Kostant function.

(2) The function $K_q^1(\mu)$ was introduced by Lusztig [1983] for finite types; see also [Kato 1982]. On the other hand, the function $K_q^\infty(\mu)$ for finite types can be found in the work of Guillemin and Rassart [2004].

We obtain from the Casselman–Shalika formula (Corollary 2-6)

$$\begin{aligned} z^{-\lambda}\chi(V(\lambda)) &= \sum_{\beta\in Q_+} (\dim V(\lambda)_{\lambda-\beta}) z^{-\beta} \\ &= z^{-\lambda-\rho}\chi_q(V(\lambda+\rho)) \prod_{\alpha\in\Delta_+} (1-q^{-1}z^{-\alpha})^{-\operatorname{mult}\alpha} \\ &= \left(\sum_{\mu\in Q_+} H_{\lambda+\rho}(\mu;q) z^{-\mu}\right)\left(\sum_{\nu\in Q_+} K_q^1(\nu) z^{-\nu}\right). \end{aligned}$$

Therefore, we have a q-deformation of the Kostant multiplicity formula:

Proposition 3-17. $\dim V(\lambda)_{\lambda-\beta} = \sum_{\mu\in Q_+} H_{\lambda+\rho}(\mu;q) K_q^1(\beta-\mu)$.

In order to see that this is indeed a q-deformation of the Kostant multiplicity formula, we need to determine the value of $H_{\lambda+\rho}(\mu; 1)$.

Lemma 3-18. *We have*

$$H_{\lambda+\rho}(\mu; 1) = \begin{cases} (-1)^{\ell(w)} & \text{if } w \circ \lambda = -\mu \text{ for some } w \in W, \\ 0 & \text{otherwise,} \end{cases}$$

where we define $w \circ \lambda = w(\lambda+\rho) - \lambda - \rho$ for $w \in W$ and $\lambda \in P_+$.

Note that such a $w \in W$ is unique if it exists, so there is no ambiguity in the assertion.

Proof. From Definition 2-2, we obtain

$$\sum_{\mu \in Q_+} H_{\lambda+\rho}(\mu; 1) z^{\lambda+\rho-\mu} = \sum_{w \in W} (-1)^{\ell(w)} z^{w(\lambda+\rho)}.$$

The condition $\lambda + \rho - \mu = w(\lambda+\rho)$ is equivalent to $w \circ \lambda = -\mu$. □

Now we take $q = 1$ in Proposition 3-17 and use Lemma 3-18 to obtain the classical Kostant multiplicity formula

$$\dim V(\lambda)_{\lambda-\beta} = \sum_{w \in W} (-1)^{\ell(w)} K(w \circ \lambda + \beta).$$

Note that the sum is actually a finite sum. Indeed, we have $w \circ \lambda < 0$ for each $w \in W$ and $w \circ \lambda + \beta \geq 0$ only for finitely many $w \in W$ for fixed $\lambda \in P_+$ and $\beta \in Q_+$. For the same reason, the sum in (3-23) below is also a finite sum.

Remark 3-19. In Section 2 we obtained (Remark 2-8 and Lemma 2-9)

$$H_{\lambda+\rho}(\mu; \infty) = \dim V(\lambda)_{\lambda-\mu}, \tag{3-20}$$

$$H_{\lambda+\rho}(\mu; -1) = \dim(V(\lambda) \otimes V(\rho))_{\lambda+\rho-\mu}. \tag{3-21}$$

When $\mathfrak{g}$ is of finite type, we define $H_\lambda(\mu; q)$ as in Definition 2-2, and we can prove the analogous results. See [Kim and Lee 2011] for details.

We next derive a formula for $H_{\lambda+\rho}(\mu; q)$:

Proposition 3-22.

$$H_{\lambda+\rho}(\mu; q) = \sum_{w \in W} (-1)^{\ell(w)} K_q^\infty(w \circ \lambda + \mu). \tag{3-23}$$

Proof. From the definitions we have

$$\begin{aligned} \chi_q(V(\lambda+\rho)) &= \sum_{\mu \in Q_+} H_{\lambda+\rho}(\mu; q) z^{\lambda+\rho-\mu} \\ &= \Big(\sum_{w \in W} (-1)^{\ell(w)} z^{w(\lambda+\rho)}\Big)\Big(\sum_{\nu \in Q_+} K_q^\infty(\nu) z^{-\nu}\Big). \end{aligned}$$

The identity comes from expanding the product and comparing the coefficients. □

If we take the limit $q \to \infty$ in (3-23), we have, from (3-20),

$$\dim V(\lambda)_{\lambda-\mu} = \sum_{w\in W} (-1)^{\ell(w)} K(w\circ\lambda+\mu),$$

which is again the classical Kostant multiplicity formula.

If we take $q=-1$ in (3-23), we obtain, from (3-21),

$$\dim(V(\lambda)\otimes V(\rho))_{\lambda+\rho-\mu} = \sum_{w\in W} (-1)^{\ell(w)} K_{-1}^{\infty}(w\circ\lambda+\mu). \tag{3-24}$$

This is a generalization of the formula in [Guillemin and Rassart 2004, Theorem 1] to the affine case.

Example 3-25. Assume that $\mathfrak{g}$ is of type $A_1^{(1)}$. We write

$$\mu = m\alpha_0 + n\alpha_1 = (m,n) \in Q_+$$

and set $\lambda=0$ in (3-23). Through standard computation, we obtain

$$\{w\rho+\mu-\rho : w\in W\} = \Big\{\Big(m-\frac{k(k+1)}{2}, n-\frac{k(k-1)}{2}\Big) \,\Big|\, k\in\mathbb{Z}\Big\}.$$

Thus we have

$$H_\rho(m,n;q) = \sum_{k\in\mathbb{Z}} (-1)^k K_q^{\infty}\Big(m-\frac{k(k+1)}{2}, n-\frac{k(k-1)}{2}\Big).$$

By taking the limit as $q\to\infty$, we obtain, for $(m,n)\neq(0,0)$,

$$0 = \sum_{k\in\mathbb{Z}} (-1)^k K\Big(m-\frac{k(k+1)}{2}, n-\frac{k(k-1)}{2}\Big).$$

In this case, $K(m,n)$ counts the number of vector partitions of (m,n) into parts of the forms (a,a), $(a-1,a)$, or $(a,a-1)$. Then we have obtained (3-9) [Carlitz 1965, p. 148].

We further investigate properties of the function $H_\lambda(\mu;q)$. From the definitions of $K_q^{\infty}(\mu)$ and $K_q^1(\mu)$, we have

$$\begin{aligned}\Big(\sum_{\mu\in Q_+} K_q^{\infty}(\mu)z^{\mu}\Big)\Big(\sum_{\nu\in Q_+} K_q^{1}(\nu)z^{\nu}\Big) \\ &= \prod_{\alpha\in\Delta_+}\Big(\frac{1-q^{-1}z^{\alpha}}{1-z^{\alpha}}\Big)^{\operatorname{mult}\alpha} \prod_{\alpha\in\Delta_+}(1-q^{-1}z^{\alpha})^{-\operatorname{mult}\alpha} \\ &= \prod_{\alpha\in\Delta_+}(1-z^{\alpha})^{-\operatorname{mult}\alpha} = \sum_{\beta\in Q_+} K(\beta)z^{\beta},\end{aligned}$$

where $K(\beta)$ is the classical Kostant function. Thus we have

$$\sum_{\mu\in Q_+} K_q^{\infty}(\mu)K_q^1(\beta-\mu)=K(\beta), \tag{3-26}$$

and we obtain, for $\beta>0$,

$$K_q^{\infty}(\beta)=K(\beta)-K_q^1(\beta)-\sum_{0<\nu<\beta} K_q^{\infty}(\nu)K_q^1(\beta-\nu), \tag{3-27}$$

and $K_q^{\infty}(0)=K_q^1(0)=K(0)=1$.

Then we obtain from Proposition 3-22

$$\begin{aligned} H_{\lambda+\rho}(\mu;q)=H_{\lambda+\rho}(\mu;1)+\sum_{w\in W}(-1)^{\ell(w)}K(w\circ\lambda+\mu)-\sum_{w\in W}(-1)^{\ell(w)}K_q^1(w\circ\lambda+\mu)\\ -\sum_{\substack{w\in W\\ w\circ\lambda+\mu>0}}(-1)^{\ell(w)}\sum_{0<\nu<w\circ\lambda+\mu}K_q^{\infty}(\nu)K_q^1(w\circ\lambda+\mu-\nu),\end{aligned}$$

where $H_{\lambda+\rho}(\mu;1)$ plays the role of a correction term for the case $w\circ\lambda+\mu=0$. See Lemma 3-18 for the value of $H_{\lambda+\rho}(\mu;1)$. We also used the fact that

$$K(\beta)=K_q^1(\beta)=K_q^{\infty}(\beta)=0$$

unless $\beta\geq 0$.

Now we apply the classical Kostant formula and get:

Proposition 3-28. *Assume that $\lambda\in P_+$ and $\mu\in Q_+$. Then we have*

$$\begin{aligned} H_{\lambda+\rho}(\mu;q)=H_{\lambda+\rho}(\mu;1)+\dim V(\lambda)_{\lambda-\mu}-\sum_{w\in W}(-1)^{\ell(w)}K_q^1(w\circ\lambda+\mu)\\ -\sum_{\substack{w\in W\\ w\circ\lambda+\mu>0}}(-1)^{\ell(w)}\sum_{0<\nu<w\circ\lambda+\mu}K_q^{\infty}(\nu)K_q^1(w\circ\lambda+\mu-\nu).\end{aligned}$$

For the rest of this section, we assume that $\mathfrak{g}$ is of finite type. We denote by $\rho^\vee$ the element of $\mathfrak{h}$ defined by $\langle\alpha_i,\rho^\vee\rangle=1$ for all the simple roots α_i. The following identity was conjectured by Lusztig [1983] and proved by S. Kato [1982].

Proposition 3-29. *For $\lambda\in P_+$ and $\mu\in Q_+$, we have*

$$\sum_{w\in W}(-1)^{\ell(w)}K_q^1(w\circ\lambda+\mu)=q^{-\langle\mu,\rho^\vee\rangle}P_{w_{\lambda-\mu},w_\lambda}(q),$$

where w_ν is the element in the affine Weyl group $\widehat{W}$ corresponding to $\nu\in P_+$, and $P_{w_{\lambda-\mu},w_\lambda}(q)$ is the Kazhdan–Lusztig polynomial.

Hence, from Proposition 3-28, we obtain:

Corollary 3-30.
$$H_{\lambda+\rho}(\mu;q)=H_{\lambda+\rho}(\mu;1)+\dim V(\lambda)_{\lambda-\mu}-q^{-\langle\mu,\rho^\vee\rangle}P_{w_{\lambda-\mu},w_\lambda}(q) - \sum_{\substack{w\in W\\ w\circ\lambda+\mu>0}}(-1)^{\ell(w)}\sum_{0<\nu<w\circ\lambda+\mu}K_q^\infty(\nu)K_q^1(w\circ\lambda+\mu-\nu).$$

Setting $q=1$, and noting that $K_1^\infty(\beta)=0$ if $\beta>0$, we obtain the famous property of the Kazhdan–Lusztig polynomial:

Corollary 3-31.
$$\dim V(\lambda)_{\lambda-\mu}=P_{w_{\lambda-\mu},w_\lambda}(1).$$

Acknowledgments

We thank M. Patnaik for explaining his results [Braverman et al. ≥ 2012]. We also thank A. Ram, P. Gunnells, S. Friedberg and B. Brubaker for useful comments.

References

[Andrews 1976] G. E. Andrews, *The theory of partitions*, Encyclopedia Math. Appl. **2**, Addison-Wesley, Reading, MA, 1976. MR 58 #27738 Zbl 0371.10001

[Beck and Nakajima 2004] J. Beck and H. Nakajima, "Crystal bases and two-sided cells of quantum affine algebras", *Duke Math. J.* **123**:2 (2004), 335–402. MR 2005e:17020 Zbl 1062.17006

[Beck et al. 1999] J. Beck, V. Chari, and A. Pressley, "An algebraic characterization of the affine canonical basis", *Duke Math. J.* **99**:3 (1999), 455–487. MR 2000g:17013 Zbl 0964.17013

[Braverman et al. ≥ 2012] A. Braverman, D. Kazhdan, and M. Patnaik, "The Iwahori–Hecke algebra for an affine Kac–Moody group", in preparation.

[Bump and Nakasuji 2010] D. Bump and M. Nakasuji, "Integration on p-adic groups and crystal bases", *Proc. Amer. Math. Soc.* **138**:5 (2010), 1595–1605. MR 2011f:22015 Zbl 05706777

[Carlitz 1965] L. Carlitz, "Generating functions and partition problems", pp. 144–169 in *Symposium on Recent Developments in the Theory of Numbers* (Pasadena, CA, 1963), edited by A. L. Whiteman, Proc. Sympos. Pure Math. **8**, Amer. Math. Soc., Providence, RI, 1965. MR 31 #72 Zbl 0142.25104

[Guillemin and Rassart 2004] V. Guillemin and E. Rassart, "Signature quantization and representations of compact Lie groups", *Proc. Natl. Acad. Sci. USA* **101**:30 (2004), 10884–10889. MR 2005k:53175 Zbl 1063.53094

[Kac 1990] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990. MR 92k:17038 Zbl 0716.17022

[Kato 1982] S.-I. Kato, "Spherical functions and a q-analogue of Kostant's weight multiplicity formula", *Invent. Math.* **66**:3 (1982), 461–468. MR 84b:22030 Zbl 0498.17005

[Kim and Lee 2011] H. H. Kim and K.-H. Lee, "Representation theory of p-adic groups and canonical bases", *Adv. Math.* **227**:2 (2011), 945–961. MR 2793028 Zbl 1228.22011

[Lusztig 1983] G. Lusztig, "Singularities, character formulas, and a q-analog of weight multiplicities", pp. 208–229 in *Analyse et topologie sur les espaces singuliers, II, III* (Luminy, 1981), edited by A. A. Beilinson et al., Astérisque **101–102**, Soc. Math. France, Paris, 1983. MR 85m:17005 Zbl 0561.22013

[Lusztig 1993] G. Lusztig, *Introduction to quantum groups*, Progr. Math. **110**, Birkhäuser, Boston, 1993. MR 94m:17016 Zbl 0788.17010

Received January 24, 2011.

HENRY H. KIM
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO, ONTARIO M5S2E4
CANADA

and

KOREA INSTITUTE FOR ADVANCED STUDY
SEOUL
KOREA

henrykim@math.toronto.edu

KYU-HWAN LEE
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CONNECTICUT
STORRS, CT 06269-3009
UNITED STATES

khlee@math.uconn.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

DIRICHLET–FORD DOMAINS AND ARITHMETIC REFLECTION GROUPS

GRANT S. LAKELAND

We show that a Fuchsian group, acting on the upper half-plane model for $\mathbb{H}^2$, admits a Ford domain which is also a Dirichlet domain, for some center, if and only if it is an index 2 subgroup of a reflection group. This is used to exhibit an example of a maximal arithmetic hyperbolic reflection group which is not congruence. Analogous results, and counterexamples, are given in the case of Kleinian groups.

1. Introduction

The action of the modular group $\mathrm{PSL}_2(\mathbb{Z})$ on the upper half-plane model for hyperbolic 2-space $\mathbb{H}^2$ has been extensively studied. It is well known that a fundamental domain for this action is given by the triangle T with vertices at $\rho = \frac{1}{2} + \frac{\sqrt{-3}}{2}$, $-\bar{\rho}$ and ∞. This domain is an example of two common constructions of fundamental domains for Fuchsian groups: it is both a Ford domain and a Dirichlet domain for the action of the modular group. Furthermore, it arises from more than one distinct choice of Dirichlet center, as taking the Dirichlet domain centered at any $z_0 = iy$, for $y > 1$, gives rise to T. One expects the Dirichlet domain to change along with the choice of center [Díaz and Ushijima 2009], so in this sense the modular group exhibits some atypical properties.

It is also well known that $\mathrm{PSL}_2(\mathbb{Z})$ is the orientation-preserving index 2 subgroup of the group generated by reflections in an ideal triangle in $\mathbb{H}^2$ with angles $\frac{\pi}{2}$, $\frac{\pi}{3}$ and 0, located at i, ρ and ∞ respectively. More generally, a *hyperbolic reflection group* is a subgroup of $\mathrm{Isom}(\mathbb{H}^2)$ generated by reflections in the sides of a polygon $Q \subset \mathbb{H}^2$. Such a group is discrete if and only if each angle of Q is either 0 or an integer submultiple of π.

The purpose of this paper is to determine exactly which Fuchsian groups admit a fundamental domain that is both a Dirichlet domain and a Ford domain (which we will call a *DF domain*) or a Dirichlet domain for multiple centers (a *double Dirichlet domain*).

It turns out that the above properties of $\mathrm{PSL}_2(\mathbb{Z})$ are very much related to that;

MSC2010: primary 20H10; secondary 19B37.
Keywords: fundamental domain, arithmetic reflection groups, congruence.

we give in Theorems 5.3 and 5.4 the following characterization for such groups:

Theorem. *A finitely generated, finite coarea Fuchsian group Γ admits a DF domain* (*or a double Dirichlet domain*) *P if and only if Γ is an index 2 subgroup of the discrete group G of reflections in a hyperbolic polygon Q.*

The condition given by this result provides a method of checking whether a given Fuchsian group is the index 2 orientation-preserving subgroup of a hyperbolic reflection group. This is particularly useful in the context of *maximal arithmetic* reflection groups. A noncocompact hyperbolic reflection group is *arithmetic* if it is commensurable with $\mathrm{PSL}_2(\mathbb{Z})$. A maximal arithmetic reflection group is then an arithmetic reflection group which is not properly contained in another arithmetic reflection group. As an application of the above theorem, we give:

Corollary. *There exists a maximal arithmetic hyperbolic reflection group which is not congruence.*

This answers a question raised by Agol, Belolipetsky, Storm, and Whyte in [Agol et al. 2008] (see also [Belolipetsky 2011]).

Any group Γ satisfying the theorem must have *genus zero* [Long et al. 2006], as well as a certain symmetrical property. Having a double Dirichlet or DF domain therefore also gives an obstruction to the group having *nontrivial cuspidal cohomology* [Grunewald and Schwermer 1981].

Motivated by this, one may ask whether there exists a similar obstruction to nontrivial cuspidal cohomology for Kleinian groups. We will show that this is not the case: in Section 7 we exhibit a Kleinian group which possesses both nontrivial cuspidal cohomology and a DF domain. However, the condition of having such a domain does impose some restrictions on Γ; perhaps the most striking is that the group possesses a generating set, all of whose elements have real trace (Theorem 7.3).

This paper is organized as follows. After the preliminaries of Section 2, Section 3 will examine Fuchsian groups with DF domains, and show that such domains are symmetrical and give rise to punctured spheres. The more general case of the double Dirichlet domain is discussed in Section 4. In Section 5, it will be shown that the main theorem follows from the previous sections and standard results on reflection groups. An example of a noncongruence maximal arithmetic hyperbolic reflection group can be found in Section 6. Section 7 contains a discussion of these domains in the setting of Kleinian groups.

2. Preliminaries

We will work in the upper half-plane model for the hyperbolic plane. The group of conformal, orientation-preserving isometries (or linear fractional transformations) of $\mathbb{H}^2$ can be identified with $\mathrm{PSL}_2(\mathbb{R})$ via the correspondence

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \longleftrightarrow \quad z \mapsto \frac{az+b}{cz+d}.$$

A *Fuchsian group* Γ is a subgroup of $\mathrm{PSL}_2(\mathbb{R})$, discrete with respect to the topology induced by regarding that group as a subset of $\mathbb{R}^4$. The *Dirichlet domain* for Γ centered at z_0 is defined to be

$$\{x \in \mathbb{H}^2 \mid d(x, z_0) \leq d(x, \alpha(z_0)) \text{ for all } \alpha \in \Gamma \text{ distinct from } 1\}.$$

It is an intersection of closed half-spaces.

Beardon [1983, Section 9.5] demonstrates an alternative definition, in terms of reflections, which allows us to define a *generalized Dirichlet domain* by taking our center to be on the boundary $\partial\mathbb{H}^2$. We will typically conjugate Γ in $\mathrm{PSL}_2(\mathbb{R})$ so that this center is placed at ∞ in the upper half-plane. We suppose Γ is *zonal*, or that ∞ is a parabolic fixed point, and so the reflections given are not uniquely determined for any parabolic isometry fixing ∞. To account for this, we define a *Ford domain* [Ford 1951] to be the intersection of the region exterior to all isometric circles with a fundamental domain for the action of the parabolic subgroup stabilizing ∞, $\Gamma_\infty < \Gamma$, which is a vertical strip.

For a given finitely generated Fuchsian group Γ, the *signature*

$$(g; n_1, \ldots, n_t; m; f)$$

of Γ records the topology of the quotient space $\mathbb{H}^2/\Gamma$, where g is the genus, t is the number of cone points of orders $n_1, \ldots, n_t$ respectively, m is the number of cusps, and f is the number of infinite area funnels. If Γ is the orientation-preserving index 2 subgroup of a reflection group, then $\mathbb{H}^2/\Gamma$ is a sphere with cusps and/or cone points, and thus in this case we have $g = 0$. If additionally Γ has finite coarea, then we also have that $f = 0$.

The group of orientation-preserving isometries of the upper half-space model of $\mathbb{H}^3$ can likewise be identified with $\mathrm{PSL}_2(\mathbb{C})$. A *Kleinian group* is a discrete subgroup of this isometry group. The definitions of Dirichlet domain and Ford domain carry over to this situation, with one small modification: instead of Γ being zonal, we assume that Γ_∞ contains a copy of $\mathbb{Z}^2$.

Throughout, we will assume that Γ is finitely generated, and hence that all fundamental domains we encounter have a finite number of sides. For simplicity, we will also suppose that $f = 0$ and that Γ has finite covolume (and thus that all fundamental domains have finite volume; that is, finitely many ideal vertices, each adjacent to two sides), although many of the arguments should extend to the case where Γ does not have finite covolume.

3. DF domains

Suppose Γ contains a nontrivial parabolic subgroup Γ_∞ fixing ∞. In $\mathbb{H}^2$, Γ_∞ must be cyclic, and after conjugation, we may take it to be generated by

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Theorem 3.1. *If Γ admits a DF domain, then the quotient space $\mathbb{H}^2/\Gamma$ is a punctured sphere, possibly with cone points.*

Before commencing the proof of this, we will prove two elementary but important lemmas. The first is given as Exercise 2 in Section 9.6 of [Beardon 1983].

Lemma 3.2. *Any vertex cycle on the boundary of a Ford domain P is contained within a horocycle based at ∞.*

Proof. Fix a vertex v. By construction of P, v lies on or exterior to all isometric circles, and necessarily lies on at least one. We first consider a $\gamma \in \Gamma$ such that $v \notin S_\gamma$. Then v lies exterior to S_γ. It follows that γ sends v into the interior of $S_{\gamma^{-1}}$. Thus $\gamma(v)$ cannot be a vertex of P. Now suppose that $v \in S_\gamma$. Then γ is the composition of reflection in S_γ, which fixes v, and reflection in a vertical line. It therefore necessarily preserves the imaginary part of v, proving the lemma. □

Remark. The lemma holds for any point on the boundary of the Ford domain P. For our purposes, it will be enough to have it for the vertices of P.

Lemma 3.3 [Greenberg 1977, p. 203]. *Let P be a Dirichlet domain for Γ with center z_0. Let $1 \neq \gamma$ be an element of Γ and suppose that $z, \gamma(z) \in \partial P \cap \mathbb{H}^2$. Then $d_{\mathbb{H}}(z, z_0) = d_{\mathbb{H}}(\gamma(z), z_0)$.*

Proof. This is an application of the definition of a Dirichlet domain stated above. Specifically, setting $x = z$ and $\alpha = \gamma^{-1}$ yields the inequality

$$d(z, z_0) \leq d(z, \gamma^{-1}(z_0)) = d(\gamma(z), z_0),$$

the latter equality holding because γ is an isometry. Setting $x = \gamma(z)$ and $\alpha = \gamma$ now gives

$$d(\gamma(z), z_0) \leq d(\gamma(z), \gamma(z_0)) = d(z, z_0).$$

Combining these two inequalities gives the required equality. □

Proof of Theorem 3.1. Suppose we are given a DF domain P for Γ. Since P is a Ford domain, it is contained in a fundamental region for Γ_∞, which is a vertical strip

$$\{z \in \mathbb{H}^2 \mid x_0 \leq \operatorname{Re}(z) \leq x_0 + 1\}$$

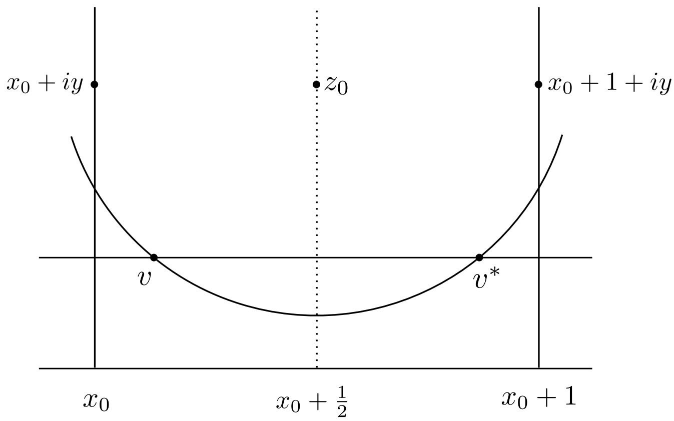


Figure 1. $\gamma(v) = v^*$.

for some $x_0 \in \mathbb{R}$. Shimizu's Lemma (see [Maskit 1988], p. 18) tells us that the radii of the isometric circles S_γ cannot exceed 1. Thus we may consider a point $z = x_0 + iy \in \partial P$, where $y > 1$. Choosing $\gamma = T$, and applying Lemma 3.3 to z and $\gamma(z)$, we find that $\mathrm{Re}(z_0) = x_0 + \frac{1}{2}$.

Next suppose that $v \in \mathbb{H}^2$ is a vertex of P, and $\gamma \in \Gamma$ a side pairing such that $\gamma(v)$ is another vertex of P. Then, by Lemma 3.2, $\mathrm{Im}(\gamma(v)) = \mathrm{Im}(v)$, and by Lemma 3.3, $d_{\mathbb{H}}(\gamma(v), z_0) = d_{\mathbb{H}}(v, z_0)$. Consider the two sets $\{z \in \mathbb{H}^2 \mid \mathrm{Im}(z) = \mathrm{Im}(v)\}$ and $\{z \in \mathbb{H}^2 \mid d_{\mathbb{H}}(z, z_0) = d_{\mathbb{H}}(v, z_0)\}$. The former is the horizontal line through v, and the latter a circle with Euclidean center located vertically above z_0 (see Figure 1). In particular, the picture is symmetrical in the vertical line $\left\{\mathrm{Re}(z) = x_0 + \frac{1}{2}\right\}$. Either $\gamma(v) = v$ or $\gamma(v) = v^*$, where v^* is the reflection of v in the line $\left\{\mathrm{Re}(z) = x_0 + \frac{1}{2}\right\}$.

Suppose that $\gamma(v) = v$. Then, by considering a point $w \in \partial P$ close to v, the fact that $d(w, z_0) = d(\gamma(w), z_0)$ means that v necessarily lies directly below the Dirichlet center z_0. The contrapositive of this states that if $\mathrm{Re}(v) \neq x_0 + \frac{1}{2}$, then any side pairing γ pairing v with a vertex of P must send v to v^*.

Suppose now that $v \in \partial\mathbb{H}^2$ is a vertex of P. Then two isometric circles meet at v. Fix one such circle S. S is the isometric circle S_γ of some element $\gamma \in \Gamma$. S_γ contains a side of P adjacent to v, and we pick two points of S_γ, $w_1, w_2 \in \partial P \cap \mathbb{H}^2$. By Lemma 3.3, γ must send both w_1 and w_2 to points the same respective distances from z_0. For each i, the point w_i is either fixed or sent to its reflection in the line $\{\mathrm{Re}(z) = \mathrm{Re}(z_0)\}$. If w_1 were fixed, w_2 would neither be fixed nor sent to its reflection, and vice versa if w_2 were fixed. Thus we conclude that γ sends points of S to their reflections in the line $\{\mathrm{Re}(z) = \mathrm{Re}(z_0)\}$.

We can now show that $\mathbb{H}^2/\Gamma$ is a punctured sphere. We first identify the two vertical sides of P, creating the cusp at ∞ and a circle awaiting identification. Consider the side of P adjacent to the side contained in the vertical line $\mathrm{Re}(z) = x_0$.

This side lies on some isometric circle S_γ. We see that γ must identify our side with a side adjacent to the side of P contained in the line $\mathrm{Re}(z) = x_0 + 1$. Working inwards toward the center and applying this argument repeatedly, we see that all sides must pair up symmetrically. In particular, there can not exist two hyperbolic generators whose axes intersect precisely once. Thus we conclude that the quotient space has genus zero. $\square$

Remarks. (1) We may take the Dirichlet center of P to be any point of the interior of P on this vertical line $\{\mathrm{Re}(z) = x_0 + \frac{1}{2}\}$. To see this, let z_0 be any such point, and $\gamma \in \Gamma \setminus \Gamma_\infty$ a side pairing of P. Since $\gamma(S_\gamma) = S_{\gamma^{-1}}$, and this pair is arranged symmetrically with respect to the line $\{\mathrm{Re}(z) = x_0 + \frac{1}{2}\}$, both of these isometric circles are geodesics of the form used to construct the Dirichlet polygon centered at z_0.

(2) The converse of Theorem 3.1 is false. The symmetrical nature of P implies a certain symmetry in the quotient space $\mathbb{H}^2/\Gamma$, namely that the surface admits an orientation-reversing involution of order 2. This is not the case for a generic punctured sphere.

4. Double Dirichlet domains

We now suppose that the same fundamental domain P is obtained when we construct the Dirichlet domains P_0 and P_1 centered at z_0 and $z_1 \in \mathbb{H}^2$ respectively. For comparison with the previous section, we will assume that we have conjugated Γ in $\mathrm{PSL}_2(\mathbb{R})$ so that the geodesic line L containing z_0 and z_1 is vertical.

Theorem 4.1. *If the Dirichlet domains P_0 and P_1 for Γ, centered at $z_0 \neq z_1 \in \mathbb{H}^2$ respectively, coincide, then the quotient space $\mathbb{H}^2/\Gamma$ is a sphere, with cone points and/or punctures.*

Proof. Much of the work in Section 3 was concerned with showing precisely how the sides of P were identified. This follows relatively swiftly here, once we have cleared up one technical point. We often think of a fundamental domain as a subset of $\mathbb{H}^2$ combined with a set of side pairings identifying its sides. We only assume that the sets P_0 and P_1 are equal, and thus we must make sure that Γ identifies their sides the same way each time.

Lemma 4.2. *If $P = P_0 = P_1$ is the Dirichlet domain centered at z_0 and at z_1, then the sides of P are identified the same way in each case.*

Proof. Suppose, for the sake of contradiction, that this is not the case. Any side of a Dirichlet domain bisects the domain's center and its image under some isometry. Here, we have a side A of P which is the bisector of both the pair z_0 and $\gamma_0^{-1}(z_0)$ and the pair z_1 and $\gamma_1^{-1}(z_1)$, where $\gamma_0 \neq \gamma_1$ are the isometries defining that side of P. It follows that γ_0 pairs A with some side B, and γ_1 pairs A with some other side $C \neq B$. Let $d := d(z_0, z_1)$ be the distance between the two centers z_0 and z_1. Since $\gamma_0^{-1}(z_0)$ and $\gamma_1^{-1}(z_1)$ are the reflections of each in A, we see that

$$d(\gamma_0^{-1}(z_0), \gamma_1^{-1}(z_1)) = d.$$

Applying the isometry γ_1 to both points, this gives that

$$d(\gamma_1(\gamma_0^{-1}(z_0)), z_1) = d.$$

Now, if $\gamma_1(\gamma_0^{-1}(z_0)) = z_0$, then the isometries γ_0 and γ_1 both send $\gamma_0^{-1}(z_0)$ to z_0 and $\gamma_1^{-1}(z_1)$ to z_1. Since they also both preserve orientation, this implies that $\gamma_0 = \gamma_1$, which is a contradiction. Thus $\gamma_1(\gamma_0^{-1}(z_0)) \neq z_0$. But then $\gamma_1(\gamma_0^{-1}(z_0))$ is a point in the orbit of z_0, and thus the construction of P_0 involves the half-space $\{x \in \mathbb{H}^2 \mid d(x, z_0) \leq d(x, \gamma_1(\gamma_0^{-1}(z_0)))\}$. As we saw above,

$$d(\gamma_1(\gamma_0^{-1}(z_0)), z_1) = d(z_0, z_1) = d.$$

Hence z_1 is equidistant from z_0 and $\gamma_1(\gamma_0^{-1}(z_0))$. Thus z_1 cannot be in the interior of P_0, contradicting the assumption that $P_0 = P_1$. □

The following result will allow us to conclude the proof of Theorem 4.1.

Lemma 4.3. *Each side of P (and each point of ∂P) is identified with its reflection in the line L.*

Proof. Given a point $v \in \partial P$, v is sent to a point of ∂P the same distance away from z_0. Put another way, v is sent somewhere on the hyperbolic circle of center z_0 and radius $d(v, z_0)$. But v is also sent to a point on the hyperbolic circle of center z_1 and radius $d(v, z_1)$. Thus we see a picture similar to Figure 1, except instead of a horizontal line, we have a second circle, centered vertically above or below z_0. These two circles intersect only at v if $v \in L$, and at v and v^*, the reflection of v in L, if $v \notin L$. If $v \in L$ then v is necessarily an elliptic fixed point and a vertex of P, and the two sides adjacent to v are identified with one another. If $v \notin L$, it suffices to show that v cannot be fixed by a side pairing, and thus must be identified with v^*. Let v be a nonvertex point, and γ the side pairing associated to the side containing v in its interior. Since v is not a vertex, it cannot be an elliptic fixed point, and so γ must identify v with v^*. From this, it follows that each vertex is also identified with its reflection. □

So we now know that our domain P has the same symmetrical property that we saw DF domains possess. If the line $L \cap \mathring{P}$ extends vertically to ∞, then the argument from the proof of Theorem 3.1 applies directly, and we are done. If the line terminates at a boundary point of P, then we observe that the two sides adjacent to this vertex are identified symmetrically, creating a cone point instead of a cusp. This creates a circle awaiting identification as in the proof of Theorem 3.1, and the rest of the argument applies from there. □

Remarks. (1) The first remark at the end of Section 3 applies here as well. That is, if we take any point $z \in L \cap \mathring{P}$ as our Dirichlet center, we will obtain the Dirichlet domain P. Thus we see that a Fuchsian group which admits a DF domain is simply one that admits Dirichlet domain with a line of centers and a cusp on the line of symmetry.

(2) The same discussion can also be used to show that these are the only Dirichlet centers giving rise to P. The Dirichlet center must be equidistant from a point of ∂P and its destination under its side pairing; in this setup, the locus of such points is always precisely L. Thus, it is impossible to find a Fuchsian group with a triangle of Dirichlet centers all giving rise to the same domain.

5. Reflection groups

The goal of this section is to prove the main theorem. As a corollary, we will show that given the signature of any sphere which can be obtained as a quotient of $\mathbb{H}^2$, then we may exhibit a Fuchsian group Γ that admits a double Dirichlet domain (and a DF domain if there is at least one puncture) and gives rise to a quotient space of the given signature.

We first recall the following results regarding reflection groups (see [Ratcliffe 2006], Section 7.1).

Theorem 5.1. *Let G be a discrete reflection group with respect to the polygon Q. Then all the angles of Q are submultiples of π, and if g_S and g_T are reflections in the adjacent sides S and T of Q with $\theta(S, T) = \pi/k$, then $g_S \circ g_T$ has order k.*

Theorem 5.2. *Let Q be a finite sided convex hyperbolic polygon of finite volume, all of whose angles are submultiples of π. Then the group G, generated by reflections of $\mathbb{H}^2$ in the sides of Q, is a discrete reflection group with respect to the polygon Q.*

We will appeal to these results, as well as to the results of Sections 3 and 4, in the following discussion.

Theorem 5.3. *If the finitely generated, orientation-preserving, finite coarea Fuchsian group Γ admits a double Dirichlet domain, or a DF domain, P, then Γ is an index 2 subgroup of the discrete group G of reflections in a hyperbolic polygon Q.*

Proof. Suppose first that Γ admits a DF domain P. We know that P has reflectional symmetry about a vertical axis L which bisects P. Since P is a fundamental domain for Γ, the side pairings of P generate Γ. Each side pairing, with the exception of the parabolic element pairing the vertical sides, has the form $\sigma_L \circ \sigma_i$, where σ_L denotes reflection in L and σ_i is reflection in the isometric circle S_i, $1 \leq i \leq m$, where S_i contains a side of P. Furthermore, since each side is paired with its mirror image in L, it suffices to consider the σ_i corresponding to sides in one half of P. The parabolic side pairing can be written $\sigma_L \circ \sigma_K$, where σ_K is reflection in K, the vertical side of P in the same half as the S_i. Thus we have a generating set for Γ of the form

$$\{\sigma_L \circ \sigma_1, \ldots, \sigma_L \circ \sigma_m, \sigma_L \circ \sigma_K\}$$

for some $m \in \mathbb{N}$. Consider the group G obtained by adding the reflection σ_L to this generating set. The set becomes

$$\{\sigma_L, \sigma_L \circ \sigma_1, \ldots, \sigma_L \circ \sigma_m, \sigma_L \circ \sigma_K\}$$

and because $\sigma_L = \sigma_L^{-1}$ has order 2, it follows that we can replace the generator $\sigma_L \circ \sigma_i$ with the element σ_i and still have a generating set. The generating set

$$\{\sigma_L, \sigma_1, \ldots, \sigma_m, \sigma_K\}$$

is precisely the set of reflections in the sides of a polygon Q with sides on K, L and S_i, $1 \leq i \leq m$. To prove that all of the angles of Q are submultiples of π, it suffices to observe that the vertices of P are paired symmetrically, and that the Poincaré polyhedron theorem gives that the sum of the angles in each cycle is $2\pi/s$, for $s \in \mathbb{N}$. Now Theorem 5.2 allows us to reach the desired conclusion.

To prove the result for the case where $L \cap \mathring{P}$ does not extend to $\partial\mathbb{H}^2$, we simply observe that, in this case, every side pairing generator of Γ can be written $\sigma_L \circ \sigma_i$, since here there are no vertical sides. Instead of the cusp at ∞ we have another finite vertex of P, but since this vertex lies on the line L, it must also be an elliptic fixed point, and the paragraph above applies. □

We now turn to the converse of Theorem 5.3.

Theorem 5.4. *If G is a discrete group of reflections in a polygon $Q \subset \mathbb{H}^2$, then G contains an index 2 subgroup of orientation-preserving isometries that admits a double Dirichlet domain* (*and a DF domain if Q has an ideal vertex at ∞*).

Proof. Let Q be such a polygon. If necessary, rotate Q so that one of its sides is vertical. Call this side L. By Theorem 5.1, all angles of Q are submultiples of π. Denote by σ_L reflection in the vertical side L of Q. If Q has another vertical side (and hence an ideal vertex at ∞), call this side K and denote reflection in K by σ_K. Denote by σ_i reflection in the (nonvertical) line S_i containing a side of Q. By definition, these reflections constitute a generating set for G. Let $\Gamma < G$ be the subgroup generated by elements of the form $\sigma_2 \circ \sigma_1$ where σ_1 and σ_2 are reflections in the generating set for G. Then Γ is a Fuchsian group. Since $\sigma_L \notin \Gamma$, we see that the set $P := Q \cup \sigma_L Q$ is contained within a fundamental domain for Γ. We will show that P is itself a fundamental domain for Γ.

To see this, denote by $T_i := \sigma_L(S_i)$ the geodesic obtained by reflecting S_i in L. Then T_i contains a side of P. Also denote $\sigma_L(K)$ by M. Then K is paired with M by the element $\sigma_L \circ \sigma_K \in \Gamma$, and S_i is paired with T_i by $\sigma_L \circ \sigma_i$. Thus the sides of P are paired by generators of Γ. To see that these side pairings generate Γ themselves, consider a generating element $\sigma_2 \circ \sigma_1 \in \Gamma$. We may write

$$\sigma_2 \circ \sigma_1 = \sigma_2 \circ (\sigma_L \circ \sigma_L) \circ \sigma_1 = (\sigma_2 \circ \sigma_L) \circ (\sigma_L \circ \sigma_1) = (\sigma_L \circ \sigma_2)^{-1} \circ (\sigma_L \circ \sigma_1),$$

which shows that together, the elements $\sigma_L \circ \sigma_i$ and $\sigma_L \circ \sigma_K$ generate Γ. We therefore have that Γ has index 2 in G, and that P is a fundamental domain for Γ.

To see that Γ admits a fundamental domain of the required type, it will suffice to check that P is one. Let z_0 be any point on the line L which lies in the interior of P. If there is a second vertical side K, then it is the line bisecting z_0 and $\sigma_K(z_0)$, so $\sigma_L(K) = M$ bisects $\sigma_L(z_0) = z_0$ and $\sigma_L(\sigma_K(z_0))$. Thus M is a line of the form found in the definition on a Dirichlet domain centered at z_0. A similar argument applied to $(\sigma_L \circ \sigma_K)^{-1} = \sigma_K \circ \sigma_L$ shows that K is also such a line. Now S_i is the line bisecting z_0 and $\sigma_i(z_0)$, so $\sigma_L(S_i) = T_i$ bisects $\sigma_L(z_0) = z_0$ and $\sigma_L(\sigma_i(z_0))$. This shows that T_i is a line of the form found in the definition on a Dirichlet domain centered at z_0. A similar argument shows that the same is true of S_i, and thus we see that P must contain a Dirichlet domain centered at z_0. But we know that P is itself a fundamental domain for Γ, so that P is a Dirichlet domain for any center $z_0 \in L \cap \mathring{P}$.

If there is a second vertical side K, we must also check that P is a Ford domain. S_i is the isometric circle of the generator $\sigma_L \circ \sigma_i$, and $T_i = \sigma_L(S_i)$ is the isometric circle of the inverse element. Since $\sigma_L \circ \sigma_K$ pairs the two vertical sides of P and generates Γ_∞, it follows that P must contain a Ford domain for Γ. But P is itself a fundamental domain, so this Ford domain cannot be a proper subset, and hence is equal to P. □

We now show that Fuchsian groups with this symmetrical property, though they necessarily have genus zero, have no other restrictions on their signature.

Corollary 5.5. *Given the signature* $(0; n_1, \ldots, n_t; m)$ *of a* (*nontrivial*, *hyperbolic*) *sphere with* $m \geq 0$ *punctures and* $t \geq 0$ *cone points of orders* $n_i \in \mathbb{N}$, *for* $1 \leq i \leq t$, *there exists a Fuchsian group* Γ *such that* Γ *admits a double Dirichlet domain* (*and a DF domain if* $m > 0$) *and* $\mathbb{H}^2/\Gamma$ *is a sphere of the given signature.*

Proof. Suppose $m > 0$. Construct Q by placing one vertex at ∞, t vertices in $\mathbb{H}^2$ of angles π/n_i ($n_i \geq 2$) for $1 \leq i \leq t$, and $m-1$ ideal vertices in $\mathbb{R}$. If $m = 0$, construct a compact t-gon with angles $\pi/n_1, \ldots, \pi/n_t$. Now let G be the group of hyperbolic isometries generated by reflections in the sides of Q. By Theorem 5.4, Γ admits a DF domain (or double Dirichlet domain if $m = 0$) $P = Q \cup \sigma Q$, where σ denotes reflection in one of the vertical sides L of Q. The symmetrical identifications, combined with the Poincaré polyhedron theorem, give that the quotient surface has the required signature. □

Remark. If $m > 0$ above, then there is a certain amount of freedom in our choice of the polygon Q. For example, we do not necessarily have to place one of the ideal vertices of Q at ∞. We do so in order to ensure that we obtain a DF domain for Γ. Instead, we could have all of the ideal vertices lie in $\mathbb{R}$, thereby placing the line of symmetry L away from any of the ideal vertices. Similarly, if $m > 1$, we could construct Q so that L meets only one of the m ideal vertices, instead of 2 in the construction above. We also do not have to construct Q so that each angle is bisected by a vertical line; we only do so in order to demonstrate that it is possible to find the required polygon.

6. A noncongruence arithmetic maximal reflection group

In this section, we will prove explicitly that there exists a noncongruence arithmetic maximal hyperbolic reflection group. Recall that a noncocompact hyperbolic reflection group $\Gamma_{\text{ref}} < \text{Isom}(\mathbb{H}^2)$ is called *arithmetic* if and only if it is commensurable with $\text{PSL}_2(\mathbb{Z})$. Such a group is then called *congruence* if it contains some principal congruence subgroup

$$\Gamma(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a \equiv d \equiv \pm 1, b \equiv c \equiv 0 \text{ mod } n \right\} \subset \text{PSL}_2(\mathbb{Z}).$$

Consider the group $\Gamma < \text{PSL}_2(\mathbb{R})$ generated by the matrices

$$\gamma_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \gamma_2 = \begin{pmatrix} 0 & \frac{-1}{\sqrt{11}} \\ \sqrt{11} & 0 \end{pmatrix}, \quad \gamma_3 = \begin{pmatrix} \sqrt{11} & \frac{5}{\sqrt{11}} \\ 2\sqrt{11} & \sqrt{11} \end{pmatrix},$$

$$\gamma_4 = \begin{pmatrix} 10 & 3 \\ 33 & 10 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 23 & 8 \\ 66 & 23 \end{pmatrix}.$$

We first wish to show that Γ is discrete. Consider the group

$$\Gamma_0(11) = \left\{ \begin{pmatrix} a & b \\ 11c & d \end{pmatrix} \;\middle|\; a, b, c, d \in \mathbb{Z},\ ad - 11bc = 1 \right\} \subset \text{PSL}_2(\mathbb{Z}).$$

It is well known that the normalizer $N(\Gamma_0(11))$ of $\Gamma_0(11)$ in $\text{PSL}_2(\mathbb{R})$ is a (maximal arithmetic) Fuchsian group generated by $\Gamma_0(11)$ and

$$\begin{pmatrix} 0 & -\frac{1}{\sqrt{11}} \\ \sqrt{11} & 0 \end{pmatrix},$$

which is $\gamma_2 \in \Gamma$ [Maclachlan 1981; Long et al. 2006]. We see then that

$$\begin{pmatrix} 0 & -\frac{1}{\sqrt{11}} \\ \sqrt{11} & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -11 & -5 \end{pmatrix} = \begin{pmatrix} \sqrt{11} & \frac{5}{\sqrt{11}} \\ 2\sqrt{11} & \sqrt{11} \end{pmatrix} = \gamma_3 \in \Gamma,$$

and since $\gamma_1, \gamma_4, \gamma_5 \in \Gamma_0(11)$, we have that $\Gamma < N(\Gamma_0(11))$, and so Γ is discrete. We next wish to construct a Ford domain for Γ. In Figure 2 we see the isometric circles corresponding to the generators listed above and their inverses.

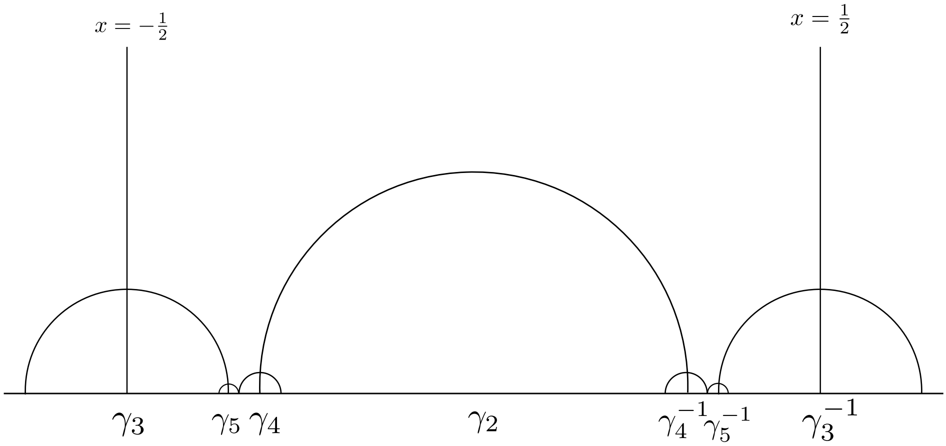


Figure 2. A Ford domain for Γ.

The claim is that this polygon is in fact the required Ford domain. To see this, observe that each generator γ_i can be decomposed into the product of two reflections $\gamma_i = \sigma_L \circ \sigma_i$, where σ_i is reflection in the isometric circle S_i of γ_i, σ_1 is reflection in the line $x = -\frac{1}{2}$, and σ_L is reflection in the line $x = 0$. Thus the elements of the generating set for Γ pair the sides of P, and each pushes $\mathring{P}$ completely off itself. This shows that P is a fundamental domain for Γ; by its construction, it is a Ford domain.

Thus we see that the quotient space $\mathbb{H}^2/\Gamma$ is a sphere of signature $(0; 2, 2, 2, 2; 2)$ and area 4π. Further, P is a DF domain, as each of these generators pairs one side S_i of P with its reflection $\sigma_L(S_i)$ in the line $x = 0$. Thus, by Theorem 5.3, we see that Γ is the index 2 orientation-preserving subgroup of the group Γ_{ref} of reflections in a hyperbolic hexagon Q with angles $(0, \frac{\pi}{2}, \frac{\pi}{2}, 0, \frac{\pi}{2}, \frac{\pi}{2})$. The claim is that this hyperbolic reflection group Γ_{ref} is arithmetic, maximal (as an arithmetic reflection group), and noncongruence.

Claim 1. Γ_{ref} *is arithmetic.*

Proof. Note that the finite area of $\mathbb{H}^2/\Gamma$ implies that the index $[N(\Gamma_0(11)) : \Gamma]$ is finite. Hence, since $N(\Gamma_0(11))$ is arithmetic, we see that Γ is also arithmetic, from which it follows that Γ_{ref} is arithmetic. □

Claim 2. Γ_{ref} *is a maximal reflection group.*

Proof. If Γ_{ref} were not maximal, it would be properly contained in another reflection group H_{ref}, which is therefore also arithmetic. Let $H < H_{\text{ref}}$ denote the orientation-preserving index 2 subgroup. Note that then we have $\Gamma < H$. Since Γ and H are both arithmetic Fuchsian groups of genus zero, they are contained in a common maximal, arithmetic, genus zero Fuchsian group M from the appropriate list in [Long et al. 2006]. As we saw above, Γ is contained in the normalizer $N(\Gamma_0(11))$, and by area considerations we find that $[N(\Gamma_0(11)) : \Gamma] = 2$. Further, Γ cannot be contained in any other of these maximal arithmetic groups; to see this, observe that if $n \neq 11$ then, if we pick some nonzero integer b coprime to n, we may find integers a, d such that $\left(\begin{smallmatrix} a & b \\ n & d \end{smallmatrix}\right) \in \Gamma_0(n)$. We then have

$$\gamma_2 \begin{pmatrix} a & b \\ n & d \end{pmatrix} \gamma_2 = \begin{pmatrix} 0 & -\frac{1}{\sqrt{11}} \\ \sqrt{11} & 0 \end{pmatrix} \begin{pmatrix} a & b \\ n & d \end{pmatrix} \begin{pmatrix} 0 & -\frac{1}{\sqrt{11}} \\ \sqrt{11} & 0 \end{pmatrix} = \begin{pmatrix} -d & \frac{n}{11} \\ 11b & -a \end{pmatrix}.$$

We wish to show that this does not belong to $\Gamma_0(n)$. If n is not divisible by 11 this is clear, so suppose $n \geq 22$ is a multiple of 11. Then, by construction, b is coprime to 11, and so $11b$ is not divisible by n. This shows that γ_2 cannot belong to any normalizer $N(\Gamma_0(n))$ except $N(\Gamma_0(11))$.

It remains to verify that we cannot have $H = M = N(\Gamma_0(11))$. Construction of the Ford domain for $N(\Gamma_0(11))$ yields the generating set

$$\delta_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \delta_2 = \begin{pmatrix} 0 & \frac{-1}{\sqrt{11}} \\ \sqrt{11} & 0 \end{pmatrix}, \quad \delta_3 = \begin{pmatrix} \sqrt{11} & \frac{5}{\sqrt{11}} \\ 2\sqrt{11} & \sqrt{11} \end{pmatrix},$$

$$\delta_4 = \begin{pmatrix} \sqrt{11} & \frac{-4}{\sqrt{11}} \\ 3\sqrt{11} & -\sqrt{11} \end{pmatrix}, \quad \delta_4' = \begin{pmatrix} -\sqrt{11} & \frac{-4}{\sqrt{11}} \\ 3\sqrt{11} & \sqrt{11} \end{pmatrix}.$$

The Ford domain corresponding to these generators is shown in Figure 3.

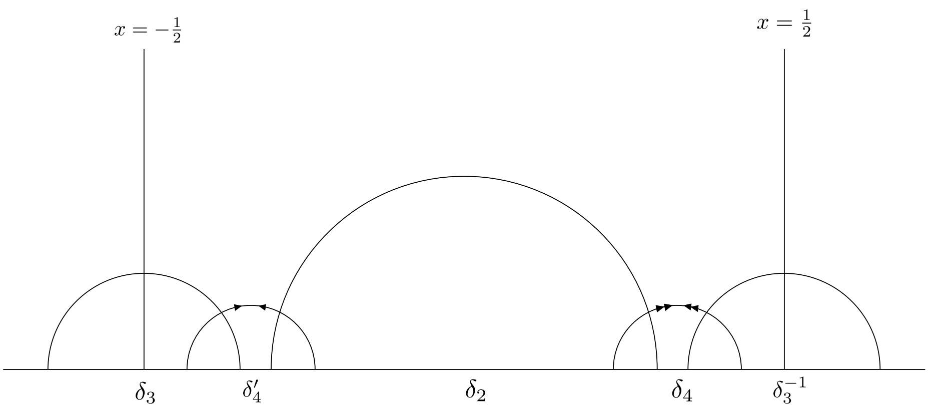


Figure 3. A Ford domain for $N(\Gamma_0(11))$. All sides except those marked are paired with their opposites.

The fact that three of the generating elements are involutions, which pair adjacent sides of the Ford domain, precludes $N(\Gamma_0(11))$ from possessing a DF domain. By Theorem 5.4, this also precludes it from being an index 2 subgroup of a reflection group. Thus Γ_{ref} is maximal. □

Remark. Since Γ_{ref} is an arithmetic maximal hyperbolic reflection group, one would expect to find it in existing lists of such groups. This example appears to be the lattice 2-fill$(L_{26.1})$ in Allcock's enumeration [2010] of rank-3 reflective Lorentzian lattices, which would correspond to the case $N = 26$ in [Nikulin 2000, Table 1]. If one could show Γ_{ref} is indeed this lattice, this would provide an alternative proof that it is maximal arithmetic; however, we omit this at present, as the proofs given above suffice for our purposes.

Claim 3. Γ_{ref} *is not congruence.*

Proof. Suppose Γ_{ref} is congruence. Then it contains some principal congruence subgroup $\Gamma(n)$. These groups all belong to the modular group, so $\Gamma(n)$ is contained in $G = \Gamma \cap \mathrm{PSL}_2(\mathbb{Z})$. By Wohlfahrt's theorem (see [Newman 1972], p. 149), G contains $\Gamma(n)$ for n equal to the level of G, that is, the smallest natural number such that G contains the normal closure of

$$T^n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$$

in $\mathrm{PSL}_2(\mathbb{Z})$.

Subclaim. *The level of G is* 11.

Proof of subclaim. The group $G = \Gamma \cap \mathrm{PSL}_2(\mathbb{Z}) = \Gamma \cap \Gamma_0(11)$ is not equal to Γ, by the presence of the nonintegral elements γ_2 and γ_3. However, it contains the matrices

$$\beta_1 = \gamma_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \beta_2 = \gamma_2\gamma_1^{-1}\gamma_2 = \begin{pmatrix} 1 & 0 \\ 11 & 1 \end{pmatrix}, \quad \beta_3 = \gamma_2\gamma_3 = \begin{pmatrix} -2 & -1 \\ 11 & 5 \end{pmatrix},$$

$$\beta_4 = \gamma_2\gamma_3^{-1} = \begin{pmatrix} 2 & -1 \\ 11 & -5 \end{pmatrix}, \quad \beta_5 = \gamma_4 = \begin{pmatrix} 10 & 3 \\ 33 & 10 \end{pmatrix}, \quad \beta_6 = \gamma_5 = \begin{pmatrix} 23 & 8 \\ 66 & 23 \end{pmatrix}.$$

The isometric circles of these elements and their inverses are shown in Figure 4. Notice that the isometric circles centered at $\frac{2}{11}$ and $-\frac{2}{11}$ are paired with those at $\frac{5}{11}$ and $-\frac{5}{11}$ respectively; with these four circles excepted, each other side is paired with its reflection in the line $x = 0$. There are four equivalence classes of ideal points: these classes are $\{\infty\}$, $\{0\}$, $\left\{\frac{1}{3}, -\frac{1}{3}\right\}$, $\left\{\frac{4}{11}, \frac{3}{11}, -\frac{3}{11}, -\frac{4}{11}\right\}$. All four finite vertices belong to the same cycle, and their angles are $\pi/3$ at $x = \pm\frac{1}{2}$, and $2\pi/3$ at $x = \pm\frac{3}{22}$, giving angle sum 2π.

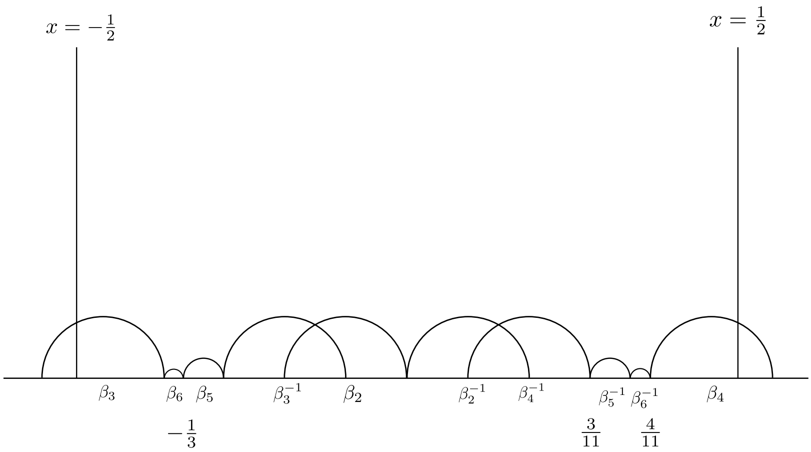


Figure 4. A Ford domain for G.

The region P_G of $\mathbb{H}^2$ bounded by these circles and the lines $x=-\frac{1}{2}$ and $x=\frac{1}{2}$ has area 8π and contains a Ford domain for G. This is enough for us to conclude that it is a Ford domain for G: since G is a proper subgroup of Γ, of finite index due to the finite area of P_G, G must have coarea a multiple $4m\pi$ of 4π, where $m=[\Gamma:G]>1$. That the area of P_G is 8π tells us that $m\le 2$, and hence that in fact $m=2$. So G has index 2 in Γ and index 24 in $\mathrm{PSL}_2(\mathbb{Z})$, and the list above is a generating set.

To prove the Subclaim, we need to show that given any $\varphi\in\mathrm{PSL}_2(\mathbb{Z})$, we have that $\varphi T^{11}\varphi^{-1}\in G$. If φ fixes ∞ this is clear, so suppose $\varphi(\infty)\neq\infty$. Topologically, $\mathbb{H}^2/G$ is a torus with four cusps, with the cusp orbits in $\mathbb{Q}\cup\{\infty\}$ represented by 0, ∞, $\frac{1}{3}$ and $\frac{3}{11}$. Therefore $\varphi(\infty)$ is G-equivalent to exactly one of these four points; let $g\in G$ be such that $g^{-1}\varphi(\infty)$ is this point. We observe that $T^{11}\in G$; we also find that

$$\begin{pmatrix}0&-1\\1&0\end{pmatrix}\begin{pmatrix}1&11\\0&1\end{pmatrix}\begin{pmatrix}0&-1\\1&0\end{pmatrix}=\begin{pmatrix}1&0\\-11&1\end{pmatrix}=\begin{pmatrix}1&0\\11&1\end{pmatrix}^{-1}\in G$$

is a parabolic element fixing 0, that

$$\begin{pmatrix}1&0\\3&1\end{pmatrix}\begin{pmatrix}1&11\\0&1\end{pmatrix}\begin{pmatrix}-1&0\\3&-1\end{pmatrix}=\begin{pmatrix}-32&11\\-99&34\end{pmatrix}=\begin{pmatrix}10&3\\33&10\end{pmatrix}\begin{pmatrix}-23&8\\66&-23\end{pmatrix}\in G$$

is a parabolic element fixing $\frac{1}{3}$, and that

$$\begin{aligned}\begin{pmatrix}3&1\\11&4\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}-4&1\\11&-3\end{pmatrix}&=\begin{pmatrix}-32&9\\-121&34\end{pmatrix}\\&=\begin{pmatrix}2&-1\\11&-5\end{pmatrix}\begin{pmatrix}23&8\\66&23\end{pmatrix}\begin{pmatrix}-5&-1\\11&-2\end{pmatrix}\begin{pmatrix}-10&3\\33&-10\end{pmatrix}\in G\end{aligned}$$

is a parabolic element fixing $\frac{3}{11}$. Note that in this last case, G also contains a conjugate of T^{11} fixing $\frac{3}{11}$, by taking the 11th power of the given element. Thus there exists a conjugate α of T^{11} such that $\alpha \in G$ and α fixes $g^{-1}\varphi(\infty)$. The element $g.\alpha.g^{-1} \in G$ is therefore a parabolic element, conjugate in $\mathrm{PSL}_2(\mathbb{Z})$ to T^{11}, with parabolic fixed point at $\varphi(\infty)$. We wish to show that $g.\alpha.g^{-1} = \varphi T^{11}\varphi^{-1}$. Since the former element is known to be a conjugate of T^{11}, we may alternatively write it as $\psi T^{11}\psi^{-1}$ for some $\psi \in \mathrm{PSL}_2(\mathbb{Z})$ with $\psi(\infty) = \varphi(\infty)$. Now $\psi^{-1}\varphi \in \mathrm{PSL}_2(\mathbb{Z})$ fixes ∞ and so must be a power of T; in particular, $\psi^{-1}\varphi$ commutes with T. It follows that $\psi^{-1}\varphi T^{11}\varphi^{-1}\psi = T^{11}$ and therefore

$$g.\alpha.g^{-1} = \psi T^{11}\psi^{-1} = \varphi T^{11}\varphi^{-1}$$

as required. Thus G contains all elements of the form $\varphi T^{11}\varphi^{-1}$, and so the level of G is at most 11. To see that it is not smaller, observe that G does not contain any element of the form $\left(\begin{smallmatrix} 1 & 0 \\ t & 1 \end{smallmatrix}\right)$ for $t = 1, 2, \ldots, 10$. □

To complete the proof of Claim 3, we note that by the Subclaim, G must contain $\Gamma(11)$. Computation in GAP Version 4.4.12 [GAP 2008] reveals that the core of G in $\mathrm{PSL}_2(\mathbb{Z})$, the largest normal subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ contained in G, has index $k = 1351680 = 2^{13} \cdot 3 \cdot 5 \cdot 11$ in $\mathrm{PSL}_2(\mathbb{Z})$. Thus G cannot contain a normal subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ of index (in $\mathrm{PSL}_2(\mathbb{Z})$) smaller than this constant. But all principal congruence subgroups are normal, and $[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma(11)] = 660 < k$. From this contradiction we conclude that Γ_{ref} is not congruence. □

Remark. Hsu [1996] gives a congruence test which can be applied to G. Since G has index 24 in $\mathrm{PSL}_2(\mathbb{Z})$, we obtain a representation in the symmetric group S_{24}. After expressing the known generators for G in terms of $L = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $R = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, we find

$$L = (2\ 4\ 9\ 15\ 8\ 5\ 11\ 13\ 7\ 3\ 6)\ (10\ 17\ 21\ 23\ 22\ 19\ 14\ 12\ 18\ 20\ 16)$$

and

$$R = (1\ 2\ 5\ 12\ 14\ 7\ 4\ 10\ 16\ 8\ 3)\ (9\ 17\ 19\ 13\ 11\ 18\ 21\ 24\ 22\ 20\ 15)$$

are both of order 11, also giving that the level of G is 11. Hsu's test is then that G is a congruence subgroup if and only if $(R^2L^{-\frac{1}{2}})^3 = 1$, where $\frac{1}{2}$ is the multiplicative inverse of 2 mod 11, in this case equal to 6. Here we find that R^2L^{-6} has order 6, and so G is noncongruence.

Remark. The example given above is not the only arithmetic maximal reflection group which is not congruence. Of the 23 noncocompact arithmetic maximal reflection groups which belong to $\mathrm{Isom}(\mathbb{H}^2)$, eight are not congruence. Furthermore, there exists a noncongruence example in $\mathrm{Isom}(\mathbb{H}^3)$. Further details will appear in the author's Ph.D. thesis.

7. Kleinian groups and DF domains

In this section, it will be shown that only one direction (the analogue of 5.4) of the main theorem holds when we consider Kleinian groups in the place of Fuchsian groups. This is because the added dimension gives new possibilities for the shape of the domains in question; in particular, they no longer have to glue up in a completely symmetrical way, although some symmetry remains. Examples will be given to demonstrate this flexibility, which extends as far as having nontrivial cuspidal cohomology. The discussion will be restricted to DF domains; as the above work demonstrates, it is not unreasonable to suppose that double Dirichlet domains share many similar properties.

Theorem 7.1. *Let $Q \subset \mathbb{H}^3$ be a finite sided, convex hyperbolic polyhedron with all dihedral angles integer submultiples of π, and let G be the discrete group of reflections in Q. Then G contains an index 2 Kleinian subgroup that admits a double Dirichlet domain (and a DF domain if Q has an ideal vertex).*

Proof. Suppose that Q is placed in upper half-space $\mathbb{H}^3$ such that one of its faces L is contained in a vertical plane. Let

$$G = \langle \tau_1, \ldots, \tau_m, \tau_L \rangle$$

be a generating set for G. Let

$$\Gamma = \langle \tau_L \circ \tau_1, \ldots, \tau_L \circ \tau_m \rangle$$

be the index 2 subgroup. Let $P = Q \cup \tau_L Q$. Let $w_0 = x_0 + y_0 i + z_0 j \in \mathring{L}$, for $z_0 > 0$. The claim is that w_0 is a Dirichlet center for Γ. Fix a generator $\gamma_i = \tau_L \circ \tau_i$. Then the plane P_i fixed by τ_i bisects w_0 and $\tau_i(w_0)$, and so $\tau_L(P_i)$, which by construction is a face of P, bisects w_0 and $\gamma_i(w_0)$. $\square$

The next result provides the first counterexamples of Theorem 5.3 by exhibiting Kleinian groups that admit DF domains and do not have index 2 in a reflection group.

Proposition 7.2. *Let Q be an all right hyperbolic polyhedron, with a vertex at ∞, and all vertices ideal. Let G be the group of reflections in Q. Then G contains a subgroup of index 4 that admits a DF domain.*

Proof. Since Q is all right, the link of each vertex is a rectangle. Rotate Q in $\mathbb{H}^3$ so that the four faces adjacent to the vertex at ∞, each lie above x- or y-lines in $\mathbb{C}$, where x-lines are parallel to the real axis, and y-lines are parallel to the imaginary axis. Let H be a side above a y-line, V a side above an x-line, and τ_H and τ_V the respective reflections. Let $P = (Q \cup \tau_H Q) \cup \tau_V(Q \cup \tau_H Q)$. Then P is the union of 4 copies of Q. Looking down from ∞ on the floor of P, label by A the nonvertical face adjacent to the top left vertex and to the vertical face opposite H. Label any nonvertical faces adjacent to this face B. Proceed to label every nonvertical face A or B, with no two adjacent faces sharing the same label. The symmetry of P implies that this labeling is symmetric in both the x- and y-directions. Define the subgroup Γ as follows. Given a nonvertical side P_i of P, if P_i has label A, let the element $\tau_H \circ \tau_i$ belong to Γ; if P_i has label B, let $\tau_V \circ \tau_i$ belong to Γ. If H' is the face of Q opposite H, and V' the face of Q opposite V, let $\tau_H \circ \tau_{H'}$ and $\tau_V \circ \tau_{V'}$ belong to Γ. Then P is a DF domain for Γ. $\square$

Remark. Given a group Γ constructed as in the above proof, note that Γ is not an index 2 subgroup of the group of reflections in the polyhedron $(Q \cup \tau_H Q)$. This is because the reflection τ_H will be absent from this group, preventing the construction of elements of Γ of the form $\tau_H \circ \tau_i$. The same is valid for the group of reflections in the polyhedron $(Q \cup \tau_V Q)$.

Since there is no direct analogue of Theorem 5.3 for Kleinian groups, the question arises as to what, if anything, is implied about a Kleinian group by it having a DF domain. For example, one might ask whether such groups must have trivial cuspidal cohomology. The following example gives a Kleinian group which admits a DF domain, but which has nontrivial cuspidal cohomology; that is, there exists a nonperipheral homology class of infinite order in the first homology of the quotient space.

Example. Let $\Gamma < \mathrm{PSL}_2(\mathbb{C})$ be generated by the matrices

$$\begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5i \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & \frac{-1}{\sqrt{2}} \\ \sqrt{2} & 0 \end{pmatrix}, \quad \begin{pmatrix} -\sqrt{2} & \frac{i}{\sqrt{2}} \\ -i\sqrt{2} & -\sqrt{2} \end{pmatrix},$$

and

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{-1}{\sqrt{2}} \\ \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & \bar{a} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2}a & \sqrt{2}a\bar{a} - \frac{1}{\sqrt{2}} \\ \sqrt{2} & \sqrt{2}\bar{a} \end{pmatrix}$$

for each $a \in \{1, 2, 1+i, 2+i, 2i, 1+2i, 2+2i, 1-i, 2-i, -2i, 1-2i, 2-2i\}$, where $\bar{a}$ is the complex conjugate of a. Then the isometric spheres of these matrices have centers at the Gaussian integers $\{x + iy \mid x, y \in \mathbb{Z}\}$ and radius $1/\sqrt{2}$. The square with vertices at $\pm\frac{5}{2} \pm \frac{5}{2}i$ is a Dirichlet domain for the action of Γ_∞. Let P be the intersection of the exterior of all these isometric spheres with the chimney above the given rectangle. Then P is a DF domain for Γ, with Dirichlet center any point of $\mathring{P}$ above 0. Every dihedral angle of P is $\pi/2$. The quotient space $\mathbb{H}^3/\Gamma$ has 14 boundary components; the cusp at ∞ gives a boundary torus, and each of the 13 cusp cycles in $\mathbb{C}$ gives a $(2, 2, 2, 2)$ or a $(2, 4, 4)$ sphere. Thus the peripheral homology has rank 1. Computation using GAP gives that $H_1(\mathbb{H}^3/\Gamma)$ has rank 2, so there is infinite nonperipheral homology.

Remarks. (1) The cuspidal cohomology of this example has rank 1, but it can be modified to give examples where this rank is arbitrarily high.

(2) This example is arithmetic. To see this, consider the Picard group $\mathrm{PSL}_2(\mathbb{O}_1)$. This group contains as a finite index subgroup the congruence subgroup $\Gamma_0(2)$, where the lower left entry is a member of the ideal in $\mathbb{O}_1$ generated by 2. This congruence subgroup is normalized by the element

$$\gamma = \begin{pmatrix} 0 & \frac{-1}{\sqrt{2}} \\ \sqrt{2} & 0 \end{pmatrix}$$

and, since γ has order 2, $\Gamma_0(2)$ is an index 2 subgroup of the group H obtained by adding γ. The group Γ given in the example is a subgroup of H, of finite index due to the finite volume of the DF domain. In turn, H is commensurable with $\mathrm{PSL}_2(\mathbb{O}_1)$, as both share $\Gamma_0(2)$ as a subgroup of finite index.

(3) The quotient space of $\mathbb{H}^3$ by this group is not a manifold, so one can thus ask whether there exists another example which has nontrivial cuspidal cohomology, and which is additionally torsion free.

Although there does not appear to be a specific condition for a Kleinian group which is equivalent to having a DF domain, we can say something about a group that admits a DF domain. We cannot always decompose an orientation-preserving isometry of $\mathbb{H}^3$ into the composition of two reflections, but Carathéodory [1954] shows that we need at most four. If $\gamma \notin \Gamma_\infty$, these can be taken to be $\gamma = \gamma_4 \circ \gamma_3 \circ \gamma_2 \circ \gamma_1$, where γ_1 is reflection in the isometric sphere S_γ, γ_2 in the vertical plane R_γ bisecting S_γ and $S_{\gamma^{-1}}$, and $\gamma_4 \circ \gamma_3$ is rotation around the vertical axis through the north pole of $S_{\gamma^{-1}}$.

Theorem 7.3. *Suppose the Kleinian group Γ admits a DF domain P. Then the planes R_γ, for side pairings $\gamma \in \Gamma \setminus \Gamma_\infty$ of P, all intersect in a vertical axis. Furthermore, for each such γ, $\gamma_4 \circ \gamma_3 = 1$, and so each element of the corresponding generating set for Γ has real trace.*

Proof. Let P be a Ford domain. Suppose there is some side pairing γ such that $\gamma_4 \circ \gamma_3 \neq 1$. By considering the north pole of S_γ and its image, the north pole of $S_{\gamma^{-1}}$, we see that if P were a Dirichlet domain, its center w_0 would have to be in the plane R_γ. But given any such choice of w_0, one can find a point $w \in P \cap S_\gamma$ such that $d(w_0, w) \neq d(w_0, \gamma(w))$. Thus P is not a Dirichlet domain. Since each $\gamma \in \Gamma \setminus \Gamma_\infty$ is then simply the composition of two reflections, it is the conjugate in $\mathrm{PSL}_2(\mathbb{C})$ of an element of $\mathrm{PSL}_2(\mathbb{R})$. It thus has real trace. Since it is assumed that any element of Γ_∞ is parabolic, these too have real trace.

Next suppose that the planes R_γ do not have a common intersection. Since we know that $\gamma_4 \circ \gamma_3 = 1$, for a given γ, the plane R_γ represents the set of potential

Dirichlet centers. If there is no common such center, P is not a Dirichlet domain. Thus if P is a DF domain, the planes R_γ have a common intersection. $\square$

The examples given earlier in this section give a flavor of the particular case with only two distinct, perpendicular planes R_γ. It is therefore possible for DF domains to be more complicated than this. This theorem provides a useful criterion for having a DF domain, which can be used to check known Ford domains. Observe that the vertical axis of intersection of the planes R_γ must correspond to a Dirichlet center for the action of Γ_∞. Thus we see that the figure-8 knot group [Riley 1975], as well as the Whitehead link group and the group of the Borromean rings [Wielenberg 1978] do not admit DF domains. Furthermore, the groups obtained from a standard Ford domain in [Wielenberg 1981] cannot admit DF domains. Although in some cases, with the right choice of Ford domain, one can generate congruence subgroups of Bianchi groups using elements of real trace, the sides of the domain are identified in a way similar to the corresponding Fuchsian congruence subgroup, and so these groups seldom admit a DF domain.

Acknowledgments

I would like to express my gratitude to my advisor, Alan Reid, for his tremendous support, encouragement and guidance. I wish also to thank Daniel Allcock, Misha Belolipetsky and Hossein Namazi for additional valuable conversation and correspondence, and the referee for helpful comments and suggestions.

References

[Agol et al. 2008] I. Agol, M. Belolipetsky, P. Storm, and K. Whyte, "Finiteness of arithmetic hyperbolic reflection groups", *Groups, Geometry, and Dynamics* **2**:4 (2008), 481–498. MR 2009m:20054 Zbl 1194.22011

[Allcock 2010] D. Allcock, "The reflective Lorentzian lattices of rank 3", preprint, 2010. To appear in *Mem. Amer. Math. Soc.* arXiv 1010.0486

[Beardon 1983] A. F. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics **91**, Springer, New York, 1983. MR 85d:22026 Zbl 0528.30001

[Belolipetsky 2011] M. Belolipetsky, "Finiteness theorems for congruence reflection groups", *Transform. Groups* **16**:4 (2011), 939–954. arXiv 1008.0934

[Carathéodory 1954] C. Carathéodory, *Theory of functions of a complex variable, 1*, Chelsea, New York, 1954. MR 15,612b Zbl 0055.30301

[Díaz and Ushijima 2009] R. Díaz and A. Ushijima, "On the properness of some algebraic equations appearing in Fuchsian groups", *Topology Proceedings* **33** (2009), 81–106. MR 2010i:30056 Zbl 1179.30039

[Ford 1951] L. R. Ford, *Automorphic functions*, Chelsea, 1951. Zbl 55.0810.04

[GAP 2008] *GAP - groups, algorithms, and programming*, Version 4.4.12, The GAP Group, St. Andrews, Fife, 2008, Available at http://www.gap-system.org.

[Greenberg 1977] L. Greenberg, "Finiteness theorems for Fuchsian and Kleinian groups", pp. 199–257 in *Discrete groups and automorphic functions* (Cambridge, 1975), edited by W. J. Harvey, Academic Press, London, 1977. MR 58 #28483

[Grunewald and Schwermer 1981] F. Grunewald and J. Schwermer, "A nonvanishing theorem for the cuspidal cohomology of SL_2 over imaginary quadratic integers", *Math. Ann.* **258**:2 (1981), 183–200. MR 83c:10032 Zbl 0458.20041

[Hsu 1996] T. Hsu, "Identifying congruence subgroups of the modular group", *Proc. Amer. Math. Soc.* **124**:5 (1996), 1351–1359. MR 96k:20100 Zbl 0846.20051

[Long et al. 2006] D. D. Long, C. Maclachlan, and A. W. Reid, "Arithmetic Fuchsian groups of genus zero", *Pure Appl. Math. Q.* **2**:2 (2006), 569–599. MR 2007f:20092 Zbl 1107.20037

[Maclachlan 1981] C. Maclachlan, "Groups of units of zero ternary quadratic forms", *Proc. Roy. Soc. Edinburgh Sect. A* **88**:1-2 (1981), 141–157. MR 83a:10030 Zbl 0467.10018

[Maskit 1988] B. Maskit, *Kleinian groups*, Grundlehren der Mathematischen Wissenschaften **287**, Springer, Berlin, 1988. MR 90a:30132 Zbl 0627.30039

[Newman 1972] M. Newman, *Integral matrices*, Pure and Applied Mathematics **45**, Academic Press, New York, 1972. MR 49 #5038 Zbl 0254.15009

[Nikulin 2000] V. V. Nikulin, "О классификации гиперболических систем корней ранга три", *Tr. Mat. Inst. Steklova* **230** (2000), 3–255. Translated as "On the classification of hyperbolic root systems of rank three" in *Proc. Steklov Inst. Math.* **230**:3 (2000), 1–241. MR 2001m:17034 Zbl 0997.17014

[Ratcliffe 2006] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*, 2nd ed., Graduate Texts in Mathematics **149**, Springer, New York, 2006. MR 2007d:57029 Zbl 1106.51009

[Riley 1975] R. Riley, "A quadratic parabolic group", *Math. Proc. Cambridge Philos. Soc.* **77**:2 (1975), 281–288. MR 54 #542 Zbl 0309.55002

[Wielenberg 1978] N. J. Wielenberg, "The structure of certain subgroups of the Picard group", *Math. Proc. Cambridge Philos. Soc.* **84**:3 (1978), 427–436. MR 80b:57010 Zbl 0399.57005

[Wielenberg 1981] N. J. Wielenberg, "Hyperbolic 3-manifolds which share a fundamental polyhedron", pp. 505–513 in *Riemann surfaces and related topics* (Stony Brook, NY, 1978), edited by I. Kra and B. Maskit, Ann. of Math. Stud. **97**, Princeton University Press, Princeton, NJ, 1981. MR 82i:57012 Zbl 0465.30039

Received March 8, 2011.

GRANT S. LAKELAND
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
AUSTIN, TX 78712
UNITED STATES
glakeland@math.utexas.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

FORMAL EQUIVALENCE OF POISSON STRUCTURES AROUND POISSON SUBMANIFOLDS

IOAN MĂRCUȚ

Let (M, π) be a Poisson manifold. A Poisson submanifold $P \subset M$ gives rise to a Lie algebroid $A_P \to P$. Formal deformations of π around P are controlled by certain cohomology groups associated to A_P. Assuming that these groups vanish, we prove that π is formally rigid around P; that is, any other Poisson structure on M, with the same first-order jet along P, is formally Poisson diffeomorphic to π. When P is a symplectic leaf, we find a list of criteria that are sufficient for these cohomological obstructions to vanish. In particular, we obtain a formal version of the normal form theorem for Poisson manifolds around symplectic leaves.

1. Introduction

A *Poisson bracket* on a manifold M is a Poisson algebra structure on the space of smooth functions on M, that is, a Lie bracket $\{\cdot, \cdot\}$ on $C^\infty(M)$ satisfying the derivation property

$$\{f, gh\} = \{f, g\}h + \{f, h\}g \quad \text{for all } f, g, h \in C^\infty(M). \tag{1}$$

Equivalently, it can be given by a bivector $\pi \in \mathfrak{X}^2(M)$ that satisfies $[\pi, \pi] = 0$. The two definitions are related by the formula

$$\langle \pi, df \wedge dg \rangle = \{f, g\} \quad \text{for all } f, g \in C^\infty(M).$$

An immersed submanifold $\iota : P \to M$ is called a *Poisson submanifold* of M if π is tangent to P. This ensures that $\pi_{|P}$ is a Poisson structure on P for which the restriction map

$$\iota^* : C^\infty(M) \to C^\infty(P)$$

is a Lie algebra homomorphism. We regard the Poisson algebra $(C^\infty(P), \{\cdot, \cdot\})$ as the 0th-order approximation of the Poisson structure on M. If P is embedded,

This research was supported by the NWO Vidi Project “Poisson topology.”
MSC2010: primary 53D17, 58H15; secondary 70K45, 17B55, 17B70.
Keywords: Poisson geometry, Lie algebroid, graded Lie algebra.

then P is a Poisson submanifold if and only if its vanishing ideal

$$I(P) = \{f \in C^\infty(M) \mid \iota^*(f) = 0\}$$

is an ideal in the Lie algebra $(C^\infty(M), \{\cdot,\cdot\})$. Assuming that P is also closed[1], we have a canonical identification of Poisson algebras

$$(C^\infty(P), \{\cdot,\cdot\}) = (C^\infty(M)/I(P), \{\cdot,\cdot\}).$$

This gives a recipe for constructing higher-order approximations. For example, the first-order approximation fits into an exact sequence of Poisson algebras

$$\begin{aligned}(2)\quad 0 \to (I(P)/I^2(P), \{\cdot,\cdot\})& \\ &\to (C^\infty(M)/I^2(P), \{\cdot,\cdot\}) \to (C^\infty(P), \{\cdot,\cdot\}) \to 0.\end{aligned}$$

The Poisson algebra structures in this sequence depend only on $j^1_{|P}\pi$, the first jet of π along P. A better way to describe (2) is using the language of Lie algebroids. As explained in Section 2, the extension (2) gives rise to a Lie algebroid structure A_P on $T^*_P M$ that fits into a short exact sequence of Lie algebroids

$$0 \to TP^\circ \to A_P \to T^*P \to 0, \tag{3}$$

where $TP^\circ \subset T^*_P M = A_P$ is the annihilator of TP and T^*P is the cotangent Lie algebroid of $(P, \{\cdot,\cdot\})$. In particular, we obtain a representation of A_P on TP°, and thus also on its symmetric powers $\mathcal{S}^k(TP^\circ)$.

We study formal rigidity of Poisson structures around Poisson submanifolds. In general, deformation and rigidity problems in Poisson geometry are controlled by the Poisson cohomology groups $H^\bullet_\pi(M)$, which are the cohomology of the complex of multivector fields $(\mathfrak{X}^\bullet(M), d_\pi)$, where

$$d_\pi := [\pi, \cdot\,].$$

For a Poisson submanifold P, this d_π induces a differential on $\mathfrak{X}^\bullet(M)_{|P}$, the complex of multivector fields along M. The corresponding cohomology, denoted by $H^\bullet_\pi(M, P)$, is called the Poisson cohomology relative to P [Ginzburg and Lu 1992]. The formal rigidity of Poisson structures around Poisson submanifolds is controlled by a version of this cohomology with coefficients. Lie algebroids provide the right setting to make this precise; that is, the relative Poisson cohomology groups can be computed as the cohomology of the Lie algebroid A_P

$$H^\bullet_\pi(M, P) = H^\bullet(A_P),$$

[1]Since we study local properties of (M, π) around P, only the condition that P is embedded is essential; closeness can be achieved by replacing M with a tubular neighborhood of P.

and the cohomology groups of A_P with coefficients in $\mathcal{S}^k(TP^\circ)$, which we denote by $H^\bullet(A_P;\mathcal{S}^k(TP^\circ))$, control formal rigidity (see Section 2 for the definition of Lie algebroid cohomology).

Our main result is the following:

Theorem 1.1. *Let π_1 and π_2 be two Poisson structures on M, such that $P \subset M$ is an embedded Poisson submanifold for both, and such that they have the same first-order jet along P. If their common algebroid A_P has the property that*

$$H^2(A_P;\mathcal{S}^k(TP^\circ)) = 0 \quad \textit{for all } k \geq 2,$$

then the two structures are formally Poisson diffeomorphic. More precisely, there exists a diffeomorphism

$$\psi : \mathcal{U} \to \mathcal{V},$$

with $d\psi_{|T_PM} = \mathrm{id}_{T_PM}$, where $\mathcal{U}$ and $\mathcal{V}$ are open neighborhoods of P, such that $\pi_{1|\mathcal{U}}$ and $\psi^*(\pi_{2|\mathcal{V}})$ have the same infinite jet along P:

$$j^\infty_{|P}(\pi_{1|\mathcal{U}}) = j^\infty_{|P}(\psi^*(\pi_{2|\mathcal{V}})).$$

Applying Theorem 1.1 to the linear Poisson structure on the dual of a compact, semisimple Lie algebra, we obtain the following result.

Corollary 1.2. *Let $\mathfrak{g}$ be a semisimple Lie algebra of compact type and consider π_{lin} the linear Poisson structure on $\mathfrak{g}^*$. Let $\mathbb{S}(\mathfrak{g}) \subset \mathfrak{g}^*$ be the sphere in $\mathfrak{g}^*$ centered at* 0, *of radius* 1 *with respect to some invariant inner product. Then $\mathbb{S}(\mathfrak{g})$ is a Poisson submanifold, and any Poisson structure π_1 defined in some open neighborhood of* $\mathbb{S}(\mathfrak{g})$, *such that*

$$j^1_{|\mathbb{S}(\mathfrak{g})}(\pi_{\mathrm{lin}}) = j^1_{|\mathbb{S}(\mathfrak{g})}(\pi_1),$$

is formally Poisson diffeomorphic to π_{lin}.

The *symplectic leaves* of (M,π) are Poisson submanifolds of a special type. Recall that a Poisson manifold carries a canonical singular foliation whose leaves are the maximal integral submanifolds of the distribution $\pi^\sharp(T^*M)$. Such a leaf S has a natural symplectic structure given by $\omega_S := \pi_{|S}^{-1}$. If $(S,\omega_S) \subset (M,\pi)$ is an embedded symplectic leaf, then the Lie algebroid extension (3) — which encodes only the first-order jet π along S — can be used to construct a second Poisson structure π^1_S, called the *first-order approximation* of π around S, defined on some open neighborhood of S and having the same first jet as π along S.

In [Crainic and Mărcuţ 2010] we obtained a normal form theorem for Poisson structures around symplectic leaves: we proved that, under some assumptions on the first jet of π along S, the Poisson structures π and π^1_S are Poisson diffeomorphic around S. Our goal is to give a formal version of this result, which we state below in its most general form (observe that it is a direct consequence of Theorem 1.1).

Theorem 1.3. *Let (M, π) be a Poisson manifold and $S \subset M$ an embedded symplectic leaf. If the cohomology groups*

$$H^2(A_S, \mathcal{S}^k(TS^\circ))$$

vanish for all $k \geq 2$, then π is formally Poisson diffeomorphic to its first-order approximation around S.

In many cases we show that these cohomological obstructions vanish, and we obtain the following corollaries.

Corollary 1.4. *Let (M, π) be a Poisson manifold and $S \subset M$ an embedded symplectic leaf. Assume that the Poisson homotopy cover of S is a smooth principal bundle with vanishing second de Rham cohomology group, and that its structure group G satisfies*

$$H^2_{\mathrm{diff}}(G, \mathcal{S}^k(\mathfrak{g})) = 0 \quad \textit{for all } k \geq 2,$$

where $\mathfrak{g}$ is the Lie algebra of G, and $H^\bullet_{\mathrm{diff}}(G, \mathcal{S}^k(\mathfrak{g}))$ denotes the differentiable cohomology of G with coefficients in the k-th symmetric power of the adjoint representation. Then π is formally Poisson diffeomorphic to its first-order approximation around S.

Since the differentiable cohomology of compact groups vanishes, we obtain the following immediate corollary.

Corollary 1.5. *Let (M, π) be a Poisson manifold and $S \subset M$ an embedded symplectic leaf. If the Poisson homotopy cover of S is a smooth principal bundle with vanishing second de Rham cohomology group and compact structure group, then π is formally Poisson diffeomorphic to its first-order approximation around S.*

The next consequence is bit more technical:

Corollary 1.6. *Let (M, π) be a Poisson manifold and $S \subset M$ an embedded symplectic leaf whose isotropy Lie algebra is reductive. If the abelianization algebroid*

$$A_S^{\mathrm{ab}} := A_S/[TS^\circ, TS^\circ]$$

is integrable by a simply connected principal bundle with compact structure group and vanishing second de Rham cohomology group, then π is formally Poisson diffeomorphic to its first-order approximation around S.

Corollary 1.7. *Let (M, π) be a Poisson manifold and $S \subset M$ an embedded symplectic leaf through $x \in M$. If the isotropy Lie algebra at x is semisimple, $\pi_1(S, x)$ is finite, and $\pi_2(S, x)$ is torsion, then π is formally Poisson diffeomorphic to its first-order approximation around S.*

Some related results. The first-order approximation of a Poisson manifold (M,π) around a one-point leaf x (a zero of π) is the linear Lie–Poisson structure on $\mathfrak{g}_x^*$, the dual of the isotropy Lie algebra at x. Formal linearization in this setup was proven by Weinstein [1983] for semisimple $\mathfrak{g}_x$. This case is also covered by our Corollary 1.7. Under the stronger assumption that $\mathfrak{g}_x$ is semisimple of compact type, Conn [1985] proved that a neighborhood of x is in fact Poisson diffeomorphic to an open neighborhood of 0 in the local model $\mathfrak{g}_x^*$.

Vorobjev [2001] constructed the first-order approximation around arbitrary symplectic leaves (see [Crainic and Mărcuţ 2010] for a more geometrical approach).

A weaker version of our Theorem 1.1 — of which we became aware only at the end of this research — was stated by Itskov et al. [1998]. They work around compact symplectic leaves instead of embedded Poisson submanifolds, proving that for each k, there exists a diffeomorphism that identifies the Poisson structures up to order k [Itskov et al. 1998, Theorem 7.1]. Compactness of the leaf is nevertheless too strong an assumption for formal equivalence. For example, they conclude in their Corollary 7.4 that hypotheses similar to those in our Corollary 1.7 imply the vanishing of the cohomology groups $H^2(A_S,\mathcal{S}^k(TS^\circ))$, but also remark that compactness of the leaf is incompatible with these assumptions (it forces S to be a point).

To prove Theorem 1.1, we reduce it to a result on the equivalence of Maurer–Cartan elements in complete graded Lie algebras, which we prove in the Appendix. The same criteria for equivalence of Maurer–Cartan elements, but in the context of differential graded algebras, can be found in [Abad et al. 2010, Appendix A].

To prove the vanishing of the cohomological obstructions, and the corollaries listed above, we use techniques such as Whitehead's Lemma for semisimple Lie algebras and spectral sequences for Lie algebroids, but also the more powerful techniques developed in [Crainic 2003], such as the Van Est map and vanishing of cohomology of proper groupoids.

Theorem [Crainic and Mărcuţ 2010, main result]. *Let (M,π) be a Poisson manifold and $S\subset M$ an embedded symplectic leaf*; *π is Poisson diffeomorphic to its first-order approximation around S if the following conditions are satisfied*:

- *the Poisson homotopy cover P of S is smooth*;
- $H^2_{\mathrm{dR}}(P)=0$;
- *the structure group of P is compact*;
- *S is compact.*

The first three conditions are the hypotheses of Corollary 1.5. So, giving up on compactness of the leaf, we still conclude that π and its first-order approximation are *formally* Poisson diffeomorphic. Nevertheless, the conditions of Corollary 1.5

are too strong in the formal setting; they force the semisimple part of the isotropy Lie algebra to be compact. Thus we consider the more technical Corollary 1.6 to be the correct analog in the formal category of the normal form theorem in [Crainic and Mărcuţ 2010]. In fact, Corollary 1.5 is a consequence of Corollary 1.6; it is precisely the case when the semisimple part of the isotropy Lie algebra is compact.

2. The first-order data

We recall some definitions; for more on Lie algebroids, see [Mackenzie 1987].

Definitions 2.1. A *Lie algebroid* over a manifold B is a vector bundle $\mathcal{A} \to B$ endowed with a Lie bracket $[\cdot,\cdot]$ on its space of sections $\Gamma(\mathcal{A})$ and a vector bundle map $\rho : \mathcal{A} \to TB$, called the *anchor*, which satisfy the Leibniz identity:

$$[\alpha, f\beta] = f[\alpha,\beta] + L_{\rho(\alpha)}(f)\beta \quad \text{for all } f \in C^\infty(B),\ \alpha,\beta \in \Gamma(\mathcal{A}).$$

A *representation* of $\mathcal{A}$ is a vector bundle $E \to B$ endowed with a bilinear map

$$\nabla : \Gamma(\mathcal{A}) \times \Gamma(E) \to \Gamma(E),$$

satisfying

$$\nabla_{f\alpha}(s) = f\nabla_\alpha(s), \quad \nabla_\alpha(fs) = f\nabla_\alpha(s) + L_{\rho(\alpha)}(f)s,$$

and the flatness condition

$$\nabla_\alpha\nabla_\beta(s) - \nabla_\beta\nabla_\alpha(s) = \nabla_{[\alpha,\beta]}(s).$$

The *cohomology of a Lie algebroid* $(\mathcal{A},[\cdot,\cdot],\rho)$ *with coefficients in a representation* (E,∇) is defined by the complex $\Omega^\bullet(\mathcal{A},E) := \Gamma(\Lambda^\bullet\mathcal{A}^*\otimes E)$ with differential given by the classical Koszul formula:

$$\begin{aligned} d_\nabla\omega(\alpha_0,\dots,\alpha_q) = &\sum_i (-1)^i \nabla_{\alpha_i}\big(\omega(\alpha_1,\dots,\widehat{\alpha}_i,\dots,\alpha_q)\big) \\ &+ \sum_{i<j}(-1)^{i+j}\omega\big([\alpha_i,\alpha_j],\dots,\widehat{\alpha}_i,\dots,\widehat{\alpha}_j,\dots,\alpha_q\big). \end{aligned}$$

The corresponding cohomology groups are denoted by $H^\bullet(\mathcal{A},E)$.

To a Poisson manifold (M,π) one can associate a Lie algebroid structure on the cotangent bundle T^*M, with anchor given by π viewed as a bundle map

$$\pi^\sharp : T^*M \to TM$$

and bracket uniquely determined by

$$[df, dg] := d\{f,g\} \quad \text{for all } f,g \in C^\infty(M);$$

see [Vaisman 1994] for details.

Let $P \subset M$ be an embedded Poisson submanifold. Since π is tangent to P, it is easy to see that the algebroid structure can be restricted to P, in the sense that there is a unique Lie algebroid structure on $A_P := T^*_P M$ with anchor $\pi^\sharp_{|P}$ and bracket such that the restriction map $\Gamma(T^*M) \to \Gamma(A_P)$ is a Lie algebra homomorphism. The dual of the inclusion $TP \subset T_P M$ gives a map $A_P \to T^*P$ that is a Lie algebroid homomorphism, where T^*P is the cotangent Lie algebroid of $(P, \pi_{|P})$. This way we obtain the extension of Lie algebroids from the introduction:

$$0 \to (TP^\circ, [\cdot\,,\cdot]) \to (A_P, [\cdot\,,\cdot]) \to (T^*P, [\cdot\,,\cdot]) \to 0. \tag{4}$$

This short exact sequence implies that TP° is an ideal in $(A_P, [\cdot\,,\cdot])$; therefore

$$\nabla : \Gamma(A_P) \times \Gamma(TP^\circ) \to \Gamma(TP^\circ), \quad \nabla_\alpha(\eta) := [\alpha, \eta]$$

defines a representation of A_P on TP°, and thus on its symmetric powers $\mathcal{S}^k(TP^\circ)$. The resulting cohomology groups are the obstructions appearing in Theorems 1.1 and 1.3. The Lie algebroid structures on A_P and the sequence (4) depend only on the first jet of π along P (that is, the brackets and anchors can be expressed in terms of $\pi_{|P}$ and the first-order derivatives of π restricted to P).

Remark 2.2. We regard the Lie algebroid A_P as the first-order approximation of the Poisson bracket at P. To justify this interpretation, fix a Poisson structure π_P on P, where $P \subset M$ is a closed embedded submanifold. Then there is a one-to-one correspondence between Poisson algebra structures on the commutative algebra $C^\infty(M)/I^2(P)$, which fit into the short exact sequence

$$\begin{aligned} 0 \to (I(P)/I^2(P), \{\cdot\,,\cdot\}) \qquad\qquad\qquad\qquad\qquad\qquad \\ \to (C^\infty(M)/I^2(P), \{\cdot\,,\cdot\}) \to (C^\infty(P), \{\cdot\,,\cdot\}) \to 0, \end{aligned} \tag{5}$$

and Lie algebroid structures on $A_P := T^*_P M$, which fit into a sequence of the form

$$0 \to (TP^\circ, [\cdot\,,\cdot]) \to (A_P, [\cdot\,,\cdot]) \to (T^*P, [\cdot\,,\cdot]) \to 0. \tag{6}$$

The exterior derivative induces a map

$$d : C^\infty(M)/I^2(P) \to \Gamma(A_P),$$

and the correspondence between the brackets is uniquely determined by the fact that this is a Lie algebra homomorphism.

Example 2.3. Consider $P := \mathbb{R}^2$ as the submanifold $\{z = 0\} \subset M := \mathbb{R}^3$. We construct a first-order extension of the trivial Poisson structure on P to M, that is, a Poisson algebra structure on the commutative algebra

$$C^\infty(M)/I^2(P) = C^\infty(M)/(z^2) = C^\infty(P) \oplus zC^\infty(P)$$

with the property that $\{f, g\} \in (z)$, for all $f, g \in C^\infty(M)/(z^2)$. Explicitly, define

$$\{f, g\} = z\Big(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} + x\frac{\partial f}{\partial x}\frac{\partial g}{\partial z} - x\frac{\partial f}{\partial z}\frac{\partial g}{\partial x}\Big) \mod (z^2).$$

A straightforward computation yields that $\{\cdot,\cdot\}$ satisfies the Jacobi identity, and therefore we have an extension of Poisson algebras

$$0 \to zC^\infty(P) \to C^\infty(P) \oplus zC^\infty(P) \to C^\infty(P) \to 0,$$

where the Poisson bracket on P is zero. The total space of the corresponding Lie algebroid A_P is $\mathbb{R}^3 \times P \to P$. The bracket is given on the global frame $dx_{|P}$, $dy_{|P}$, $dz_{|P}$ by

$$[dx_{|P}, dy_{|P}] = dz_{|P}, \quad [dy_{|P}, dz_{|P}] = 0, \quad [dx_{|P}, dz_{|P}] = x dz_{|P},$$

and extended bilinearly to all sections, since the anchor is zero.

Nevertheless, there is no Poisson structure on M (nor on any open neighborhood of P) that has this Poisson algebra as its first-order approximation. Assume, to the contrary, that on some open neighborhood $\mathcal{U}$ of P such a Poisson structure exists. Then it must have the form

$$\{x, y\} = z + z^2 h, \quad \{y, z\} = z^2 k, \quad \{x, z\} = xz + z^2 l,$$

for some smooth functions h, k, l defined on $\mathcal{U}$. Computing the Jacobiator of x, y, and z, we obtain

$$J = \{x, \{y, z\}\} + \{z, \{x, y\}\} + \{y, \{z, x\}\} = z^2\big((2-x)k(x, y, 0) + 1\big) + z^3 a(x, y, z),$$

where a is a smooth function. In particular, we see that J cannot vanish, since

$$\frac{\partial^2 J}{\partial z^2}(2, y, 0) = 2 \neq 0.$$

This example shows that not everything that looks like the first jet of a Poisson structure around P (that is, an extension of the form (6) or (5)) comes from an actual Poisson structure.

On the other hand, if P is a symplectic manifold, the situation changes for the better; every "first jet" of a Poisson structure can be extended to a Poisson structure around P. More precisely, consider (S, ω_S) a symplectic manifold, with $S \subset M$ embedded, and an algebroid structure on $A_S := T^*_S M$ that fits into the exact sequence

$$0 \to TS^\circ \to A_S \to T^*S \to 0.$$

Then, using a tubular neighborhood $\mathscr{E} : T_S M / TS \to M$, one can construct a Poisson structure $\pi^1_S = \pi^1_S(A_S, \omega_S, \mathscr{E})$ on some open neighborhood of S, from which we recover the first-order data: it has (S, ω_S) as a symplectic leaf, and the algebroid structure induced on $T^*_S M$ is A_S. This Poisson structure was first constructed by Vorobjev [2001]; we also recommend [Crainic and Mărcuţ 2010] for some different approaches. Applied to different tubular neighborhoods, this construction produces Poisson structures which, when restricted to small enough neighborhoods of S, are Poisson diffeomorphic [Vorobjev 2001]. So the isomorphism class of the germ around S of π^1_S doesn't depend on $\mathscr{E}$.

We can view the whole story from a different perspective; start with a Poisson structure π on M, for which (S, ω_S) is an embedded symplectic leaf, and denote as usual by A_S the Lie algebroid on $T^*_S M$. For $\mathscr{E}$ a tubular neighborhood of S, we call $\pi^1_S = \pi^1_S(A_S, \omega_S, \mathscr{E})$ the *first-order approximation* of π around S. The first-order approximation is defined on some open neighborhood of S in M, and it plays the role of a local normal form for π around S.

3. The formal equivalence theorem

The algebra of formal vector fields. Take the graded Lie algebra $(\mathfrak{X}^\bullet(M), [\,\cdot\,,\cdot\,])$ of multivector fields on M, with the Nijenhuis–Schouten bracket and $\deg(W) = k-1$ for $W \in \mathfrak{X}^k(M)$. For a closed, embedded submanifold $P \subset M$, denote by $\mathfrak{X}^\bullet_P(M)$ the following subalgebra of multivector fields tangent to P:

$$\mathfrak{X}^\bullet_P(M) := \{u \in \mathfrak{X}^\bullet(M) \mid u_{|P} \in \mathfrak{X}^\bullet(P)\}.$$

The vanishing ideal $I(P) \subset C^\infty(M)$ of P induces a filtration $\mathscr{F}$ on $\mathfrak{X}^\bullet_P(M)$:

$$\mathfrak{X}^\bullet_P(M) \supset \mathscr{F}^\bullet_0 \supset \mathscr{F}^\bullet_1 \supset \ldots \mathscr{F}^\bullet_k \supset \mathscr{F}^\bullet_{k+1} \supset \ldots,$$

$$\mathscr{F}^\bullet_k = I^{k+1}(P)\mathfrak{X}^\bullet(M), \quad k \geq 0.$$

It is readily checked that

$$[\mathscr{F}_k, \mathscr{F}_l] \subset \mathscr{F}_{k+l}, \quad [\mathfrak{X}^\bullet_P(M), \mathscr{F}_k] \subset \mathscr{F}_k. \tag{7}$$

Let $\hat{\mathfrak{X}}^\bullet_P(M)$ be the completion of $\mathfrak{X}^\bullet_P(M)$ with respect to the filtration $\mathscr{F}$, defined by the projective limit

$$\hat{\mathfrak{X}}^\bullet_P(M) := \varprojlim \mathfrak{X}^\bullet_P(M)/\mathscr{F}^\bullet_k.$$

By (7), it follows that $\hat{\mathfrak{X}}^\bullet_P(M)$ inherits a graded Lie algebra structure, such that, for $k \geq 0$, the natural maps

$$j^k_{|P} : \hat{\mathfrak{X}}^\bullet_P(M) \to \mathfrak{X}^\bullet_P(M)/\mathscr{F}^\bullet_k$$

are Lie algebra homomorphisms. The algebra $(\hat{\mathfrak{X}}^\bullet_P(M), [\,\cdot\,,\cdot\,])$ is called the *algebra of formal multivector fields along P*. Consider also the homomorphism

$$j^\infty_{|P} : \mathfrak{X}^\bullet_P(M) \to \hat{\mathfrak{X}}^\bullet_P(M).$$

From a version of Borel's Theorem (see, for example, [Moerdijk and Reyes 1991]) about the existence of a smooth section with a specified infinite jet along a submanifold, it follows that $j^\infty_{|P}$ is surjective. Observe that $\hat{\mathfrak{X}}^\bullet_P(M)$ inherits a filtration $\hat{\mathcal{F}}$ from $\mathfrak{X}^\bullet_P(M)$, given by

$$\hat{\mathcal{F}}^\bullet_k = j^\infty_{|P}\mathcal{F}^\bullet_k,$$

that satisfies the corresponding equations (7).

The adjoint action of an element $X \in \hat{\mathcal{F}}^1_1$

$$\mathrm{ad}_X : \hat{\mathfrak{X}}^\bullet_P(M) \to \hat{\mathfrak{X}}^\bullet_P(M), \quad \mathrm{ad}_X(Y) := [X, Y]$$

increases the degree of the filtration by 1. Therefore the partial sums

$$\sum_{i=0}^{n} \frac{\mathrm{ad}^i_X}{i!}(Y)$$

are constant modulo $\hat{\mathcal{F}}_k$ for $n \geq k$ and all $Y \in \hat{\mathfrak{X}}^\bullet_P(M)$. This and the completeness of the filtration on $\hat{\mathcal{F}}$ show that the exponential of ad_X

$$e^{\mathrm{ad}_X} : \hat{\mathfrak{X}}^\bullet_P(M) \to \hat{\mathfrak{X}}^\bullet_P(M), \quad e^{\mathrm{ad}_X}(Y) := \sum_{n\geq 0} \frac{\mathrm{ad}^n_X}{n!}(Y)$$

is well-defined. It is readily checked that e^{ad_X} is a graded Lie algebra isomorphism with inverse $e^{-\mathrm{ad}_X}$ and that it preserves the filtration. We need the following geometric interpretation of these isomorphisms.

Lemma 3.1. *For every $X \in \hat{\mathcal{F}}^1_1$, there exists $\psi : M \to M$ a diffeomorphism of M, with $\psi_{|P} = \mathrm{id}_P$ and $d\psi_{|P} = \mathrm{id}_{T_PM}$, such that for every $W \in \mathfrak{X}^\bullet_P(M)$, we have*

$$j^\infty_{|P}(\psi^*(W)) = e^{\mathrm{ad}_X}(j^\infty_{|P}(W)).$$

Proof. By Borel's Theorem, there is a vector field V on M such that $X = j^\infty_{|P}(V)$. We claim that V can be chosen to be complete. Let g be a complete metric on M and let $\phi : M \to [0, 1]$ be a smooth function that satisfies $\phi = 1$ on the set $\{x \mid g_x(V_x, V_x) \leq \frac{1}{2}\}$ and $\phi = 0$ on the set $\{x \mid g_x(V_x, V_x) \geq 1\}$. Since $V_{|P} = 0$, it follows that ϕV has the same germ as V around P, and therefore $j^\infty_{|P}(\phi V) = X$. On the other hand, since ϕV is bounded, it is complete, so replace V by ϕV.

We show that $\psi := \mathrm{Fl}_V$, the flow of V at time 1, satisfies all requirements. Since $j^1_{|P}(V) = 0$, it is clear that $\psi_{|P} = \mathrm{id}_P$ and $d\psi_{|P} = \mathrm{id}_{T_PM}$.

Consider $W \in \mathfrak{X}^\bullet_P(M)$, and denote by $W_s := \mathrm{Fl}^*_{sV}(W)$ the pullback of W by the flow of V at time s. Since W_s satisfies the differential equation $dW_s/ds = [V, W_s]$, a simple computation gives

$$\frac{d}{ds}\Big(\sum_{i=0}^{k}\frac{(-s)^i\,\mathrm{ad}_V^i}{i!}(W_s)\Big)=\frac{(-s)^k\,\mathrm{ad}_V^{k+1}}{k!}(W_s).$$

This shows that the sum

$$\sum_{i=0}^{k}\frac{(-s)^i\,\mathrm{ad}_V^i}{i!}(W_s)$$

modulo $\mathscr{F}_{k+1}$ is independent of s, and therefore

$$W-\sum_{i=0}^{k}\frac{(-1)^i\,\mathrm{ad}_V^i}{i!}(\psi^*(W))\in\mathscr{F}_{k+1}.$$

Applying $j^\infty_{|P}$ to this equation yields

$$j^\infty_{|P}(W)-\sum_{i=0}^{k}\frac{(-1)^i\,\mathrm{ad}_X^i}{i!}j^\infty_{|P}(\psi^*(W))\in\hat{\mathscr{F}}_{k+1},$$

and hence we conclude

$$j^\infty_{|P}(W)=e^{-\mathrm{ad}_X}j^\infty_{|P}(\psi^*(W)).$$ □

The cohomology of the restricted algebroid. Let (M,π) be a Poisson manifold and $P\subset M$ a closed, embedded Poisson submanifold. The cohomologies we are considering are all versions of the Poisson cohomology $H^\bullet_\pi(M)$, computed by the complex $\mathfrak{X}^\bullet(M)$ of multivector fields on M and differential $d_\pi=[\pi,\cdot\,]$. Since P is a Poisson submanifold, we have that $[\pi,I(P)\mathfrak{X}^\bullet(M)]\subset I(P)\mathfrak{X}^\bullet(M)$, and more generally, it follows that $I^k(P)\mathfrak{X}^\bullet(M)$ forms a subcomplex. Taking consecutive quotients, we obtain the complexes

$$\big(I^k(P)\mathfrak{X}^\bullet(M)/I^{k+1}(P)\mathfrak{X}^\bullet(M),\,d^k_\pi\big),$$

with differential d^k_π induced by $[\pi,\cdot\,]$. For $k=0$, we obtain the Poisson cohomology relative to P. Observe that the differential on these complexes depends only on the first jet of π along P, and therefore, following the philosophy of Section 2, it can be described only in terms of the algebroid A_P.

Proposition 3.2. *The following two complexes are isomorphic*:

$$\big(I^k(P)\mathfrak{X}^\bullet(M)/I^{k+1}(P)\mathfrak{X}^\bullet(M),\,d^k_\pi\big)\cong\big(\Omega^\bullet(A_P,\mathcal{S}^k(TP^\circ)),\,d_{\nabla^k}\big)\quad\textit{for all } k\geq 0.$$

Proof. Since the space of sections of TP° is spanned by differentials of elements in $I(P)$, it is easy to see that the map given by

$$\tau_k: I^k(P)\mathfrak{X}^\bullet(M)\to\Omega^\bullet(A_P,\mathcal{S}^k(TP^\circ))=\Gamma(\Lambda^\bullet(T_PM)\otimes\mathcal{S}^k(TP^\circ)),$$

$$\tau_k(f_1 \dots f_k W) = W_{|P} \otimes df_{1|P} \odot \cdots \odot df_{k|P},$$

where $f_1, \dots, f_k \in I(P)$ and $W \in \mathfrak{X}^\bullet(M)$, is well-defined and surjective. Also, its kernel is precisely $I^{k+1}(P)\mathfrak{X}^\bullet(M)$. Hence, it remains to prove that

$$\tau_k([\pi, W]) = d_{\nabla^k}(\tau_k(W)) \quad \text{for all } W \in I^k(P)\mathfrak{X}^\bullet(M). \tag{8}$$

Recall that the algebroid A_P has anchor $\rho = \pi^\sharp_{|P}$ and bracket determined by

$$[d\phi_{|P}, d\psi_{|P}]_P := d\{\phi, \psi\}_{|P} \quad \text{for all } \phi, \psi \in C^\infty(M).$$

Also, for $k = 0$, we have that ∇^0 is given by

$$\nabla^0 : \Gamma(A_P) \times C^\infty(P) \to C^\infty(P), \quad \nabla^0_\eta(h) = L_{\rho(\eta)}(h).$$

Since both differentials d_π and d_{∇^k} act by derivations and ∇^k is obtained by extending ∇^1 by derivations, it suffices to prove (8) for $\phi \in C^\infty(M)$ and $X \in \mathfrak{X}^1(M)$ (with $k = 0$), and for $f \in I(P)$ (with $k = 1$).

Let $\phi \in C^\infty(M)$ and $\eta \in \Gamma(A_P)$. Since π is tangent to P, we obtain

$$\tau_0([\pi, \phi])(\eta) = [\pi, \phi]_{|P}(\eta) = d\phi_{|P}(\pi^\sharp_{|P}(\eta)) = L_{\rho(\eta)}(\tau_0(\phi)) = d_{\nabla^0}(\tau_0(\phi))(\eta).$$

Let $X \in \mathfrak{X}^1(M)$ and $\phi, \psi \in C^\infty(M)$, and define $\eta := d\phi_{|P}$ for $\theta := d\psi_{|P} \in \Gamma(A_P)$. Then

$$\begin{aligned}
\tau_0([\pi, X])(\eta, \theta) &= [\pi, X]_{|P}(d\phi_{|P}, d\psi_{|P}) \\
&= \big(\{X(\phi), \psi\} + \{\phi, X(\psi)\} - X(\{\phi, \psi\})\big)_{|P} \\
&= \pi^\sharp_{|P}(d\phi_{|P})(X_{|P}(d\psi_{|P})) \\
&\qquad - \pi^\sharp_{|P}(d\psi_{|P})(X_{|P}(d\phi_{|P})) - X_{|P}(d\{\phi, \psi\}_{|P}) \\
&= L_{\rho(\eta)}(\tau_0(X)(\theta)) - L_{\rho(\theta)}(\tau_0(X)(\eta)) - \tau_0(X)([\eta, \theta]_P) \\
&= d_{\nabla^0}(\tau_0(X))(\eta, \theta),
\end{aligned}$$

and thus (8) holds for X.

Consider now $f \in I(P)$ and $\eta := d\phi_{|P} \in \Gamma(A_P)$, with $\phi \in C^\infty(M)$. The formula defining τ_k implies that for every $W \in I^k(P)\mathfrak{X}^\bullet(M)$, we have

$$\tau_k(i_{d\phi}(W)) = i_{d\phi_{|P}} \tau_k(W).$$

Using this, the following computation finishes the proof:

$$\begin{aligned}
\tau_1([\pi, f])(\eta) &= \tau_1([\pi, f](d\phi)) = \tau_1(\{\phi, f\}) = d\{\phi, f\}_{|P} \\
&= [\eta, df_{|P}]_P = \nabla^1_\eta(\tau(f)) = d_{\nabla^1}(\tau(f))(\eta).
\end{aligned}$$

□

Proof of Theorem 1.1. By replacing M with a tubular neighborhood of P, we can assume that P is closed in M. Write

$$\gamma := j^{\infty}_{|P}\pi_1, \quad \gamma' := j^{\infty}_{|P}\pi_2 \in \hat{\mathfrak{X}}^2_P(M).$$

By Proposition 3.2, we can recast the hypothesis as

$$[\gamma,\gamma]=0, \quad [\gamma',\gamma']=0, \quad \gamma-\gamma' \in \hat{\mathcal{F}}_1, \quad H^2(\hat{\mathcal{F}}^\bullet_k/\hat{\mathcal{F}}^\bullet_{k+1}, d_\gamma)=0,$$

for all $k \geq 1$, where $d_\gamma := \mathrm{ad}_\gamma$. All these conditions are expressed in terms of the graded Lie algebra $\mathcal{L}^\bullet := \hat{\mathfrak{X}}^{\bullet+1}_P(M)$, with a complete filtration $\hat{\mathcal{F}}$. Theorem A.5 in the Appendix shows that there exists a formal vector field $X \in \hat{\mathcal{F}}^1_1$ such that $\gamma = e^{\mathrm{ad}_X}(\gamma')$. By Lemma 3.1, there exists a diffeomorphism ψ of M, such that $j^{\infty}_{|P}(\psi^*(W)) = e^{\mathrm{ad}_X} j^{\infty}_{|P}(W)$, for all $W \in \mathfrak{X}^\bullet_P(M)$. This concludes the proof, since

$$j^{\infty}_{|P}(\psi^*(\pi_2)) = e^{\mathrm{ad}_X} j^{\infty}_{|P}(\pi_2) = e^{\mathrm{ad}_X}(\gamma') = \gamma = j^{\infty}_{|P}(\pi_1).$$

Existence of Poisson structures with a specified infinite jet. This proof can be applied to obtain a result on existence of Poisson bivectors with a specified infinite jet. Let S be a closed embedded submanifold of M. An element $\hat{\pi} \in \hat{\mathfrak{X}}^2_S(M)$, satisfying $[\hat{\pi},\hat{\pi}]=0$, is called a formal Poisson bivector. Observe that

$$\hat{\pi}_{|S} := \hat{\pi} \mod \hat{\mathcal{F}}_0 \in \mathfrak{X}^2(S)$$

is a Poisson structure on S. We call S a symplectic leaf on $\hat{\pi}$ if $\hat{\pi}_{|S}$ is nondegenerate. Assuming that S is a symplectic leaf of $\hat{\pi}$, by the discussion in Section 2, the first jet of $\hat{\pi}$,

$$j^1_{|S}(\hat{\pi}) = \hat{\pi} \mod \hat{\mathcal{F}}_1,$$

determines a Lie algebroid A_S on $T^*_S M$, and thus can be used to construct a Poisson bivector π^1_S on some open neighborhood $\mathcal{U}$ of S, whose first jet coincides with that of $\hat{\pi}$. If the cohomology groups

$$H^2(A_S; \mathcal{S}^k(TS^\circ))$$

vanish for all $k \geq 2$, then by the proof of Theorem 1.1, there exists a formal vector field $X \in \hat{\mathcal{F}}^1_1$ such that $e^{\mathrm{ad}_X}(j^{\infty}_{|S}\pi^1_S) = \hat{\pi}$. By Lemma 3.1, we find a diffeomorphism $\psi : \mathcal{U} \to \mathcal{U}$ such that

$$j^{\infty}_{|S}(\psi^*(\pi^1_S)) = e^{\mathrm{ad}_X}(j^{\infty}_{|S}\pi^1_S) = \hat{\pi}.$$

Thus $\pi := \psi^*(\pi^1_S)$ gives a Poisson structure defined on an open neighborhood of S whose infinite jet is $\hat{\pi}$. Hence we have proved the following statement.

Corollary 3.3. *Let $\hat{\pi} \in \hat{\mathfrak{X}}^2_S(M)$ be a formal Poisson structure for which S is a symplectic leaf. If for any $k \geq 2$, the algebroid A_S induced by $j^1_{|S}\hat{\pi}$ satisfies*

$$H^2(A_S; \mathcal{S}^k(TS^\circ)) = 0,$$

then there exists a Poisson structure π defined on some open neighborhood of S such that $\hat{\pi} = j^\infty_{|S}\pi$.

4. Proofs of the criteria

Here we explain and prove the corollaries from the Introduction.

Integration of Lie algebroids and differentiable cohomology. We recall some properties of Lie groupoids and Lie algebroids; see [Mackenzie 1987; Moerdijk and Mrčun 2003] for the general theory. A Lie groupoid over a manifold B is denoted by $\mathcal{G}$, the source and target maps by $s, t : \mathcal{G} \to B$, and the unit map by $u : B \to \mathcal{G}$. To a Lie groupoid $\mathcal{G}$ one can associate a Lie algebroid $A(\mathcal{G}) \to B$, which is the infinitesimal counterpart of $\mathcal{G}$. A Lie algebroid $\mathcal{A}$ is called *integrable* if $\mathcal{A} \cong A(\mathcal{G})$ for some Lie groupoid $\mathcal{G}$. The relation between Lie algebroids and Lie groupoids is similar to that between Lie algebras and Lie groups, the most significant difference being that not every Lie algebroid is integrable.

Recall that a *transitive Lie algebroid* is a Lie algebroid $\mathcal{A} \to B$ with surjective anchor. For example, if $S \subset M$ is a symplectic leaf of a Poisson manifold (M, π), then the Lie algebroid A_S is transitive. A Lie groupoid $\mathcal{G}$ is called *transitive* if the map $(s, t) : \mathcal{G} \to M \times M$ is a surjective submersion. The Lie algebroid of a transitive Lie groupoid is transitive. Conversely, if the base B of a transitive Lie algebroid $\mathcal{A}$ is connected, and $\mathcal{A}$ is integrable, then any Lie groupoid $\mathcal{G}$ integrating it is transitive. Every transitive Lie groupoid is a *gauge groupoid*; that is, it is of the form $P \times_G P$, where G is a Lie group and $p : P \to B$ is a principal G-bundle. For P one can take any s-fiber $s^{-1}(x)$ of $\mathcal{G}$ for $x \in B$, and $G := s^{-1}(x) \cap t^{-1}(x)$. We can recover $\mathcal{A}$ from P as follows: as a bundle $\mathcal{A} = TP/G$, the Lie bracket is induced by the identification

$$\Gamma(\mathcal{A}) = \mathfrak{X}(P)^G,$$

and the anchor is given by dp. We will also say, about a principle G-bundle P for which $A \approxeq TP/G$, that it *integrates* A. As for Lie algebras, if a transitive Lie algebroid with connected base is integrable, then, up to isomorphism, there exists a unique 1-connected principal bundle integrating it [Mackenzie 1987].

Let $S \subset M$ be a symplectic leaf of a Poisson manifold (M, π), and assume that the transitive algebroid A_S is integrable. The connected and simply connected principal bundle $P \to S$ for which $P \times_G P$ integrates A_S is called *the Poisson homotopy cover* of S. We say that P is smooth if A_S is integrable; this terminology is justified by the fact that P exists also in the nonintegrable case as a topological principal bundle over S [Crainic and Fernandes 2003].

Let $\mathcal{A}$ be a transitive Lie algebroid with connected base space B, and denote by $\mathfrak{g} \subset \mathcal{A}$ the kernel of the anchor. On each fiber of $\mathfrak{g}$, the Lie bracket restricts to a Lie algebra structure $(\mathfrak{g}_x, [\cdot, \cdot]_x)$, and this Lie algebra is called *the isotropy Lie algebra at* x. In the integrable case, when $\mathcal{A} = A(\mathcal{G})$, the isotropy Lie algebra coincides with the Lie algebra of the *isotropy group* $G_x := s^{-1}(x) \cap t^{-1}(x)$. In the case of a symplectic leaf $S \subset M$ of a Poisson manifold, the kernel of the anchor of the Lie algebroid A_S is given by $\mathfrak{g} := TS^\circ$.

A Lie groupoid $\mathcal{G}$ is called *proper* if $(s, t) : \mathcal{G} \to B \times B$ is a proper map.

A *representation* of a Lie groupoid $\mathcal{G}$ over B is a vector bundle $E \to B$ and a smooth linear action $g : E_x \to E_y$ for every arrow $g : x \to y$ satisfying the obvious identities. A representation E of $\mathcal{G}$ can be differentiated to a representation of its Lie algebroid $A(\mathcal{G})$ on the same vector bundle E. If the s-fibers of $\mathcal{G}$ are connected and simply connected, then every representation of $A(\mathcal{G})$ comes from a representation of $\mathcal{G}$ [Crainic and Fernandes 2003, Proposition 2.2], and in our applications this is usually the case.

The *differentiable cohomology* of a Lie groupoid $\mathcal{G}$ with coefficients in a representation $E \to B$ is computed by the complex $\mathcal{C}^p_{\mathrm{diff}}(\mathcal{G}; E)$ of smooth maps $c : \mathcal{G}^{(p)} \to E$, where

$$\mathcal{G}^{(p)} := \{(g_1, \dots, g_p) \in \mathcal{G}^p \mid s(g_i) = t(g_{i+1}), i = 1, \dots, p-1\}$$

with $c(g_1, \dots, g_p) \in E_{t(g_1)}$, and with differential given by

$$\begin{aligned} dc(g_1, \dots, g_{p+1}) &= g_1 c(g_2, \dots, g_{p+1}) \\ &\quad + \sum_{i=1}^{p} (-1)^i c(g_1, \dots, g_i g_{i+1}, \dots, g_{p+1}) + (-1)^{p+1} c(g_1, \dots, g_p). \end{aligned}$$

The resulting cohomology groups are denoted $H^\bullet_{\mathrm{diff}}(\mathcal{G}, E)$. For more details on this subject, see [Haefliger 1979].

In the following proposition we list some results from [Crainic 2003] that are needed in the proofs of the corollaries from the Introduction.

Proposition 4.1. *Let $\mathcal{G}$ be a Lie groupoid over B with Lie algebroid $\mathcal{A}$, and let $E \to B$ be a representation of $\mathcal{G}$.*

(1) *If the s-fibers of $\mathcal{G}$ are cohomologically* 2*-connected, then*

$$H^2(\mathcal{A}; E) \cong H^2_{\mathrm{diff}}(\mathcal{G}; E).$$

(2) *If $\mathcal{G}$ is proper, then $H^2_{\mathrm{diff}}(\mathcal{G}; E) = 0$.*

(3) *If $\mathcal{G}$ is transitive, then $H^2_{\mathrm{diff}}(\mathcal{G}; E) \cong H^2_{\mathrm{diff}}(\mathcal{G}_x; E_x)$, where $x \in B$ and $\mathcal{G}_x := s^{-1}(x) \cap t^{-1}(x)$.*

Proof. (1) is a particular case of [Crainic 2003, Theorem 4], and (2) follows from [Crainic 2003, Proposition 1]. Since $\mathcal{G}$ is transitive, it is Morita equivalent to $\mathcal{G}_x$ [Moerdijk and Mrčun 2003]; by [Crainic 2003, Theorem 1], a Morita equivalence induces an isomorphism in cohomology, and this proves (3). □

Proof of Corollary 1.2. Recall that the cotangent Lie algebroid of $(\mathfrak{g}^*, \pi_{\text{lin}})$ is isomorphic to the action Lie algebroid $\mathfrak{g} \ltimes \mathfrak{g}^* \to \mathfrak{g}^*$ for the coadjoint action of $\mathfrak{g}$ on $\mathfrak{g}^*$, and that it is integrable by the action groupoid $G \ltimes \mathfrak{g}^*$, where G denotes the compact, connected and simply connected Lie group of $\mathfrak{g}$. Also, the symplectic leaves of $(\mathfrak{g}^*, \pi_{\text{lin}})$ are the orbits of the action of G. So, because $\mathbb{S}(\mathfrak{g})$ is G-invariant, it is a union of symplectic leaves, and therefore a Poisson submanifold. The algebroid $A_{\mathbb{S}(\mathfrak{g})}$ is isomorphic to the action algebroid $\mathfrak{g} \ltimes \mathbb{S}(\mathfrak{g})$, and therefore it is integrable by the action groupoid $G \ltimes \mathbb{S}(\mathfrak{g})$. Since G is simply connected, it follows that $H^2_{dR}(G) = 0$ [Duistermaat and Kolk 2000, Theorem 1.14.2]. On the other hand, all s-fibers of $G \ltimes \mathbb{S}(\mathfrak{g})$ are diffeomorphic to G, and so the assumptions of Proposition 4.1(1) are satisfied, and therefore, for any representation $E \to \mathbb{S}(\mathfrak{g})$ of $G \ltimes \mathbb{S}(\mathfrak{g})$, we have

$$H^2(\mathfrak{g} \ltimes \mathbb{S}(\mathfrak{g}); E) \cong H^2_{\text{diff}}(G \ltimes \mathbb{S}(\mathfrak{g}); E).$$

Since $G \ltimes \mathbb{S}(\mathfrak{g})$ is compact, it is proper, and hence by Proposition 4.1(2), we have $H^2_{\text{diff}}(G \ltimes \mathbb{S}(\mathfrak{g}); E) = 0$ for every representation E. Now the corollary follows from Theorem 1.1. □

Proof of Corollary 1.4. Denote by P the Poisson homotopy cover of S with structure group G. By hypothesis, P is smooth, simply connected and with vanishing second de Rham cohomology group. Let $\mathcal{G} := P \times_G P$ be the gauge groupoid of P. Since every s-fiber of $\mathcal{G}$ is diffeomorphic to P, $\mathcal{G}$ satisfies the assumptions of Proposition 4.1(1), and therefore

$$H^2(A_S; \mathcal{S}^k(TS^\circ)) \cong H^2_{\text{diff}}(\mathcal{G}; \mathcal{S}^k(TS^\circ)).$$

Since $\mathcal{G}$ is transitive, by Proposition 4.1(3), we have

$$H^2_{\text{diff}}(\mathcal{G}; \mathcal{S}^k(TS^\circ)) \cong H^2_{\text{diff}}(G; \mathcal{S}^k(T_xS^\circ)).$$

Since $T_xS^\circ \cong \mathfrak{g}$ as G representations (both integrate the adjoint representation of $\mathfrak{g}$), the proof follows from Theorem 1.3. □

Proof of Corollary 1.5. This follows from Corollary 1.4, because the differentiable cohomology of compact Lie groups vanishes, by Proposition 4.1(2). □

Proof of Corollary 1.6. Let $x \in S$, and denote by $\mathfrak{g}_x := T_xS^\circ$ the isotropy Lie algebra of the transitive algebroid A_S. By hypothesis, $\mathfrak{g}_x$ is reductive; that is, it splits as a direct product of a semisimple Lie algebra and its center $\mathfrak{g}_x = \mathfrak{s}_x \oplus \mathfrak{z}_x$, where $\mathfrak{s}_x = [\mathfrak{g}_x, \mathfrak{g}_x]$ and $\mathfrak{z}_x = Z(\mathfrak{g}_x)$ is the center of $\mathfrak{g}_x$. Since $\mathfrak{g} = TS^\circ$ is a Lie algebra bundle, it follows that this splitting is in fact global:

$$\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \oplus Z(\mathfrak{g}) = \mathfrak{s} \oplus \mathfrak{z}.$$

Since $\mathfrak{s} = [\mathfrak{g}, \mathfrak{g}]$ is an ideal of A_S, we obtain a short exact sequence of algebroids

$$0 \to \mathfrak{s} \to A_S \to A_S^{\mathrm{ab}} \to 0,$$

with $A_S^{\mathrm{ab}} = A_S/[\mathfrak{g}, \mathfrak{g}]$. Similar to the spectral sequence for Lie algebra extensions [Hochschild and Serre 1953], there is a spectral sequence for extensions of Lie algebroids [Mackenzie 1987, Theorem 5.5 and the remark following it], which in our case converges to $H^\bullet(A_S; \mathcal{S}^k(\mathfrak{g}))$, with

$$E_2^{p,q} = H^p(A_S^{\mathrm{ab}}; H^q(\mathfrak{s}; \mathcal{S}^k(\mathfrak{g}))) \Rightarrow H^{p+q}(A_S; \mathcal{S}^k(\mathfrak{g})).$$

Since $\mathfrak{s}$ is in the kernel of the anchor, $H^q(\mathfrak{s}; \mathcal{S}^k(\mathfrak{g}))$ is indeed a vector bundle, with fiber $H^q(\mathfrak{s}; \mathcal{S}^k(\mathfrak{g}))_x = H^q(\mathfrak{s}_x; \mathcal{S}^k(\mathfrak{g}_x))$, and it inherits a representation of A_S^{ab}. Since $\mathfrak{s}_x$ is semisimple, by the Whitehead Lemma we have that $H^1(\mathfrak{s}_x; \mathcal{S}^k(\mathfrak{g}_x)) = 0$ and $H^2(\mathfrak{s}_x; \mathcal{S}^k(\mathfrak{g}_x)) = 0$. Therefore,

$$H^2(A_S; \mathcal{S}^k(\mathfrak{g})) \cong H^2(A_S^{\mathrm{ab}}; \mathcal{S}^k(\mathfrak{g})^{\mathfrak{s}}), \tag{9}$$

where $\mathcal{S}^k(\mathfrak{g}_x)^{\mathfrak{s}_x}$ is the $\mathfrak{s}_x$-invariant part of $\mathcal{S}^k(\mathfrak{g}_x)$. By hypothesis, A_S^{ab} is integrable by a principle bundle P^{ab} that is simply connected and that has vanishing second de Rham cohomology and compact structure group T. Therefore, by (9) and by applying Proposition 4.1(1), (2) and (3), we obtain that

$$\begin{aligned} H^2(A_S; \mathcal{S}^k(\mathfrak{g})) &\cong H^2(A_S^{\mathrm{ab}}; \mathcal{S}^k(\mathfrak{g})^{\mathfrak{s}}) \cong H^2_{\mathrm{diff}}(P^{\mathrm{ab}} \times_T P^{\mathrm{ab}}; \mathcal{S}^k(\mathfrak{g})^{\mathfrak{s}}) \\ &\cong H^2_{\mathrm{diff}}(T; \mathcal{S}^k(\mathfrak{g}_x)^{\mathfrak{s}_x}) = 0. \end{aligned}$$

Theorem 1.3 finishes the proof. □

Proof of Corollary 1.7. Assume that $\mathfrak{g}_x$ is semisimple, $\pi_1(S, x)$ is finite, and $\pi_2(S, x)$ is a torsion group. With the notation above, we have $A_S^{\mathrm{ab}} \cong TS$. Also, TS is integrable, and the simply connected principal bundle integrating it is $\widetilde{S}$, the universal cover of S. Finiteness of $\pi_1(S)$ is equivalent to compactness of the structure group of $\widetilde{S}$. By the Hurewicz theorem, we have $H_2(\widetilde{S}, \mathbb{Z}) \cong \pi_2(\widetilde{S})$, and since $\pi_2(\widetilde{S}) = \pi_2(S)$ is torsion, we have $H^2_{\mathrm{dR}}(\widetilde{S}) = 0$. So the result follows from Corollary 1.6. □

Appendix: Equivalence of MC-elements in complete GLAs

Here we discuss some general facts about graded Lie algebras endowed with a complete filtration, with the aim of proving a criterion for equivalence of Maurer–Cartan elements (Theorem A.5), which is used in the proof of Theorem 1.1. Some of the constructions given here can be also found in [Bursztyn et al. 2009, Appendix B.1] in the more general setting of differential graded Lie algebras with a complete filtration. In fact, all our constructions can be adapted to this setup, including in particular Theorem A.5. The analog of Theorem A.5 in the case of differential graded associative algebras is in [Abad et al. 2010, Appendix A].

Definitions A.1. (1) A *graded Lie algebra* $(\mathscr{L}^\bullet, [\cdot\,,\cdot])$ (or GLA) consists of a $\mathbb{Z}$-graded vector space $\mathscr{L}^\bullet$ endowed with a graded bracket $[\cdot\,,\cdot] : \mathscr{L}^p \times \mathscr{L}^q \to \mathscr{L}^{p+q}$, which is graded commutative and satisfies the graded Jacobi identity:

$$[X, Y] = -(-1)^{|X||Y|}[Y, X], \quad [X, [Y, Z]] = [[X, Y], Z] + (-1)^{|X||Y|}[Y, [X, Z]].$$

(2) An element $\gamma \in \mathscr{L}^1$ satisfying $[\gamma, \gamma] = 0$ is called a *Maurer–Cartan element*.

(3) A *filtration* on a GLA is a decreasing sequence of homogeneous subspaces

$$\mathscr{L}^\bullet \supset \mathscr{F}_0\mathscr{L}^\bullet \supset \cdots \supset \mathscr{F}_n\mathscr{L}^\bullet \supset \mathscr{F}_{n+1}\mathscr{L}^\bullet \supset \cdots$$

satisfying

$$[\mathscr{F}_n\mathscr{L}, \mathscr{F}_m\mathscr{L}] \subset \mathscr{F}_{n+m}\mathscr{L}, \quad [\mathscr{L}, \mathscr{F}_n\mathscr{L}] \subset \mathscr{F}_n\mathscr{L}.$$

(4) A filtration $\mathscr{F}\mathscr{L}$ is called *complete* if $\mathscr{L}$ is isomorphic to the projective limit $\varprojlim \mathscr{L}/\mathscr{F}_n\mathscr{L}$.

An example of a GLA with a complete filtration appeared in Section 3: starting from a manifold M and a closed embedded submanifold $P \subset M$, we constructed $(\hat{\mathfrak{X}}^{\bullet+1}_P(M), [\cdot\,,\cdot])$, the algebra of formal vector fields along P, with filtration given by the powers of the vanishing ideal of P. So, the index of the filtration plays the role of the order to which elements vanish along P.

For a general GLA with a complete filtration $\mathscr{F}\mathscr{L}$, define the *order* of an element as follows:

$$\mathbb{O} : \mathscr{L} \to \{0, 1, \ldots, \infty\},$$
$$\mathbb{O}(X) = \begin{cases} 0 & \text{if } X \in \mathscr{L}\backslash\mathscr{F}_1\mathscr{L}, \\ n & \text{if } X \in \mathscr{F}_n\mathscr{L}\backslash\mathscr{F}_{n+1}\mathscr{L}, \\ \infty & \text{if } X = 0. \end{cases}$$

The order has the following properties, which follow from those of the filtration:

- $\mathbb{O}(X) = \infty$ if and only if $X = 0$,
- $\mathbb{O}(X + Y) \geq \mathbb{O}(X) \wedge \mathbb{O}(Y)$[2],
- $\mathbb{O}(\alpha X) \geq \mathbb{O}(X)$ for all $\alpha \in \mathbb{R}$,
- $\mathbb{O}([X, Y]) \geq \mathbb{O}(X) + \mathbb{O}(Y)$.

[2] $u \wedge v$ denotes $\min\{u, v\}$.

The completeness assumption on the filtration implies the following property:

Lemma A.2. *Let $\{X_n\}_{n\geq 0} \in \mathscr{L}$ be a sequence of elements such that*

$$\lim_{n\to\infty} \mathbb{O}(X_n) = \infty.$$

Then there exists a unique element $X \in \mathscr{L}$, denoted $X := \sum_{n\geq 0} X_n$, such that

$$X - \sum_{k=0}^{n} X_k \in \mathscr{F}_m \mathscr{L},$$

for all n big enough.

Note that $\mathfrak{g}(\mathscr{L}) := \mathscr{F}_1\mathscr{L}^0$ forms a Lie subalgebra of $\mathscr{L}^0$. Elements $X \in \mathfrak{g}(\mathscr{L})$ satisfy $\mathbb{O}(\mathrm{ad}_X(Y)) \geq \mathbb{O}(Y) + 1$ for all $Y \in \mathscr{L}$, and therefore, by Lemma A.2, the exponential of ad_X is well defined, and it is a GLA-automorphism of $\mathscr{L}^\bullet$, written

$$\mathrm{Ad}(e^X) : \mathscr{L}^\bullet \to \mathscr{L}^\bullet, \quad \mathrm{Ad}(e^X)Y := e^{\mathrm{ad}_X}(Y) = \sum_{n\geq 0} \frac{\mathrm{ad}_X^n}{n!}(Y).$$

By Lemma A.2, the Campbell–Hausdorff formula converges for all $X, Y \in \mathfrak{g}(\mathscr{L})$:

$$X * Y = X + Y + \sum_{k\geq 1} \frac{(-1)^k}{k+1} D_k(X, Y), \tag{10}$$

where

$$D_k(X, Y) = \sum_{l_i+m_i>0} \frac{\mathrm{ad}_X^{l_1}}{l_1!} \circ \frac{\mathrm{ad}_Y^{m_1}}{m_1!} \circ \dots \circ \frac{\mathrm{ad}_X^{l_k}}{l_k!} \circ \frac{\mathrm{ad}_Y^{m_k}}{m_k!}(X).$$

We use the notation $\mathscr{G}(\mathscr{L}) = \{e^X \mid X \in \mathfrak{g}(\mathscr{L})\}$; that is, $\mathscr{G}(\mathscr{L})$ is the same space as $\mathfrak{g}(\mathscr{L})$, but we denote its elements by e^X. The universal properties of the Campbell–Hausdorff formula (10) imply that $\mathscr{G}(\mathscr{L})$ endowed with the product $e^X e^Y = e^{X*Y}$ forms a group. Also, Ad gives an action of $\mathscr{G}(\mathscr{L})$ on $\mathscr{L}$ by graded Lie algebra automorphisms, which preserves the order:

- $\mathrm{Ad}(e^{X*Y}) = \mathrm{Ad}(e^X e^Y) = \mathrm{Ad}(e^X) \circ \mathrm{Ad}(e^Y)$,
- $\mathrm{Ad}(e^X)([U, V]) = [\mathrm{Ad}(e^X)U, \mathrm{Ad}(e^X)V]$,
- $\mathbb{O}(\mathrm{Ad}(e^X)(U)) = \mathbb{O}(U)$,

for all $X, Y \in \mathfrak{g}(\mathscr{L})$ and all $U, V \in \mathscr{L}$.

For later use, we give the following straightforward estimates:

Lemma A.3. *For all $X, Y, X', Y' \in \mathfrak{g}(\mathscr{L})$ and $U \in \mathscr{L}$, we have*

(1) $\mathbb{O}(X * Y - X' * Y') \geq \mathbb{O}(X - X') \wedge \mathbb{O}(Y - Y')$ *and*

(2) $\mathbb{O}(\mathrm{Ad}(e^X)U - \mathrm{Ad}(e^Y)U) \geq \mathbb{O}(X - Y)$.

Let γ be an MC-element. Notice that $[\gamma,\gamma]=0$ implies that $d_\gamma := \mathrm{ad}_\gamma$ is a differential on $\mathcal{L}^\bullet$. The fact that $\mathcal{F}_k\mathcal{L}$ are ideals implies that $(\mathcal{F}_k\mathcal{L}^\bullet, d_\gamma)$ are subcomplexes of $(\mathcal{L}^\bullet, d_\gamma)$. The induced differential on the consecutive complexes depends only on γ modulo $\mathcal{F}_1$, and their cohomology groups are denoted

$$H^n_\gamma(\mathcal{F}_k\mathcal{L}^\bullet/\mathcal{F}_{k+1}\mathcal{L}^\bullet).$$

$\mathrm{Ad}(e^X)\gamma$ is again an MC-element for $e^X \in \mathcal{G}(\mathcal{L})$, and we call γ and $\mathrm{Ad}(e^X)\gamma$ *gauge equivalent*. The next Lemma gives a linear approximation of the action $\mathcal{G}(\mathcal{L})$ on MC-elements.

Lemma A.4. *For γ an MC-element and $e^X \in \mathcal{G}(\mathcal{L})$, we have*

$$\mathbb{O}(\mathrm{Ad}(e^X)\gamma - \gamma + d_\gamma X) \geq 2\mathbb{O}(X).$$

We have the following criterion for gauge equivalence.

Theorem A.5. *Let $(\mathcal{L}^\bullet, [\cdot,\cdot])$ be a GLA with a complete filtration $\mathcal{F}_n\mathcal{L}$, and let γ,γ' be two Maurer–Cartan elements. If $\mathbb{O}(\gamma-\gamma')\geq 1$, and if for all $q \geq \mathbb{O}(\gamma-\gamma')$ we have*

$$H^1_\gamma(\mathcal{F}_q\mathcal{L}^\bullet/\mathcal{F}_{q+1}\mathcal{L}^\bullet)=0,$$

then γ and γ' are gauge equivalent; that is, there exists an element $e^X \in \mathcal{G}(\mathcal{L})$ such that $\gamma = \mathrm{Ad}(e^X)\gamma'$.

Proof. Define $p := \mathbb{O}(\gamma-\gamma')$. By hypothesis, for $q \geq p$, we can find homotopy operators

$$h^q_1 : \mathcal{F}_q\mathcal{L}^1 \to \mathcal{F}_q\mathcal{L}^0 \quad \text{and} \quad h^q_2 : \mathcal{F}_q\mathcal{L}^2 \to \mathcal{F}_q\mathcal{L}^1$$

such that $h^q_1(\mathcal{F}_{q+1}\mathcal{L}^1) \subset \mathcal{F}_{q+1}\mathcal{L}^0$, $h^q_2(\mathcal{F}_{q+1}\mathcal{L}^2) \subset \mathcal{F}_{q+1}\mathcal{L}^1$ and

$$(d_\gamma h^q_1 + h^q_2 d_\gamma - Id)(\mathcal{F}_q\mathcal{L}^1) \subset \mathcal{F}_{q+1}\mathcal{L}^1.$$

We first prove an estimate. Let $q \geq p$, and let $\tilde\gamma$ be an MC-element such that $\mathbb{O}(\tilde\gamma-\gamma)\geq q$. Then for $\widetilde{X} := h^q_1(\tilde\gamma-\gamma)$, we claim that the following estimates hold:

$$\mathbb{O}(\widetilde{X}) \geq q, \quad \mathbb{O}(\mathrm{Ad}(e^{\widetilde{X}})\tilde\gamma - \gamma) \geq q+1. \tag{11}$$

The first follows by the properties of h^q_1, and to prove the second we compute:

$$\begin{aligned}\mathbb{O}(\mathrm{Ad}(e^{\widetilde{X}})\tilde\gamma-\gamma) &\geq \mathbb{O}(\mathrm{Ad}(e^{\widetilde{X}})\tilde\gamma - \tilde\gamma + d_{\tilde\gamma}(\widetilde{X})) \wedge \mathbb{O}(\tilde\gamma - d_{\tilde\gamma}(\widetilde{X})-\gamma)\\ &\geq 2\mathbb{O}(\widetilde{X}) \wedge \mathbb{O}([\gamma-\tilde\gamma,\widetilde{X}]) \wedge \mathbb{O}(\tilde\gamma-\gamma-d_\gamma(\widetilde{X}))\\ &\geq 2q \wedge (\mathbb{O}(\gamma-\tilde\gamma)+\mathbb{O}(\widetilde{X})) \wedge \mathbb{O}(\tilde\gamma-\gamma-d_\gamma(\widetilde{X}))\\ &\geq 2q \wedge \mathbb{O}((Id-d_\gamma h^q_1)(\tilde\gamma-\gamma)),\end{aligned}$$

where for the second inequality we use Lemma A.4. The last term can be evaluated as follows:

$$\begin{aligned}\mathbb{O}((Id-d_\gamma h_1^q)(\tilde{\gamma}-\gamma)) &\geq \mathbb{O}((Id-d_\gamma h_1^q-h_2^q d_\gamma)(\tilde{\gamma}-\gamma))\wedge\mathbb{O}(h_2^q(d_\gamma(\tilde{\gamma}-\gamma)))\\ &\geq (q+1)\wedge\mathbb{O}(h_2^q(d_\gamma(\tilde{\gamma}-\gamma))).\end{aligned}$$

Since $d_\gamma(\tilde{\gamma}-\gamma)=-\frac{1}{2}[\tilde{\gamma}-\gamma,\tilde{\gamma}-\gamma]$, we have $\mathbb{O}(d_\gamma(\tilde{\gamma}-\gamma))\geq 2q\geq q+1$, so

$$\mathbb{O}((Id-d_\gamma h_1^q)(\tilde{\gamma}-\gamma))\geq q+1,$$

and this proves (11).

We construct a sequence of MC-elements $\{\gamma_k\}_{k\geq 0}$ and a sequence of group elements $\{e^{X_k}\}_{k\geq 1}\in\mathcal{G}(\mathcal{L})$ by the following recursive formulas:

$$\begin{aligned}&\gamma_0 := \gamma',\\ &X_k := h_1^{p+k-1}(\gamma_{k-1}-\gamma) \quad \text{for } k\geq 1,\\ &\gamma_k := \mathrm{Ad}(e^{X_k})\gamma_{k-1} \quad \text{for } k\geq 1.\end{aligned}$$

To show that these formulas do indeed give well-defined sequences, we have to check that $\gamma_{k-1}-\gamma\in\mathcal{F}_{p+k-1}\mathcal{L}^1$. This holds for $k=1$, and in general it follows by applying the estimate (11) inductively at each step $k\geq 1$ to $\widetilde{\gamma}=\gamma_{k-1}$ and $q=p+k-1$, to obtain

$$\mathbb{O}(X_k)\geq p+k-1, \quad \mathbb{O}(\gamma_k-\gamma)\geq p+k.$$

Using Lemma A.3(1), we obtain

$$\mathbb{O}(X_k * X_{k-1}\cdots * X_1 - X_{k-1}\cdots * X_1)\geq\mathbb{O}(X_k)\geq p+k-1,$$

and therefore by Lemma A.2, the product $X_k * X_{k-1} * \cdots * X_1$ converges to some element X. Applying Lemma A.3(1) k times, we obtain

$$\mathbb{O}(X_k * X_{k-1}\cdots * X_1)\geq\mathbb{O}(X_k)\wedge\mathbb{O}(X_{k-1})\wedge\cdots\wedge\mathbb{O}(X_1)\geq 1,$$

and thus $X\in\mathfrak{g}(\mathcal{L})$. On the other hand, we have

$$\begin{aligned}\mathbb{O}(\mathrm{Ad}(e^X)\gamma'-\gamma) &\geq \mathbb{O}(\mathrm{Ad}(e^X)\gamma'-\gamma_k)\wedge\mathbb{O}(\gamma_k-\gamma)\\ &\geq \mathbb{O}(\mathrm{Ad}(e^X)\gamma'-\mathrm{Ad}(e^{X_k*\cdots*X_1})\gamma')\wedge(p+k)\\ &\geq \mathbb{O}(X-X_k*\cdots*X_1)\wedge(p+k),\end{aligned}$$

where for the last estimate we used Lemma A.3(2). If we let $k\to\infty$, we obtain the conclusion: $\mathrm{Ad}(e^X)\gamma'=\gamma$. □

Acknowledgment

I would like to thank Marius Crainic for his very useful suggestions and comments.

References

[Abad et al. 2010] C. A. Abad, M. Crainic, and B. Dherin, "Tensor products of representations up to homotopy", preprint, 2010. arXiv 1009.5852v1

[Bursztyn et al. 2009] H. Bursztyn, V. Dolgushev, and S. Waldmann, "Morita equivalence and characteristic classes of star products", preprint, 2009. arXiv 0909.4259

[Conn 1985] J. F. Conn, "Normal forms for smooth Poisson structures", *Ann. of Math.* (2) **121**:3 (1985), 565–593. MR 86m:58050 Zbl 0592.58025

[Crainic 2003] M. Crainic, "Differentiable and algebroid cohomology, Van Est isomorphisms, and characteristic classes", *Comment. Math. Helv.* **78**:4 (2003), 681–721. MR 2004m:58034 Zbl 1041.58007

[Crainic and Fernandes 2003] M. Crainic and R. L. Fernandes, "Integrability of Lie brackets", *Ann. of Math.* (2) **157**:2 (2003), 575–620. MR 2004h:58027 Zbl 1037.22003

[Crainic and Mărcuţ 2010] M. Crainic and I. T. Mărcuţ, "A normal form theorem around symplectic leaves", preprint, 2010. arXiv 1009.2090v1

[Duistermaat and Kolk 2000] J. J. Duistermaat and J. A. C. Kolk, *Lie groups*, Springer, Berlin, 2000. MR 2001j:22008 Zbl 0955.22001

[Ginzburg and Lu 1992] V. L. Ginzburg and J.-H. Lu, "Poisson cohomology of Morita-equivalent Poisson manifolds", *Int. Math. Res. Notices* **1992**:10 (1992), 199–205. MR 93k:58091 Zbl 0783.58026

[Haefliger 1979] A. Haefliger, "Differentiable cohomology", pp. 19–70 in *Differential topology* (Varenna, 1976), edited by V. Villani, Liguori, Naples, 1979. MR 83k:17014 Zbl 0467.57007

[Hochschild and Serre 1953] G. Hochschild and J.-P. Serre, "Cohomology of Lie algebras", *Ann. of Math.* (2) **57**:3 (1953), 591–603. MR 14,943c Zbl 0053.01402

[Itskov et al. 1998] V. Itskov, M. Karasev, and Y. Vorobjev, "Infinitesimal Poisson cohomology", pp. 327–360 in *Coherent transform, quantization, and Poisson geometry*, edited by M. V. Karasev, Amer. Math. Soc. Transl. Ser. 2 **187**, Amer. Math. Soc., Providence, RI, 1998. MR 2001k:58031 Zbl 0922.58028

[Mackenzie 1987] K. Mackenzie, *Lie groupoids and Lie algebroids in differential geometry*, London Mathematical Society Lecture Note Series **124**, Cambridge University Press, Cambridge, 1987. MR 89g:58225 Zbl 0683.53029

[Moerdijk and Mrčun 2003] I. Moerdijk and J. Mrčun, *Introduction to foliations and Lie groupoids*, Cambridge Studies in Advanced Mathematics **91**, Cambridge University Press, Cambridge, 2003. MR 2005c:58039 Zbl 1029.58012

[Moerdijk and Reyes 1991] I. Moerdijk and G. E. Reyes, *Models for smooth infinitesimal analysis*, Springer, New York, 1991. MR 91m:18017 Zbl 0715.18001

[Vaisman 1994] I. Vaisman, *Lectures on the geometry of Poisson manifolds*, Progress in Mathematics **118**, Birkhäuser, Basel, 1994. MR 95h:58057 Zbl 0810.53019

[Vorobjev 2001] Y. Vorobjev, "Coupling tensors and Poisson geometry near a single symplectic leaf", pp. 249–274 in *Lie algebroids and related topics in differential geometry* (Warsaw, 2000), edited by J. Kubarski et al., Banach Center Publ. **54**, Polish Acad. Sci. Inst. Math., Warsaw, 2001. MR 2003d:53145 Zbl 1007.53062

[Weinstein 1983] A. Weinstein, "The local structure of Poisson manifolds", *J. Differential Geom.* **18**:3 (1983), 523–557. MR 86i:58059 Zbl 0524.58011

Received November 27, 2010. Revised June 29, 2011.

IOAN MĂRCUȚ
DEPARTMENT OF MATHEMATICS
UTRECHT UNIVERSITY
3508 TA UTRECHT
THE NETHERLANDS
I.T.Marcut@uu.nl

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

A REGULARITY THEOREM FOR GRAPHIC SPACELIKE MEAN CURVATURE FLOWS

BENJAMIN STUART THORPE

For mean curvature flows in Euclidean spaces, Brian White proved a regularity theorem which gives $C^{2,\alpha}$ estimates in regions of spacetime where the Gaussian density is close enough to 1. This is proved by applying Huisken's monotonicity formula. Here we will consider mean curvature flows in semi-Euclidean spaces, where each spatial slice is an m-dimensional graph in $\mathbb{R}^{m+n}_n$ satisfying a gradient bound stronger than the spacelike condition. By defining a suitable quantity to replace the Gaussian density ratio, we prove monotonicity theorems similar to Huisken's and use them to prove a regularity theorem similar to White's.

1. Introduction

A mean curvature flow can roughly be described as a family of submanifolds $\mathcal{M} = \{\mathcal{M}(t)\}_{t\in I}$ evolving with velocity equal to the mean curvature vector on each $\mathcal{M}(t)$. Let $\mathcal{M}$ be such a flow, where each spatial slice $\mathcal{M}(t)$ is assumed to be an m-dimensional submanifold of a Euclidean space. For spacetime points (y, s), the *Gaussian density ratio* is given by

$$\int_{\mathcal{M}(t)} \frac{1}{(4\pi(s-t))^{m/2}} \exp\left(-\frac{|x-y|^2}{4(s-t)}\right) dx$$

for times $t < s$. Huisken [1990] proved an important monotonicity formula, which roughly says that this will be nonincreasing with respect to t on mean curvature flows. A local version of this formula was proved by Ecker [2004, Proposition 4.17]. One application of these monotonicity formulas is the proof of Brian White's [2005] local regularity theorem for mean curvature flows in Euclidean spaces. This theorem says that such a flow will be smooth in regions of spacetime where the Gaussian density ratios are close enough to 1.

Our goal is to prove a similar regularity theorem, but now for spacelike mean curvature flows in semi-Euclidean spaces. We will assume that these flows are

MSC2010: 35B65, 35K93.
Keywords: mean curvature flow, semi-Riemannian geometry.

graphs and that they satisfy some uniform gradient bound stronger than the spacelike condition. Roughly, we will prove that if such a flow is smooth on an interval $(0, T)$ then it can be extended smoothly to time T (see Theorem 16). This should be compared to [White 2005, Theorem 3.5]. We prove this by defining a quantity that has similar properties to the Gaussian density ratio. This quantity is chosen in such a way that the evolution equations for spacelike mean curvature flows will allow us to prove monotonicity formulas similar to Huisken's and Ecker's. The proof of the regularity theorem itself is then similar to the proofs in [White 2005] and [Ecker 2004], with some adjustments.

The main differences between this case and the Euclidean case are caused by the semi-Euclidean metric. Obviously the mean curvature flow system is only parabolic when the spacelike condition is satisfied. Therefore any gradient estimates are only useful if they are stronger than the spacelike condition. This is why we will always assume such a bound on the gradient.[1] This assumption is also useful when defining our replacement for the Gaussian density ratio. For example, we need the gradient bound to guarantee that this quantity is finite on a smooth flow (since we need the eigenvalues of the induced metric to stay uniformly away from zero). We will frequently need this assumption, used with inequality (4), to get the uniform bounds needed to use the dominated convergence theorem (such arguments here are more difficult than in the Euclidean case, and therefore will be explained in more detail).

Other difficulties due to the semi-Euclidean metric appear in the proofs of the monotonicity and regularity theorems. For example, Ecker's local formula involves a nice localisation function which is not useful in the semi-Euclidean case, thus making our proof of local monotonicity slightly more awkward (see Theorem 10 and compare to [Ecker 2004, Proposition 4.17]). We also get different signs in the evolution equations for various quantities, so that the inequalities seen in the Euclidean case are often reversed here (see Equation (8), for example). The results of this are seen in the monotonicity theorems, where we see that certain quantities are nondecreasing, but where the corresponding quantities in the Euclidean case would be nonincreasing (also see Theorem 13, where the inequality in the assumption is the reverse of what we get in the Euclidean case).

The results proved in this paper formed part of the author's Ph.D. thesis at Durham University, under the supervision of Wilhelm Klingenberg.

2. Preliminaries

Notation. We will attempt to keep our notation as close to the notation in [White 2005] as possible, so that the similarities are clear. When $N \geq 2$, $\mathbb{R}^N$ will be

[1] For a case where such an estimate holds, see the appendix of [Thorpe 2011].

Euclidean space with elements denoted by x and with the usual norm $|x|$. $B^N_R(x)$ will be the ball of radius R and centre x. We will denote by $\mathbb{R}^{N,1}$ the spacetime $\mathbb{R}^N \times \mathbb{R}$ with elements $X=(x,t)$ and parabolic norm $\|X\| = \max\{|x|, |t|^{1/2}\}$. We write $B^{N,1}_R(X) = B^N_R(x)\times(t-R^2, t+R^2)$ and $U^{N,1}_R(X) = B^N_R(x)\times(t-R^2, t]$. The function $\tau : \mathbb{R}^{N,1} \to \mathbb{R}$ will be the projection $\tau(x,t) = t$ onto the time axis. For any $\lambda > 0$, we define the parabolic dilation $D_\lambda : \mathbb{R}^{N,1} \to \mathbb{R}^{N,1}$ by $D_\lambda(x,t) = (\lambda x, \lambda^2 t)$. Note that $\|D_\lambda X\| = \lambda \|X\|$. For subsets U of $\mathbb{R}^{N,1}$ and functions f from U into some Euclidean space, we define

$$d(X,U) = \inf\{\|X-Y\| \mid Y \notin U\} \quad \text{and} \quad \|f\|_{p,\alpha} = \sum_{k+2h\le p} \|D^k(\partial_t)^h f\|_{0,\alpha},$$

(for $0<\alpha<1$ and nonnegative integers p) where $[f]_\alpha = \sup_{X\ne Y\in U} |f(X) - f(Y)|/\|X-Y\|^\alpha$ and $\|f\|_{0,\alpha} = \sup_U |f| + [f]_\alpha$, and where we have used the notation $\partial_t f = \partial f/\partial t$, $\partial_A f = \partial f/\partial x^A$, $D = (\partial_1, \dots, \partial_N)$. In the obvious way, we also define the parabolic C^p norm by $\|f\|_p = \sum_{k+2h\le p} \sup_U |D^k(\partial_t)^h f|$. If we say that a sequence of functions converges in C^p or $C^{p,\alpha}$ on some set, we just mean that it converges on that set with respect to the corresponding norm.

Semi-Euclidean spaces. For integers $m\ge 2$ and $n \ge 1$, it will be convenient here for us to consider the space $\mathbb{R}^{m+n}$ with elements denoted by $x = (\hat{x}, \tilde{x})$, where $\hat{x} \in \mathbb{R}^m$ and $\tilde{x}\in\mathbb{R}^n$. With this notation, we can define the *semi-Euclidean spaces* $\mathbb{R}^{m+n}_n = (\mathbb{R}^{m+n}, \langle\cdot,\cdot\rangle)$ with metric tensor $\langle x,y\rangle = \hat{x}\cdot\hat{y} - \tilde{x}\cdot\tilde{y}$. If we use the summation convention with indices $i, j = 1, \dots, m$ and $\nu, \gamma = m+1, \dots, m+n$, then $\langle x, y\rangle = x^i y^i - x^\gamma y^\gamma$ and we denote by $\bar{g}$ the corresponding diagonal matrix with $\bar{g}_{ij} = \delta_{ij}$, $\bar{g}_{\nu\gamma} = -\delta_{\nu\gamma}$.

Let M be a submanifold of $\mathbb{R}^{m+n}_n$; then we can take the induced metric g on M in the usual way, and we say that M is *spacelike* if g is positive definite. The corresponding Levi-Civita connections (denoted $\bar\nabla$ and ∇) are defined in the usual way, and the second fundamental form on M is given by $B(V,W) = \bar\nabla_V W - \nabla_V W = (\bar\nabla_V W)^\perp$ for tangent vector fields V, W on M (where $\perp$ denotes projection to normal spaces of M in $\mathbb{R}^{m+n}_n$). Taking the trace of this (with respect to the induced metric g) gives the *mean curvature vector* $H = \operatorname{trace}_g B$ of this submanifold. We can also define the gradient (grad_M), divergence (div_M) and induced Laplace operator (Δ_M) on this submanifold, all taken with respect to the induced metric; see [O'Neill 1983] for details.

If Ω is a domain in $\mathbb{R}^m$ and $F : \Omega \to \mathbb{R}^{m+n}_n$ is an embedding such that $M = F(\Omega)$ is a spacelike submanifold of $\mathbb{R}^{m+n}_n$, then it is not difficult to check that the mean curvature is given by

$$H = (g^{ij}\partial_{ij}F)^\perp = \frac{1}{\sqrt{\det g}}\partial_i\big(\sqrt{\det g}\, g^{ij}\partial_j F\big) = \Delta_M F, \tag{1}$$

where $g_{ij} = \langle \partial_i F, \partial_j F \rangle$ gives the induced metric. This is proved as in the Euclidean case; see [Ecker 2004, Appendix A].

3. Graphic mean curvature flows

We will consider graphic flows of the form

$$\mathcal{M} = \{(\hat{x}, u(\hat{x}, t), t) \mid \hat{x} \in \Omega, t \in I\} \subset \mathbb{R}^{m+n}_n \times \mathbb{R}, \tag{2}$$

where Ω is some domain in $\mathbb{R}^m$, I is some time interval in $\mathbb{R}$ (not necessarily open) and $u : \Omega \times I \to \mathbb{R}^n$. When we say that such a flow $\mathcal{M}$ is smooth (or locally $C^{2,\alpha}$, etc.), we mean that the function u has that property. We will also discuss sequences $\mathcal{M}_J$ of such flows (where $J = 1, 2, \dots$). When we talk about convergence of $\mathcal{M}_J$ in some space of functions, we actually mean convergence of the corresponding u_J.

On each *spatial slice* $\mathcal{M}(t) = \{x \in \mathbb{R}^{m+n}_n \mid (x, t) \in \mathcal{M}\}$, we take the metric induced from $\mathbb{R}^{m+n}_n$ and assume that it is spacelike. It will be convenient for us to use the following norm for the differential map $Du(x, t) : \mathbb{R}^m \to \mathbb{R}^n$,

$$|||Du|||(x, t) = \sup_{v \in \mathbb{R}^m, |v|=1} |Du(x, t)(v)|.$$

Here $|\cdot|$ denotes the usual Euclidean norm, and D is taken with respect to the space variables only (as usual). It is possible to show that $|||Du|||^2$ will be equal to the largest eigenvalue of $Du^T Du$ at each point, and that $|||Du||| \le |Du| \le \sqrt{m}|||Du|||$. Using the obvious relationship between $|||Du|||$ and the eigenvalues of the induced metric, we see that the graph will be spacelike if and only if $|||Du||| < 1$. If Ω is convex and $|||Du|||^2 < 1 - \kappa$ then it is easy to check that, for any $t \in I$,[2]

$$|u(\hat{x}, t) - u(\hat{y}, t)| \le \sup_\Omega |||Du(\,\cdot\,, t)|||\, |\hat{x} - \hat{y}| \le (1 - \kappa)^{1/2} |\hat{x} - \hat{y}|, \tag{3}$$

and then, whenever $s \ge t$ are both in I,

$$\begin{aligned} |u(\hat{x}, t) - u(\hat{y}, s)| &\le |u(\hat{x}, t) - u(\hat{y}, t)| + |u(\hat{y}, t) - u(\hat{y}, s)| \\ &\le (1 - \kappa)^{1/2} |\hat{x} - \hat{y}| + (s - t) \sup\nolimits_{(t,s)} |\partial_t u(\hat{y}, \cdot\,)|. \end{aligned} \tag{4}$$

We denote by $H(x, t)$ the mean curvature vector at the point x of the spatial slice $\mathcal{M}(t)$. We will consider graphic flows that satisfy the quasilinear system

$$\partial_t u = \hat{g}^{ij}(Du) \partial_{ij} u \tag{5}$$

[2] To prove this, take $\hat{x} \in \Omega$ and let $h \in \mathbb{R}^m$ be such that $\hat{x} + \delta h \in \Omega$ for all $\delta \in [0, 1]$. Then $|u(\hat{x}+h) - u(\hat{x})| = |(\int_0^1 Du(\hat{x}+\delta h)\, d\delta) \cdot h| \le \int_0^1 |Du(\hat{x}+\delta h) \cdot h|\, d\delta = |h| \int_0^1 |Du(\hat{x}+\delta h) \cdot h|/|h|\, d\delta \le |h| \sup_\Omega |||Du|||$.

on $\Omega \times I$, where $\hat{g}_{ij} = \delta_{ij} - \partial_i u^\nu \partial_j u^\nu$ gives the components of the induced metric on each spatial slice. This system will be parabolic because the spacelike condition implies that $\hat{g} = I - Du^T Du$ will be positive definite.

Proposition 1. *Let $\mathcal{M}$ be a graph as in* (2), *and let I be open. Then $\mathcal{M}$ is a mean curvature flow in $\mathbb{R}^{m+n}_n \times \mathbb{R}$ if and only if the function u satisfies the system* (5).

Proof. If (5) holds then, to show that we have a mean curvature flow, it is enough to get parametrisations F of our spatial slices with $\partial_t F = H$. In other words, for each $s \in I$, we hope to find ϕ such that $F(\hat{x}, t) = (\phi(\hat{x}, t), u(\phi(\hat{x}, t), t))$ satisfies $\partial_t F(\hat{x}, t) = H(F(\hat{x}, t), t)$ for times t close to s. But we know that the mean curvature of our graph is $(0, \hat{g}^{ij}\partial_{ij}u)^\perp$, and that $\partial_t u = \hat{g}^{ij}\partial_{ij}u$. These facts and the chain rule applied to F imply that we need $\partial_t F = (\partial_t\phi, Du\partial_t\phi) + (0, \partial_t u)$ to be equal to $(0, \partial_t u)^\perp$. This is equivalent to the system $\partial_t \phi^j = \partial_t u \cdot \partial_j u \hat{g}^{ij}(Du)|_{(\phi(\hat{x},t),t)}$ for $j = 1, \dots, m$. By thinking of $\hat{x}$ as being fixed, we can think of this as a system of *ordinary* differential equations and solve for some $\phi(t)$ with initial condition $\phi(s) = \hat{x}$, for any $\hat{x} \in \Omega$. By the usual existence and uniqueness theorems [Lee 2003, Theorem 17.15], solutions $\phi_{\hat{x},s}(t)$ will exist for each $\hat{x} \in \Omega$ and $s \in I$. If we write $\phi_{\hat{x},s}(t) = \phi_s(\hat{x}, t)$, then $\phi_s(\,\cdot\,, s)$ is the identity, ϕ_s is defined on some open set $\mathscr{E} \subset \Omega \times I$ containing $\Omega \times \{s\}$, and each $\phi_s(\,\cdot\,, t)$ will be a diffeomorphism [Lee 2003, Problem 17-15]. Then ϕ_s is the required function, so we have a mean curvature flow.

Conversely, if $\mathcal{M}$ is a mean curvature flow then we take $F = (\hat{F}, \tilde{F})$ such that $\partial_t F = H$ and $F(\hat{x}, t) = (\hat{F}(\hat{x}, t), u(\hat{F}(\hat{x}, t), t))$. By the chain rule, this gives $\partial_t F(\hat{x}, t) = (I, Du(\hat{F}(\hat{x}, t), t)) \cdot \partial_t \hat{F}(\hat{x}, t) + (0, \partial_t u(\hat{F}(\hat{x}, t), t))$. The left-hand side is a normal vector and the first term on the right-hand side is tangential, therefore $\partial_t F(\hat{x}, t) = (0, \partial_t u)^\perp|_{(\hat{F}(\hat{x},t),t)}$. We already know that $\partial_t F(\hat{x}, t) = H(F(\hat{x}, t), t)$, but the mean curvature at $F(\hat{x}, t)$ is given by $(0, \hat{g}^{ij}\partial_{ij}u)^\perp|_{(\hat{F}(\hat{x},t),t)}$. Hence $(0, \partial_t u)^\perp = (0, \hat{g}^{ij}\partial_{ij}u)^\perp$, and from here it is easy to check that we must have $\partial_t u = \hat{g}^{ij}\partial_{ij}u$. □

Assumption 2. *$\mathcal{M}$ is a graphic flow, as in* (2), *where Ω is a convex, smooth domain in $\mathbb{R}^m$, and where the smooth function $u : \Omega \times I \to \mathbb{R}^n$ satisfies the system* (5) *and the inequality $|||Du|||^2 \le 1 - \kappa$ for some constant $\kappa > 0$.*

For such flows, and for times t on the interior of I, we can use the parametrisation F from the proof of Proposition 1 to prove the following facts. Note that we will repeatedly use the fact that $\partial_t F = H = \Delta_{\mathcal{M}(t)} F$, by Equation (1), and we will write $g = (g_{ij}) = (\langle \partial_i F, \partial_j F \rangle)$ for the induced metric on spatial slices. The first fact is a version of the divergence theorem on mean curvature flows,

$$\int_{\mathcal{M}(t)} \langle H, V \rangle = \int_{\mathcal{M}(t)} \langle \Delta_{\mathcal{M}(t)} F, V \rangle = -\int_{\mathcal{M}(t)} \operatorname{div}_{\mathcal{M}(t)} V$$

for vector fields V with compact support on $\mathcal{M}(t)$, where the integrals are taken over the spatial slice with respect to the induced metric g.[3] If $f(x,t)$ is a real-valued function defined on the flow then

$$\frac{df}{dt} = \partial_t f + \langle \bar{g} Df, H\rangle \quad \text{and} \quad \Delta_{\mathcal{M}(t)} f = \langle H, \bar{g} Df\rangle + \operatorname{div}_{\mathcal{M}(t)} (\bar{g} Df), \tag{6}$$

where $\bar{g}$ is the matrix defined in the previous section. The second equation here, along with the divergence theorem above, gives

$$\int_{\mathcal{M}(t)} \big(\phi \Delta_{\mathcal{M}(t)} \eta - \eta \Delta_{\mathcal{M}(t)} \phi\big) = 0, \tag{7}$$

whenever ϕ and η are C^2 on $\mathcal{M}(t)$ with ϕ having compact support. Finally, using the usual formula for differentiating determinants, we have the following evolution equation on mean curvature flows,

$$\frac{d}{dt}\sqrt{\det g} = -\sqrt{\det g}\, \langle H, H\rangle \geq 0. \tag{8}$$

Definition 3. Let $X_0 = (x_0, t_0) \in \mathbb{R}^{m+n,1}$. Define $\Phi_{X_0} : \mathbb{R}^{m+n} \times (-\infty, t_0) \to \mathbb{R}$ by

$$\Phi_{X_0}(x,t) = \frac{1}{(4\pi(t_0 - t))^{m/2}} \exp\left(-\frac{\langle x - x_0, x - x_0\rangle}{4(t_0 - t)}\right).$$

Let $\mathcal{M}$ be a graphic flow in $\mathbb{R}^{m+n}_n \times \mathbb{R}$, as in (2). For times $t < t_0$, we define

$$\Theta(\mathcal{M}, X_0, t) = \int_{x \in \mathcal{M}(t)} \Phi_{X_0}(x,t).$$

We see that

$$\frac{\partial \Phi_{X_0}}{\partial t} = \frac{m\Phi_{X_0}}{2(t_0 - t)} - \frac{\langle x - x_0, x - x_0\rangle\, \Phi_{X_0}}{4(t_0 - t)^2}, \quad \bar{g} D\Phi_{X_0} = -\frac{(x - x_0)\Phi_{X_0}}{2(t_0 - t)}. \tag{9}$$

These equations, combined with (6), give

$$\begin{aligned}
&\left(\frac{d}{dt} + \Delta_{\mathcal{M}(t)}\right)\Phi_{X_0} \\
&\quad = \partial_t \Phi_{X_0} + 2\big\langle \bar{g} D\Phi_{X_0}, H\big\rangle + \operatorname{div}_{\mathcal{M}(t)}(\bar{g} D\Phi_{X_0}) \\
&\quad = \partial_t \Phi_{X_0} + \operatorname{div}_{\mathcal{M}(t)}(\bar{g} D\Phi_{X_0}) + \frac{\big\langle (\bar{g} D\Phi_{X_0})^\perp, (\bar{g} D\Phi_{X_0})^\perp\big\rangle}{\Phi_{X_0}} \\
&\qquad\qquad - \left\langle H - \frac{(\bar{g} D\Phi_{X_0})^\perp}{\Phi_{X_0}}, H - \frac{(\bar{g} D\Phi_{X_0})^\perp}{\Phi_{X_0}}\right\rangle \Phi_{X_0} + \langle H, H\rangle\, \Phi_{X_0}.
\end{aligned} \tag{10}$$

[3] Whenever it will not cause confusion, we will write integrals of the form $\int_{x \in \mathcal{M}(t)} f(x,t)\, dx$ as $\int_{\mathcal{M}(t)} f$ to save space. Such integrals are *always* taken relative to the induced metric from $\mathbb{R}^{m+n}_n$. Similarly, we write $\Delta_{\mathcal{M}(t)} f(x,t)$ as $\Delta_{\mathcal{M}(t)} f$ when the meaning is clear.

But the first three terms on the right-hand side of this equation add up to 0 since, by (9),

$$\operatorname{div}_{\mathcal{M}(t)}(\bar{g}D\Phi_{X_0}) = \frac{-m\Phi_{X_0}}{2(t_0-t)} + \frac{\Phi_{X_0}}{4(t_0-t)^2}\left\langle (x-x_0)^\top, (x-x_0)^\top \right\rangle,$$

$$\frac{\left\langle (\bar{g}D\Phi_{X_0})^\perp, (\bar{g}D\Phi_{X_0})^\perp \right\rangle}{\Phi_{X_0}} = \frac{\Phi_{X_0}}{4(t_0-t)^2}\left\langle (x-x_0)^\perp, (x-x_0)^\perp \right\rangle.$$

Now we use this, and the evolution equation for $\sqrt{\det g}$, to differentiate the integral $\int_{x\in\mathcal{M}(t)} \Phi_{X_0}(x,t)\phi(x,t)$ when ϕ is some nonnegative C^2 function where each $\phi(\,\cdot\,,t)$ has compact support on $\mathcal{M}(t)$.

$$\begin{aligned}
\frac{d}{dt}\int_{\mathcal{M}(t)} &\Phi_{X_0}\phi \\
&= \int_{\mathcal{M}(t)} \left(\phi\frac{d\Phi_{X_0}}{dt} + \Phi_{X_0}\frac{d\phi}{dt} - \langle H, H\rangle \phi\Phi_{X_0}\right) \\
&= \int_{\mathcal{M}(t)} \Phi_{X_0}\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\phi + \left(\left(\frac{d}{dt} + \Delta_{\mathcal{M}(t)}\right)\Phi_{X_0} - \langle H, H\rangle\,\Phi_{X_0}\right)\phi \\
&= \int_{\mathcal{M}(t)} \Phi_{X_0}\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\phi - \left\langle H - \frac{(\bar{g}D\Phi_{X_0})^\perp}{\Phi_{X_0}}, H - \frac{(\bar{g}D\Phi_{X_0})^\perp}{\Phi_{X_0}}\right\rangle \phi\Phi_{X_0},
\end{aligned}$$

where we have used (7) and then (10). By (9) this gives:

$$\begin{aligned}
\text{(11)}\quad \frac{d}{dt}\int_{\mathcal{M}(t)} \Phi_{X_0}\phi &= \int_{\mathcal{M}(t)} \Phi_{X_0}\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\phi \\
&\quad - \int_{\mathcal{M}(t)} \left\langle H - \frac{(x-x_0)^\perp}{2(t_0-t)}, H - \frac{(x-x_0)^\perp}{2(t_0-t)}\right\rangle \phi\Phi_{X_0}.
\end{aligned}$$

This will be very useful later, and it is our first step towards the proof of monotonicity formulas. It is important to remember that the second term on the right-hand side is nonnegative (since the flow is spacelike, which means that normal vectors will be timelike or zero). This is unlike the Euclidean case, where the corresponding term would be nonpositive.

Proposition 4. *Let $\mathcal{M}$ be as in Assumption 2, but with $\Omega = \mathbb{R}^m$ and $I = (-\infty, T]$ for $T > 0$. If, for every point (x,t) on the flow, we have*

$$\text{(12)}\qquad H(x,t) = \frac{x^\perp}{2t},$$

then $\mathcal{M} \cap \{X \mid \tau(X) \le 0\}$ is invariant under parabolic dilations.

Proof. The idea (as for a similar result in [Ilmanen 1997]) is to assume that there is some point $Y = (y,t)$ on $\mathcal{M}' = \mathcal{M} \cap \{X \mid \tau(X) \le 0\}$, but not on $D_\lambda\mathcal{M}'$ for some λ. Then we take a compactly supported C^2 function ϕ with $\phi(y) = 1$ and $\phi = 0$ on $D_\lambda\mathcal{M}(t)$.

The transformation formula for integrals gives $\int_{D_\lambda \mathcal{M}(t)} \phi = \lambda^m \int_{\mathcal{M}(t/\lambda^2)} \phi(\lambda x)$. Then our evolution equation for $\sqrt{\det g}$ implies that

$$\frac{d}{d\lambda}\int_{D_\lambda \mathcal{M}(t)} \frac{\phi}{\lambda^m}$$
$$= \int_{x\in\mathcal{M}(t/\lambda^2)} \left(\frac{2t}{\lambda^3}\phi(\lambda x)\left\langle H, \frac{x^\perp}{2t/\lambda^2}\right\rangle + D\phi(\lambda x)\cdot x - \frac{2t}{\lambda^2}D\phi(\lambda x)\cdot H + \frac{m}{\lambda}\phi(\lambda x)\right),$$

where we have used Equation (12) to get $H = x^\perp/(2t/\lambda^2)$ on $\mathcal{M}(t/\lambda^2)$, and the fact that $\partial_t F = H$. We can deal with the first term by using the divergence theorem (and the fact that H is a normal vector) to get

$$\int_{x\in\mathcal{M}(t/\lambda^2)} \langle H, \phi(\lambda x) x^\perp\rangle = -\int_{x\in\mathcal{M}(t/\lambda^2)} \operatorname{div}_{\mathcal{M}(t/\lambda^2)}(\phi(\lambda x)x),$$

and by using the fact that $\operatorname{div}_{\mathcal{M}(t/\lambda^2)}(\phi(\lambda x)x) = m\phi(\lambda x) + \lambda D\phi(\lambda x)\cdot x^\top$. It follows that

$$\frac{d}{d\lambda}\int_{D_\lambda \mathcal{M}(t)} \phi = 0,$$

so $\int_{D_\lambda \mathcal{M}} \phi$ remains constant as λ varies. The contradiction proves our claim. □

Proposition 5. *Let* $X, Y \in \mathbb{R}^{m+n,1}$, $s < \tau(Y)$ *and* $\lambda > 0$; *then*

$$\Theta(D_\lambda(\mathcal{M} - X), Y, s) = \Theta(\mathcal{M}, X + D_{1/\lambda}Y, \tau(X) + s/\lambda^2).$$

Proof. If $\mathcal{M}$ is given by $u : \Omega \times I \to \mathbb{R}^n$ and if $X = (\hat{x}, \tilde{x}, t)$, then $D_\lambda(\mathcal{M} - X)$ is given by $u_{\lambda,X}(\cdot,\cdot) = \lambda(u(\cdot/\lambda + \hat{x}, \cdot/\lambda^2 + t) - \tilde{x})$ on $D_\lambda(\Omega \times I - (\hat{x}, t))$. Then the transformation rule for integrals gives the expected result. □

4. Monotonicity for entire flows

Given a flow satisfying Assumption 2, we say that it is an *entire* flow if $\Omega = \mathbb{R}^m$ and $I = (-\infty, T]$ for some $T \in (-\infty, \infty]$. If $\mathcal{M}$ is such an entire flow, then $\Theta(\mathcal{M}, X_0, t)$ is finite at points $X_0 = (x_0, t_0) = (\hat{x}_0, u(\hat{x}_0, t_0), t_0)$ on $\mathcal{M}$ for times $t < t_0$. To prove this, we use $\sqrt{\det \hat{g}} < 1$ and the fact that inequality (4) gives a bound on the exponent in Φ_{X_0} on the flow,

$$(13)\quad -\frac{\langle x - x_0, x - x_0\rangle}{4(t_0 - t)} \le \frac{-\kappa|\hat{x} - \hat{x}_0|^2}{4(t_0 - t)} + \frac{(1-\kappa)^{1/2}\sup_{(t,t_0)}|\partial_t u(\hat{x}_0,\cdot)||\hat{x} - \hat{x}_0|}{2} + \frac{(t_0 - t)^2 \sup_{(t,t_0)}|\partial_t u(\hat{x}_0,\cdot)|^2}{4(t_0 - t)}.$$

Here we can use the fact that u is smooth, so the time derivative in this inequality will be bounded on (t, t_0) by some constant. Also, the fact that $t < t_0$ is fixed means that $t_0 - t > 0$ will be constant. This means that, for large $|\hat{x} - \hat{x}_0|$, the first term in the right-hand side of (13) will dominate. So we have a bound on $\Theta(\mathcal{M}, X_0, t)$ by some integral known to be finite, being given by the usual formula for Gaussian integrals, $\int_{\mathbb{R}^m} \exp(-A_{ij} y^i y^j/2) dy = \sqrt{(2\pi)^m/\det(A_{ij})}$. (Here the matrix A_{ij} is constant, symmetric and positive definite. Almost all of the bounds on integrals that we use will follow from this.)

The simplest example is a *nonmoving plane*, where each spatial slice is a spacelike plane (independent of time). Then Du is constant and $\partial_t u = 0$. Obviously this implies that $|u(\hat{x}, t) - u(\hat{x}_0, t_0)|^2 = |Du \cdot (\hat{x} - \hat{x}_0)|^2 = (\hat{x} - \hat{x}_0)^T Du^T Du(\hat{x} - \hat{x}_0)$, where we know that $\hat{g} = I - Du^T Du$. For any point $X_0 = (\hat{x}_0, u(\hat{x}_0, t_0), t_0)$ on the flow, we then see that

$$\Theta(\mathcal{M}, X_0, t) = \int_{\mathbb{R}^m} \frac{1}{(4\pi(t_0 - t))^{m/2}} \exp\left(-\frac{(\hat{x} - \hat{x}_0)^T \hat{g}(\hat{x} - \hat{x}_0)}{4(t_0 - t)}\right) \sqrt{\det \hat{g}}\, d\hat{x} = 1,$$

where we again use the Gaussian integral formula. Therefore Θ is equal to 1 on nonmoving planes.

Theorem 6. *Let $\mathcal{M}$ be an entire flow satisfying Assumption 2, and let the mean curvature H be uniformly bounded on $\mathcal{M}$. Then*

$$\frac{d}{dt}\Theta(\mathcal{M}, X_0, t) = -\int_{x \in \mathcal{M}(t)} \left\langle H(x, t) + \frac{(x - x_0)^\perp}{2(t_0 - t)}, H(x, t) + \frac{(x - x_0)^\perp}{2(t_0 - t)} \right\rangle \Phi_{X_0}$$

when $X_0 = (x_0, t_0) \in \mathcal{M}$ and $t < t_0$.

This theorem gives us a monotonicity formula, similar to Huisken's, for entire spacelike mean curvature flows. It tells us that Θ will be *nondecreasing* with respect to the time variable on such flows (since the right-hand side in the formula is nonnegative, by the spacelike condition). This is different to the Euclidean case, where the Gaussian density ratio would be *nonincreasing*. The proof of this theorem should be compared to the one in [Ecker 2004, p. 55]. We could even weaken the assumption on H, but for now it is enough to assume that it is bounded.

Proof. For each $R > 0$ we can choose (as in [Ecker 2004, proof of Theorem 4.13]) functions $\chi_R^m : \mathbb{R}^m \to \mathbb{R}$ such that $\chi_{B_R^m(0)} \leq \chi_R^m \leq \chi_{B_{2R}^m(0)}$ and $R|D\chi_R^m| + R^2|D^2\chi_R^m| \leq C$ for some constant C.[4] Using these functions, we define $\chi_R : \mathbb{R}_n^{m+n} \to \mathbb{R}$ by taking $\chi_R(x) = \chi_R(\hat{x}, \tilde{x}) = \chi_R^m(\hat{x})$ for any $x = (\hat{x}, \tilde{x})$. We apply (11) with $\phi = \chi_R$ to get

$$\frac{d}{dt}\int_{\mathcal{M}(t)} \Phi_{X_0}\chi_R = -\int_{\mathcal{M}(t)} \left\langle H + \frac{(x - x_0)^\perp}{2(t_0 - t)}, H + \frac{(x - x_0)^\perp}{2(t_0 - t)} \right\rangle \Phi_{X_0}\chi_R + \int_{\mathcal{M}(t)} \Phi_{X_0}\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\chi_R. \tag{14}$$

[4] For a set K, we denote by χ_K the characteristic function of K.

Using (6), the Schwarz inequality and the bounds on the eigenvalues of $\hat{g}$ (from the assumed bound on the gradient), we have[5]

$$\left|\left(\frac{d}{dt}-\Delta_{\mathcal{M}(t)}\right)\chi_R\right| \le |\hat{g}^{-1}(Du)|\cdot|D^2\chi_R^m| \le C_0(\kappa)\frac{C}{R^2}\chi_{B_{2R}^m(0)-B_R^m(0)}, \tag{15}$$

where we have also used the fact that χ_R^m is constant outside $B_{2R}^m(0)-B_R^m(0)$.

Now we will restrict to any fixed bounded time interval $I'=[a,b]\subset(-\infty,t_0)$, considering only times $t\in I'$. The first thing to note here is that we have positive upper and lower bounds, independent of t (but depending on I'), on both t_0-t and $1/(t_0-t)$. Next we note that the flow is smooth on $(-\infty,t_0]$ (by our assumptions in the statement of the theorem) and $X_0=(\hat{x}_0,u(\hat{x}_0,t_0),t_0)$ lies on the flow, so we have $\sup_{[t,t_0]}|\partial_t u(\hat{x}_0,\cdot)|\le\sup_{[a,t_0]}|\partial_t u(\hat{x}_0,\cdot)|$, where $\sup_{[a,t_0]}|\partial_t u(\hat{x}_0,\cdot)|$ is a finite constant independent of $t\in I'$. We can use this to apply inequality (4) to bound the exponent of Φ_{X_0} on our flow, getting

$$-\frac{\langle x-x_0,x-x_0\rangle}{4(t_0-t)} \le \frac{-\kappa|\hat{x}-\hat{x}_0|^2}{4(t_0-a)}+\frac{(t_0-a)\sup_{[a,t_0]}|\partial_t u(\hat{x}_0,\cdot)|^2}{4} + \frac{(1-\kappa)^{1/2}\sup_{[a,t_0]}|\partial_t u(\hat{x}_0,\cdot)|}{2}|\hat{x}-\hat{x}_0|.$$

We denote the right-hand side of this inequality by $Q(|\hat{x}-\hat{x}_0|)$, where the coefficients of the polynomial Q depend on I' and $\hat{x}_0$ but are independent of $t\in I'$. Also, using the Schwarz and triangle inequalities, with the assumed bounds on H and $|||Du|||$ (and hence on the eigenvalues of $\hat{g}$), it is not difficult to see that we have $-\big\langle H+(x-x_0)^\perp/2(t_0-t), H+(x-x_0)^\perp/2(t_0-t)\big\rangle\le P(|\hat{x}-\hat{x}_0|)$ on our flow, where P is some polynomial with coefficients again independent of $t\in I'$. Now we recall (14) and use it to get

$$\begin{aligned}
&\left|\frac{d}{dt}\int_{\mathcal{M}(t)}\Phi_{X_0}\chi_R+\int_{\mathcal{M}(t)}\left\langle H+\frac{(x-x_0)^\perp}{2(t_0-t)},H+\frac{(x-x_0)^\perp}{2(t_0-t)}\right\rangle\Phi_{X_0}\right| \\
&\quad\le\left|\int_{\mathcal{M}(t)}\Phi_{X_0}\left(\frac{d}{dt}-\Delta_{\mathcal{M}(t)}\right)\chi_R\right| \\
&\qquad\qquad+\left|\int_{\mathcal{M}(t)}\left\langle H+\frac{(x-x_0)^\perp}{2(t_0-t)},H+\frac{(x-x_0)^\perp}{2(t_0-t)}\right\rangle\Phi_{X_0}(1-\chi_R)\right| \\
&\quad\le\int_{\mathbb{R}^m}\frac{C_0C}{R^2}\frac{\chi_{B_{2R}^m(0)-B_R^m(0)}}{(4\pi(t_0-b))^{m/2}}\exp\big(Q(|\hat{x}-\hat{x}_0|)\big)\,d\hat{x} \\
&\qquad\qquad+\int_{\mathbb{R}^m}P(|\hat{x}-\hat{x}_0|)\frac{(1-\chi_{B_R^m(0)})}{(4\pi(t_0-b))^{m/2}}\exp\big(Q(|\hat{x}-\hat{x}_0|)\big)\,d\hat{x},
\end{aligned}$$

[5] From now on, $C(\cdot,\dots,\cdot)$ will always denote a positive constant depending on the quantities in parentheses.

where we have used all of the inequalities above, as well as $\sqrt{\det \hat{g}} \le 1$. Both integrands in the right-hand side are bounded by an integrable function independent of R (since Q is dominated by the $-|\hat{x} - \hat{x}_0|^2$ term and P is just a polynomial). Both integrands converge pointwise to zero on $\mathbb{R}^m$ as $R \to \infty$, which allows us to apply the dominated convergence theorem to see that the right-hand side of this inequality converges to zero. Since the right-hand side is independent of $t \in I'$, this convergence is uniform. So we have

$$\lim_{R\to\infty} \frac{d}{dt} \int_{\mathcal{M}(t)} \Phi_{X_0} \chi_R = -\int_{\mathcal{M}(t)} \left\langle H + \frac{(x - x_0)^\perp}{2(t_0 - t)}, H + \frac{(x - x_0)^\perp}{2(t_0 - t)} \right\rangle \Phi_{X_0}.$$

The uniform convergence allows us to swap the order of the limit and the derivative on the left-hand side to get

$$-\int_{\mathcal{M}(t)} \left\langle H + \frac{(x - x_0)^\perp}{2(t_0 - t)}, H + \frac{(x - x_0)^\perp}{2(t_0 - t)} \right\rangle \Phi_{X_0} = \frac{d}{dt} \int_{\mathcal{M}(t)} \Phi_{X_0},$$

where we have again used a dominated convergence argument (involving Q, etc.) and the fact that χ_R^m converges to 1 pointwise. Since we can do this for any such I', the equation above holds for all $t < t_0$. □

Corollary 7. *Let $\mathcal{M}$ be as in Theorem 6, then $\Theta(\mathcal{M}, X, t) \le 1$ for all $X \in \mathcal{M}$ and all $t < \tau(X)$. Also, $\Theta(\mathcal{M}, X, t) = 1$ for all $X \in \mathcal{M}$ and all $t < \tau(X)$ if and only if $\mathcal{M}$ is a nonmoving plane.*

Proof. Let $Y = (y, s) \in \mathcal{M}$, then we claim that $\lim_{t\to s} \Theta(\mathcal{M}, Y, t) = 1$. We prove this by considering dilations of the flow using Proposition 5.

$$\Theta(\mathcal{M}, Y, t) = \Theta(D_{1/(s-t)^{1/2}}(\mathcal{M} - Y), 0, -1) \tag{16}$$

and, since the flow is smooth at Y, the flows $D_{1/(s-t)^{1/2}}(\mathcal{M} - Y)$ converge to a nonmoving plane as $t \to s$. To understand why, write $\lambda = \sqrt{s - t}$ and let each $D_{1/\lambda}(\mathcal{M} - Y)$ be given by the graph of a function u_λ. If $(\mathcal{M} - Y)$ is the graph of a function u, then $u_\lambda(\hat{z}, r) = u(\lambda\hat{z}, \lambda^2 r)/\lambda$ and the definition of the derivative (with respect to λ) gives $\lim_{\lambda\to 0} u_\lambda(\hat{z}, r) = Du(0, 0) \cdot \hat{z} + 0 \cdot 2r\partial_t u(0, 0)$. Therefore $D_{1/\lambda}(\mathcal{M} - Y)$ converges pointwise to a nonmoving plane as $\lambda \to 0$.

We easily see that

$$Du_\lambda(\hat{z}, r) = Du(\lambda\hat{z}, \lambda^2 r) \to Du(0, 0),$$

so $\det \hat{g}(Du_\lambda)$ converges to $\det \hat{g}(Du(0, 0))$. Also,

$$\sup_{[-1,0]} |\partial_t u_\lambda(0, \cdot)| = \lambda \sup_{[-\lambda^2,0]} |\partial_t u(0, \cdot)| \to 0$$

as $\lambda \to 0$. We can use these facts to apply the dominated convergence theorem to $\Theta(D_{1/\lambda}(\mathcal{M} - Y), 0, -1)$, by again using inequality (4) in the usual way to get an upper bound on the exponent of $\Phi_0(\,\cdot\,,-1)$ on each $D_{1/\lambda}(\mathcal{M}-Y)$,

$$\frac{-|\hat{x}|^2+|u_\lambda(\hat{x},-1)|^2}{4} \le \frac{-|\hat{x}|^2+\big((1-\kappa)^{1/2}|\hat{x}|+\sup_{[-1,0]}|\partial_t u_\lambda(0,\cdot\,)|\big)^2}{4}$$
$$\le \frac{-\kappa|\hat{x}|^2+2(1-\kappa)^{1/2}|\hat{x}|+1}{4},$$

whenever λ is small enough that $\sup_{[-1,0]}|\partial_t u_\lambda(0,\cdot\,)|\le 1$. Now we have a bound (for all small λ) on the integrands of each $\Theta(D_{1/\lambda}(\mathcal{M}-Y),0,-1)$ by some integrable function. We can therefore apply the dominated convergence theorem to get $\Theta(D_{1/\lambda}(\mathcal{M}-Y),0,-1)\to 1$ as $\lambda\to 0$, since Θ is always equal to 1 on nonmoving planes. This fact and (16) give $\Theta(\mathcal{M},Y,t)\to 1$ as $t\to s$. The monotonicity theorem tells us that $\Theta(\mathcal{M},Y,t)$ is nondecreasing with respect to $t<s$ and therefore must be ≤ 1.

For the second part of the corollary, if $\Theta(\mathcal{M},Y,t)\equiv 1$ then the monotonicity formula gives

$$0=\frac{d}{dt}\Theta(\mathcal{M},Y,t)=-\int_{\mathcal{M}(t)}\left\langle H+\frac{(x-y)^\perp}{2(s-t)},\,H+\frac{(x-y)^\perp}{2(s-t)}\right\rangle\Phi_Y,$$

and therefore (since normal vectors are timelike or zero) we have

$$H(x,t)=-(x-y)^\perp/2(s-t).$$

This means that the flow $\mathcal{M}'=(\mathcal{M}-Y)\cap\{X\mid\tau(X)\le 0\}$ satisfies (12) and must be invariant under parabolic dilations. As $\lambda\to\infty$, the flows $D_\lambda\mathcal{M}'$ again converge to a nonmoving plane, which must be equal to $\mathcal{M}'$. This is true for all $Y\in\mathcal{M}$ with $\tau(T)<T=\sup I$, so $\mathcal{M}$ must be a nonmoving plane. ☐

5. Local monotonicity

If a flow satisfying Assumption 2 has $I=[a,b)$, then Proposition 17 (given in the Appendix) implies that we can extend it continuously to $[a,b]$. Taking a subset of Ω if necessary (remember that we are interested in local theorems here), the following assumption will hold.

Assumption 8. *With $\mathcal{M}$ as in Assumption 2, $\Omega\times I$ is bounded and u is continuous on its closure.*

Now we will prove a kind of local monotonicity theorem, which will be used to prove a local regularity theorem later. We will need to define a local version of Θ. We can choose a C^2 function $\phi:\mathbb{R}^m\to\mathbb{R}$ which satisfies $\chi_{B^m_{1/2}(0)}\le\phi\le\chi_{B^m_1(0)}$ and $|D^2\phi|\le C_1$, where C_1 is some positive constant depending only on m. Then, for any spacetime point $X_0=(\hat{x}_0,\tilde{x}_0,t_0)$ and any $\rho>0$, we define a function on

$\mathbb{R}^{m+n}$ by

$$\phi_{\rho,X_0}(x)=\phi_{\rho,X_0}(\hat{x},\tilde{x})=\phi\Big(\frac{\hat{x}-\hat{x}_0}{\rho}\Big),$$

which will satisfy $\chi_{B^m_{\rho/2}(\hat{x}_0)\times\mathbb{R}^n}\le\phi_{\rho,X_0}\le\chi_{B^m_\rho(\hat{x}_0)\times\mathbb{R}^n}$ and $|D^2\phi_{\rho,X_0}|\le C_1/\rho^2$. It will also be convenient now for us to define the sets $Q^{m,n,1}_\rho(X)=B^m_\rho(\hat{x})\times\mathbb{R}^n\times(t-\rho^2,t)$ and $P^{m,n,1}_\rho(X)=B^m_\rho(\hat{x})\times\mathbb{R}^n\times(t-\rho^2,t+\rho^2)$ for any spacetime point $X=(\hat{x},\tilde{x},t)$.

Definition 9. Let $\mathcal{M}$ be a graphic flow in $\mathbb{R}^{m+n}_n\times\mathbb{R}$, as in (2). If $X_0\in\mathbb{R}^{m+n,1}$ and $\rho>0$ are such that $Q^{m,n,1}_\rho(X_0)\subset\Omega\times\mathbb{R}^n\times I$ then we define

$$\Theta(\mathcal{M},X_0,t,\rho)=\int_{x\in\mathcal{M}(t)}\Phi_{X_0}(x,t)\phi_{\rho,X_0}(x)$$

for $t<\tau(X_0)$ in I.

As in Proposition 5, we can prove

$$\Theta(D_\lambda(\mathcal{M}-X),Y,t,\rho)=\Theta(\mathcal{M},X+D_{1/\lambda}Y,\tau(X)+t/\lambda^2,\rho/\lambda). \tag{17}$$

By the dominated convergence theorem, we easily see that $\Theta(\mathcal{M},X,s,\rho)$ is continuous with respect to $X\in\mathcal{M}$. Now we can prove a local monotonicity theorem. We will use the notation $\bar{\mathcal{M}}$ for the closure of $\mathcal{M}$.

Theorem 10. *Let $\mathcal{M}$ satisfy Assumption 8, and let $\rho>0$. Then there exist positive constants C_2 and $\delta<\rho^2$ such that, whenever $X_0\in\bar{\mathcal{M}}$ is such that $Q^{m,n,1}_\rho(X_0)\subset\Omega\times\mathbb{R}^n\times I$, the function $t\mapsto\Theta(\mathcal{M},X_0,t,\rho)+C_2t$ will be nondecreasing with respect to $t\in(\tau(X_0)-\delta,\tau(X_0))$.*

Note that C_2 and δ will be independent of such points X_0, but will depend on $\mathcal{M}$ and ρ.

Proof. We know from (11) that

$$\frac{d}{dt}\Theta(\mathcal{M},X_0,t,\rho)\ge\int_{\mathcal{M}(t)}\Phi_{X_0}\left(\frac{d}{dt}-\Delta_{\mathcal{M}(t)}\right)\phi_{\rho,X_0}. \tag{18}$$

As in the proof of Theorem 6, it is easy to check that $\left|(d/dt-\Delta_{\mathcal{M}(t)})\phi_{\rho,X_0}\right|\le C_3\chi_{B^m_\rho(\hat{x}_0)\times\mathbb{R}^n-B^m_{\rho/2}(\hat{x}_0)\times\mathbb{R}^n}$, where $C_3=C_3(\kappa,\rho)$ is constant. Let $\hat{x},\hat{y}\in\bar{\Omega}$ and $t<s$ in $\bar{I}$ be such that $\rho/2<|\hat{x}-\hat{y}|<\rho$. Then, by (3) and the triangle inequality, we have

$$\begin{aligned}(19)\quad & -|\hat{x}-\hat{y}|^2+|u(\hat{x},s)-u(\hat{y},t)|^2\\ &\le-\kappa|\hat{x}-\hat{y}|^2+2(1-\kappa)^{1/2}|\hat{x}-\hat{y}||u(\hat{x},s)-u(\hat{x},t)|+|u(\hat{x},s)-u(\hat{x},t)|^2\\ &\le-\kappa\rho^2/4+2(1-\kappa)^{1/2}\rho|u(\hat{x},s)-u(\hat{x},t)|+|u(\hat{x},s)-u(\hat{x},t)|^2.\end{aligned}$$

But, by uniform continuity of u (since it is continuous on the closure of $\Omega \times I$), we can take $\delta > 0$ (not depending on $\hat{x}, \hat{y}, s, t$) such that the right-hand side of (19) will be $\leq -\kappa\rho^2/8$ whenever $|t-s| < \delta$.[6] Taking $s = \tau(X_0)$ and combining the above inequalities with the fact that $\sqrt{\det \hat{g}} \leq 1$ gives

$$\left| \int_{\mathcal{M}(t)} \Phi_{X_0} \left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)} \right) \phi_{\rho, X_0} \right| \leq \int_\Omega \frac{C_3 \chi_{B^m_\rho(\hat{x}_0) - B^m_{\rho/2}(\hat{x}_0)}}{(4\pi(\tau(X_0)-t))^{m/2}} \exp\left(\frac{-\rho^2\kappa/32}{\tau(X_0)-t} \right),$$

for $0 < \tau(X_0) - t < \delta$. Taking $t \to \tau(X_0)$ in the right-hand side shows that it is bounded by some finite constant C_4 for these values of t. Therefore

$$\frac{d}{dt}\Theta(\mathcal{M}, X_0, t, \rho) \geq -C_4$$

for $t \in (\tau(X_0) - \delta, \tau(X_0))$, proving the theorem. □

Corollary 11. *Let $\mathcal{M}$ be as in Theorem 10. If $X_0 \in \bar{\mathcal{M}}$ and $\rho_0 > 0$ are such that $Q^{m,n,1}_{\rho_0}(Y) \subset \Omega \times \mathbb{R}^n \times I$ for all $Y \in Q^{m,n,1}_{\rho_0}(X_0)$, and if*

$$\lim_{t \to \tau(X_0)} \Theta(\mathcal{M}, X_0, t, \rho_0) > 1 - \epsilon$$

for some $\epsilon > 0$, then there exists $\rho \in (0, \rho_0)$ such that

$$\Theta(\mathcal{M}, Y, t, \rho_0) \geq 1 - \epsilon$$

for all $Y \in Q^{m,n,1}_\rho(X_0) \cap \mathcal{M}$ and all $t \in (\tau(Y) - \rho^2, \tau(Y))$.

Proof. Let $\lim_{t \to \tau(X_0)} \Theta(\mathcal{M}, X_0, t, \rho_0) \geq 1 - \epsilon + \eta$ for some $\eta > 0$ (the limit exists in $\mathbb{R} \cup \{\infty\}$ by the local monotonicity theorem). Then there must exist $\rho_1 \in (0, \rho_0]$ such that $\Theta(\mathcal{M}, X_0, \tau(X_0) - \rho_1^2, \rho_0) > 1 - \epsilon + \eta/2$. We can choose ρ_1 to be as small as we like, so we take ρ_1^2 to be less than both $\delta(\mathcal{M}, \rho_0)$ and $\eta/4C_2(\mathcal{M}, \rho_0)$ (with δ and C_2 as in Theorem 10). By continuity, there will exist $\rho \in (0, \rho_1)$ such that, for all $Y \in Q^{m,n,1}_\rho(X_0) \cap \mathcal{M}$,

$$\Theta(\mathcal{M}, Y, \tau(X_0) - \rho_1^2, \rho_0) > 1 - \epsilon + \eta/4$$

and $(\tau(Y) - \rho^2, \tau(Y)) \subset (\tau(X_0) - \rho_1^2, \tau(X_0)) \subset (\tau(X_0) - \delta, \tau(X_0))$. So we can apply Theorem 10 to $\Theta(\mathcal{M}, Y, t, \rho_0)$ for $t \in (\tau(Y) - \rho^2, \tau(Y))$ to get

$$\Theta(\mathcal{M}, Y, \tau(X_0) - \rho_1^2, \rho_0) + C_2(\tau(X_0) - \rho_1^2) \leq \Theta(\mathcal{M}, Y, t, \rho_0) + C_2 t$$

for all such Y and t, which in turn implies

$$\Theta(\mathcal{M}, Y, t, \rho_0) \geq C_2(\tau(X_0) - t - \rho_1^2) + 1 - \epsilon + \eta/4 \geq 1 - \epsilon + (\eta/4 - \rho_1^2 C_2),$$

where the last term is positive by our choice of ρ_1. □

[6]Uniform continuity implies that, for any $\epsilon > 0$, there exists $\delta > 0$ such that $\|(\hat{x}, s) - (\hat{y}, t)\| < \delta$ implies $|u(\hat{x}, s) - u(\hat{y}, t)| < \epsilon$. Taking $\hat{x} = \hat{y}$ and a small enough ϵ here proves our claim.

Proposition 12. *Let $\mathcal{M}$ satisfy Assumption 8, and let X_0 and ρ be as in Theorem 10. Then* $\lim_{t\to\tau(X_0)} \Theta(\mathcal{M}, X_0, t, \rho) = \lim_{t\to\tau(X_0)} \Theta(\mathcal{M}, X_0, t)$. *In particular, the limit on the left-hand side is independent of ρ.*

Proof. It is easy to see that, if we write $X_0 = (\hat{x}_0, u(\hat{x}_0, t_0), t_0)$,

$$\begin{aligned}
&0 \le \Theta(\mathcal{M}, X_0, t) - \Theta(\mathcal{M}, X_0, t, \rho)\\
&= \int_\Omega \frac{\exp\left(\dfrac{-|\hat{x}-\hat{x}_0|^2 + |u(\hat{x}, t) - u(\hat{x}_0, t_0)|^2}{4(t_0-t)}\right)}{(4\pi(t_0-t))^{m/2}} \left(1-\phi\left(\frac{\hat{x}-\hat{x}_0}{\rho}\right)\right)\sqrt{\det \hat{g}}\, d\hat{x}.
\end{aligned}$$

But $\sqrt{\det \hat{g}} < 1$ and $1-\phi((\hat{x}-\hat{x}_0)/\rho)$ is at most 1 and vanishes for $\hat{x} \in B^m_{\rho/2}(\hat{x}_0)$. Thus we only need to consider $|\hat{x}-\hat{x}_0| \ge \rho/2$ and, as in inequality (19), we get

$$\begin{aligned}
&-|\hat{x}-\hat{x}_0|^2 + |u(\hat{x}, t) - u(\hat{x}_0, t_0)|^2\\
&\quad \le -\kappa|\hat{x}-\hat{x}_0| + 2(1-\kappa)^{1/2}|\hat{x}-\hat{x}_0||u(\hat{x}_0, t) - u(\hat{x}_0, t_0)| + |u(\hat{x}_0, t) - u(\hat{x}_0, t_0)|^2\\
&\quad \le -\kappa\rho^2/4 + 2(1-\kappa)^{1/2} \operatorname{diam}\Omega |u(\hat{x}_0, t) - u(\hat{x}_0, t_0)| + |u(\hat{x}_0, t) - u(\hat{x}_0, t_0)|^2,
\end{aligned}$$

which is $\le -\kappa\rho^2/8$ when we take $|u(\hat{x}_0, t) - u(\hat{x}_0, t_0)|$ small enough (by continuity) by taking t close enough to t_0. Therefore, for such t, we have

$$\Theta(\mathcal{M}, X_0, t) - \Theta(\mathcal{M}, X_0, t, \rho) \le \int_\Omega \frac{\exp\left((-\kappa\rho^2/8)/4(t_0-t)\right)}{(4\pi(t_0-t))^{m/2}}\, d\hat{x},$$

which converges to 0 as $t \to t_0$. □

6. Local regularity

In [White 2005], a regularity theorem for mean curvature flows in Euclidean spaces is proved. To do this, a kind of local $C^{2,\alpha}$ norm is used (defined at each point of a flow and denoted by $K_{2,\alpha}$). For a sequence of $C^{2,\alpha}$ flows, denoted by $\mathcal{M}_J$, if this norm is uniformly bounded on compact subsets as $J \to \infty$ then a version of the Arzelà–Ascoli theorem [White 2005, Theorem 2.6] gives local parabolic C^2 convergence of a subsequence to some locally $C^{2,\alpha}$ flow. However, the definition of this norm involves rotations, which would cause problems in the semi-Euclidean case (because of the spacelike condition and because the mean curvature flow system is not preserved by such rotations). It is convenient for us to define a slightly different quantity with similar properties. The idea will be to use the gradient bound (from the spacelike assumption) to ignore the first few terms in the $C^{2,\alpha}$ norm, thus removing the need to translate and rotate in the definition of $K_{2,\alpha}$.

Suppose that we have a spacelike, graphic flow $\mathcal{M}$ (as in (2), not necessarily a mean curvature flow) and $X \in \Omega \times \mathbb{R}^n \times I$. For any $\alpha \in (0, 1)$, we define $G_{2,\alpha}(\mathcal{M}, X)$

to be the infimum of the numbers $\lambda > 0$ such that

$$(20) \qquad [Du_{\lambda,X}|_{U^{m,1}}]_\alpha + \|D^2 u_{\lambda,X}|_{U^{m,1}}\|_{0,\alpha} + \|\partial_t u_{\lambda,X}|_{U^{m,1}}\|_{0,\alpha} \le 1,$$

where $u_{\lambda,X}$ is the function whose graph gives the flow $D_\lambda(\mathcal{M} - X)$, and $u_{\lambda,X}|_{U^{m,1}}$ is the restriction to $U^{m,1} = B_1^m(0) \times (-1, 0]$. This will be finite when the flow is smooth (to understand why, see how each term in (20) is affected by dilations). It is important to note that, for any $X = (\hat{x}, \tilde{x}, t)$, $G_{2,\alpha}(\mathcal{M}, X)$ is independent of $\tilde{x}$ (since the definition only involves derivatives of u). We will also need the obvious facts that this quantity will be zero on nonmoving planes and that $G_{2,\alpha}(D_\lambda(\mathcal{M}-X), 0) = G_{2,\alpha}(\mathcal{M}, X)/\lambda$.

The most important property of $G_{2,\alpha}$ is a version of the Arzelà–Ascoli theorem. Roughly, if we have a sequence of smooth spacelike flows $\mathcal{M}_J$, each containing the origin and with $G_{2,\alpha}(\mathcal{M}_J, \cdot\,)$ uniformly bounded on compact subsets of spacetime as $J \to \infty$, then we have local parabolic C^2 convergence of some subsequence to a locally $C^{2,\alpha}$ limit flow. Comparing $G_{2,\alpha}$ to $K_{2,\alpha}$ and applying Theorem 2.6 of [White 2005] gives us this fact, but we will still explain in detail in Proposition 19 in a special case (the only case that we need). Furthermore, if each of the flows satisfies the system (5) then so will the limit (by the C^2 convergence). This limit must then be smooth by induction, since a $C^{k,\alpha}$ solution to the system must be $C^{k+1,\alpha}$, by the usual theorems for linear equations; see [Friedman 1964, Chapter 3], for example.

Theorem 13. *Let $\alpha, \kappa \in (0, 1)$ be given. Then there exist positive constants ϵ and C_5 such that if*

(a) *$\mathcal{M}$ is as in Assumption 2, with $\sup I = 0 \in I$ and with $u(0, 0) = 0$, and*

(b) *$\rho_0 > 1$ is such that $Q^{m,n,1}_{\rho_0}(Y) \subset \Omega \times \mathbb{R}^n \times I$ and*

$$\Theta(\mathcal{M}, Y, t, \rho_0) \ge 1 - \epsilon$$

for all $Y \in Q^{m,n,1}_1(0) \cap \mathcal{M}$ and all $t \in (\tau(Y) - 1, \tau(Y))$,

then $\sup_{X \in Q^{m,n,1}_1(0)} G_{2,\alpha}(\mathcal{M}, X) d(X, P^{m,n,1}_1(0)) \le C_5$.

It is important to notice that the constants ϵ and C_5 will depend on κ, α, m, n, but will be independent of $\mathcal{M}$. Also, since $G_{2,\alpha}$ scales like the reciprocal of parabolic distance, the inequality in the conclusion of the theorem is invariant under parabolic dilations. This is the most important theorem of this section and is a version of White's local regularity theorem. The proof should be compared to those of [White 2005, Theorem 3.1] and [Ecker 2004, Theorem 5.6]. As in this latter reference, we use the local version of Θ. As in [White 2005], we aim for bounds on the $C^{2,\alpha}$

norm and use the Schauder estimates,[7] rather than aiming for bounds on the second fundamental form and using related interior estimates as in [Ecker 2004].

Proof. Let $\bar{\epsilon}$ be the infimum of numbers $\epsilon > 0$ for which the theorem fails (i.e., for which no such C_5 exists). We need $\bar{\epsilon} > 0$, so we assume $\bar{\epsilon} = 0$ to get a contradiction. We take a sequence $\epsilon_J \to \bar{\epsilon}$ with $\epsilon_J > \bar{\epsilon}$. Then there exist sequences $\mathcal{M}_J$ and $\rho_J > 1$, satisfying all of the assumptions of the theorem (with the same α and κ), but with $\epsilon_J, \mathcal{M}_J, \rho_J$ in place of $\epsilon, \mathcal{M}, \rho_0$, and with

$$\gamma_J = \sup_{X \in Q_1^{m,n,1}(0)} d(X, P_1^{m,n,1}(0))\, G_{2,\alpha}(\mathcal{M}_J, X) \to \infty$$

as $J \to \infty$. Each γ_J is finite since $\mathcal{M}_J$ is smooth. For each J we can choose Y_J in $Q_1^{m,n,1}(0)$ such that $G_{2,\alpha}(\mathcal{M}_J, Y_J) d(Y_J, P_1^{m,n,1}(0)) \geq \gamma_J/2$, and we can assume that $Y_J \in \mathcal{M}_J$.[8] We define $\lambda_J = G_{2,\alpha}(\mathcal{M}_J, Y_J)$ and consider the flows[9]

$$\tilde{\mathcal{M}}_J = D_{\lambda_J}(\mathcal{M}_J - Y_J),$$

which all contain the origin (in spacetime). Then $G_{2,\alpha}(\tilde{\mathcal{M}}_J, 0) = 1$ for all J and $D_{\lambda_J}(P_1^{m,n,1}(0) - Y_J) = P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)$. But now

$$\frac{\gamma_J}{2} \leq G_{2,\alpha}(\mathcal{M}_J, Y_J) d(Y_J, P_1^{m,n,1}(0)) = 1 \times d(0, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)),$$

so $d(0, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)) \to \infty$ since $\gamma_J \to \infty$ as $J \to \infty$. Let X be a point in $Q_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)$. Then

$$d(X, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)) G_{2,\alpha}(\tilde{\mathcal{M}}_J, X) \leq \gamma_J \leq 2d(0, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)),$$

from which we obtain

$$G_{2,\alpha}(\tilde{\mathcal{M}}_J, X) \leq \frac{2d(0, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J))}{d(X, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J))}.$$

The triangle inequality gives $\|0-Y\| \leq \|0-X\| + \|Y-X\|$, and taking the supremum over all $Y \notin P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)$ gives

$$d(X, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)) \geq d(0, P_{\lambda_J}^{m,n,1}(-D_{\lambda_J} Y_J)) - \|X\|,$$

[7] Note that White uses the Schauder estimates for the heat equation but, since we do not want to rotate our flows, we have to use a more general version of the Schauder estimates.

[8] Remember that $G_{2,\alpha}(\mathcal{M}, (\hat{x}, \tilde{x}, t))$ is independent of $\tilde{x}$, and so is $d((\hat{x}, \tilde{x}, t), P_1^{m,n,1}(0))$.

[9] Note that the flows $\mathcal{M}_J$ and $\tilde{\mathcal{M}}_J$ will be graphs of functions u_J and $\tilde{u}_J$ on sets $\Omega_J \times I_J$ and $\tilde{\Omega}_J \times \tilde{I}_J$ respectively, where $\sup I_J = 0 \Rightarrow \sup \tilde{I}_J = \tau(-D_{\lambda_J} Y_J) > 0$.

which leads to

$$G_{2,\alpha}(\tilde{\mathcal{M}}_J, X) \le \frac{2}{1 - \|X\|/d(0, P^{m,n,1}_{\lambda_J}(-D_{\lambda_J} Y_J))}, \tag{21}$$

whenever the right-hand side is positive. Since $d(0, P^{m,n,1}_{\lambda_J}(-D_{\lambda_J} Y_J)) \to \infty$, this inequality tells us that $G_{2,\alpha}(\tilde{\mathcal{M}}_J, X)$ is uniformly bounded (as $J \to \infty$) on compact subsets of spacetime with $\tau(X) \le 0$.[10] This allows us to apply Proposition 19 to the sequence $\tilde{\mathcal{M}}_J \cap \{X \mid \tau(X) \le 0\}$ to get parabolic C^2 convergence, on compact subsets of $\mathbb{R}^m \times (-\infty, 0]$, of a subsequence to a limit flow $\mathcal{M}'$. We can assume that this subsequence is our original sequence, and will therefore continue to use the notation $\tilde{\mathcal{M}}_J$. The limit $\mathcal{M}'$ will be a smooth entire graphic flow defined on $\mathbb{R}^m \times (-\infty, 0]$ (since $\lambda_J \to \infty$). It will be the graph of a function u' satisfying the system (5) (since the convergence is C^2). Also, since the gradient bound is unaffected by parabolic dilations, $\sup |||Du'|||^2 \le 1 - \kappa$. Proposition 19 tells us that $\mathcal{M}'$ has uniformly bounded mean curvature. This allows us to apply the monotonicity theorem and related results to the flow.

Now we use the assumption that $\Theta(\mathcal{M}_J, Y, s, \rho_J) \ge 1 - \epsilon_J$ for $Y \in Q^{m,n,1}_1(0) \cap \mathcal{M}_J$ and $s \in (\tau(Y) - 1, \tau(Y))$. By (17), this is equivalent to the inequality

$$\Theta(\tilde{\mathcal{M}}_J, Y, s, \lambda_J \rho_J) \ge 1 - \epsilon_J$$

for $Y \in Q^{m,n,1}_{\lambda_J}(-D_{\lambda_J} Y_J) \cap \tilde{\mathcal{M}}_J$ and $s \in (\tau(Y) - \lambda_J^2, \tau(Y))$. Given $Z = (\hat{z}, u'(\hat{z}, t), t)$ in $\mathcal{M}'$, with $s < t < 0$, we can take a sequence $Z_J = (\hat{z}, \tilde{u}_J(\hat{z}, t), t) \in \tilde{\mathcal{M}}_J$ with $Z_J \to Z$. Then, for large enough J, the fact that $d(0, P^{m,n,1}_{\lambda_J}(-D_{\lambda_J} Y_J)) \to \infty$ implies that Z_J (which is bounded since it converges) will be in $Q^{m,n,1}_{\lambda_J}(-D_{\lambda_J} Y_J)$. Obviously we will have $s \in (\tau(Z_J) - \lambda_J^2, \tau(Z_J))$ for all large J. This gives $\Theta(\tilde{\mathcal{M}}_J, Z_J, s, \lambda_J \rho_J) \ge 1 - \epsilon_J$. We see easily that $\Theta(\tilde{\mathcal{M}}_J, Z_J, s, \lambda_J \rho_J)$ equals

$$\int_{\tilde{\Omega}_J} \frac{\exp\left(\dfrac{-|\hat{x} - \hat{z}|^2 + |\tilde{u}_J(\hat{x}, s) - \tilde{u}_J(\hat{z}, t)|^2}{4(t-s)}\right)}{(4\pi(t-s))^{m/2}} \phi\left(\frac{\hat{x} - \hat{z}}{\lambda_J \rho_J}\right) \sqrt{\det \hat{g}(D\tilde{u}_J(\hat{x}, s))}\, d\hat{x}, \tag{22}$$

where the integral can be thought of as an integral over $\mathbb{R}^m$ since ϕ has compact support. By the C^2 convergence $\tilde{u}_J \to u'$ and the fact that $\rho_J \lambda_J \to \infty$ with $\phi \equiv 1$ in some ball with centre 0, the integrands above will converge pointwise to the integrand in $\Theta(\mathcal{M}', Z, s)$. But we have $\phi \le 1$, $\sqrt{\det \hat{g}} \le 1$ and $t - s > 0$ independent

[10] For example, for any such compact set we can assume $G_{2,\alpha}(\tilde{\mathcal{M}}_J, X) \le 4$ for all X in this set by assuming $\|X\| \le R$ and taking J so large that $d(0, P^{m,n,1}_{\lambda_J}(-D_{\lambda_J} Y_J)) \ge 2R$.

of J, as well as

$$-|\hat{x}-\hat{z}|^2+|\tilde{u}_J(\hat{x},s)-\tilde{u}_J(\hat{z},t)|^2 \le -\kappa|\hat{x}-\hat{z}|^2+(t-s)^2\sup_{(s,t)}|\partial_t\tilde{u}_J(\hat{z},\cdot)|^2 \\ +2(1-\kappa)^{1/2}|\hat{x}-\hat{z}|(t-s)\sup_{(s,t)}|\partial_t\tilde{u}_J(\hat{z},\cdot)|$$

by inequality (4). By the parabolic C^2 convergence, we can assume for large J that $\sup_{(s,t)}|\partial_t\tilde{u}_J(\hat{z},\cdot)|$ is arbitrarily close to $\sup_{(s,t)}|\partial_t u'(\hat{z},\cdot)|$, which is finite (by smoothness of u') and independent of J. These inequalities combine to give a bound on the integrands of (22) by some function that is independent of J and integrable over $\mathbb{R}^m$. This allows us to apply the dominated convergence theorem to get $\Theta(\mathcal{M}',Z,s)\leftarrow\Theta(\tilde{\mathcal{M}}_J,Z_J,s,\lambda_J\rho_J)\ge 1-\epsilon_J\to 1-\bar{\epsilon}$. So, for all $Z\in\mathcal{M}'$ with $s<\tau(Z)<0$, we have $\Theta(\mathcal{M}',Z,s)\ge 1-\bar{\epsilon}$. Now, since we assumed $\bar{\epsilon}=0$, the fact that $\mathcal{M}'$ is entire with $\Theta(\mathcal{M}',Z,s)\ge 1$ implies by Corollary 7 that $\Theta(\mathcal{M}',Z,s)\equiv 1$. Therefore $\mathcal{M}'$ must be a nonmoving plane.

Let u' be as above and consider the linear operator with constant coefficients $\partial_t-\hat{g}^{ij}(Du')\partial_{ij}$ applied to $\tilde{u}_J$. The system (5) and the fact that $\partial_{ij}u'=\partial_t u'=0$ then give $(\partial_t-\hat{g}^{ij}(Du')\partial_{ij})(\tilde{u}_J-u')=(\hat{g}^{ij}(D\tilde{u}_J)-\hat{g}^{ij}(Du'))\partial_{ij}\tilde{u}_J$. For $U_2^{m,1}(0)\subset\mathbb{R}^m\times(-\infty,0]$, the Schauder estimates for linear parabolic equations [Krylov 1996, Theorem 8.11.1] tell us that

$$\begin{aligned}\|(\tilde{u}_J-u')|_{U_2^{m,1}(0)}\|_{2,\alpha} &\le C_6\big(\|(\partial_t-\hat{g}^{ij}(Du')\partial_{ij})(\tilde{u}_J-u')|_{U_4^{m,1}(0)}\|_{0,\alpha}+\sup_{U_4^{m,1}(0)}|\tilde{u}_J-u'|\big)\\ &= C_6\big(\|(\hat{g}^{ij}(D\tilde{u}_J)-\hat{g}^{ij}(Du'))\partial_{ij}\tilde{u}_J|_{U_4^{m,1}(0)}\|_{0,\alpha}+\sup_{U_4^{m,1}(0)}|\tilde{u}_J-u'|\big),\end{aligned}$$

whenever J is large enough that $U_6^{m,1}(0)\subset\tilde{\Omega}_J\times\tilde{I}_J$, and where the constant C_6 will depend on m,n,α,κ. But both terms on the right-hand side converge to 0 as $J\to\infty$, since $\partial_{ij}\tilde{u}_J$ is bounded in $C^{0,\alpha}$ on compact subsets (by inequality (21)) and since $(\hat{g}^{ij}(D\tilde{u}_J)-\hat{g}^{ij}(Du'))\to 0$ in C^1 on compact sets. This means that, on $U_2^{m,1}(0)$, the convergence $\tilde{u}_J\to u'$ is $C^{2,\alpha}$. In particular, the terms of the $C^{2,\alpha}$ norm of $\tilde{u}_J$ involved in the definition of $G_{2,\alpha}(\tilde{\mathcal{M}}_J,0)$ will converge to 0 (since these terms are zero on u'). This finally gives a contradiction because we dilated in such a way that $G_{2,\alpha}(\tilde{\mathcal{M}}_J,0)=1$, which implies that $[D\tilde{u}_J]_\alpha+\|D^2\tilde{u}_J\|_{0,\alpha}+\|\partial_t\tilde{u}_J\|_{0,\alpha}$ is bounded from below, independently of J, on $U_2^{m,1}(0)$. Therefore $\bar{\epsilon}>0$. □

Corollary 14. *Let ϵ and C_5 be as in Theorem 13. Let $\mathcal{M}$ satisfy Assumption 2, with $X_0\in\mathcal{M}$ and $\tau(X_0)=\sup I$.*[11] *Suppose that $\rho_0>\rho>0$ are such that $Q_{\rho_0}^{m,n,1}(Y)\subset\Omega\times\mathbb{R}^n\times I$ and*

$$\Theta(\mathcal{M},Y,s,\rho_0)\ge 1-\epsilon,$$

[11] By these assumptions, the flow will be smooth at time $\tau(X_0)$, since we are taking X_0 to be a point *on* the flow.

for all $Y \in Q_\rho^{m,n,1}(X_0) \cap \mathcal{M}$ *and all* $s \in (\tau(Y) - \rho^2, \tau(Y))$*. Then*

$$\sup_{\mathcal{M} \cap Q_\rho^{m,n,1}(X_0)} G_{2,\alpha}(\mathcal{M}, \cdot\,) d(\,\cdot\,, P_\rho^{m,n,1}(X_0)) \le C_5.$$

Proof. This follows easily from Theorem 13 after taking the dilation $D_{1/\rho}(\mathcal{M} - X_0)$ and applying (17). □

The next corollary should be compared to Theorem 3.5 of [White 2005].

Corollary 15. *Let* $\mathcal{M}$ *satisfy Assumption 8. Let* X_0 *lie in the closure* $\bar{\mathcal{M}}$ *such that* $\tau(X_0) = \sup I$.[12] *Suppose* $\rho_0 > \rho > 0$ *are such that* $Q_{\rho_0}^{m,n,1}(Y) \subset \Omega \times \mathbb{R}^n \times I$ *and*

$$\Theta(\mathcal{M}, Y, s, \rho_0) \ge 1 - \epsilon,$$

for all $Y \in Q_\rho^{m,n,1}(X_0) \cap \mathcal{M}$ *and all* $s \in (\tau(Y) - \rho^2, \tau(Y))$*. Then* $\bar{\mathcal{M}}$ *will be smooth in some spacetime neighbourhood of* X_0*.*

Proof. We take a sequence $X_J \to X_0$ (as $J \to \infty$) in $\mathcal{M}$ with $\tau(X_J) < \tau(X_0)$ and with $\hat{x}_J = \hat{x}_0$. For large J, $\|X_J - X_0\| < \rho/2$ and we define

$$\mathcal{M}_J = \{Y \in \mathcal{M} \mid \tau(Y) \le \tau(X_J)\}.$$

Now $\Theta(\mathcal{M}_J, Y, s, \rho_0) \ge 1 - \epsilon$ for $Y \in Q_{\rho/2}^{m,n,1}(X_J) \cap \mathcal{M}_J \subset Q_\rho^{m,n,1}(X_0) \cap \mathcal{M}$ and $s \in (\tau(Y) - \rho^2/4, \tau(Y)) \subset (\tau(Y) - \rho^2, \tau(Y))$. Then, by Corollary 14,

$$\sup_{\mathcal{M}_J \cap Q_{\rho/2}^{m,n,1}(X_J)} G_{2,\alpha}(\mathcal{M}_J, \cdot\,) d(\,\cdot\,, P_{\rho/2}^{m,n,1}(X_J)) \le C_5$$

for large J. This gives a $C^{2,\alpha}$ bound on each $\mathcal{M}_J$ in some fixed spacetime neighbourhood of X_0. Then, since $\tau(X_J) \to \tau(X_0)$, we see that $\bar{\mathcal{M}}$ is $C^{2,\alpha}$ in this neighbourhood and therefore smooth. □

Theorem 16. *Let* $\mathcal{M}$ *be a spacelike graphic mean curvature flow in* $\mathbb{R}_n^{m+n} \times \mathbb{R}$*, given by a smooth function* $u : \Omega \times (0, T) \to \mathbb{R}^n$ *with* $|||Du|||^2 \le 1 - \kappa$ *for some positive constant* κ*. Then* $\mathcal{M}$ *can be extended smoothly to the time* T*.*

Proof. We can extend u continuously to T, thanks to Proposition 17, and let $X_0 = (\hat{x}_0, u(\hat{x}_0, T), T)$ for any $\hat{x}_0 \in \Omega$. By Proposition 1, u satisfies system (5). We can take a convex, bounded neighbourhood $\Omega_0 \subset \Omega$ of $\hat{x}_0$ and some $t_0 \in (0, T)$. Then the flow $\mathcal{M}_0$ given by the restriction of u to $\Omega_0 \times (t_0, T)$ will satisfy Assumption 8. Choosing $\rho_0 > 0$ to be sufficiently small, we first apply Theorem 18 and Proposition 12 to get $\lim_{t\to T} \Theta(\mathcal{M}_0, X_0, t, \rho_0) > 1 - \epsilon$. Then we can apply Corollary 11, which allows us to use Corollary 15 to get smoothness of $\bar{\mathcal{M}}_0$ in a neighbourhood of X_0. We can do this at any $\hat{x}_0 \in \Omega$, and therefore $\mathcal{M}$ can be extended smoothly to T. □

[12] By these assumptions, $\bar{\mathcal{M}}$ is continuous at time $\tau(X_0)$ but not necessarily smooth, since we only assume X_0 to be on the closure and not necessarily on the flow itself.

Appendix

Proposition 17. *Suppose that $\mathcal{M}$ is as in Assumption 2, with $I = [a, b)$. Then u can be extended to a continuous function on $\Omega \times [a, b]$.*

Proof. Take the linear operator $P = \partial_t - \hat{g}^{ij}(Du)\partial_{ij}$. Using $Pu = 0$, applying Theorem 2.14 of [Lieberman 1996] (in particular, the comment that follows it) on cylinders in $\Omega \times (a, b)$ tells us that, for any $\hat{x} \in \Omega$, $u(\hat{x}, \cdot)$ is uniformly continuous on some interval with supremum b. It can therefore be extended continuously to $[a, b]$. On $\Omega \times [a, b)$ we have $|u(\hat{x}, t) - u(\hat{y}, t)| \le (1-\kappa)^{1/2}|\hat{x} - \hat{y}|$, so taking the limit of this as $t \to b$ gives continuity of the extension with respect to $\hat{x}$. □

Theorem 18. *Suppose that $\mathcal{M}$ satisfies Assumption 8, with $I = (0, T)$. For any $X_0 = (\hat{x}_0, u(\hat{x}_0, T), T)$ with $\hat{x}_0 \in \Omega$, we will have* $\lim_{t\to T} \Theta(\mathcal{M}, X_0, t) \ge 1$.

It is important to remember that we are not assuming the flow to be smooth on $\Omega \times (0, T]$, only continuous. The proof of this theorem is roughly the same as the proofs of similar results in [Wang 2001].

Proof. We will first define a function on the flow, $\zeta = 1 + \log(1/\kappa^{m/2}) - \log(\cosh\theta)$, where θ is the *hyperbolic angle* defined on page 3 of [Li and Salavessa 2011].[13] An evolution equation discussed in Sections 4 and 5 of the same work tells us that

$$\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\zeta \ge \kappa|B|^2,$$

where $|B|^2$ is the norm of the second fundamental form on the spatial slices. We note that there exist constants $C_7, C_8 > 0$ (depending on κ) such that $C_7|B|^2 \le |D^2u|^2 \le C_8|B|^2$.[14] Another useful fact is that, by the assumption $|||Du|||^2 \le 1-\kappa$, there exists a constant $C_9(\kappa) > 0$ such that if $v \in \mathbb{R}^{m+n}_n$ is any tangent vector to $\mathcal{M}(t)$ then $\langle v, v\rangle \le |v|^2 \le C_9\,\langle v, v\rangle$. If we use ϕ_{ρ,X_0} from Definition 9 for small enough ρ, (11) gives

$$\frac{d}{dt}\int_{\mathcal{M}(t)} \Phi_{X_0}\zeta\phi_{\rho,X_0} \ge \int_{\mathcal{M}(t)} \Phi_{X_0}\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)(\zeta\phi_{\rho,X_0}).$$

It is easy to check, as in [Ecker 2004, Lemma 3.14], that we have the product rule

$$\begin{aligned}&\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)(\phi_{\rho,X_0}\zeta)\\ &\quad= \zeta\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\phi_{\rho,X_0} + \phi_{\rho,X_0}\left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\zeta - 2\left\langle \mathrm{grad}_{\mathcal{M}(t)}\phi_{\rho,X_0}, \mathrm{grad}_{\mathcal{M}(t)}\zeta\right\rangle.\end{aligned}$$

[13] At a point on $\mathcal{M}$, $\cosh\theta$ is just the value of $1/\sqrt{\det\hat{g}}$ at the corresponding point in $\Omega \times I$.

[14] We can write $|B|^2 = |\langle B_{ij}, B_{kl}\rangle\hat{g}^{ik}\hat{g}^{jl}|$; see [Li and Salavessa 2011] for details. $|D^2u|$ just denotes the Euclidean norm of D^2u, and to prove the inequality we need the fact that the eigenvalues of $Du^T Du$ are bounded above and below. Compare to [Ilmanen 1997, p. 31].

By Young's inequality,

$$\begin{aligned}\left\langle \mathrm{grad}_{\mathcal{M}(t)}\phi_{\rho,X_0}, \mathrm{grad}_{\mathcal{M}(t)}\zeta\right\rangle &= \left\langle \frac{\mathrm{grad}_{\mathcal{M}(t)}\phi_{\rho,X_0}}{\sqrt{\phi_{\rho,X_0}}}, \sqrt{\phi_{\rho,X_0}}\,\mathrm{grad}_{\mathcal{M}(t)}\zeta\right\rangle \\ &\le \frac{1}{2\epsilon}\frac{|\mathrm{grad}_{\mathcal{M}(t)}\phi_{\rho,X_0}|^2}{\phi_{\rho,X_0}} + \frac{\epsilon}{2}\phi_{\rho,X_0}|\mathrm{grad}_{\mathcal{M}(t)}\zeta|^2 \\ &\le \frac{1}{\epsilon}C_{10}\frac{|D\phi_{\rho,X_0}|^2}{\phi_{\rho,X_0}} + \frac{\epsilon}{2}C_9\phi_{\rho,X_0}\left\langle \mathrm{grad}_{\mathcal{M}(t)}\zeta, \mathrm{grad}_{\mathcal{M}(t)}\zeta\right\rangle,\end{aligned}$$

where $C_{10}(\kappa) > 0$ and ϵ is any positive number. Since ϕ_{ρ,X_0} is compactly supported on the flow, Example 3.16 of [Ecker 2004][15] implies that

$$|D\phi_{\rho,X_0}|^2/\phi_{\rho,X_0} \le 2\max|D^2\phi_{\rho,X_0}|,$$

where we remember that $|D^2\phi_{\rho,X_0}| < C_1/\rho^2$. Using facts from [Li and Salavessa 2011] (see Equation 3.9 and the first inequality for $|B|^2$ in the proof of Proposition 5.2 there), we see that $\left\langle \mathrm{grad}_{\mathcal{M}(t)}\zeta, \mathrm{grad}_{\mathcal{M}(t)}\zeta\right\rangle \le C_{11}|B|^2$ for some constant $C_{11}(\kappa)$. So there exist constants $C_{12}, C_{13}, C_{14} > 0$ (depending on κ, ρ) such that

$$\begin{aligned}2\left\langle \mathrm{grad}_{\mathcal{M}(t)}\phi_{\rho,X_0}, \mathrm{grad}_{\mathcal{M}(t)}\zeta\right\rangle &\le \frac{C_{12}}{\epsilon} + \epsilon C_{13}\phi_{\rho,X_0}|B|^2, \\ \left(\frac{d}{dt} - \Delta_{\mathcal{M}(t)}\right)\phi_{\rho,X_0} &\le C_{14},\end{aligned}$$

where we prove the second inequality as in Theorem 6. Combining all of the inequalities above,

$$\begin{aligned}\frac{d}{dt}\int_{\mathcal{M}(t)}\Phi_{X_0}\zeta\phi_{\rho,X_0} &\ge \int_{\mathcal{M}(t)}\Phi_{X_0}\Big(\kappa\phi_{\rho,X_0}|B|^2 - C_{15}C_{14} - \frac{C_{12}}{\epsilon} - \epsilon C_{13}\phi_{\rho,X_0}|B|^2\Big) \\ &= \frac{\kappa}{2}\int_{\mathcal{M}(t)}\Phi_{X_0}\phi_{\rho,X_0}|B|^2 - C_{16}\Theta(\mathcal{M}, X_0, t),\end{aligned}$$

where we use the fact that ζ is clearly less than or equal to some constant $C_{15}(\kappa)$ and choose $\epsilon = \kappa/2C_{13}$ and $C_{16} = C_{15}C_{14} + C_{12}/\epsilon$. We can use this to prove the theorem. We assume that $\lim_{t\to T}\Theta(\mathcal{M}, X_0, t) < 1$ and hope to get a contradiction. So for t close enough to $\tau(X_0) = T$ (say $t \in (T-\delta, T)$ for some $\delta > 0$) we can assume that

$$\frac{d}{dt}\int_{\mathcal{M}(t)}\Phi_{X_0}\zeta\phi_{\rho,X_0} \ge \frac{\kappa}{2}\int_{\mathcal{M}(t)}\Phi_{X_0}\phi_{\rho,X_0}|B|^2 - C_{16}. \tag{23}$$

We can see how this inequality is affected by parabolic dilations, D_λ for $\lambda > 1$, by using the transformation formula for integrals, and by noting that ζ involves first

[15] $|D\phi|^2/\phi \le 2\max|D^2\phi|$ for compactly supported C^2 functions.

derivatives and $|B|$ involves second derivatives. We get

$$\frac{d}{ds}\int_{D_\lambda(\mathcal{M}-X_0)(s)} \Phi_0\zeta\phi_{\lambda\rho,0} \ge -\frac{C_{16}}{\lambda^2}+\frac{\kappa}{2}\int_{D_\lambda(\mathcal{M}-X_0)(s)} \Phi_0|B|^2\phi_{\lambda\rho,0}$$

for $s\in(-\lambda^2\delta,0)$, remembering that $\lambda>1$. We now take $\tau<\delta/2$ and integrate with respect to s over the interval $(-\delta/2-\tau,-\delta/2)$ to get

$$\left[\int_{D_\lambda(\mathcal{M}-X_0)(s)} \Phi_0\zeta\phi_{\lambda\rho,0}\right]_{-\delta/2-\tau}^{-\delta/2} \ge -\frac{C_{16}\tau}{\lambda^2}+\frac{\kappa}{2}\int_{-\delta/2-\tau}^{-\delta/2}\int_{D_\lambda(\mathcal{M}-X_0)(s)} \Phi_0|B|^2\phi_{\lambda\rho,0}.$$

The left-hand side and the first term on the right-hand side clearly have limit zero as $\lambda\to\infty$. Therefore we must have $\int_{-\delta/2-\tau}^{-\delta/2}\int_{D_\lambda(\mathcal{M}-X_0)(s)} \Phi_0|B|^2\phi_{\lambda\rho,0}\to 0$. As in [Wang 2001, p. 26], we can use the integral mean value theorem to choose sequences $\lambda_J\to\infty$, $\tau_J\to 0$ and $s_J\in[-\delta/2-\tau_J,-\delta/2]$ such that

$$\int_{D_{\lambda_J}(\mathcal{M}-X_0)(s_J)} \Phi_0|B|^2\phi_{\lambda_J\rho,0}\to 0 \text{ as } J\to\infty.$$

We have $\delta/2\le|s_J|\le\delta$, so

$$\Phi_0(\hat{x},\tilde{x},s_J)=\frac{\exp\left((-|\hat{x}|^2+|\tilde{x}|^2)/4|s_J|\right)}{(4\pi|s_J|)^{m/2}}\ge\frac{\exp\left(-|\hat{x}|^2/2\delta\right)}{(4\pi\delta)^{m/2}}.$$

The function $\phi_{\lambda_J\rho,0}$ is zero outside $B^m_{\lambda_J\rho}(0)\times\mathbb{R}^n$ and equals 1 inside $B^m_{\lambda_J\rho/2}(0)\times\mathbb{R}^n$. For any $R>0$ we can take J large enough that $B^m_R(0)\times\mathbb{R}^n\subset B^m_{\lambda_J\rho/2}(0)\times\mathbb{R}^n$, implying that

$$\frac{\exp(-R^2/2\delta)}{(4\pi\delta)^{m/2}}\int_{D_{\lambda_J}(\mathcal{M}-X_0)(s_J)\cap B^m_R(0)\times\mathbb{R}^n}|B|^2\le\int_{D_{\lambda_J}(\mathcal{M}-X_0)(s_J)}\Phi_0\phi_{\rho\lambda_J,0}|B|^2.$$

We therefore have

$$\frac{\exp(-R^2/2\delta)}{(4\pi\delta)^{m/2}}\int_{D_{\lambda_J}(\mathcal{M}-X_0)(s_J)\cap B^m_R(0)\times\mathbb{R}^n}|B|^2\to 0 \quad\text{as } J\to\infty. \tag{24}$$

Now let us consider the functions $\tilde{u}_J(\hat{x})$ whose graphs give the spatial slices $D_{\lambda_J}(\mathcal{M}-X_0)(s_J)$. The fact that $\lambda_J\to\infty$ tells us that, for any $R>0$, we can take J large enough that $B^m_R(0)$ is contained in the domain of $\tilde{u}_J$. Since we also have a uniform bound on the gradients $D\tilde{u}_J$, the usual Arzelà–Ascoli theorem argument gives a subsequence (which we continue to denote by $\tilde{u}_J$) converging pointwise on $\mathbb{R}^m$, and uniformly on each $B^m_R(0)$, to some limit $\tilde{u}$. Define

$$v_J^{k\gamma}=\partial_k\tilde{u}_J^\gamma \quad\text{and}\quad c_J^{k\gamma}=\frac{1}{\mathrm{vol}(B^m_R(0))}\int_{B^m_R(0)}v_J^{k\gamma},$$

and take a convergent subsequence $c_J^{k\gamma} \to c^{k\gamma}$ (since the sequence is bounded, by the gradient bound on $\tilde{u}_J$). Apply the Poincaré inequality to get

$$\int_{B_R^m(0)} |v_J^{k\gamma} - c_J^{k\gamma}|^2 \leq C_{17} \int_{B_R^m(0)} |Dv_J^{k\gamma}|^2 \leq C_{17} \int_{B_R^m(0)} |D^2\tilde{u}_J|^2 \to 0,$$

where the last step uses (24) and $|D^2\tilde{u}_J|^2 \leq C_8|B|^2$. So $v_J^{k\gamma} - c_J^{k\gamma} \to 0$ with respect to the L^2 norm on $B_R^m(0)$. Now we can assume that the derivatives of our sequence converge pointwise almost everywhere to constants. These constants will be the weak derivatives of $\tilde{u}$, which therefore must be linear. Since this holds for any R, and since Θ is equal to 1 on nonmoving planes, we can use this to apply the dominated convergence theorem to see that

$$\begin{aligned} 1 &\leq \lim_{J\to\infty} \Theta(D_{\lambda_J}(\mathcal{M} - X_0), 0, s_J) \\ &= \lim_{J\to\infty} \Theta(\mathcal{M}, X_0, T + s_J/\lambda_J^2) = \lim_{t\to T} \Theta(\mathcal{M}, X_0, t), \end{aligned}$$

contradicting the assumption that $\lim_{t\to T} \Theta(\mathcal{M}, X_0, t) < 1$. □

We only need the next proposition in the proof of Theorem 13. First we note that a bound on $G_{2,\alpha}$ on some subset of spacetime implies a bound on $[Du]_\alpha + [D^2u]_\alpha + \sup|D^2u| + [\partial_t u]_\alpha + \sup|\partial_t u|$ on a subset of $\mathbb{R}^{m,1}$. If the flow contains the origin, then inequality (4) and the spacelike assumption give a bound on $|Du|$ and on $|u|$ on this subset, and therefore a bound on $\|u\|_{2,\alpha}$.

Proposition 19. *The sequence $\tilde{\mathcal{M}}_J \cap \{X \mid \tau(X) \leq 0\}$, from the proof of Theorem 13, has a convergent subsequence (this is parabolic C^2 convergence on compact subsets). The limit is a smooth entire flow $\mathcal{M}'$, defined on $\mathbb{R}^m \times (-\infty, 0]$, satisfying Assumption 2, with uniformly bounded mean curvature vector.*

Proof. Let $\tilde{u}_J$, $\tilde{\mathcal{M}}_J$, etc. be exactly as in the proof of Theorem 13. Then, since $\lambda_J \to \infty$ and $\sup \tilde{I}_J \geq 0$, any compact subset of $\mathbb{R}^m \times (-\infty, 0]$ will be contained in the domain of $\tilde{u}_J$ for large enough J. By inequality (21), $G_{2,\alpha}(\tilde{\mathcal{M}}_J, \cdot)$ will be uniformly bounded on compact subsets of spacetime with $\tau(X) \leq 0$, as $J \to \infty$. Therefore we get uniform bounds on $\|\tilde{u}_J\|_{2,\alpha}$ on compact subsets of $\mathbb{R}^m \times (-\infty, 0]$. We can use this to prove convergence of a subsequence by following the same steps as in the proof of the Arzelà–Ascoli theorem. We use the Cantor diagonalization process to choose a pointwise convergent subsequence on $\mathbb{R}^m \times (-\infty, 0]$, and then the $C^{2,\alpha}$ estimates imply C^2 convergence on compact subsets. The limit $\mathcal{M}'$ is clearly $C^{2,\alpha}$ and therefore smooth by the usual induction argument (since the system (5) holds on $\mathcal{M}'$). Also, the $C^{2,\alpha}$ bound implies a uniform bound on the mean curvature of $\mathcal{M}'$. □

References

[Ecker 2004] K. Ecker, *Regularity theory for mean curvature flow*, Progress in Nonlinear Differential Equations and their Applications **57**, Birkhäuser, Boston, 2004. MR 2005b:53108 Zbl 1058.53054

[Friedman 1964] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, Englewood Cliffs, NJ, 1964. MR 31 #6062 Zbl 0144.34903

[Huisken 1990] G. Huisken, "Asymptotic behavior for singularities of the mean curvature flow", *J. Differential Geom.* **31**:1 (1990), 285–299. MR 90m:53016 Zbl 0694.53005

[Ilmanen 1997] T. Ilmanen, "Singularities of mean curvature flow of surfaces", preprint, 1997, Available at www.math.ethz.ch/~ilmanen/papers/sing.ps.

[Krylov 1996] N. V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Math. **12**, American Mathematical Society, Providence, RI, 1996. MR 97i:35001 Zbl 0865.35001

[Lee 2003] J. M. Lee, *Introduction to smooth manifolds*, Graduate Texts in Math. **218**, Springer, New York, 2003. MR 2003k:58001 Zbl 1030.53001

[Li and Salavessa 2011] G. Li and I. Salavessa, "Mean curvature flow of spacelike graphs", *Math. Z.* **269**:3-4 (2011). Zbl 1231.53057

[Lieberman 1996] G. M. Lieberman, *Second order parabolic differential equations*, World Scientific, River Edge, NJ, 1996. MR 98k:35003 Zbl 0884.35001

[O'Neill 1983] B. O'Neill, *Semi-Riemannian geometry, with applications to relativity*, Pure and Applied Mathematics **103**, Academic Press, New York, 1983. MR 85f:53002 Zbl 0531.53051

[Thorpe 2011] B. S. Thorpe, "The maximal graph Dirichlet problem in semi-Euclidean spaces", preprint, 2011. To appear in *Comm. Anal. Geom.* arXiv 1112.4219

[Wang 2001] M.-T. Wang, "Mean curvature flow of surfaces in Einstein four-manifolds", *J. Differential Geom.* **57**:2 (2001), 301–338. MR 2003j:53108 Zbl 1035.53094

[White 2005] B. White, "A local regularity theorem for mean curvature flow", *Ann. of Math.* (2) **161**:3 (2005), 1487–1519. MR 2006i:53100 Zbl 1091.53045

Received March 19, 2011. Revised July 11, 2011.

BENJAMIN STUART THORPE
DEPARTMENT OF MATHEMATICAL SCIENCES
DURHAM UNIVERSITY
SCIENCE LABORATORIES
SOUTH ROAD
DURHAM, DH1 3LE
UNITED KINGDOM
b.s.thorpe@durham.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 2, 2012

ANALOGUES OF LEVEL-N EISENSTEIN SERIES

HIROFUMI TSUMURA

We consider certain analogues of level-N Eisenstein series involving hyperbolic functions. By developing the method used in our previous work, we prove some relation formulas for these series at positive integers which include our previous results corresponding to the cases of level 1 and 2. Furthermore, using these results, we evaluate certain two-variable analogues of level-N Eisenstein series.

1. Introduction

In [Tsumura 2008], we considered an analogue of the Eisenstein series defined by

$$\mathcal{G}_k(i) = \sum_{m\in\mathbb{Z}\setminus\{0\}} \sum_{n\in\mathbb{Z}} \frac{(-1)^n}{\sinh(m\pi)(m+ni)^k} \tag{1-1}$$

$$= \sum_{m\in\mathbb{Z}\setminus\{0\}} \sum_{n\in\mathbb{Z}} \frac{1}{\sinh((m+ni)\pi)(m+ni)^k} \quad (k\in\mathbb{N}), \tag{1-2}$$

where $i=\sqrt{-1}$ and $\sinh x = (e^x - e^{-x})/2$. We evaluated $\mathcal{G}_{2p-1}(i)$ $(p\in\mathbb{N})$ in terms of π and the lemniscate constant ϖ defined by

$$\varpi = 2\int_0^1 \frac{dx}{\sqrt{1-x^4}} = \frac{\Gamma(1/4)^2}{2\sqrt{2\pi}} = 2.6220575542921\ldots.$$

More precisely we gave

$$\mathcal{G}_{2p-1}(i) = \frac{2(-1)^p}{\pi} \sum_{j=1}^{p} (1-2^{1-2p+2j})\zeta(2p-2j)\big((-1)^j G_{2j}(i) - 2\zeta(2j)\big), \tag{1-3}$$

where $\zeta(s)$ is Riemann's zeta function and $G_{2j}(\tau)$ is the ordinary Eisenstein series defined by

$$G_{2j}(\tau) = \sum_{m\in\mathbb{Z}} \sum_{\substack{n\in\mathbb{Z}\\ (m,n)\neq(0,0)}} \frac{1}{(m+n\tau)^{2j}} \tag{1-4}$$

MSC2010: primary 11M41; secondary 11M99.
Keywords: Eisenstein series, Riemann zeta function, hyperbolic functions.

for $j \in \mathbb{N}$ and $\tau \in \mathbb{C}$ with $\operatorname{Im} \tau > 0$. Note that $G_2(\tau)$ is conditionally convergent with respect to the order of summation as above. We can view (1-3) as a double series analogue of the following formula given by Cauchy [1889] and Mellin [1902]:

$$\sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{(-1)^m}{\sinh(m\pi) m^{4k-1}} = \frac{2}{\pi} \sum_{j=0}^{2k} (1 - 2^{1-4k+2j}) \zeta(4k-2j)(-1)^j (2^{1-2j} - 1) \zeta(2j),$$

and similar formulas for the Dirichlet series involving hyperbolic functions; see, for example, [Berndt 1977; 1978; Meyer 2000].

As another type analogue of $G_{2j}(i)$, we considered

$$\mathfrak{C}_l^{\langle r \rangle}(i) = \sum_{\substack{m \in \mathbb{Z} \\ m \neq 0}} \sum_{n \in \mathbb{Z}} \frac{(\coth(m\pi))^r}{(m+ni)^l} \qquad (l \in \mathbb{N}_{\geq 3},\ r \in \mathbb{Z}_{\geq -1}), \tag{1-5}$$

where $\coth x = (e^x + e^{-x})/(e^x - e^{-x})$, and evaluated them in the case $l \equiv r \mod 2$; see [Tsumura 2009].

In [Komori et al. 2010], using a method completely different from the one in [Tsumura 2008; 2009], Komori, Matsumoto and the author evaluated

$$\mathcal{G}_k^{\langle r \rangle}(\tau) = \sum_{m \in \mathbb{Z} \setminus \{0\}} \sum_{n \in \mathbb{Z}} \frac{1}{\sinh((m+n\tau)\pi i/\tau)^r (m+n\tau)^k} \quad (r, k \in \mathbb{N}) \tag{1-6}$$

(and more generalized double series) for any $\tau \in \mathbb{C}$ with $\operatorname{Im} \tau > 0$.

In [Tsumura 2010] we considered analogues of level-2 Eisenstein series such as

$$\sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \frac{1}{\sinh((2m+1+(2n+1)i)\pi/2)(2m+1+(2n+1)i)^k}, \tag{1-7}$$

$$\sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \frac{1}{\cosh((2m+1+(2n+1)i)\pi/2)(2m+1+(2n+1)i)^k}, \tag{1-8}$$

$$\sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \frac{\tanh((2m+1+(2n+1)i)\pi/2)}{(2m+1+(2n+1)i)^l}, \tag{1-9}$$

$$\sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \frac{\coth((2m+1+(2n+1)i)\pi/2)}{(2m+1+(2n+1)i)^l} \tag{1-10}$$

for $k, l \in \mathbb{N}$ with $l \geq 3$, and evaluated them in terms of π and ϖ. Note that the level-N Eisenstein series is defined by

$$G_k(\tau; \underline{a} \bmod n) = \sum_{\substack{m \in \mathbb{Z} \\ m \equiv a_1 \bmod N}} \sum_{\substack{n \in \mathbb{Z} \\ n \equiv a_2 \bmod N \\ (m,n) \neq (0,0)}} \frac{1}{(m+n\tau)^k} \tag{1-11}$$

for $k \in \mathbb{N}_{\geq 2}$ and $\underline{a} = (a_1, a_2) \in \mathbb{Z}^2$ with $0 \leq a_1, a_2 < N$, which was studied by Hecke [Hecke 1937, Section 1] (see also, for example, [Koblitz 1993, Chapter III]).

In this paper, by developing the method used in [Tsumura 2008; 2009; 2010], we consider analogues of level-N Eisenstein series involving hyperbolic functions, namely

$$\mathfrak{C}_k^{\langle r\rangle}(\tau; \underline{a} \bmod n) = \sum_{\substack{m\in\mathbb{Z}\setminus\{0\} \\ m\equiv a_1 \bmod N}} \sum_{\substack{n\in\mathbb{Z} \\ n\equiv a_2 \bmod N}} \frac{\coth((m+n\tau)\pi i/N\tau)^r}{(m+n\tau)^k} \tag{1-12}$$

for $k \in \mathbb{N}_{\geq 2}$, $r \in \mathbb{Z}$ and $\underline{a} = (a_1, a_2) \in \mathbb{Z}^2$ with $0 \leq a_1, a_2 < N$. Note that (1-12) in the case $k = 2$ and $r = 2$ is conditionally convergent with respect to the order of summation as above. In fact, since $(\coth x)^2 = 1 + 1/(\sinh x)^2$, we have

$$\mathfrak{C}_2^{\langle 2\rangle}(\tau; \underline{a} \bmod n) = \sum_{\substack{m\in\mathbb{Z}\setminus\{0\} \\ m\equiv a_1 \bmod N}} \sum_{\substack{n\in\mathbb{Z} \\ n\equiv a_2 \bmod N}} \left(1 + \frac{1}{\sinh((m+n\tau)\pi i/N\tau)^2}\right) \frac{1}{(m+n\tau)^2}.$$

If we divide this double series into two parts, the first is conditionally convergent and the second is absolutely convergent. Considering $(\coth x)^{2\nu}$, we can inductively confirm that $\mathfrak{C}_2^{\langle 2\nu\rangle}(\tau; \underline{a} \bmod n)$ ($\nu \in \mathbb{N}$) is also conditionally convergent.

Outline of article. In Section 2, we state evaluation formulas for some quantities of the form (1-12) (see Theorem 2.1, whose proof is given in Section 3). We also evaluate (1-12) in terms of (1-11) and certain partial zeta values which will be defined by (2-4) (see Examples 2.5 and 2.6). This subsumes previous results on (1-5) corresponding to the case $(r, N) = (1, 1)$ [Tsumura 2009] and on (1-9) and (1-10) corresponding to the cases $(r, N) = (\pm 1, 2)$ [Tsumura 2010]. Here, for example, we give a new formula corresponding to the case $r = 2$:

$$\mathfrak{C}_4^{\langle 2\rangle}(i; (1, 1) \bmod 2) = -\frac{5\varpi^4 + 2\pi^3}{360}.$$

More generally, we give explicit formulas for level-N versions of these expressions (see Example 2.6). From these results, we evaluate the level-N version of (1-6), defined by

$$\mathcal{G}_k^{\langle r\rangle}(\tau; \underline{a} \bmod n) = \sum_{\substack{m\in\mathbb{Z}\setminus\{0\} \\ m\equiv a_1 \bmod N}} \sum_{\substack{n\in\mathbb{Z} \\ n\equiv a_2 \bmod N}} \frac{1}{\sinh((m+n\tau)\pi i/N\tau)^r (m+n\tau)^k} \tag{1-13}$$

(see Proposition 2.4; also Remark 3.9).

In Section 4, based on the results above, we evaluate a two-variable analogue of (1-11) defined by

$$\widetilde{G}_{j,k}(\tau;\ \underline{a} \bmod n) = \sum_{\substack{m\in\mathbb{Z}\\ m\equiv a_1 \bmod N}} \sum_{\substack{n\in\mathbb{Z}\\ n\equiv a_2 \bmod N\\ (m,n)\neq(0,0)}} \sum_{\substack{l\in\mathbb{Z}\\ l\equiv a_2 \bmod N\\ (m,l)\neq(0,0)}} \frac{1}{(m+l\tau)^j(m+n\tau)^k} \tag{1-14}$$

for $j, k \in \mathbb{N}_{\geq 2}$. Note that in the case $j = 2$ or $k = 2$, (1-14) is conditionally convergent with respect to the order of summation as above. We prove some relation formulas among $\widetilde{G}_{j,k}(\tau;\ \underline{a} \bmod n)$ and $\mathscr{G}_l^{\langle r\rangle}(\tau;\ \underline{a} \bmod n)$ (see Theorems 4.1 and 4.2), and evaluate $\widetilde{G}_{2p,2q}(i;\ \underline{a} \bmod n)$ (see Examples 4.3 and 4.4). For example, we obtain

$$\begin{aligned}\widetilde{G}_{4,4}(i; (1,1) \bmod 2) &= \sum_{\substack{m\in\mathbb{Z}\\ m\equiv 1 \bmod 2}} \sum_{\substack{n\in\mathbb{Z}\\ n\equiv 1 \bmod 2}} \sum_{\substack{l\in\mathbb{Z}\\ l\equiv 1 \bmod 2}} \frac{1}{(m+li)^4(m+ni)^4}\\ &= \frac{\varpi^8}{8960} - \frac{\varpi^4\pi^4}{17280} + \frac{\pi^7}{6048}.\end{aligned}$$

This paper contains a lot of examples of evaluation formulas. They were checked numerically using Mathematica 7.

2. Relation formulas for $\mathfrak{C}_k^{\nu}(\alpha)$

From now on, we set $N \in \mathbb{N}$, $\underline{a} = (a_1, a_2) \in \mathbb{Z}^2$ with $0 \leq a_1, a_2 < N$ and $\tau \in \mathbb{C}$ with $\operatorname{Im} \tau > 0$. For convenience, we set

$$\mathfrak{a} = \underline{a} \bmod N.$$

Theorem 2.1. *For $r \in \mathbb{Z}$ and $p \in \mathbb{N}$, we have*

$$\begin{aligned}&(N\tau)^{2p+1}\mathfrak{C}_{2p+1}^{\langle r+1\rangle}(\tau;\mathfrak{a})\\ &\qquad = \frac{2i}{\pi}\sum_{\omega=1}^{p}\zeta(2p-2\omega)(N\tau)^{2\omega+2}\mathfrak{C}_{2\omega+2}^{\langle r\rangle}(\tau;\mathfrak{a}) + 2\zeta(2p)\frac{(N\tau)^3}{\pi^2}\mathfrak{C}_3^{\langle r-1\rangle}(\tau;\mathfrak{a})\end{aligned} \tag{2-1}$$

and

$$(N\tau)^{2p+2}\mathfrak{C}_{2p+2}^{\langle r+1\rangle}(\tau;\mathfrak{a}) = \frac{2i}{\pi}\sum_{\omega=0}^{p}\zeta(2p-2\omega)(N\tau)^{2\omega+3}\mathfrak{C}_{2\omega+3}^{\langle r\rangle}(\tau;\mathfrak{a}). \tag{2-2}$$

We will prove this theorem in the next section. Note that if we know the values $\mathfrak{C}_3^{\langle -1\rangle}(\tau;\mathfrak{a})$ and $\mathfrak{C}_4^{\langle 0\rangle}(\tau;\mathfrak{a})$, then we can inductively evaluate $\mathfrak{C}_k^{\langle r\rangle}(\tau;\mathfrak{a})$ for $k \in \mathbb{N}_{\geq 3}$ and $r \in \mathbb{Z}_{\geq -1}$ with $k \equiv r \bmod 2$, as follows.

By the definition (1-12), we can see that

$$\mathfrak{E}_{2k}^{\langle 0\rangle}(\tau;\mathfrak{a})=\begin{cases} G_{2k}(\tau;\mathfrak{a}) & \text{if } a_1\neq 0,\\ G_{2k}(\tau;\mathfrak{a})-\tau^{-2k}\displaystyle\sum_{\substack{n\in\mathbb{Z}\\ n\equiv a_2 \bmod N}}\frac{1}{n^{2k}} & \text{if } a_1=0 \text{ and } a_2\neq 0,\\ N^{-2k}\big(G_{2k}(\tau)-2\tau^{-2k}\zeta(2k)\big) & \text{if } a_1=a_2=0, \end{cases}\tag{2-3}$$

for $k\in\mathbb{N}_{\geq 2}$. For simplicity, we define a certain partial zeta value by

$$\tilde{\zeta}(l;a \bmod N):=\sum_{\substack{n\in\mathbb{Z}\setminus\{0\}\\ n\equiv a \bmod N}}\frac{1}{n^l}\qquad (l\in\mathbb{N}_{\geq 2}).\tag{2-4}$$

The proof of the next proposition will be given in Section 3 as well.

Proposition 2.2. *With the same notation*,

$$\mathfrak{E}_{3}^{\langle -1\rangle}(\tau;\mathfrak{a})=\begin{cases} \dfrac{i\pi}{N\tau}G_2(\tau;\mathfrak{a}) & \textit{if } a_1\neq 0,\\ \dfrac{i\pi}{N\tau}\big(G_2(\tau;\mathfrak{a})-\tau^{-2}\tilde{\zeta}(2;a_2 \bmod N)\big) & \textit{if } a_1=0 \text{ and } a_2\neq 0,\\ \dfrac{i\pi}{N^3\tau}\big(G_2(\tau)-2\tau^{-2}\zeta(2)\big) & \textit{if } a_1=a_2=0. \end{cases}$$

From Theorem 2.1 and Proposition 2.2, we derive:

Theorem 2.3. *For $r\in\mathbb{Z}_{\geq -1}$ and $k\in\mathbb{N}_{\geq 3}$ with $k\equiv r \bmod 2$,*

$$\tau^k\pi^r\mathfrak{E}_k^{\langle r\rangle}(\tau;\mathfrak{a})\in\mathbb{Q}\big[\tau,\ \pi,\ \big\{\tilde{\zeta}(2j;a_2 \bmod N),\ G_{2j}(\tau;\mathfrak{a})\big\}_{j\in\mathbb{N}}\big].\tag{2-5}$$

Proof. We prove (2-5) by induction on $r\geq -1$. First we assume $r=-1$. Since $k\equiv r \bmod 2$ with $k\geq 3$, we can write $k=2p+3$ $(p\geq 0)$. Hence we further use induction on p. When $p=0$, namely $k=3$, we immediately obtain the assertion from Proposition 2.2. Furthermore, by (2-2) with $r=-1$, we have

$$\frac{\pi}{i}(N\tau)^{2p+3}\mathfrak{E}_{2p+3}^{\langle -1\rangle}(\tau;\mathfrak{a}) = -(N\tau)^{2p+2}\mathfrak{E}_{2p+2}^{\langle 0\rangle}(\tau;\mathfrak{a})+\frac{2i}{\pi}\sum_{\omega=0}^{p-1}\zeta(2p-2\omega)(N\tau)^{2\omega+3}\mathfrak{E}_{2\omega+3}^{\langle -1\rangle}(\tau;\mathfrak{a}).$$

Hence, by (2-3), we obtain the assertion by induction on p in the case $r=-1$.

Next we assume that the induction hypotheses hold for r. By multiplying the both sides of (2-1) and of (2-2) by π^{r+1}, we obtain the assertion in the case of $r+1$. Thus we complete the proof. □

As we noted in Section 1, using the relation $1/(\sinh x)^2=(\coth x)^2-1$ and the binomial theorem, we have the following relation between $\mathfrak{E}_k^{\langle r\rangle}(\tau;\mathfrak{a})$ and $\mathcal{G}_k^{\langle r\rangle}(\tau;\mathfrak{a})$ defined by (1-13).

Proposition 2.4. *For $\nu \in \mathbb{N}$,*

$$\mathscr{G}_k^{\langle 2\nu\rangle}(\tau;\mathfrak{a}) = \sum_{j=0}^{\nu} \binom{\nu}{j} (-1)^{\nu-j} \mathfrak{C}_k^{\langle 2j\rangle}(\tau;\mathfrak{a}). \tag{2-6}$$

Therefore, for $l \in \mathbb{N}$ and $\nu \in \mathbb{N}$,

$$\tau^{2l}\pi^{2\nu}\mathscr{G}_{2l}^{\langle 2\nu\rangle}(\tau;\mathfrak{a}) \in \mathbb{Q}\big[\tau,\ \pi,\ \big\{\tilde{\zeta}(2k;a_2 \bmod N),\ G_{2k}(\tau;\mathfrak{a})\big\}_{k\in\mathbb{N}}\big]. \tag{2-7}$$

Hence we can evaluate $\mathscr{G}_{2l}^{\langle 2\nu\rangle}(\tau;\mathfrak{a})$ by using the result on $\mathfrak{C}_{2l}^{\langle 2j\rangle}(\tau;\mathfrak{a})$ (see below). We will consider $\mathscr{G}_{2l+1}^{\langle 2\nu+1\rangle}(\tau;\mathfrak{a})$ in Remark 3.9.

Example 2.5. In the case $N = 1$, we simply denote (1-12) by $\mathfrak{C}_k^{\langle r\rangle}(\tau)$. Then, combining Theorem 2.1, Proposition 2.2 and (2-3), we obtain

$$\begin{aligned}
\mathfrak{C}_3^{\langle -1\rangle}(\tau) &= i(-\pi^3+3\pi\tau^2G_2(\tau))/(3\tau^3),\\
\mathfrak{C}_5^{\langle -1\rangle}(\tau) &= i(-2\pi^5+5\pi^3\tau^2G_2(\tau)+15\pi\tau^4G_4(\tau))/(15\tau^5),\\
\mathfrak{C}_3^{\langle 1\rangle}(\tau) &= i(-4\pi^4+15\tau^2G_2(\tau)\pi^2-45\tau^4G_4(\tau))/(45\tau^3\pi),\\
\mathfrak{C}_5^{\langle 1\rangle}(\tau) &= i(-4\pi^6+7\tau^2G_2(\tau)\pi^4+105\tau^4G_4(\tau)\pi^2-315\tau^6G_6(\tau))/(315\tau^5\pi),\\
\mathfrak{C}_4^{\langle 2\rangle}(\tau) &= (16\pi^6-84\tau^2G_2(\tau)\pi^4+630\tau^4G_4(\tau)\pi^2-945\tau^6G_6(\tau))/(945\tau^4\pi^2),\\
\mathfrak{C}_6^{\langle 2\rangle}(\tau) &= (64\pi^8-180\tau^2G_2(\tau)\pi^6-945\tau^4G_4(\tau)\pi^4+9450\tau^6G_6(\tau)\pi^2\\
&\qquad\qquad -14175\tau^8G_8(\tau))/(14175\tau^6\pi^2),\\
\mathfrak{C}_3^{\langle 3\rangle}(\tau) &= i(-44\pi^6+189\tau^2G_2(\tau)\pi^4-945\tau^4G_4(\tau)\pi^2+945\tau^6G_6(\tau))/(945\tau^3\pi^3),\\
\mathfrak{C}_5^{\langle 3\rangle}(\tau) &= i(-4\pi^8-45\tau^2G_2(\tau)\pi^6+1260\tau^4G_4(\tau)\pi^4-4725\tau^6G_6(\tau)\pi^2\\
&\qquad\qquad +4725\tau^8G_8(\tau))/(4725\tau^5\pi^3),\\
\mathfrak{C}_4^{\langle 4\rangle}(\tau) &= (208\pi^8-1080\tau^2G_2(\tau)\pi^6+8505\tau^4G_4(\tau)\pi^4-18900\tau^6G_6(\tau)\pi^2\\
&\qquad\qquad +14175\tau^8G_8(\tau))/(14175\tau^4\pi^4),\\
\mathfrak{C}_6^{\langle 4\rangle}(\tau) &= (1024\pi^{10}-2376\tau^2G_2(\tau)\pi^8-30690\tau^4G_4(\tau)\pi^6+270270\tau^6G_6(\tau)\pi^4\\
&\qquad\qquad -623700\tau^8G_8(\tau)\pi^2+467775\tau^{10}G_{10}(\tau))/(467775\tau^6\pi^4).
\end{aligned}$$

The case $\tau = i$ was studied in [Tsumura 2009], and we recover the results found there. For example,

$$\begin{aligned}
\mathfrak{C}_4^{\langle 2\rangle}(i) &= \frac{42\varpi^4+16\pi^4-84\pi^3}{945},\\
\mathfrak{C}_4^{\langle 4\rangle}(i) &= \frac{27\varpi^8+567\varpi^4\pi^4+208\pi^8-1080\pi^7}{14175\pi^4}.
\end{aligned}$$

By Proposition 2.4, we can inductively evaluate $\mathscr{G}_{2\nu}^{\langle 2l\rangle}(\tau)$ in terms of $G_{2j}(\tau)$ and $\zeta(2k)$. This fact was already given in [Komori et al. 2010] by a totally different method. Here we recover, for example,

$$\mathscr{G}_4^{\langle 2\rangle}(i) = \frac{-21\varpi^4 + 37\pi^4 - 84\pi^3}{945}, \tag{2-8}$$

$$\mathscr{G}_4^{\langle 4\rangle}(i) = \frac{27\varpi^8 + 252\varpi^4\pi^4 - 587\pi^8 + 1440\pi^7}{14175\pi^4}. \tag{2-9}$$

Next we consider the case $\tau = \rho = e^{2\pi i/3}$. We recall the properties of $G_{2k}(\rho)$. For the details, see [Koblitz 1993; Nesterenko and Philippon 2001; Serre 1970; Waldschmidt 1999]; also [Komori et al. 2010]. Let

$$\widetilde{\varpi} = \frac{\Gamma(1/3)^3}{2^{4/3}\pi} = 2.42865064788758\cdots \tag{2-10}$$

which is an analogue of the lemniscate constant ϖ. Then we obtain $G_2(\rho) = 2\pi\rho/\sqrt{3}$,

$$G_6(\rho) = \frac{\widetilde{\varpi}^6}{35}, \quad G_{12}(\rho) = \frac{\widetilde{\varpi}^{12}}{7007}, \quad G_{18}(\rho) = \frac{\widetilde{\varpi}^{18}}{1440257}, \ldots \tag{2-11}$$

and $G_k(\rho) = 0$ for $k \geq 3$ with $6 \nmid k$. Using these results, we can evaluate $\mathfrak{C}_k^{\langle r\rangle}(\rho)$, similarly to the case $\tau = i$, for example,

$$\mathfrak{C}_4^{\langle 2\rangle}(\rho) = \frac{-27\widetilde{\varpi}^6 + 16\pi^6 - 56\sqrt{3}\pi^5}{945\rho\pi^2},$$

$$\mathfrak{C}_4^{\langle 4\rangle}(\rho) = \frac{-18900\widetilde{\varpi}^6\pi^2 + 7280\pi^8 - 25200\sqrt{3}\pi^7}{496125\rho\pi^4}.$$

From these results, we recover these formulas from [Komori et al. 2010]:

$$\mathscr{G}_4^{\langle 2\rangle}(\rho) = \frac{-27\widetilde{\varpi}^6 + 37\pi^6 - 56\sqrt{3}\pi^5}{945\rho\pi^2}, \tag{2-12}$$

$$\mathscr{G}_4^{\langle 4\rangle}(\rho) = \frac{270\widetilde{\varpi}^6 - 587\pi^6 + 960\sqrt{3}\pi^5}{14175\rho\pi^2}. \tag{2-13}$$

Example 2.6. We consider the case $N > 1$, $a_1 \neq 0$ and $a_2 \neq 0$. We simply denote the level-N Eisenstein series by $G_{2j}^{\underline{a}}(\tau)$ instead of $G_{2j}(\tau;\mathfrak{a})$. Then we have the following formulas which are explicit examples of the main result in this paper:

$$\mathfrak{C}_3^{\langle -1\rangle}(\tau;\mathfrak{a}) = iG_2^{\underline{a}}(\tau)\pi/(N\tau),$$

$$\mathfrak{C}_5^{\langle -1\rangle}(\tau;\mathfrak{a}) = i(G_2^{\underline{a}}(\tau)\pi^3 + 3N^2\tau^2G_4^{\underline{a}}(\tau)\pi)/(3N^3\tau^3),$$

$$\mathfrak{C}_3^{\langle 1\rangle}(\tau;\mathfrak{a}) = i(G_2^{\underline{a}}(\tau)\pi^2 - 3N^2\tau^2G_4^{\underline{a}}(\tau))/(3N\tau\pi),$$

$$\mathfrak{C}_5^{\langle 1\rangle}(\tau;\mathfrak{a}) = i(G_2^{\underline{a}}(\tau)\pi^4 + 15iN^2\tau^2 G_4^{\underline{a}}(\tau)\pi^2 - 45N^4\tau^4 G_6^{\underline{a}}(\tau))/(45N^3\tau^3\pi),$$

$$\mathfrak{C}_4^{\langle 2\rangle}(\tau;\mathfrak{a}) = (-4G_2^{\underline{a}}(\tau)\pi^4 + 30N^2\tau^2 G_4^{\underline{a}}(\tau)\pi^2 - 45N^4\tau^4 G_6^{\underline{a}}(\tau))/(45N^2\tau^2\pi^2),$$

$$\mathfrak{C}_6^{\langle 2\rangle}(\tau;\mathfrak{a}) = (-4G_2^{\underline{a}}(\tau)\pi^6 - 21N^2\tau^2 G_4^{\underline{a}}(\tau)\pi^4 + 210N^4\tau^4 G_6^{\underline{a}}(\tau)\pi^2 - 315N^6\tau^6 G_8^{\underline{a}}(\tau))/(315N^4\tau^4\pi^2),$$

$$\mathfrak{C}_3^{\langle 3\rangle}(\tau;\mathfrak{a}) = i(9G_2^{\underline{a}}(\tau)\pi^4 - 45N^2\tau^2 G_4^{\underline{a}}(\tau)\pi^2 + 45N^4\tau^4 G_6^{\underline{a}}(\tau))/(45N\tau\pi^3),$$

$$\mathfrak{C}_5^{\langle 3\rangle}(\tau;\mathfrak{a}) = i(-G_2^{\underline{a}}(\tau)\pi^6 + 28N^2\tau^2 G_4^{\underline{a}}(\tau)\pi^4 - 105N^4\tau^4 G_6^{\underline{a}}(\tau)\pi^2 + 105N^6\tau^6 G_8^{\underline{a}}(\tau))/(105N^3\tau^3\pi^3),$$

$$\mathfrak{C}_4^{\langle 4\rangle}(\tau;\mathfrak{a}) = (-8G_2^{\underline{a}}(\tau)\pi^6 + 63N^2\tau^2 G_4^{\underline{a}}(\tau)\pi^4 - 140N^4\tau^4 G_6^{\underline{a}}(\tau)\pi^2 + 105N^6\tau^6 G_8^{\underline{a}}(\tau))/(105N^2\tau^2\pi^4),$$

$$\mathfrak{C}_6^{\langle 4\rangle}(\tau;\mathfrak{a}) = (-24G_2^{\underline{a}}(\tau)\pi^8 - 310N^2\tau^2 G_4^{\underline{a}}(\tau)\pi^6 + 2730N^4\tau^4 G_6^{\underline{a}}(\tau)\pi^4 - 6300N^6\tau^6 G_8^{\underline{a}}(\tau)\pi^2 + 4725N^8\tau^8 G_{10}^{\underline{a}}(\tau))/(4725N^4\tau^4\pi^4).$$

In [Tsumura 2010], we studied the case when $(N, a_1, a_2, \tau) = (2, 1, 1, i)$ and $r = \pm 1$, based on [Katayama 1978]. In this case, as mentioned in both of these papers, we see $G_2^{(1,1)}(i) = -\pi/4$, $G_{4k+2}^{(1,1)}(i) = 0$ and $G_{4k}^{(1,1)}(i) \in \mathbb{Q}\cdot\varpi^{4k}$ $(k \in \mathbb{N})$, which can be concretely calculated; for example,

$$G_4^{(1,1)}(i) = -\frac{\varpi^4}{48}, \quad G_8^{(1,1)}(i) = \frac{\varpi^8}{8960}, \quad G_{12}^{(1,1)}(i) = -\frac{\varpi^{12}}{1689600}.$$

Hence, by the formulas above, we can explicitly evaluate $\mathfrak{C}_k^{\langle r\rangle}(\tau; (1,1) \bmod 2)$ when $k \equiv r \bmod 2$. In particular, when $r = \pm 1$, these coincide with the results given in [Tsumura 2010]. As examples in the cases $r = 2, 4$, we give

$$\mathfrak{C}_4^{\langle 2\rangle}(i; (1,1) \bmod 2) = -\frac{5\varpi^4 + 2\pi^3}{360},$$

$$\mathfrak{C}_4^{\langle 4\rangle}(i; (1,1) \bmod 2) = \frac{3\varpi^8 - 21\varpi^4\pi^4 - 8\pi^7}{1680\pi^4},$$

and

$$\mathcal{G}_4^{\langle 2\rangle}(i; (1,1) \bmod 2) = \frac{5\varpi^8 - 4\pi^3}{720}, \tag{2-14}$$

$$\mathcal{G}_4^{\langle 4\rangle}(i; (1,1) \bmod 2) = \frac{9\varpi^8 - 28\varpi^4\pi^4 + 32\pi^7}{5040\pi^4}. \tag{2-15}$$

3. Proofs of Theorem 2.1 and Proposition 2.2

For $\underline{a} = (a_1, a_2) \in \mathbb{Z}^2$ with $0 \le a_1, a_2 < N$, we set $\beta = (a_1 + a_2\tau)/N$ for simplicity. We fix a small $\varepsilon > 0$. For $u \in [1, 1+\varepsilon]$, $r \in \mathbb{Z}$ and $k \in \mathbb{N}$, we define

$$\widehat{\mathfrak{D}}_k^{\langle r\rangle}(\tau;\beta;u)=\sideset{}{^*}\sum_{m\in\mathbb{Z}}\sum_{n=1}^{\infty}\frac{u^{-n}\coth((m+\beta+n\tau)\pi i/\tau)^r}{\sinh((m+\beta+n\tau)\pi i/\tau)(m+\beta+n\tau)^k}+\sideset{}{^*}\sum_{m\in\mathbb{Z}}\sum_{n=1}^{\infty}\frac{u^{-n}\coth((m+\beta-n\tau)\pi i/\tau)^r}{\sinh((m+\beta-n\tau)\pi i/\tau)(m+\beta-n\tau)^k}+\sideset{}{^*}\sum_{m\in\mathbb{Z}}\frac{\coth((m+\beta)\pi i/\tau)^r}{\sinh((m+\beta)\pi i/\tau)(m+\beta)^k}, \tag{3-1}$$

where $\sum^*_{m\in\mathbb{Z}}$ stands for the sum over $m\in\mathbb{Z}\setminus\{0\}$ if $a_1=0$ and over $m\in\mathbb{Z}$ if $a_1\neq 0$.

When $u>1$, we define $\widehat{\mathfrak{D}}_{-k}^{\langle r\rangle}(\tau;\beta;u)$ for $k\in\mathbb{Z}_{\geq 0}$ by (3-1). This is well-defined in the following sense. Since $\sinh(x)=0$ implies $x\in\pi i\mathbb{Z}$, the equality

$$\sinh\Bigl((m+\beta+n\tau)\frac{\pi i}{\tau}\Bigr)=\sinh\Bigl(\bigl(Nm+a_1+(Nn+a_2)\,\tau\bigr)\frac{\pi i}{N\tau}\Bigr)=0\quad (m,n\in\mathbb{Z})$$

implies $(a_1,a_2)=(0,0)$ and $m=0$. Similarly, $\cosh(x)=0$ implies $x\in\pi i/2+\pi i\mathbb{Z}$, so the equality

$$\cosh\Bigl((m+\beta+n\tau)\,\frac{\pi i}{\tau}\Bigr)=\cosh\Bigl(\bigl(Nm+a_1+(Nn+a_2)\,\tau\bigr)\frac{\pi i}{N\tau}\Bigr)=0\quad (m,n\in\mathbb{Z})$$

implies $(a_1,a_2)=(0,N/2)$ and $m=0$. Hence, by the definition of $\sum^*$ a few lines above, we see that (3-1) is absolutely convergent under the conditions above, that is, well-defined.

Since $\cosh(n\pi i)=(-1)^n$ and $\sinh(n\pi i)=0$, we can rewrite (3-1) as

$$\widehat{\mathfrak{D}}_k^{\langle r\rangle}(\tau;\beta;u)=\sideset{}{^*}\sum_{m\in\mathbb{Z}}\frac{\coth((m+\beta)\pi i/\tau)^r}{\sinh((m+\beta)\pi i/\tau)}\times\Biggl(\sum_{n=1}^{\infty}(-u)^{-n}\biggl(\frac{1}{(m+\beta+n\tau)^k}+\frac{1}{(m+\beta-n\tau)^k}\biggr)+\frac{1}{(m+\beta)^k}\Biggr). \tag{3-2}$$

When $k\geq 2$, we see that $\widehat{\mathfrak{D}}_k^{\langle r\rangle}(\tau;\beta;u)$ converges absolutely and uniformly for u in $[1,1+\varepsilon]$. Furthermore, when $k=1$, we have

$$\widehat{\mathfrak{D}}_1^{\langle r\rangle}(\tau;\beta;u)=\sideset{}{^*}\sum_{m\in\mathbb{Z}}\frac{2(m+\beta)\coth((m+\beta)\pi i/\tau)^r}{\sinh((m+\beta)\pi i/\tau)}\sum_{n=1}^{\infty}\frac{(-u)^{-n}}{(m+\beta)^2-n^2\tau^2}+\sideset{}{^*}\sum_{m\in\mathbb{Z}}\frac{\coth((m+\beta)\pi i/\tau)^r}{\sinh((m+\beta)\pi i/\tau)(m+\beta)}, \tag{3-3}$$

which converges absolutely and uniformly for u in $[1,1+\varepsilon]$. Hence, for any $k\in\mathbb{N}$, we have

$$\lim_{u\to 1}\widehat{\mathfrak{D}}_k^{\langle r\rangle}(\tau;\beta;u)=\widehat{\mathfrak{D}}_k^{\langle r\rangle}(\tau;\beta;1) =\sum_{m\in\mathbb{Z}}^{*}\sum_{n\in\mathbb{N}}\frac{\coth((m+\beta+n\tau)\pi i/\tau)^r}{\sinh((m+\beta+n\tau)\pi i/\tau)(m+\beta+n\tau)^k}. \tag{3-4}$$

Now we let

$$\mathcal{S}_r(\theta;\tau;\beta)=\sum_{m\in\mathbb{Z}}^{*}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)}. \tag{3-5}$$

Set $A=\mathrm{Re}(i/\tau)$ and $B=\mathrm{Im}(i/\tau)$. Then $A>0$ because $\mathrm{Im}\,\tau>0$. We further let $D(R):=\{\theta\in\mathbb{C}:|\theta|<R\}$ be the closed disk of radius R, where $R>0$.

Lemma 3.1. *$\mathcal{S}_r(\theta;\tau;\beta)$ converges absolutely for $\theta\in D\,(A\pi/(A+|B|))$.*

Proof. Let $\theta\in D(A\pi/(A+|B|))$ and set $(a,b)=(\mathrm{Re}\,\theta,\mathrm{Im}\,\theta)$. Then

$$|a|,\ |b|<\frac{A\pi}{A+|B|}. \tag{3-6}$$

Here we consider the order of $\mathcal{S}_r(\theta;\tau;\beta)$, namely

$$\mathcal{S}_r(\theta;\tau;\beta)=O\left(e^{|m|\,\mathrm{Re}((\pm\theta-\pi)i/\tau)}\right)\qquad(|m|\to\infty),$$

which implies the maximum of two cases corresponding to $\pm\theta$. By (3-6), we have

$$\begin{aligned}\mathrm{Re}((\pm\theta-\pi)i/\tau)&=\mathrm{Re}((\pm a-\pi\pm bi)(A+Bi))=(\pm a-\pi)A\mp bB\\ &\le(|a|-\pi)A+|b||B|<\left(\frac{A\pi}{A+|B|}-\pi\right)A+\frac{A|B|\pi}{A+|B|}=0.\end{aligned}$$

Therefore we have the assertion. □

As in [Tsumura 2008, § 2], we set

$$\mathfrak{H}(\theta;u):=-\frac{1}{2}\left(\frac{e^{\theta}}{e^{\theta}+u}+\frac{e^{-\theta}}{e^{-\theta}+u}\right) \tag{3-7}$$

for $\theta\in\mathbb{C}$ and $u\in[1,1+\varepsilon]$. This function is holomorphic for $\theta\in D(\pi)$, and satisfies

$$\lim_{u\to 1}\mathfrak{H}(\theta;u)=-\tfrac{1}{2}\qquad(\theta\in D(\pi)). \tag{3-8}$$

We also set

$$J_r(\theta;\tau;\beta;u):=\mathcal{S}_r(\theta;\tau;\beta)(2\mathfrak{H}(i\theta;u)+1). \tag{3-9}$$

Since $A\pi/(A+|B|)\le\pi$, it follows from Lemma 3.1 that $J_r(\theta;\tau;\beta;u)$ is holomorphic for $\theta\in D(A\pi/(A+|B|))$. Hence, for each $u\in[1,1+\varepsilon]$, we can expand $J_r(\theta;\tau;\beta;u)$ as

$$J_r(\theta;\tau;\beta;u)=\sum_{n=0}^{\infty}\Lambda_n^{\langle r\rangle}(\tau;\beta;u)\frac{\theta^n}{n!}\quad(\theta\in D(A\pi/(A+|B|))). \tag{3-10}$$

By Cauchy's integral theorem, for any $\gamma\in\mathbb{R}$ with $0<\gamma<A\pi/(A+|B|)$, we have

$$\frac{|\Lambda_n^{\langle r\rangle}(\tau;\beta;u)|}{n!}\le\frac{1}{2\pi}\int_{C_\gamma}|J_r(\theta;\tau;\beta;u)||z|^{-n-1}|dz|\le\frac{M_\gamma}{\gamma^n}\quad(n\in\mathbb{Z}_{\ge 0}), \tag{3-11}$$

where $C_\gamma : z=\gamma e^{it}$ $(0\le t\le 2\pi)$ and

$$M_\gamma:=\max_{(z,u)\in C_\gamma\times[1,1+\varepsilon]}|J_r(z;\tau;\beta;u)|.$$

Hence the right-hand side of (3-10) is uniformly convergent in $u\in[1,1+\varepsilon]$ if $\theta\in D(A\pi/(A+|B|))$. By (3-8) and (3-9), we have $J_r(\theta;\tau;\beta;u)\to 0$ as $u\to 1$. Therefore we see that

$$\Lambda_n^{\langle r\rangle}(\tau;\beta;u)\to 0\quad(u\to 1;\ n\in\mathbb{Z}_{\ge 0}). \tag{3-12}$$

Lemma 3.2. *For $u\in(1,1+\varepsilon]$ and $\theta\in D(A\pi/(A+|B|))$,*

$$J_r(\theta;\tau;\beta;u)=\sum_{j=0}^{\infty}\widehat{\mathfrak{D}}_{-j}^{\langle r\rangle}(\tau;\beta;u)\frac{\theta^j}{j!}, \tag{3-13}$$

that is, $\widehat{\mathfrak{D}}_{-j}^{\langle r\rangle}(\tau;\beta;u)=\Lambda_j^{\langle r\rangle}(\tau;\beta;u)$, for $j\in\mathbb{Z}_{\ge 0}$.

Proof. When $u>1$, from (3-7), we have (see [Tsumura 2008, Lemma 2.1])

$$\mathfrak{H}(i\theta;u)=\sum_{n=1}^{\infty}(-u)^{-n}\cos(n\theta).$$

Therefore, from (3-5) and (3-9), we have

$$\begin{aligned}J_r(\theta;\tau;\beta;u)&=\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n=1}^{\infty}(-u)^{-n}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta+n\tau)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)}\\&\quad+\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n=1}^{\infty}(-u)^{-n}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta-n\tau)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)}\\&\quad+\sum_{m\in\mathbb{Z}}{}^{*}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)}.\end{aligned} \tag{3-14}$$

Using the Maclaurin expansion of e^x and the definition of $\widehat{\mathfrak{D}}_{-k}^{\langle r\rangle}(\tau;\beta;u)$ in (3-1), namely in (3-2), we complete the proof. □

Lemma 3.3. *For $r \in \mathbb{Z}$ and $k \in \mathbb{N}$,*

$$N^{k+2}\mathfrak{C}^{\langle r\rangle}_{k+2}(\tau;\mathfrak{a}) = \sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}} \frac{\coth((m+\beta+n\tau)\pi i/\tau)^r}{(m+\beta+n\tau)^{k+2}} = \sum_{j=0}^{[k/2]} \widehat{\mathfrak{D}}^{\langle r\rangle}_{k+1-2j}(\tau;\beta;1)\frac{(i\pi/\tau)^{2j+1}}{(2j+1)!}. \tag{3-15}$$

Proof. The first equality comes from the definition (1-12) and $\beta = (a_1 + a_2\tau)/N$. We prove the second equality. We first assume $k \in \mathbb{Z}_{\geq 0}$. For $u \in [1, 1+\varepsilon]$, we set

$$\begin{aligned}\Phi_r(\theta;k;\tau;\beta;u) &= \sum_{m\in\mathbb{Z}}{}^{*}\sum_{n=1}^{\infty}(-u)^{-n}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta+n\tau)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)(m+\beta+n\tau)^{k+2}} \\ &\quad+\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n=1}^{\infty}(-u)^{-n}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta-n\tau)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)(m+\beta-n\tau)^{k+2}} \\ &\quad+\sum_{m\in\mathbb{Z}}{}^{*}\frac{\coth((m+\beta)\pi i/\tau)^r e^{(m+\beta)i\theta/\tau}}{\sinh((m+\beta)\pi i/\tau)(m+\beta)^{k+2}},\end{aligned} \tag{3-16}$$

which converges absolutely and uniformly in $u \in [1, 1+\varepsilon]$ if $\theta \in D(A\pi/(A+|B|))$. If $u > 1$, it follows from Lemma 3.2 that

$$\begin{aligned}\Phi_r(\theta;k;\tau;\beta;u) &= \sum_{j=0}^{\infty}\widehat{\mathfrak{D}}^{\langle r\rangle}_{k+2-j}(\tau;\beta;u)\frac{(i\theta/\tau)^j}{j!} \\ &= \sum_{j=0}^{k+1}\widehat{\mathfrak{D}}^{\langle r\rangle}_{k+2-j}(\tau;\beta;u)\frac{(i\theta/\tau)^j}{j!} + \sum_{j=k+2}^{\infty}\Lambda^{\langle r\rangle}_{j-k-2}(\tau;\beta;u)\frac{(i\theta/\tau)^j}{j!}.\end{aligned} \tag{3-17}$$

By considering

$$\lim_{u\to 1}\tfrac{1}{2}\left\{\Phi_r(\theta;k;\tau;\beta;u) - \Phi_r(-\theta;k;\tau;\beta;u)\right\},$$

and using (3-4) and (3-11), we can let $u \to 1$ on the both sides of (3-17) if θ lies in $D(A\pi/(A+|B|))$. By (3-12), we have

$$\frac{1}{2}\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\frac{(-1)^n\coth((m+\beta)\pi i/\tau)^r\left(e^{(m+\beta+n\tau)i\theta/\tau} - e^{-(m+\beta+n\tau)i\theta/\tau}\right)}{\sinh((m+\beta)\pi i/\tau)(m+\beta+n\tau)^{k+2}} = \sum_{\nu=0}^{[k/2]}\widehat{\mathfrak{D}}^{\langle r\rangle}_{k+1-2\nu}(\tau;\beta;1)\frac{(i\theta/\tau)^{2\nu+1}}{(2\nu+1)!} \tag{3-18}$$

for $\theta \in D(A\pi/(A+|B|))$. Moreover, we claim that the left-hand side of (3-18) is absolutely convergent on the region $\Omega(\tau) := \bigcup_{n\geq 1} \mathfrak{X}_n(\tau)$, where

$$\mathfrak{X}_n(\tau) = \left\{\theta \in \mathbb{C} : \left|\theta - \left(1-\frac{1}{n}\right)\pi\right| < \frac{A\pi}{(A+|B|)n}\right\}.$$

Actually we know that the left-hand side of (3-18) is

$$O\left(e^{|m|\mathrm{Re}((\pm\theta-\pi)i/\tau)}|m+\beta+n\tau|^{-k-2}\right) \qquad (|m|, |n| \to \infty).$$

Hence we aim to prove $\mathrm{Re}((\pm\theta-\pi)i/\tau) < 0$ for any $\theta \in \Omega(\tau)$. In fact, for any n and any $\theta \in \mathfrak{X}_n$, we set $(a,b) = (\mathrm{Re}\,\theta, \mathrm{Im}\,\theta)$. Then

$$|a| < \left(1-\frac{1}{n}\right)\pi + \frac{A\pi}{(A+|B|)n} \quad \text{and} \quad |b| < \frac{A\pi}{(A+|B|)n}.$$

Hence, by recalling that $A = \mathrm{Re}(i/\tau)$ and $B = \mathrm{Im}(i/\tau)$, we obtain the claim, since

$$\begin{aligned}\mathrm{Re}((\pm\theta-\pi)i/\tau) &= \mathrm{Re}((\pm a-\pi\pm bi)(A+Bi)) = (\pm a-\pi)A \mp Bb \\ &\leq (|a|-\pi)A+|B|\,|b| < -\frac{A\pi}{n}+\frac{A^2\pi}{(A+|B|)n}+\frac{A|B|\pi}{(A+|B|)n} = 0.\end{aligned}$$

On the other hand, it is clear that the right-hand side of (3-18) is holomorphic for $\theta \in \Omega(\tau)$, so (3-18) holds for $\theta \in \Omega(\tau)$.

Finally we claim that $\Omega(\tau) \supset [(1-1/L)\pi, \pi)$, where $L = \max(1, |B|/2A)$. In order to prove this, we only have to prove $\mathfrak{X}_n(\tau) \cap \mathfrak{X}_{n+1}(\tau) \neq \varnothing$ for all $n \geq L$, because any $\mathfrak{X}_n(\tau)$ is the disk whose center is on the real axis. More precisely, we have to prove

$$\left(1-\frac{1}{n}\right)\pi + \frac{A\pi}{(A+|B|)n} \geq \left(1-\frac{1}{n+1}\right)\pi - \frac{A\pi}{(A+|B|)(n+1)},$$

if $n \geq L$. In fact, this can be easily verified. Hence we obtain the claim. Therefore (3-18) holds for $\theta \in [(1-1/L)\pi, \pi)$. If we set $\theta = \pi$ on the left-hand side of (3-18), we have

$$\sum_{m\in\mathbb{Z}}^{*} \sum_{n\in\mathbb{Z}} \frac{\coth((m+\beta+n\tau)\pi i/\tau)^r}{(m+\beta+n\tau)^{k+2}},$$

which is absolutely convergent if $k \geq 1$. Hence, by Abel's theorem, (3-18) holds for $\theta = \pi$, which implies (3-15). Thus we complete the proof. □

Remark 3.4. As stated in the proof, (3-18) holds for $k = 0$ if $\theta \in [(1-1/L)\pi, \pi)$, because the left-hand side of (3-18) converges absolutely even if $k = 0$ and θ is in $[(1-1/L)\pi, \pi)$. We claim that (3-18) holds for $\theta = \pi$ when $(k, r) = (0, 0)$.

In fact, by setting $(k, r, \theta) = (0, 0, \pi)$ on the left-hand side of (3-18), we have

$$\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\frac{1}{(m+\beta+n\tau)^2} = N^2 G_2(\tau;\mathfrak{a}) - \frac{\delta_{a_1,0}N^2}{\tau^2}\times\begin{cases} N^2\tilde{\zeta}(2; a_2 \bmod N) & \text{if } a_2 \neq 0,\\ 2\zeta(2) & \text{if } a_2 = 0,\end{cases}$$

where $\delta_{p,q}$ is the Kronecker delta. Therefore it follows from Abel's theorem that (3-18) holds for $k = 0$ and $\theta = \pi$. Hence we obtain

$$\widehat{\mathfrak{D}}_1^{\langle 0\rangle}(\tau;\beta;1) = \frac{N^2\tau}{i\pi}G_2(\tau;\mathfrak{a}) - \frac{\delta_{a_1,0}}{i\pi\tau}\begin{cases} N^2\tilde{\zeta}(2; a_2 \bmod N) & \text{if } a_2 \neq 0,\\ 2\zeta(2) & \text{if } a_2 = 0.\end{cases} \tag{3-19}$$

For $k \in \mathbb{N}$, we differentiate (3-18) in $\theta \in [(1 - 1/L)\pi, \pi)$. Then

$$\frac{1}{2}\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\frac{(-1)^n\coth((m+\beta)\pi i/\tau)^r\left(e^{(m+\beta+n\tau)i\theta/\tau}+e^{-(m+\beta+n\tau)i\theta/\tau}\right)}{\sinh((m+\beta)\pi i/\tau)(m+\beta+n\tau)^{k+1}} = \sum_{\nu=0}^{[k/2]}\widehat{\mathfrak{D}}_{k+1-2\nu}^{\langle r\rangle}(\tau;\beta;1)\frac{(i\theta/\tau)^{2\nu}}{(2\nu)!}. \tag{3-20}$$

If $k \geq 2$, both sides on (3-20) converge absolutely and uniformly in $[(1-1/L)\pi, \pi]$. Hence, by letting $\theta \to \pi$, we have:

Lemma 3.5. *For $r \in \mathbb{Z}$ and $k \in \mathbb{N}$ with $k \geq 2$,*

$$N^{k+1}\mathfrak{C}_{k+1}^{\langle r+1\rangle}(\tau;\mathfrak{a}) = \sum_{j=0}^{[k/2]}\widehat{\mathfrak{D}}_{k+1-2j}^{\langle r\rangle}(\tau;\beta;1)\frac{(i\pi/\tau)^{2j}}{(2j)!}. \tag{3-21}$$

Letting $k = 2p + \mu$ for $p \in \mathbb{N}$ and $\mu \in \{0, 1\}$ in (3-15) and (3-21), we have

$$N^{2p+2+\mu}\mathfrak{C}_{2p+2+\mu}^{\langle r\rangle}(\tau;\mathfrak{a}) = \sum_{j=0}^{p}\widehat{\mathfrak{D}}_{2p+1+\mu-2j}^{\langle r\rangle}(\tau;\beta;1)\frac{(i\pi/\tau)^{2j+1}}{(2j+1)!}, \tag{3-22}$$

$$N^{2p+1+\mu}\mathfrak{C}_{2p+1+\mu}^{\langle r+1\rangle}(\tau;\mathfrak{a}) = \sum_{j=0}^{p}\widehat{\mathfrak{D}}_{2p+1+\mu-2j}^{\langle r\rangle}(\tau;\beta;1)\frac{(i\pi/\tau)^{2j}}{(2j)!}. \tag{3-23}$$

Note that (3-22) also holds for $p = 0$ if $\mu = 1$, because (3-15) holds for $k = 1$. Here we use the following result given in our previous work.

Lemma 3.6 [Tsumura 2007, Lemma 4.4]. *Let $\{P_{2h}\}$, $\{Q_{2h}\}$, $\{R_{2h}\}$ be sequences satisfying*

$$P_{2h} = \sum_{j=0}^{h}R_{2h-2j}\frac{(i\pi)^{2j}}{(2j)!},\quad Q_{2h} = \sum_{j=0}^{h}R_{2h-2j}\frac{(i\pi)^{2j}}{(2j+1)!}\qquad (h \in \mathbb{Z}_{\geq 0}). \tag{3-24}$$

Then

$$P_{2h} = -2\sum_{\omega=0}^{h} \zeta(2h-2\omega) Q_{2\omega} \qquad (h \in \mathbb{Z}_{\geq 0}). \tag{3-25}$$

Multiply the both sides of (3-22) and (3-23) by $\tau^{2p+2+\mu}$ and $\tau^{2p+1+\mu}$, respectively. Then apply Lemma 3.5 with $P_0 = Q_0 = R_0 = \tau^{1+\mu}\widehat{\mathfrak{D}}^{\langle r\rangle}_{1+\mu}(\tau;\beta;1)$ and

$$\begin{aligned}
P_{2h} &= (N\tau)^{2p+1+\mu}\mathfrak{C}^{\langle r+1\rangle}_{2p+1+\mu}(\tau;\mathfrak{a}),\\
Q_{2h} &= \frac{1}{i\pi}(N\tau)^{2p+2+\mu}\mathfrak{C}^{\langle r\rangle}_{2p+2+\mu}(\tau;\mathfrak{a}),\\
R_{2h} &= \tau^{2h+1+\mu}\widehat{\mathfrak{D}}^{\langle r\rangle}_{2h+1+\mu}(\tau;\beta;1)
\end{aligned}$$

for $h \in \mathbb{N}$. Then it follows from (3-25) that

$$\begin{aligned}
(N\tau)^{2p+1+\mu}&\mathfrak{C}^{\langle r+1\rangle}_{2p+1+\mu}(\tau;\mathfrak{a})\\
&= -2\sum_{\omega=1}^{p}\zeta(2p-2\omega)\frac{1}{i\pi}(N\tau)^{2\omega+2+\mu}\mathfrak{C}^{\langle r\rangle}_{2\omega+2+\mu}(\tau;\mathfrak{a})\\
&\qquad - 2\zeta(2p)\tau^{1+\mu}\widehat{\mathfrak{D}}^{\langle r\rangle}_{1+\mu}(\tau;\beta;1)
\end{aligned} \tag{3-26}$$

for $p \in \mathbb{N}$. In order to complete the proof of Theorem 2.1, we have to determine $\widehat{\mathfrak{D}}^{\langle r\rangle}_{1+\mu}(\tau;\beta;1)$ for $\mu = 0, 1$. As noted above, (3-22) holds for $p = 0$ when $\mu = 1$, namely

$$N^3\mathfrak{C}^{\langle r\rangle}_{3}(\tau;\mathfrak{a}) = \widehat{\mathfrak{D}}^{\langle r\rangle}_{2}(\tau;\beta;1)\frac{i\pi}{\tau}. \tag{3-27}$$

Moreover, we obtain the following.

Lemma 3.7. *For* $r \in \mathbb{Z}$,

$$\frac{i\pi}{\tau}\widehat{\mathfrak{D}}^{\langle r\rangle}_{1}(\tau;\beta;1) = \widehat{\mathfrak{D}}^{\langle r-1\rangle}_{2}(\tau;\beta;1) = \frac{N^3\tau}{i\pi}\mathfrak{C}^{\langle r-1\rangle}_{3}(\tau;\mathfrak{a}). \tag{3-28}$$

Proof. The second equality comes from (3-27) by replacing r with $r-1$. So we will prove the first equality.

As we stated in Remark 3.4, (3-18) holds for $k = 0$ if $\theta \in [(1-1/L)\pi, \pi)$. Hence we see that

$$\begin{aligned}
\frac{1}{2}{\sum_{m\in\mathbb{Z}}}^{*}\sum_{n\in\mathbb{Z}} &\frac{(-1)^n\coth((m+\beta)\pi i/\tau)^r\left(e^{(m+\beta+n\tau)i\theta/\tau} - e^{-(m+\beta+n\tau)i\theta/\tau}\right)}{\sinh((m+\beta)\pi i/\tau)(m+\beta+n\tau)^2}\\
&= \widehat{\mathfrak{D}}^{\langle r\rangle}_{1}(\tau;\beta;1)\frac{i\theta}{\tau}
\end{aligned} \tag{3-29}$$

holds for $\theta \in [(1-1/L)\pi, \pi)$. On the other hand, (3-20) with r replaced by $r-1$ and k by 1 becomes

$$\frac{1}{2}\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}} \frac{(-1)^n \coth((m+\beta)\pi i/\tau)^{r-1}\left(e^{(m+\beta+n\tau)i\theta/\tau}+e^{-(m+\beta+n\tau)i\theta/\tau}\right)}{\sinh((m+\beta)\pi i/\tau)(m+\beta+n\tau)^2} = \widehat{\mathfrak{D}}_2^{\langle r-1\rangle}(\tau;\beta;1), \tag{3-30}$$

which also holds for $\theta \in [(1-1/L)\pi, \pi)$. Now we subtract (3-30) from (3-29) of each side. Then we have

$$\frac{1}{2}\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}} \frac{(-1)^n \coth((m+\beta)\pi i/\tau)^{r-1}\Delta(\theta)}{\sinh((m+\beta)\pi i/\tau)(m+\beta+n\tau)^2} = \widehat{\mathfrak{D}}_1^{\langle r\rangle}(\tau;\beta;1)\frac{i\theta}{\tau} - \widehat{\mathfrak{D}}_2^{\langle r-1\rangle}(\tau;\beta;1), \tag{3-31}$$

where $\Delta(\theta)$ is equal to

$$\begin{aligned}
&\coth\Bigl((m+\beta)\frac{i\pi}{\tau}\Bigr)\Bigl(e^{(m+\beta+n\tau)\frac{i\theta}{\tau}} - e^{-(m+\beta+n\tau)\frac{i\theta}{\tau}}\Bigr) - \Bigl(e^{(m+\beta+n\tau)\frac{i\theta}{\tau}} + e^{-(m+\beta+n\tau)\frac{i\theta}{\tau}}\Bigr)\\
&= \frac{1}{2\sinh\Bigl((m+\beta)\frac{i\pi}{\tau}\Bigr)}\Bigl(\Bigl(e^{(m+\beta)\frac{i\theta}{\tau}}+e^{-(m+\beta)\frac{i\theta}{\tau}}\Bigr)\Bigl(e^{(m+\beta)\frac{i\theta}{\tau}}e^{ni\theta}-e^{-(m+\beta)\frac{i\theta}{\tau}}e^{ni\theta}\Bigr)\\
&\qquad -\Bigl(e^{(m+\beta)\frac{i\theta}{\tau}} - e^{-(m+\beta)\frac{i\theta}{\tau}}\Bigr)\Bigl(e^{(m+\beta)\frac{i\theta}{\tau}}e^{ni\theta} + e^{-(m+\beta)\frac{i\theta}{\tau}}e^{ni\theta}\Bigr)\Bigr)\\
&= \frac{i\sin(n\theta)}{\sinh((m+\beta)\pi i/\tau)}.
\end{aligned}$$

Therefore the left-hand side of (3-31) is absolutely and uniformly convergent in $\theta \in [(1-1/L)\pi, \pi]$. Hence, letting $\theta \to \pi$ on the both sides of (3-31) and noting $\sin(n\pi) = 0$, we have

$$0 = \widehat{\mathfrak{D}}_1^{\langle r\rangle}(\tau;\beta;1)\frac{i\pi}{\tau} - \widehat{\mathfrak{D}}_2^{\langle r-1\rangle}(\tau;\beta;1). \qquad \square$$

Proofs of Theorem 2.1 and Proposition 2.2. Combining (3-26) and (3-28), we obtain the proof of Theorem 2.1. Combining (3-19) and (3-28), we obtain the proof of Proposition 2.2. $\square$

Remark 3.8. The left-hand side of (3-29) in the case $\theta = \pi$ and $r = 2\nu$ ($\nu \in \mathbb{Z}_{\geq 0}$) coincides with $\mathfrak{C}_2^{\langle 2\nu\rangle}(\tau;\mathfrak{a})$, which is conditionally convergent as we noted in Section 1. Therefore, by Abel's theorem, we can let $\theta \to \pi$ on the both sides of (3-29). Hence we have

$$N^2\mathfrak{C}_2^{\langle 2\nu\rangle}(\tau;\mathfrak{a}) = \widehat{\mathfrak{D}}_1^{\langle 2\nu\rangle}(\tau;\beta;1)\frac{i\pi}{\tau} \qquad (\nu \in \mathbb{Z}_{\geq 0}). \tag{3-32}$$

Therefore, by (3-28), we have

$$\mathfrak{C}_2^{\langle 2\nu\rangle}(\tau;\mathfrak{a}) = \frac{i\pi}{N\tau}\mathfrak{C}_3^{\langle 2\nu-1\rangle}(\tau;\mathfrak{a}). \tag{3-33}$$

Remark 3.9. Combining Lemmas 3.3 and 3.7, and using Examples 2.5 and 2.6, we can inductively evaluate

$$\begin{aligned}\widehat{\mathfrak{D}}_{2p+1}^{\langle 2\nu\rangle}(\tau;\beta;1) &= \sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{N}} \frac{\coth((m+\beta+n\tau)\pi i/\tau)^{2\nu}}{\sinh((m+\beta+n\tau)\pi i/\tau)(m+\beta+n\tau)^{2p+1}} \\ &= N^{2p+1} \sum_{\substack{j\in\mathbb{Z}\setminus\{0\}\\ j\equiv a_1 \bmod N}} \sum_{\substack{l\in\mathbb{Z}\\ l\equiv a_2 \bmod N}} \frac{(\coth((j+l\tau)\pi i/N\tau))^{2\nu}}{\sinh((j+l\tau)\pi i/N\tau)(j+l\tau)^{2p+1}}\end{aligned}$$

in terms of $G_{2k}(\tau;\mathfrak{a})$ and $\tilde{\zeta}(2d; a_2 \bmod N)$ $(k, d \in \mathbb{N})$. Therefore, by using the relation $1/(\sinh x)^2 = (\coth x)^2 - 1$ repeatedly, we see that

$$\mathcal{G}_{2j+1}^{\langle 2\nu+1\rangle}(\tau;\mathfrak{a}) = \sum_{\mu=0}^{\nu}\binom{\nu}{\mu}(-1)^{\nu-\mu}N^{-2j-1}\widehat{\mathfrak{D}}_{2j+1}^{\langle 2\nu\rangle}(\tau;\beta;1),$$

which can be evaluated in terms of $G_{2k}(\tau;\mathfrak{a})$ and $\tilde{\zeta}(2d; a_2 \bmod N)$.

4. Two-variable analogues of level-N Eisenstein series

In this section, we aim to evaluate two-variable analogues of level-N Eisenstein series $\widetilde{G}_{j,k}(\tau;\ \mathfrak{a})$ $(j, k \in \mathbb{N}_{\geq 2})$ defined by (1-14).

As well as in the previous section, we set $\beta = (a_1 + a_2\tau)/N$ $(0 \leq a_1, a_2 < N)$. Since $\operatorname{Im}\tau > 0$, namely $\operatorname{Re}(i/\tau) > 0$, it follows from the binomial theorem that

$$\begin{aligned}\frac{1}{(\sinh((m+\beta)\pi i/\tau))^{2\nu}} &= 2^{2\nu}\frac{e^{-2\nu(m+\beta)\pi i/\tau}}{(1-e^{-2(m+\beta)\pi i/\tau})^{2\nu}} \\ &= 2^{2\nu}e^{-2\nu(m+\beta)\pi i/\tau}\sum_{j=0}^{\infty}\binom{j+2\nu-1}{2\nu-1}e^{-2j(m+\beta)\pi i/\tau},\end{aligned}$$

if $m > 0$. By putting $\mu = j + \nu$, we conclude that this equals

$$\begin{aligned}&\frac{2^{2\nu}e^{-2\pi i\nu(m+\beta)/\tau}}{(2\nu-1)!}\sum_{\mu=\nu}^{\infty}(\mu+\nu-1)\cdots(\mu+1)\mu(\mu-1)\cdots(\mu-\nu+1)e^{-2\pi i(\mu-\nu)(m+\beta)/\tau} \\ &\qquad = \frac{2^{2\nu}}{(2\nu-1)!}\sum_{\mu=1}^{\infty}\mu\prod_{l=1}^{\nu-1}(\mu-l)(\mu+l)e^{-2\pi i\mu(m+\beta)/\tau}.\end{aligned} \tag{4-1}$$

Recall the Stirling numbers of the first kind, $\{c(n,k)\}$, defined by

$$F_n(X) = X(X-1)(X-2)\cdots(X-n+1) = \sum_{k=0}^{n} c(n,k)X^k$$

(see, for example, [Stanley 1997]). Using these numbers, we define $\{\alpha(n,k)\}$ by

$$\widetilde{\mathscr{F}}_n(X) = X\prod_{l=1}^{n-1}(X-l)(X+l)\left(= \frac{(-1)^n F_n(X)F_n(-X)}{X}\right) = \sum_{k=0}^{2n-1}\alpha(n,k)X^k. \tag{4-2}$$

Hence we have

$$\alpha(n,j) = (-1)^n \sum_{\omega=0}^{j+1}(-1)^{\omega} c(n, j+1-\omega)c(n,\omega)$$

for $0 \le j \le 2n-1$. Since $\widetilde{\mathscr{F}}_n(-X) = -\widetilde{\mathscr{F}}_n(X)$, we have $\alpha(n,2j) = 0$ for $0 \le j < n$. By (4-1), we have

$$\frac{1}{(\sinh((m+\beta)\pi i/\tau))^{2\nu}} = \frac{2^{2\mu}}{(2\nu-1)!}\sum_{j=1}^{\nu}\alpha(\nu,2j-1)\sum_{\mu=1}^{\infty}\mu^{2j-1}e^{-2\pi i\mu(m+\beta)/\tau}, \tag{4-3}$$

when $m > 0$. Here we recall the summation formula from [Lipschitz 1889]:

$$\sum_{l\in\mathbb{Z}}\frac{1}{(z+l)^k} = (-1)^k\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty}n^{k-1}e^{2\pi i n z} \tag{4-4}$$

for $k \in \mathbb{N}$ with $k \ge 2$ and $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$. This formula also holds for $k = 1$ as follows:

$$\lim_{L\to\infty}\sum_{l=-L}^{L}\frac{1}{z+l} = -\pi i - 2\pi i\sum_{n=1}^{\infty}e^{2\pi i n z} \tag{4-5}$$

for $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$ (see [Pribitkin 2002, Section 5]).

We can set $z = -(m+\beta)/\tau$ in (4-4), because we have $\operatorname{Im}(-(m+\beta)/\tau) > 0$. Then we see that (4-3) is equal to

$$\begin{aligned}\frac{1}{(\sinh((m+\beta)\pi i/\tau))^{2\nu}} &\\ &= \frac{2^{2\mu}}{(2\nu-1)!}\sum_{j=1}^{\nu}\alpha(\nu,2j-1)\frac{(2j-1)!}{(2\pi i)^{2j}}\sum_{l\in\mathbb{Z}}\frac{1}{(-(m+\beta)/\tau+l)^{2j}}\\ &= \frac{2^{2\mu}}{(2\nu-1)!}\sum_{j=1}^{\nu}\alpha(\nu,2j-1)\frac{(2j-1)!}{(2\pi i/\tau)^{2j}}\sum_{l\in\mathbb{Z}}\frac{1}{(m+\beta+l\tau)^{2j}},\end{aligned} \tag{4-6}$$

by replacing l by $-l$. This holds for $m > 0$. When $m < 0$, by replacing (m, l, β) by $(-m, -l, -\beta)$ in (4-6), we have

$$\frac{1}{(\sinh((-m-\beta)\pi i/\tau))^{2\nu}} = \frac{2^{2\mu}}{(2\nu-1)!}\sum_{j=1}^{\nu}\alpha(\nu, 2j-1)\frac{(2j-1)!}{(2\pi i/\tau)^{2j}}\sum_{l\in\mathbb{Z}}\frac{1}{(-m-\beta-l\tau)^{2j}},$$

which coincides with (4-6). This implies that (4-6) also holds for $m < 0$.

On the other hand, by (1-13), we have

$$\begin{aligned}\mathcal{G}_{2q}^{\langle 2p\rangle}(\tau;\mathfrak{a}) &= \sum_{\substack{k\in\mathbb{Z}\setminus\{0\}\\ k\equiv a_1 \bmod N}}\sum_{\substack{l\in\mathbb{Z}\\ l\equiv a_2 \bmod N}}\frac{1}{\sinh((k+l\tau)\pi i/N\tau)^{2p}(k+l\tau)^{2q}}\\ &= N^{-2q}\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\frac{1}{\sinh((m+\beta+n\tau)\pi i/\tau)^{2p}(m+\beta+n\tau)^{2q}}\\ &= N^{-2q}\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\frac{1}{\sinh((m+\beta)\pi i/\tau)^{2p}(m+\beta+n\tau)^{2q}}\end{aligned}$$

for $p, q \in \mathbb{N}$. Therefore, by (4-6) for any $m \in \mathbb{Z}\setminus\{0\}$, we have

$$\mathcal{G}_{2q}^{\langle 2p\rangle}(\tau;\mathfrak{a}) = \frac{2^{2p}N^{-2q}}{(2p-1)!}\times\sum_{j=1}^{p}\alpha(p, 2j-1)\frac{(2j-1)!}{(2\pi i/\tau)^{2j}} \sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\frac{1}{(m+\beta+l\tau)^{2j}(m+\beta+n\tau)^{2q}}.$$

By (1-14) and $\beta = (a_1 + a_2\tau)/N$, we have

$$\sum_{m\in\mathbb{Z}}{}^{*}\sum_{n\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\frac{1}{(m+\beta+l\tau)^{2j}(m+\beta+n\tau)^{2q}} = N^{2j+2q}\Big(\widetilde{G}_{2j,2q}(\tau;\ \mathfrak{a}) - \frac{\delta_{a_1,0}}{\tau^{2j+2q}}\tilde{\zeta}(2j; a_2 \bmod N)\tilde{\zeta}(2q; a_2 \bmod N)\Big),$$

where $\tilde{\zeta}$ is defined by (2-4). Combining these relations, we obtain:

Theorem 4.1. *For $p, q \in \mathbb{N}$,*

$$\mathcal{G}_{2q}^{\langle 2p\rangle}(\tau;\mathfrak{a}) = \frac{2^{2p}}{(2p-1)!}\sum_{j=1}^{p}\alpha(p, 2j-1)\frac{(2j-1)!}{(2\pi i/N\tau)^{2j}} \times\Big(\widetilde{G}_{2j,2q}(\tau;\ \mathfrak{a}) - \frac{\delta_{a_1,0}}{\tau^{2j+2q}}\tilde{\zeta}(2j; a_2 \bmod N)\tilde{\zeta}(2q; a_2 \bmod N)\Big). \tag{4-7}$$

By multiplying the both sides of (4-7) by $\tau^{2q}\pi^{2p}$, we can inductively obtain the following theorem by Proposition 2.4 and the fact $G_{2k}(\tau) \in \mathbb{Q}[G_4(\tau), G_6(\tau)]$ for $k \in \mathbb{N}_{\geq 2}$ (see [Koblitz 1993, Chapter III, § 2]).

Theorem 4.2. *For $p, q \in \mathbb{N}$,*

$$\tau^{2(p+q)}\widetilde{G}_{2p,2q}(\tau;\mathfrak{a}) \in \mathbb{Q}\big[\tau,\ \pi,\ \big\{\tilde{\zeta}(2k; a_2 \bmod N),\ G_{2k}(\tau;\mathfrak{a})\big\}_{k\in\mathbb{N}}\big].$$

In particular when $N = 1$, put $\widetilde{G}_{2p,2q}(\tau) = \widetilde{G}_{2p,2q}(\tau; (0,0) \bmod 1)$. Then

$$\tau^{2(p+q)}\widetilde{G}_{2p,2q}(\tau) \in \mathbb{Q}\big[\tau,\ \pi,\ G_2(\tau),\ G_4(\tau),\ G_6(\tau)\big].$$

Actually, combining (4-7) and the results given in Section 2, we can concretely evaluate $\widetilde{G}_{2p,2q}(\tau;\ \mathfrak{a})$ as follows.

Example 4.3. We set $N = 1$, $(a_1, a_2) = (0, 0)$, $p = 1, 2$, $q = 2$ and $\tau = i$. By (4-2), we see that $\alpha(1, 1) = 1$, $\alpha(2, 1) = -1$ and $\alpha(2, 3) = 1$. By substituting (2-8) and (2-9) into (4-7), we obtain

$$\begin{aligned}\widetilde{G}_{2,4}(i) &= -\frac{\varpi^4\pi^2}{45} + \frac{2}{63}\pi^6 - \frac{4}{45}\pi^5,\\ \widetilde{G}_{4,4}(i) &= \frac{1}{525}\varpi^8 + \frac{2}{675}\varpi^4\pi^4 - \frac{2}{135}\pi^8 + \frac{8}{189}\pi^7.\end{aligned}$$

Set $\tau = \rho$. Then, by substituting (2-12) and (2-13) into (4-7), we obtain

$$\widetilde{G}_{2,4}(\rho) = \frac{\widetilde{\varpi}^6}{35} - \frac{2}{63}\pi^6 + \frac{8\sqrt{3}}{135}\pi^5, \quad \widetilde{G}_{4,4}(\rho) = \rho\left(-\frac{2}{135}\pi^8 + \frac{16\sqrt{3}}{567}\pi^7\right).$$

Example 4.4. We set $N = 2$, $(a_1, a_2) = (1, 1)$, $p = 1, 2$, $q = 2$ and $\tau = i$. By substituting (2-14) and (2-15) into (4-7), we obtain

$$\begin{aligned}\widetilde{G}_{2,4}(i; (1,1) \bmod 2) &= \frac{\varpi^4\pi^2}{576} - \frac{\pi^5}{720},\\ \widetilde{G}_{4,4}(i; (1,1) \bmod 2) &= \frac{\varpi^8}{8960} - \frac{\varpi^4\pi^4}{17280} + \frac{\pi^7}{6048}.\end{aligned}$$

Remark 4.5. Pasles and Pribitkin [2001] studied two-variable Lipschitz summation formulas. At present, it is unclear whether or not the results stated above can be obtained from their formula.

Acknowledgements

The author wishes to express his sincere gratitude to Professors Kohji Matsumoto and Yasushi Komori for fruitful discussions. Also the author wishes to express his sincere gratitude to the referee for valuable comments and important suggestions.

This research was partially supported by Grant-in-Aid for Science Research (No. 23540022), Japan Society for the Promotion of Science.

References

[Berndt 1977] B. C. Berndt, "Modular transformations and generalizations of several formulae of Ramanujan", *Rocky Mountain J. Math.* **7**:1 (1977), 147–190. MR 55 #2714 Zbl 0365.10021

[Berndt 1978] B. C. Berndt, "Analytic Eisenstein series, theta-functions, and series relations in the spirit of Ramanujan", *J. Reine Angew. Math.* **303/304** (1978), 332–365. MR 80b:10035 Zbl 0384.10011

[Cauchy 1889] A. Cauchy, *Oeuvres complètes D'Augustin Cauchy*, Gauthier-Villars, Paris, 1889. Série II, tome VII. Zbl 21.0015.04

[Hecke 1937] E. Hecke, "Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung, I", *Math. Ann.* **114**:1 (1937), 1–28. MR 1513122 Zbl 0015.40202

[Katayama 1978] K. Katayama, "On the values of Eisenstein series", *Tokyo J. Math.* **1**:1 (1978), 157–188. MR 80f:10029 Zbl 0391.10022

[Koblitz 1993] N. Koblitz, *Introduction to elliptic curves and modular forms*, 2nd ed., Graduate Texts in Mathematics **97**, Springer, New York, 1993. MR 94a:11078 Zbl 0804.11039

[Komori et al. 2010] Y. Komori, K. Matsumoto, and H. Tsumura, "Hyperbolic-sine analogues of Eisenstein series, generalized Hurwitz numbers, and q-zeta functions", 2010. To appear in *Forum Math.* arXiv 1006.3339

[Lipschitz 1889] R. Lipschitz, "Untersuchung der Eigenschaften einer Gattung von unendlichen Reihen", *J. Reine. Angew. Math.* **105** (1889), 127–156. Zbl 21.0176.01

[Mellin 1902] H. Mellin, "Eine Formel für den Logarithmus transcendenter Functionen von endlichem Geschlecht", *Acta Math.* **25**:1 (1902), 165–183. MR 1554942 Zbl 34.0469.03

[Meyer 2000] J. L. Meyer, "Character analogues of Dedekind sums and transformations of analytic Eisenstein series", *Pacific J. Math.* **194**:1 (2000), 137–164. MR 2002b:11059 Zbl 1009.11029

[Nesterenko and Philippon 2001] Y. V. Nesterenko and P. Philippon (editors), *Introduction to algebraic independence theory*, Lecture Notes in Mathematics **1752**, Springer, Berlin, 2001. MR 2002g:11104 Zbl 0966.11032

[Pasles and Pribitkin 2001] P. C. Pasles and W. d. A. Pribitkin, "A generalization of the Lipschitz summation formula and some applications", *Proc. Amer. Math. Soc.* **129**:11 (2001), 3177–3184. MR 2002e:11059 Zbl 1125.11322

[Pribitkin 2002] W. d. A. Pribitkin, "Laplace's integral, the gamma function, and beyond", *Amer. Math. Monthly* **109**:3 (2002), 235–245. MR 2003c:44002 Zbl 1029.44001

[Serre 1970] J.-P. Serre, *Cours d'arithmétique*, SUP Le Mathématicien **2**, Presses Universitaires de France, Paris, 1970. In French; translated as *A course in arithmetic*, Graduate Texts in Mathematics **7**, Springer, New York, 1973. MR 41 #138 Zbl 0225.12002

[Stanley 1997] R. P. Stanley, *Enumerative combinatorics, 1*, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, Cambridge, 1997. MR 98a:05001 Zbl 0889.05001

[Tsumura 2007] H. Tsumura, "On functional relations between the Mordell–Tornheim double zeta functions and the Riemann zeta function", *Math. Proc. Cambridge Philos. Soc.* **142**:3 (2007), 395–405. MR 2008d:11103 Zbl 1149.11044

[Tsumura 2008] H. Tsumura, "On certain analogues of Eisenstein series and their evaluation formulas of Hurwitz type", *Bull. Lond. Math. Soc.* **40**:2 (2008), 289–297. MR 2009c:11141 Zbl 1168.11038

[Tsumura 2009] H. Tsumura, "Evaluation of certain classes of Eisenstein-type series", *Bull. Aust. Math. Soc.* **79**:2 (2009), 239–247. MR 2010e:11080 Zbl 1210.11098

[Tsumura 2010] H. Tsumura, “Analogues of the Hurwitz formulas for level 2 Eisenstein series”, *Results Math.* **58**:3-4 (2010), 365–378. MR 2012a:11139 Zbl 05816643

[Waldschmidt 1999] M. Waldschmidt, “Transcendance et indépendance algébrique de valeurs de fonctions modulaires”, pp. 353–375 in *Number theory: Fifth Conference of the Canadian Number Theory Association* (Ottawa, ON, 1996), edited by R. Gupta and K. S. Williams, CRM Proc. Lecture Notes **19**, American Mathematical Society, Providence, RI, 1999. MR 2000c:11122 Zbl 0930.11055

Received March 18, 2011. Revised December 2, 2011.

HIROFUMI TSUMURA
DEPARTMENT OF MATHEMATICS AND INFORMATION SCIENCES
TOKYO METROPOLITAN UNIVERSITY
1-1, MINAMI-OHSAWA, HACHIOJI
TOKYO 192-0397
JAPAN

tsumura@tmu.ac.jp

CONTENTS

Volume 255, no. 1 and no. 2

Guidelines for Authors

Authors may submit manuscripts at msp.berkeley.edu/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 255 No. 2 February 2012

0030-8730(201202)255:2;1-B