Pacific Journal of Mathematics

Volume 255 No. 1 January 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. 1, 2012

AVERAGING SEQUENCES

FERNANDO ALCALDE CUESTA AND ANA RECHTMAN

In the spirit of the Goodman–Plante average condition for the existence of a transverse invariant measure for foliations, we give an averaging condition to find tangentially smooth measures with prescribed Radon–Nikodým cocycle. Harmonic measures are examples of tangentially smooth measures for foliations and laminations. We also present a sufficient hypothesis under which the tangentially smooth measure is harmonic.

1. Introduction

Averaging sequences for foliations were introduced in the pioneering work of Plante [1975] on the influence that the existence of transverse invariant measures exerts on the structure of a foliation. Although that work dealt only with the case of subexponential growth, his approach is clearly reminiscent of the classic work of E. Følner [1955] on groups. Using the same kind of ideas, S. E. Goodman and Plante [1979] exhibited an averaging condition which guarantees the existence of transverse invariant measures for foliations of compact manifolds.

In this paper we formulate a more general averaging condition which gives rise to a tangentially smooth measure for a compact laminated space $(M, \mathcal{F})$. This condition may be related to the η-Følner condition in [Alcalde Cuesta and Rechtman 2011], in the same spirit as Følner, but using a modified Riemannian metric along the leaves. The modification is done by replacing any complete Riemannian metric along the leaves with the product of the metric with some density function. Namely, given a compact laminated space and a positive cocycle defined on the equivalence relation induced by the lamination on a total transversal, we prove that an η-Følner sequence gives rise to the existence of a tangentially smooth measure whose Radon–Nikodým cocycle is the given one. Moreover, we describe a sufficient hypothesis for obtaining a harmonic measure. This is the content of Theorem 4.10.

This work was partially supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT) in Mexico and the Xunta de Galicia INCITE09E2R207023ES in Spain.
MSC2010: 37A20, 43A07, 57R30.
Keywords: lamination, discrete equivalence relation, measure.

Before proving Theorem 4.10, we analyze the discrete case. We define an averaging condition for any equivalence relation $\mathcal{R}$ defined by a finitely generated pseudogroup acting on a compact space and any continuous cocycle $\delta : \mathcal{R} \to \mathbb{R}_+^*$ that we call a δ-averaging condition. In Theorem 3.6 we prove that the existence of a δ-averaging sequence gives a quasi-invariant measure with Radon–Nikodým cocycle δ. Under some additional conditions, in particular if δ is harmonic, the measure obtained is harmonic. In this case, our result is reminiscent of Kaimanovich's [1997] characterization of amenable equivalence relations.

The paper is organized as follows. In Section 2 we review some preliminaries. In particular Section 2C contains the proof of Goodman and Plante's theorem. The discussion of the discrete and continuous settings is split into two separate sections, Section 3 and Section 4, respectively, which can be read independently. In Section 5 we analyze some explicit examples. The relation between the two types of averaging sequences will be briefly discussed in Section 6.

2. Preliminaries

2A. ***Laminations and equivalence relations.*** A compact space M admits a d-dimensional lamination $\mathcal{F}$ of class C^r with $1 \le r \le \infty$ if there exists a cover of M by open sets U_i homeomorphic to the product of an open disc P_i in $\mathbb{R}^d$ centered at the origin and a locally compact separable metrizable space T_i. Thus, if we denote the corresponding foliated chart by $\varphi_i : U_i \to P_i \times T_i$, each U_i splits into *plaques* $\varphi_i^{-1}(P_i \times \{y\})$. Each point $y \in T_i$ can also be identified with the point $\varphi_i^{-1}(0, y)$ in the *local transversal* $\varphi_i^{-1}(\{0\} \times T_i)$. In addition, the change of charts $\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is given by

$$\varphi_j \circ \varphi_i^{-1}(x, y) = (\varphi_{ij}^y(x), \gamma_{ij}(y)), \tag{2-1}$$

where γ_{ij} is a homeomorphism between open subsets of T_i and T_j, and φ_{ij}^y is a C^r-diffeomorphism depending continuously on y in the C^r-topology. We say that $\mathcal{A} = \{(U_i, \varphi_i)\}_{i \in I}$ is a *good foliated atlas* if it satisfies the following conditions.

(i) The cover $\mathcal{U} = \{U_i\}_{i \in I}$ is locally finite, hence finite.

(ii) Each open set U_i is a relatively compact subset of a foliated chart.

(iii) If $U_i \cap U_j \neq \varnothing$, there is a foliated chart containing $\overline{U_i \cap U_j}$, implying that each plaque of U_i intersects at most one plaque of U_j.

Each foliated chart U_i admits a tangentially C^r-smooth Riemannian metric $g_i = \varphi_i^* g_0$ induced from a C^r-smooth Riemannian metric g_0 on $\mathbb{R}^p$. We can glue together these local Riemannian metrics g_i to obtain a global one g using a tangentially C^r-smooth partition of unity. From [Alcalde Cuesta et al. 2009, Lemma 2.6], we know that any C^r lamination of a compact space equipped with a C^r foliated atlas $\mathcal{A}$ admits a C^∞ foliated atlas C^r-equivalent to $\mathcal{A}$.

A discrete equivalence relation $\mathcal{R}$ is defined by $\mathcal{F}$ on the total transversal $T = \bigsqcup T_i$; the equivalence classes are the traces of the leaves on T. We can see $\mathcal{R}$ as the orbit equivalence relation defined by the *holonomy pseudogroup* Γ of $\mathcal{F}$, generated by the local diffeomorphisms γ_{ij}. These homeomorphisms form a finite generating set, which we will denote by $\Gamma^{(1)}$, that defines a *graphing* of $\mathcal{R}$. This means that each equivalence class $\mathcal{R}[y]$ is the set of vertices of a graph, and there is an edge joining two vertices z and w if there is $\gamma \in \Gamma^{(1)}$ such that $\gamma(z) = w$. We can define a graph metric $d_\Gamma(z, w) = \min\{n : g(z) = w \text{ for some } \gamma \in \Gamma^{(n)}\}$, where $\Gamma^{(n)}$ are the elements that can be expressed as words of length at most n in terms of $\Gamma^{(1)}$. A *transverse invariant measure* for $\mathcal{F}$ is a measure on T that is invariant under the action of Γ. It is quite rare for a measure of this kind to exist.

Remark 2.1. If $\mathcal{F}$ has no holonomy (that is, $\Gamma_y = \{\gamma \in \Gamma : \gamma(y) = y\}$ is trivial for all $y \in T$), we can endow $\mathcal{R}$ with the topology generated by the graphs of the elements of Γ. Then $\mathcal{R}$ becomes an *étale equivalence relation*, that is, the partial multiplication $\big((y, \gamma(y)), (\gamma(y), \gamma'(\gamma(y)))\big) \in \mathcal{R} * \mathcal{R} \mapsto (y, \gamma' \circ \gamma(y)) \in \mathcal{R}$ and the inversion $(y, \gamma(y)) \in \mathcal{R} \mapsto (\gamma(y), y) \in \mathcal{R}$ are continuous, and the left and right projections $\beta : (y, z) \in \mathcal{R} \mapsto y \in T$ and $\alpha : (y, z) \in \mathcal{R} \mapsto z \in T$ are local homeomorphisms. In general, by considering the germs of the elements of Γ at the points of their domains, we can replace $\mathcal{R}$ with the *transverse holonomy groupoid* [Haefliger 1984] that similarly becomes an *étale groupoid* [Renault 1980].

2B. ***Compactly generated pseudogroups.*** In the last section, we obtained a pseudogroup from a foliated atlas. Here we will recall the *Haefliger equivalence* for pseudogroups obtained from different atlases and its metric counterpart in the compact case, which we will need later in Section 2C. For any compact laminated space $(M, \mathcal{F})$ the holonomy pseudogroup Γ is *compactly generated* in the sense of [Haefliger 2002], meaning that

(i) T contains a relatively compact open set T_1 meeting all the orbits, and

(ii) the reduced pseudogroup $\Gamma|_{T_1}$ (whose elements have domain and range in T_1) admits a finite generating set $\Gamma^{(1)}$ (called a *compact generation system* of Γ on T_1) so that each element $\gamma : A \to B$ of $\Gamma^{(1)}$ is the restriction of an element $\bar{\gamma}$ of Γ whose domain contains the closure of A.

Any probability measure ν_K on the compact set $K = \overline{T}_1$ that is preserved by the action of $\Gamma|_K$ extends to a unique Borel measure ν on T which is Γ-invariant and finite on compact sets. We refer to [Plante 1975, Lemma 3.2].

Also, notice that T is covered by the domains of a family of elements of Γ with range in T_1. The union of these elements and their inverses defines the *fundamental* *equivalence* between the holonomy pseudogroup Γ and the reduced pseudogroup $\Gamma|_{T_1}$. This is the base concept to define the *Haefliger equivalence* of pseudogroups [Haefliger 1984; 2002].

Definition 2.2. Two pseudogroups Γ_1 and Γ_2 acting on the spaces T_1 and T_2 are *Haefliger equivalent* if they are reductions of a same pseudogroup Γ acting on the disjoint union $T = T_1 \sqcup T_2$, and both T_1 and T_2 meet all the orbits of Γ.

The choice of generators for Γ_1 and Γ_2 defines a metric graph structure on the orbits, but the Haefliger equivalence between Γ_1 and Γ_2 may not preserve their quasi-isometry type. Let us recall this concept introduced by M. Gromov [1993]:

Definition 2.3. Two metric spaces (M, d) and (M', d') are *quasi-isometric* if there exists a map $f : M \to M'$ and constants $\lambda \geq 1$ and $C \geq 0$ such that

$$\frac{1}{\lambda} d(y, z) - C \leq d'(f(y), f(z)) \leq \lambda d(y, z) + C$$

for all $y, z \in M$ and $d'(y', f(M)) \leq C$ for all $y' \in M'$.

Definition 2.4 [Hurder and Katok 1987; Ghys 1995]. A Haefliger equivalence between two pseudogroups Γ_1 and Γ_2 acting on T_1 and T_2, respectively, is a *Kakutani equivalence* if Γ_1 and Γ_2 admit finite generating systems such that their orbits, endowed with the graph metric, are quasi-isometric.

According to [Lozano Rojo 2006, Theorem 2.7] and [Álvarez López and Candel 2009, Theorem 4.6], if two compactly generated pseudogroups Γ_1 and Γ_2 are Haefliger equivalent, then there are compact generating systems on T_1 and T_2, respectively, such that the pseudogroups become Kakutani equivalent. These compact generating systems are called *good* by Lozano Rojo and *recurrent* by Álvarez López and Candel. The relevance of this is that the existence of averaging sequences depends on the quasi-isometric type of the orbits; see [Álvarez López and Candel 2009] and [Kanai 1985] for the details.

2C. ***Existence of transverse invariant measures.*** In this section we will discuss a sufficient condition for the existence of a transverse invariant measure, which serves as motivation for Theorems 3.6 and 4.10. Goodman and Plante [1979] formulate the following proposition. Let us start with some definitions.

Definition 2.5. Let A be a finite subset of T and γ an element of Γ. We define the difference set

$$\Delta_\gamma A = \{x \in T : x \in A, \gamma(x) \notin A\} \cup \{x \in T : x \notin A, \gamma(x) \in A\},$$

with the convention that $\gamma(x) \notin A$ holds if $\gamma(x)$ is not defined. We denote the cardinality of A by $|A|$.

Definition 2.6. A sequence of finite subsets A_n of T is an *averaging sequence* for Γ if for all $\gamma \in \Gamma^{(1)}$ (and then for all $\gamma \in \Gamma$),

$$\lim_{n\to\infty} \frac{|\Delta_\gamma A_n|}{|A_n|} = 0.$$

Proposition 2.7 [Goodman and Plante 1979]. *An averaging sequence* $\{A_n\}$ *gives rise to a transverse invariant measure* ν *whose support is contained in the limit set* $\lim_{n\to\infty} A_n = \{y \in T : \exists y_{n_k} \in A_{n_k},\, y = \lim_{k\to\infty} y_{n_k}\}$.

The idea of the proof is the following. Assuming that T is compact, we may construct a Γ-invariant probability measure on T from the sequence of probability measures ν_n defined by $\nu_n(B) = |B \cap A_n|/|A_n|$ for every Borel set $B \subset T$. By Riesz's representation theorem, each measure ν_n can be identified with a functional I_n on the space $C(T)$ of continuous real-valued functions on T. The functionals I_n are

$$I_n(f) = \frac{1}{|A_n|} \sum_{y\in A_n} f(y).$$

By passing to a subsequence, if necessary, I_n converges in the weak topology to a positive functional I which determines a unique Borel regular measure ν such that $I(f) = \int_T f\,d\nu$ for every $f \in C(T)$. The averaging condition implies that I and ν are Γ-invariant, since for every $\gamma \in \Gamma$ and every $f \in C(T)$ with support on the range of γ, we have

$$|I(f\circ\gamma) - I(f)| \le \|f\|_\infty \lim_{n\to\infty} \frac{|\Delta_\gamma A_n|}{|A_n|} = 0.$$

Finally, it is clear that $\nu(T) = 1$ and $\operatorname{supp}(\nu) = \lim_{n\to\infty} A_n$.

In the noncompact case, by replacing Γ and Γ_1 with suitable reductions, we can assume, without loss of generality, that the fundamental equivalence between the holonomy pseudogroup Γ and its reduction Γ_1 to a relatively compact open subset T_1 of T becomes a Kakutani equivalence for some compact generation systems on T and T_1. Then any averaging sequence A_n for Γ defines an averaging sequence $A_n \cap K$ for $\Gamma|_K$, where $K = \overline{T}_1$ is a compact subset of T. Hence we obtain a probability measure ν_K on K that is invariant under $\Gamma|_K$. Now we can extend ν_K to a unique Borel measure ν on T which is Γ-invariant and finite on compact sets.

Example 2.8. Consider a graph with bounded geometry, like any orbit $\Gamma(x)$ of the holonomy pseudogroup of a compact laminated space. This graph is said to be *Følner* if it contains a sequence of finite subsets of vertices A_n such that $|\partial A_n|/|A_n|$ tends to 0, where ∂A_n denotes the boundary set with respect to the graph structure. Since $\Delta_\gamma A \subset \partial A \cup \gamma^{-1}(\partial A)$ for any $\gamma \in \Gamma^{(1)}$, we get $|\Delta_\gamma A_n| \le 2|\partial A_n|$, and we have an averaging sequence. In particular, any orbit $\Gamma(x)$ having subexponential growth is an example of a Følner graph, since

$$\liminf_{n\to\infty} \frac{|A_{n+1} - A_{n-1}|}{|A_n|} = 0,$$

where $A_n = \Gamma^{(n)}(x)$.

Using the one-to-one correspondence between foliated cycles and transverse invariant measures established by D. Sullivan [1976], it is not difficult to show the following continuous version of Goodman and Plante's result:

Proposition 2.9 [Goodman and Plante 1979]. *Let $\{V_n\}$ be an averaging sequence for $\mathcal{F}$, that is, a sequence of compact domains V_n (of dimension d) in the leaves such that*

$$\lim_{n\to\infty} \frac{\mathrm{area}(\partial V_n)}{\mathrm{vol}(V_n)} = 0,$$

where area denotes the $(d-1)$-volume and vol the d-volume with respect to the complete Riemannian metric along the leaves. Then $\{V_n\}$ gives rise to a transverse invariant measure ν whose support is contained in the saturated limit set $\lim_{n\to\infty} V_n = \{p \in M : \exists p_{n_k} \in V_{n_k} : p = \lim_{k\to\infty} p_{n_k}\}$.

Recall that a *foliated d-form $\alpha \in \Omega^d(\mathcal{F})$* is a family of differentiable d-forms over the plaques of $\mathcal{A}$ depending continuously on the transverse parameter and agreeing on the intersection of each pair of foliated charts. A *foliated r-cycle* is a continuous linear functional $\xi : \Omega^d(\mathcal{F}) \to \mathbb{R}$ strictly positive on strictly positive forms, and null on exact forms with respect to the leafwise exterior derivative $d_{\mathcal{F}}$. Any averaging sequence V_n defines the sequence of foliated currents

$$\xi_n(\alpha) = \frac{1}{\mathrm{vol}(V_n)} \int_{V_n} \alpha,$$

where α is a foliated d-form. By passing to a subsequence, if necessary, we have a limit current $\xi = \lim_{n\to\infty} \xi_n$. Since the boundaries of the domains V_n vanish asymptotically, Stokes' theorem implies that ξ is a foliated d-cycle [Sullivan 1976].

3. Averaging sequences in the discrete setting

The main objective of this section is to prove the existence of a harmonic measure for an étale equivalence relation $\mathcal{R}$ that contains a modified averaging sequence. Initially, we will assume that $\mathcal{R}$ is given by a free action of a pseudogroup Γ on a compact space T, but some generalizations will be discussed later. In Section 3A, we will define a weighted measure on the equivalence classes that will allow us to recall the notion of a *modified averaging sequence* introduced by V. A. Kaimanovich [1997; 2001]. Given a continuous cocycle $\delta : \mathcal{R} \to \mathbb{R}^*_+$, the *Radon–Nikodým problem* is to determine the set of probability measures ν on T

which are quasi-invariant and admit δ as their Radon–Nikodým derivative [Renault 2005]. Theorem 3.6 gives a positive answer to this problem in the presence of a modified averaging sequence.

3A. *Quasi-invariant measures.* Let ν be a quasi-invariant measure on T. As usual, we will assume that ν is a regular Borel measure that is finite on compact sets. Integrating the counting measures on the fibers of the left projection $\beta(y,z)=y$ with respect to ν gives the *left counting measure* $d\tilde{\nu}(y,z)=d\nu(y)$. Indeed, for each Borel set $A\subset\mathcal{R}$, we define

$$\tilde{\nu}(A)=\int|A^y|\,d\mu(y),$$

where $|A^y|$ is the cardinal of the set $A^y=\{z\in T:(y,z)\in A\}\subset\mathcal{R}[y]$. The same is valid for the right projection $\alpha(y,z)=z$, and we get the *right counting measure* $d\tilde{\nu}^{-1}(y,z)=d\tilde{\nu}(z,y)=d\nu(z)$. Then $\tilde{\nu}$ and $\tilde{\nu}^{-1}$ are equivalent measures if and only if ν is quasi-invariant, in which case the Radon–Nikodým derivative is given by $\delta(y,z)=d\tilde{\nu}/d\tilde{\nu}^{-1}(y,z)$. We refer to [Moore and Schochet 2006; Kaimanovich 1997; Renault 1980; 2005].

Definition 3.1. A *cocycle with values in* $\mathbb{R}_+^*$ is a map $\delta:\mathcal{R}\to\mathbb{R}_+^*$ satisfying $\delta(x,y)\delta(y,z)=\delta(x,z)$ for all $(x,y),(y,z)\in\mathcal{R}$.

The map δ is known as the *Radon–Nikodým cocycle* of $(\mathcal{R},T,\nu)$.

Definition 3.2. Given a cocycle $\delta:\mathcal{R}\to\mathbb{R}_+^*$, the measure $|\cdot|_y$ on $\mathcal{R}[y]$ is given by $|z|_y=\delta(z,y)$ for all $z\in\mathcal{R}[y]$. Then, for a finite subset $A\subset\mathcal{R}[y]$,

$$|A|_y=\sum_{z\in A}\delta(z,y).$$

3B. *Discrete averaging sequences.* We want to give a sufficient condition to solve the Radon–Nikodým problem in the discrete setting. We state this condition using the notion of a modified averaging sequence; see [Kaimanovich 1997; 2001]:

Definition 3.3. Let $\delta:\mathcal{R}\to\mathbb{R}_+^*$ be a cocycle of $\mathcal{R}$. Let $\{A_n\}$ be a sequence of finite subsets of T such that $A_n\subset\mathcal{R}[y_n]$ for each $n\in\mathbb{N}$. We will say that $\{A_n\}$ is a *δ-averaging sequence for* Γ if

$$\lim_{n\to\infty}\frac{|\Delta_\gamma A_n|_{y_n}}{|A_n|_{y_n}}=0$$

for all $\gamma\in\Gamma^{(1)}$. An equivalence class $\mathcal{R}[y]$ is *δ-Følner* if $\mathcal{R}[y]$ contains a δ-averaging sequence $\{A_n\}$ such that $|\partial A_n|_y/|A_n|_y\to 0$ as $n\to+\infty$.

By choosing a finite generating set for Γ, we can realize each equivalence class $\mathcal{R}[y]$ as the set of vertices of a graph. We will write $z\sim w$ for each pair of neighboring vertices z and w joined by an edge in $\mathcal{R}[y]$, and $\deg z$ for the number of neighbors of $z \in \mathcal{R}[y]$. We will use $\mathcal{D}$ to denote the set of discontinuities of the degree function deg. Let ν be a quasi-invariant measure on T, and denote by $D : L^\infty(T, \nu) \to L^\infty(T, \nu)$ the Markov operator defined by

$$Df(y) = \frac{1}{\deg y} \sum_{z \sim y} f(z).$$

We use D^* to denote the dual operator acting on the space of positive Borel measures on T, and

$$\Delta : L^\infty(T, \nu) \to L^\infty(T, \nu)$$

to denote the Laplace operator defined by $\Delta f(y) = Df(y) - f(y)$.

Definition 3.4. A quasi-invariant measure ν on T is *harmonic* or *stationary* (for the simple random walk on $\mathcal{R}$) if for every bounded measurable function $f : T \to \mathbb{R}$, we have $\int \Delta f \, d\nu = 0$.

Proposition 3.5 [Paulin 1999]. *For a quasi-invariant measure ν on T, the following are equivalent*:

(i) *ν is harmonic.*

(ii) *$D^*\nu = \nu$.*

(iii) *The Radon–Nikodým cocycle $\delta : \mathcal{R} \to \mathbb{R}^*_+$ is harmonic, that is, for ν-almost every $y \in T$ and every $z \in \mathcal{R}[y]$, we have*

$$\delta(z, y) = \frac{1}{\deg z} \sum_{w \sim z} \delta(w, y).$$

Theorem 3.6. *Let $\mathcal{R}$ be the orbit equivalence relation defined by a finitely generated pseudogroup Γ acting freely on a compact space T. Let $\delta : \mathcal{R} \to \mathbb{R}^*_+$ be a continuous cocycle.*

(i) *Any δ-averaging sequence $\{A_n\}$ gives rise to a positive Borel measure ν on T whose support is contained in the limit set of $\{A_n\}$, which is quasi-invariant and has δ as its Radon–Nikodým cocycle.*

(ii) *If δ is harmonic and $\nu(\mathcal{D}) = 0$, then ν is a harmonic measure.*

Proof. We start by constructing a sequence of probability measures ν_n given by

$$\nu_n(B) = \frac{|B \cap A_n|_{y_n}}{|A_n|_{y_n}}$$

for every Borel subset B of T. By passing to a subsequence, the sequence ν_n converges in the weak topology to a positive Borel measure ν on T. First we will prove that ν is a quasi-invariant measure having a Radon–Nikodým cocycle equal to δ. For every local transformation $\gamma \in \Gamma$ and every function $f \in C(T)$ with support on the range of γ, we have

$$\int f(z)\,d(\gamma_*\nu)(z) = \int f(\gamma(y))\,d\nu(y) = \lim_{n\to\infty} \frac{1}{|A_n|_{y_n}} \sum_{y\in A_n} f(\gamma(y))\delta(y, y_n)$$

and

$$\begin{aligned}\int f(y)\delta(z, y)\,d\nu(y) &= \lim_{n\to\infty} \frac{1}{|A_n|_{y_n}} \sum_{y\in A_n} f(y)\delta(\gamma(y), y)\delta(y, y_n)\\ &= \lim_{n\to\infty} \frac{1}{|A_n|_{y_n}} \sum_{y\in A_n} f(y)\delta(\gamma(y), y_n),\end{aligned}$$

where $z = \gamma(y)$. Therefore

$$\begin{aligned}0 &\le \left| \int f(z)\,d(\gamma_*\nu)(z) - \int f(y)\delta(z, y)\,d\nu(y) \right|\\ &\le \lim_{n\to\infty} \frac{1}{|A_n|_{y_n}} \left| \sum_{y\in A_n} f(\gamma(y))\delta(y, y_n) - f(y)\delta(\gamma(y), y_n) \right|\\ &\le \lim_{n\to\infty} \|f\|_\infty \frac{|\Delta_\gamma A_n|_{y_n}}{|A_n|_{y_n}} = 0,\end{aligned}$$

and thus

$$\int f(z)\,d(\gamma_*\nu)(z) = \int f(y)\delta(z, y)\,d\nu(y),$$

proving (i).

We now prove that if δ is harmonic and $\nu(\mathcal{D}) = 0$, then ν is a harmonic measure. Observe that if $\nu(\mathcal{D}) = 0$, then Δf is continuous ν-almost everywhere, and therefore

$$\int \Delta f\,d\nu = \lim_{n\to\infty} \int \Delta f\,d\nu_n$$

for all $f \in C(T)$. If δ is harmonic, we have

$$\begin{aligned}&\int \Delta f(y)\,d\nu_n(y)\\ &\qquad= \frac{1}{|A_n|_{y_n}} \sum_{y\in A_n} \Big(\frac{1}{\deg y} \sum_{z\sim y} f(z) - f(y) \Big) \delta(y, y_n)\\ &\qquad= \frac{1}{|A_n|_{y_n}} \sum_{y\in A_n} \frac{1}{\deg y} \sum_{z\sim y} f(z)\delta(y, y_n) - f(y) \Big(\frac{1}{\deg y} \sum_{z\sim y} \delta(z, y_n) \Big)\\ &\qquad= \frac{1}{|A_n|_{y_n}} \sum_{y\in A_n} \frac{1}{\deg y} \sum_{z\sim y} f(z)\delta(y, y_n) - f(y)\delta(z, y_n)\end{aligned}$$

and then

$$\begin{aligned} 0 &\leq \left| \int \Delta f(y)\, d\nu(y) \right| \\ &\leq \lim_{n\to\infty} \frac{1}{|A_n|_{y_n}} \left| \sum_{y\in A_n} \sum_{z\sim y} f(z)\delta(y, y_n) - f(y)\delta(z, y_n) \right| \\ &\leq \lim_{n\to\infty} \|f\|_\infty \sum_{\gamma\in\Gamma^{(1)}} \frac{|\Delta_\gamma A_n|_{y_n}}{|A_n|_{y_n}} \leq \lim_{n\to\infty} 2\,\|f\|_\infty\, |\Gamma^{(1)}| \frac{|\partial A_n|_{y_n}}{|A_n|_{y_n}} = 0; \end{aligned}$$

that is, ν is a harmonic measure. □

A similar result can be found in [Schapira 2003]. In general, the second part of Theorem 3.6 remains valid when the Laplace operator Δ preserves continuous functions. This is always true when $\mathcal{D} = \varnothing$, as in the following case:

Corollary 3.7. *Let $\mathcal{R}$ be the orbit equivalence relation defined by a group of finite type Γ acting freely on a compact space T. Let $\delta : \mathcal{R} \to \mathbb{R}^*_+$ be a continuous harmonic cocycle. Any δ-averaging sequence $\{A_n\}$ gives rise to a harmonic measure ν on T supported by the limit set of $\{A_n\}$.* □

Arguing as for usual averaging sequences, we can extend Theorem 3.6 to any compactly generated pseudogroup Γ acting freely on a locally compact Polish space T. Moreover, in the 0-dimensional case, the degree function is again continuous. This applies in particular to solenoids [Benedetti and Gambaudo 2003] and laminations defined by repetitive graphs, which were introduced in [Ghys 1999] and studied in [Alcalde Cuesta et al. 2009; Blanc 2001; Lozano Rojo 2011]:

Corollary 3.8. *Let $\mathcal{R}$ be the orbit equivalence relation defined by a compactly generated pseudogroup Γ acting freely on a locally compact separable 0-dimensional space T. Let $\delta : \mathcal{R} \to \mathbb{R}^*_+$ be a continuous harmonic cocycle. Any δ-averaging sequence $\{A_n\}$ gives rise to a harmonic measure ν on T supported by the limit set of $\{A_n\}$.* □

In order to extend Theorem 3.6 to non-free actions, we can adopt two different strategies. Let us first recall that the notion of an equivalence relation is enough to describe the transverse structure of a lamination in the Borel context. More precisely, any Borel or topological lamination $\mathcal{F}$ induces a Borel equivalence relation $\mathcal{R}$ on a total transversal T (compare to Remark 2.1) defined by the action of the holonomy pseudogroup. We refer to the Ph.D. thesis of M. Bermúdez [2004] for the definition of a Borel lamination. If $\mathcal{R}$ is a discrete Borel equivalence relation defined by the action of a Borel pseudogroup Γ acting on a compact space T and $\delta : \mathcal{R} \to \mathbb{R}^*_+$ is a Borel cocycle, then the proof of Theorem 3.6 remains valid. In the topological context, Theorem 3.6 is not exactly equivalent to the situation above because the transverse holonomy groupoid and the equivalence relation are only Borel isomorphic on the residual set of leaves without holonomy. Another strategy consists of replacing étale equivalence relations with étale groupoids and proving that averaging sequences for stationary cocycles define stationary measures on groupoids. Details will be reported elsewhere.

4. Averaging sequences in the continuous setting

We are interested in stating Theorem 3.6 in the continuous setting, namely for a compact laminated space $(M, \mathcal{F})$. Instead of working with quasi-invariant measures, we are going to use tangentially smooth measures. These form a larger class than harmonic measures. As previously mentioned, transverse invariant measures for foliations are rather rare, but harmonic measures always exist. Harmonic measures were introduced by L. Garnett [1983]. In Sections 4A and 4B we will study these measures and recall some notation. In Section 4C we will construct a differential foliated 1-form from a given cocycle. Finally, in Section 4D we will use this foliated form to prove the continuous analogue of Theorem 3.6.

4A. ***Tangentially smooth measures.*** Consider a regular Borel measure μ on M. Using a C^r foliated atlas $\mathcal{A}$, we can give a local decomposition $\mu = \int \lambda_i^y d\nu_i(y)$ on each foliated chart U_i, where λ_i^y is a measure on the plaque $\varphi_i^{-1}(P_i \times \{y\})$ and ν_i a measure on T_i. In order to define the foliated Laplace operator $\Delta_{\mathcal{F}}$, we can always assume that $r \geq 3$ up to C^1-equivalence of foliated atlases, and we fix a tangentially C^r-smooth Riemannian metric g along the leaves of $\mathcal{F}$.

Definition 4.1 [Alcalde Cuesta and Rechtman 2011]. A measure μ on M is *tangentially smooth* if for every $i \in I$ and ν_i-almost every $y \in T_i$, the measures λ_i^y are absolutely continuous with respect to the Riemannian volume $d\mathrm{vol}$ restricted to the plaque passing through y, and the density functions $h_i(x, y) = d\lambda_i^y / d\mathrm{vol}(x, y)$ are smooth functions of class C^{r-1} on the plaques.

Observe that the local decomposition of μ is not necessarily unique. Let

$$\mu|_{U_i} = \int \lambda_i^y d\nu_i(y) = \int \bar{\lambda}_i^y d\bar{\nu}_i(y)$$

be two decompositions. Then we obtain

$$\int_{T_i} \int_{P_i \times \{y\}} h_i(x, y)\, d\mathrm{vol}(x, y)\, d\nu_i(y) = \int_{T_i} \int_{P_i \times \{y\}} \bar{h}_i(x, y)\, d\mathrm{vol}(x, y)\, d\bar{\nu}_i(y),$$

and we can consider the Radon–Nikodým derivative $\delta_i(y) = d\nu_i / d\bar{\nu}_i(y)$ such that $\bar{h}_i(x, y) = \delta_i(y) h_i(x, y)$. This situation arises naturally in the intersection of two foliated charts U_i and U_j. Indeed, if $U_i \cap U_j \neq \varnothing$, we have

$$\mu|_{U_i \cap U_j} = \int \lambda_i^y d\nu_i(y) = \int \lambda_j^y d\nu_j(y).$$

Thus, as before, we deduce that

$$\delta_{ij}(y) = \frac{d\nu_i}{d((\gamma_{ji})_*\nu_j)}(y) = \frac{h_j(\varphi_{ij}^y(x), \gamma_{ij}(y))}{h_i(x, y)}. \tag{4-1}$$

Then the functions h_i satisfy $\log h_j - \log h_i = \log \delta_{ij}$ on $U_i \cap U_j$. Since δ_{ij} is a function on T_i, we have that $d_{\mathcal{F}} \log h_i = d_{\mathcal{F}} \log h_j$. Then $\eta = d_{\mathcal{F}} \log h_i$ is a well-defined foliated 1-form of class C^{r-2} along the leaves, which makes it possible to estimate the transverse measure distortion under the holonomy.

Definition 4.2. The foliated 1-form η is the *modular form* of μ.

4B. *Harmonic measures.*

Definition 4.3 [Garnett 1983]. We will say that μ is *harmonic* if $\int \Delta_{\mathcal{F}} f \, d\mu = 0$ for every continuous tangentially C^{r-1}-smooth function $f : M \to \mathbb{R}$.

According to [Garnett 1983, Theorem 1], any harmonic measure is an example of a tangentially smooth measure since the densities h_i are positive harmonic functions of class C^{r-1} on the plaques. In particular, any transverse invariant measure combined with the Riemannian volume on the leaves gives a harmonic measure which is called *completely invariant*. A harmonic measure μ is completely invariant if and only if $\eta = 0$; we refer to [Candel 2003, Corollary 5.5]. In the general harmonic case, the following proposition states some properties of the modular form. This proposition is a refined version of [Deroin 2003, Lemma 4.19].

Proposition 4.4 [Deroin 2003]. *If μ is a harmonic measure, then η is a bounded foliated* 1*-form which admits a uniformly tangentially Lipschitz primitive* $\log h$ *on the residual set of leaves without holonomy.*

Proof. Let $\mathcal{A} = \{(U_i, \phi_i)\}_{i \in I}$ be a good C^r foliated atlas of $(M, \mathcal{F})$, and h_i the local density functions of μ. Let us first observe that since the functions h_i coincide on the intersections of the plaques modulo multiplication by a constant, they define a primitive of the induced 1-form on the holonomy covering of each leaf L. If $\mathcal{F}$ has no essential holonomy, the functions $\log h_i$ can be glued together to obtain a measurable global primitive $\log h$ of η. In general, the modular form η admits a continuous primitive $\log h$ on the residual set of leaves without holonomy. Now let us assume that $\mathcal{A}$ is a refinement of a good atlas $\mathcal{A}' = \{(U'_i, \phi'_i)\}_{i \in I}$, and h'_i are the corresponding local densities. Thus, every plaque of U_i is relatively compact in a plaque of U'_i. In fact, using a vertical reparametrization, we can suppose that $\phi_i^{-1}(P_i \times \{y\}) \subset (\phi'_i)^{-1}(P'_i \times \{y\})$ for every $y \in T_i$. There exists a relatively compact open set $V \subset P'_i$ such that $\phi_i^{-1}(P_i \times \{y\}) \subset (\phi'_i)^{-1}(V \times \{y\})$ for every $y \in T_i$. Since h_i is harmonic, the Harnack inequality implies the existence of a constant $C_i > 0$

such that

$$\frac{1}{C_i} \le \frac{h_i(x, y)}{h_i(x_0, y)} \le C_i \tag{4-2}$$

for all $x, x_0 \in P_i$ and for all $y \in T_i$. Since the atlases $\mathscr{A}$ and $\mathscr{A}'$ are finite, the primitive $\log h$ is uniformly Lipschitz in the tangential coordinate x. □

4C. ***Modular form associated to a cocycle.*** We now describe how to construct a modular 1-form $\eta \in \Omega^1(\mathscr{F})$ from a Borel or continuous cocycle $\delta : \mathscr{R} \to \mathbb{R}_+^*$. For simplicity, $\mathscr{R}$ is endowed here with the natural Borel or topological structure induced by the Borel or topological groupoid structure on the transverse holonomy groupoid G formed by the germs $\langle\gamma\rangle_y$ of the elements γ of Γ at the points y of their domains; see [Moore and Schochet 2006]. The natural projection

$$(\beta, \alpha) : \langle\gamma\rangle_y \in G \mapsto (y, \gamma(y)) \in \mathscr{R}$$

becomes an isomorphism of Borel or topological groupoids in restriction to the residual set of leaves without holonomy. Equivalently, we can consider a Borel or continuous cocycle $\delta : G \to \mathbb{R}_+^*$ projectable on $\mathscr{R}$.

We start by considering tangentially C^r-smooth Borel or continuous functions $c_{ki} : U_i \cap U_k \to \mathbb{R}$ given by

$$c_{ki}(\varphi_k^{-1}(x, y)) = \log \delta_{ki}(y),$$

where $\delta_{ki}(y) = \delta(y, \gamma_{ki}(y))$ for all $(x, y) \in P_k \times T_k$. By choosing a tangentially C^r-smooth partition of unity $\{\rho_i\}_{i=1}^m$ subordinated to the foliated atlas $\mathscr{A}$, we can glue the functions c_{ki} obtaining tangentially C^r-smooth Borel or continuous functions $c_i : U_i \to \mathbb{R}$ given by

$$c_i = \sum_{k=1}^m \rho_k c_{ki}.$$

The cocycle condition implies that $c_{ij} = c_{kj} - c_{ki}$, so that

$$c_j - c_i = \sum_{k=1}^m \rho_k c_{kj} - \sum_{k=1}^m \rho_k c_{ki} = \left(\sum_{k=1}^m \rho_k\right) c_{ij} = c_{ij}.$$

Hence, for each $i = 1, \ldots, m$, we can define a tangentially C^{r-1}-smooth Borel or continuous foliated 1-form

$$\eta_i = \sum_{k=1}^m (d_{\mathscr{F}}\rho_k) c_{ki}$$

on U_i. Each local 1-form η_i is exact:

$$\eta_i = \sum_{k=1}^m (d_{\mathscr{F}}\rho_k) c_{ki} = d_{\mathscr{F}} c_i = d_{\mathscr{F}} \log h_i,$$

where $h_i = e^{c_i} : U_i \to \mathbb{R}^*_+$ is a Borel or continuous function of class C^r along the leaves.

Proposition 4.5. *There is a well defined Borel or continuous closed foliated* 1*-form* $\eta \in \Omega^1(\mathcal{F})$ *such that* $\eta|_{U_i} = \eta_i$.

Proof. For each pair $i, j \in \{1, \dots, m\}$, we have that

$$\eta_j - \eta_i = \sum_{k=1}^{m} (d_{\mathcal{F}} \rho_k) c_{kj} - \sum_{k=1}^{m} (d_{\mathcal{F}} \rho_k) c_{ki} = \left(\sum_{k=1}^{m} d_{\mathcal{F}} \rho_k \right) c_{ij} = 0$$

on $U_i \cap U_j$. So the 1-form η is well defined, Borel, or continuous, and closed. □

Definition 4.6. The foliated 1-form η is the modular form of δ.

Remark 4.7. (i) The modular form η depends on the choice of the partition of unity, but its cohomology class does not.

(ii) As for harmonic measures, the modular form η of a Borel or continuous cocycle δ admits a Borel or continuous primitive $\log h$ on the residual set of leaves without holonomy. Thus, assuming that $\mathcal{F}$ has no holonomy (or passing to the holonomy covers of the leaves), we may find a global Borel or continuous primitive on M (respectively, a Borel or continuous primitive on the holonomy groupoid Hol($\mathcal{F}$)); see [Alcalde Cuesta and Rechtman 2011].

4D. *Continuous averaging sequences.* In the present setting, we can reformulate the *Radon–Nikodým problem* as the problem of determining tangentially smooth measures μ on M which admit η as their modular form. The aim of this section is to establish Theorem 3.6 for laminations. First we need a continuous analog of Definition 3.3. Consider a d-dimensional lamination $\mathcal{F}$ of class C^r on a compact space M, endowed with a tangentially C^r-smooth Riemannian metric g, and a continuous cocycle $\delta : \mathcal{R} \to \mathbb{R}^*_+$. The modular form η admits a continuous tangentially C^r-smooth primitive $\log h$ on the residual set of leaves without holonomy. On each leaf L_y without holonomy and passing through $y \in T$, we can multiply g by the normalized density function $h/h(y)$ to obtain a *modified metric* $(h/h(y))g$.

Definition 4.8. Let $\{V_n\}$ be a sequence of compact domains with boundary contained in a sequence of leaves without holonomy L_{y_n}. We will say that $\{V_n\}$ is a *η-averaging sequence for $\mathcal{F}$* if

$$\lim_{n\to\infty} \frac{\text{area}_\eta(\partial V_n)}{\text{vol}_\eta(V_n)} = 0$$

where area_η denotes the $(d-1)$-volume and vol_η the d-volume with respect to the modified metric along L_{y_n}. A leaf L_y is *η-Følner* if it contains an η-averaging sequence $\{V_n\}$ such that $\text{area}_\eta(\partial V_n)/\text{vol}_\eta(V_n) \to 0$ as $n \to +\infty$.

Remark 4.9. (i) The isoperimetric ratio $\text{area}_\eta(\partial V_n)/\text{vol}_\eta(V_n)$ does not depend on the choice of y or h in the second definition. This justifies our notation here, which differs slightly from that used in [Alcalde Cuesta and Rechtman 2011].

(ii) When μ is a completely invariant harmonic measure, the normalized density function is equal to one, and thus the modified volume and the Riemannian volume coincide. Hence we recover the common definition of an averaging sequence.

(iii) For harmonic measures, Harnack's inequalities (4-2) imply that the modified volume of the plaques and the modified area of their boundaries remain uniformly bounded.

Theorem 4.10. *Let $(M, \mathcal{F})$ be a C^r lamination of a compact space M, $1 \leq r \leq \infty$, and let $\mathcal{R}$ be the equivalence relation induced by $\mathcal{F}$ on a total transversal T. Consider a continuous cocycle $\delta : \mathcal{R} \to \mathbb{R}^*_+$, and let η be the modular form of δ. Assume that $\mathcal{F}$ admits a foliated atlas such that the modified volume of the plaques is bounded.*

(i) *Any η-averaging sequence $\{V_n\}$ for $\mathcal{F}$ gives rise to a tangentially smooth measure μ whose support is contained in the limit set of $\{V_n\}$ and whose modular form is equal to η.*

(ii) *If η has a primitive $\log h$ such that h is a harmonic function, then μ is a harmonic measure.*

Proof. As in the discrete case, we will start by constructing a sequence of foliated d-currents

$$\xi_n(\alpha) = \frac{1}{\text{vol}_\eta(V_n)} \int_{V_n} \frac{h}{h(y_n)} \alpha,$$

where α is a foliated d-form. By passing to a subsequence, the sequence ξ_n converges to a foliated d-current ξ. Let μ be the measure on M associated with the current ξ. For every function $f \in C(T)$, we have $\int f \, d\mu = \xi(f\omega)$, where $\omega = d\text{vol}$ is the volume form along the leaves.

Now, we will prove that μ is a tangentially smooth measure with modular form η. Consider a good C^r foliated atlas $\mathcal{A} = \{(U_i, \phi_i)\}_{i \in I}$ obtained by refinement from a given good atlas, and whose plaques have bounded modified volume. As we mentioned before, up to C^1-equivalence, we can now assume that $r \geq 3$. Since the modified volume of the plaques of $\mathcal{A}$ and the modified area of their boundaries remain bounded, the traces $A_n = V_n \cap T$ of the domains V_n on the total transversal T form a δ-averaging sequence, as in Definition 3.3. In fact, since V_n is covered by the plaques P_y of $\mathcal{A}$ centered at the points y of A_n, we have

$$\text{vol}_\eta(V_n) = \int_{V_n} \omega_\eta \leq \sum_{y \in A_n} \int_{P_y} \omega_\eta = \sum_{y \in A_n} \left(\int_{P_y} \frac{h(x, y)}{h(0, y)} \, d\text{vol}(x, y) \right) \delta(y, y_n),$$

where ω_η is the modified volume form along the leaves and $h(x, y)$ denotes the density function restricted to a foliated chart U_y containing the plaque P_y. Then there is a constant $C > 0$ such that $\mathrm{vol}_\eta(V_n) \leq C|A_n|_{y_n}$. Actually, we can choose $C > 0$ such that $1/C \leq \mathrm{vol}_\eta(V_n)/|A_n|_{y_n} \leq C$. Thus, by passing to a subsequence, we may assume that the ratio $\mathrm{vol}_\eta(V_n)/|A_n|_{y_n}$ converges to a constant $c > 0$. Now, as stated in the proof of Theorem 3.6, we may also assume that the sequence of measures $\nu_n(B) = |B \cap A_n|_{y_n}/|A_n|_{y_n}$ converge to a quasi-invariant measure ν on T whose Radon–Nikodým derivative is equal to δ. Combined with the modified Riemannian volume along the leaves, this transverse measure gives us a tangentially smooth measure μ' on M. Thus, for every function $f \in C(M)$ with support in U_i, we have

$$\int f \, d\mu' = \int_{T_i} \int_{P_i \times \{y\}} f(x, y) \frac{h_i(x, y)}{h_i(0, y)} \, d\mathrm{vol}(x, y) \, d\nu(y).$$

Then

$$\text{(4-3)} \quad \int f \, d\mu' = \lim_{n \to +\infty} \frac{1}{|A_n|_{y_n}} \sum_{y \in V_n \cap T_i} \left(\int_{P_i \times \{y\}} f(x,y) \frac{h_i(x,y)}{h_i(0,y)} dvol(x,y) \right) \delta(y, y_n)$$
$$= \lim_{n \to +\infty} \frac{1}{|A_n|_{y_n}} \sum_{y \in V_n \cap T_i} \int_{P_i \times \{y\}} f \omega_\eta.$$

On the other hand, by definition, we have

$$\text{(4-4)} \quad \int f \, d\mu = \xi(f\omega) = \lim_{n \to +\infty} \frac{1}{\mathrm{vol}_\eta(V_n)} \int_{V_n} f \omega_\eta$$
$$= \lim_{n \to +\infty} \frac{1}{\mathrm{vol}_\eta(V_n)} \sum_{y \in V_n \cap T_i} \int_{P_i \times \{y\}} f \omega_\eta.$$

Comparing identities (4-3) and (4-4), we deduce that $\mu = (1/c)\mu'$ is a tangentially smooth measure with modular form η.

To conclude, we will prove that μ is harmonic when h is harmonic. We will start by denoting the normalized density functions on the leaves L_{y_n} by $h_n = h/h(y_n)$. Since the Laplace operator $\Delta_{\mathcal{F}}$ preserves continuous functions, we have

$$\int \Delta_{\mathcal{F}} f \, d\mu = \lim_{n \to \infty} \frac{1}{\mathrm{vol}_h(V_n)} \int_{V_n} (\Delta_{\mathcal{F}} f) h_n \omega$$

for all $f \in C(T)$. Green's formula implies that

$$\int_{V_n} (\Delta_{\mathcal{F}} f) h_n \omega = \int_{V_n} ((\Delta_{\mathcal{F}} f) h_n - f(\Delta_{\mathcal{F}} h_n) \omega = \int_{\partial V_n} h_n \iota_{\mathrm{grad}(f)} \omega - f \iota_{\mathrm{grad}(h_n)} \omega.$$

Since h_n is harmonic, we have

$$\int_{\partial V_n} \iota_{\mathrm{grad}(h_n)} \omega = \int_{V_n} \mathrm{div}(\mathrm{grad}(h_n)) \omega = \int_{V_n} (\Delta_{\mathcal{F}} h_n) \omega = 0$$

and then

$$0 \leq \left| \int_{\partial V_n} f \iota_{\mathrm{grad}(h_n)} \omega \right| \leq \|f\|_\infty \int_{\partial V_n} \iota_{\mathrm{grad}(h_n)} \omega = 0$$

for all $n \in \mathbb{N}$. On the other hand, since f is bounded, there exists a constant $k > 0$ depending only on f, such that

$$0 \leq \left| \frac{1}{\mathrm{vol}_h(V_n)} \int_{\partial V_n} h_n \iota_{\mathrm{grad}(f)} \omega \right| \leq \lim_{n\to\infty} k \frac{\mathrm{area}_\eta(\partial V_n)}{\mathrm{vol}_\eta(V_n)} = 0.$$

Therefore

$$\int \Delta_{\mathcal{F}} f \, d\mu = \lim_{n\to\infty} \frac{1}{\mathrm{vol}_h(V_n)} \int_{V_n} (\Delta_{\mathcal{F}} f) h_n \omega = 0,$$

that is, μ is a harmonic measure. □

Remark 4.11. (i) If $\delta : \mathcal{R} \to \mathbb{R}_+^*$ is a Borel cocycle with modular form η, Theorem 4.10 also remains valid. So any η-averaging sequence for $\mathcal{F}$ gives rise to a tangentially smooth measure μ that is harmonic when η admits a primitive $\log h$ such that h is a harmonic function.

(ii) According to Remark 4.7(ii), the notion of η-Følner may be applied to the holonomy covers of the leaves of $\mathcal{F}$. Thus it suffices to replace $\mathcal{F}$ with the lifted lamination in the holonomy groupoid Hol($\mathcal{F}$) in order to globalize the previous result. As in the discrete setting, details will be discussed elsewhere.

5. Examples

5A. ***Discrete averaging sequences for amenable non-Følner actions.*** There are amenable actions of nonamenable discrete groups whose orbits contain averaging sequences [Kaimanovich 2001]. For example, let $\partial\Gamma$ be the space of ends of the free group Γ with two generators α and β whose elements are infinite words $x = \gamma_1\gamma_2 \dots$ with letters γ_n in $\Phi = \{\alpha^{\pm 1}, \beta^{\pm 1}\}$. If ν denotes the equidistributed probability measure on $\partial\Gamma$ (such that all cylinders consisting of infinite words with fixed first n letters have the same measure), then Γ acts essentially freely on $\partial\Gamma$ by sending each generator γ and each infinite word $x = \gamma_1\gamma_2 \dots$ to $\gamma.x = \gamma\gamma_1\gamma_2 \dots$. Since this action is amenable, according to [Kaimanovich 1997, Theorem 2], we know that ν-almost every orbit is δ-Følner (where δ is the Radon–Nikodým derivative of ν); see also [Alcalde Cuesta and Rechtman 2011, Proposition 4.1]. We recall here an explicit construction by Kaimanovich [2001].

For each $x \in \partial\Gamma$, let $b_x : \Gamma \to \mathbb{R}$ be the *Busemann function* defined by

$$b_x(\gamma) = \lim_{n\to+\infty} (d_\Gamma(\gamma, x_{[n]}) - d_\Gamma(1, x_{[n]})),$$

where d_Γ is the Cayley graph metric, $x_{[n]}$ is the word consisting of the first n letters of x, and 1 is the identity element. The level sets $H_k(x) = \{\gamma \in \Gamma : b_x(\gamma) = k\}$, are the *horospheres* centered at x. The Radon–Nikodým derivative of ν is given by

$$\delta(\gamma^{-1}.x, x) = \frac{d\gamma.\nu}{d\nu}(x) = 3^{-b_x(\gamma)},$$

where $\gamma.\nu$ is the translation of ν by γ. Since $|\cdot|_x = \delta(\cdot, x)$ is a harmonic measure on $\Gamma.x$, ν is also a harmonic measure. In fact, as stated in [Kaimanovich 2000, Theorem 17.4], ν is the unique harmonic probability measure on $\partial\Gamma$.

Let A_n^x be the set of all points $\gamma^{-1}.x$ in $\Gamma.x$ such that $0 \le b_x(\gamma) = d_\Gamma(1, \gamma) \le n$. Since

$$|A_n^x \cap H_k(x)|_x = \sum_{b_x(\gamma)=d_\Gamma(1,\gamma)=k} \delta(\gamma^{-1}.x, x) = 3^k \frac{1}{3^k} = 1$$

for all $0 \le k \le n$, we have that $|A_n|_x = n+1$. But $\partial A_n^x = \{1\} \cup (A_n^x \cap H_n(x))$, and so $|\partial A_n^x|_x = 2$. The δ-averaging sequence $\{A_n^x\}$ defines a harmonic measure (which is equal to ν up to multiplication by a constant).

5B. ***Averaging sequences for hyperbolic surfaces.*** The geodesic and horocycle flows are classical examples of flows on the unitary tangent bundle of a compact hyperbolic surface. They are given by the right actions of the diagonal subgroup

$$D = \left\{ \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} : t \in \mathbb{R} \right\}$$

and the unipotent subgroup

$$H^+ = \left\{ \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} : s \in \mathbb{R} \right\}$$

of $G = \mathrm{PSL}(2, \mathbb{R})$ on the quotient $\Gamma\backslash G$ by the left action of a uniform lattice Γ. If $\mathbb{H}$ denotes the hyperbolic plane, we can identify $\Gamma\backslash G$ with the unitary tangent bundle of the compact hyperbolic surface $\Gamma\backslash\mathbb{H}$. The right action of the normalizer A of H^+ in $\mathrm{PSL}(2, \mathcal{R})$ defines a foliation $\mathcal{F}$ by Riemann surfaces on $\Gamma\backslash G$. Since A is an amenable group, $\mathcal{F}$ is an amenable non-Følner foliation. Moreover, there is an A-invariant measure μ on $\Gamma\backslash G$. Garnett [1983] proved that μ is a harmonic measure by describing its density function on a foliated chart.

We can identify G/A with the boundary $\partial\mathbb{H}$ by sending each coset of A in G to the center of the horocycle defined by the corresponding coset of H^+ in G. For each point $z \in \mathbb{H}$, there is a unique probability measure ν_z on $\partial\mathbb{H}$ which is invariant by the action of all isometries of $\mathbb{H}$ fixing z. This measure is the image of the normalized Lebesgue measure on the circle of the tangent plane at z under the exponential map, and is called the *visual measure* at z. According to [Garnett 1983, Proposition 2], the normalized density function is given by $d\nu_z/d\nu_{z_0}(x)$ where

$z, z_0 \in \mathbb{H}$ and $x \in \partial\mathbb{H}$. In particular, for $x = \infty$, we have

$$\frac{d\nu_z}{d\nu_{z_0}}(\infty) = \frac{y}{y_0},$$

where $z = x + iy$ and $z_0 = x_0 + iy_0$. In the leaf passing through $x = \infty$, the sequence $V_n^\infty = \{z \in \mathbb{H} : -1 \le x \le 1,\ e^{-n} \le y \le 1\}$ becomes an η-averaging sequence (where η is the modular form of μ). Indeed, on the one hand, we have

$$\mathrm{area}_\eta(V_n^\infty) = \int_{V_n^\infty} \frac{d\nu_z}{d\nu_i}(\infty)\, d\mathrm{vol}(z) = \int_{V_n^\infty} y \frac{dx \wedge dy}{y^2} = \int_1^1 dx \int_{e^{-n}}^1 \frac{dy}{y} = 2n.$$

On the other hand, the modified length of a smooth curve $\sigma(t) = x(t) + iy(t)$ (with $0 \le t \le l$) is given by $\mathrm{length}_\eta(\sigma) = \int_0^l \sqrt{x'(t)^2 + y'(t)^2}\, dt$, and so we have

$$\mathrm{length}_\eta(\partial V_n^\infty) = 2(2 + (1 - e^n)) \le 6.$$

As before, this η-averaging sequence defines a harmonic measure (which is equal to μ up to multiplication by a constant). In fact, all leaves are η-Følner since for each point $x \in \partial\mathbb{H}$ obtained as the image of ∞ under $g \in G$, the sets $V_n^x = g(V_n^\infty)$ form an η-averaging sequence in the leaf passing through x.

5C. ***Averaging sequences for torus bundles over the circle.*** In conclusion, we will now present other examples of foliations on homogeneous spaces studied by É. Ghys and V. Sergiescu [1980]. Each matrix $A \in \mathrm{SL}(2, \mathbb{Z})$ with $|\mathrm{tr}(A)| > 2$ defines a natural representation $\varphi : \mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}^2)$ which extends to a representation $\Phi : \mathbb{R} \to \mathrm{Aut}(\mathbb{R}^2)$ given by $\Phi(t) = A^t$. If $\lambda > 1$ and $\lambda^{-1} < 1$ are the eigenvalues of A, then Φ is conjugated to the representation Φ_0 given by

$$\Phi_0(t) = \begin{pmatrix} \lambda^t & 0 \\ 0 & \lambda^{-t} \end{pmatrix}.$$

Let T_A^3 be the homogeneous space obtained as the quotient of the Lie group $G = \mathbb{R}^2 \rtimes_\Phi \mathbb{R}$ with group law $(x, y, t).(x', y', t') = ((x, y) + A^t(x', y'), t + t')$ by the uniform lattice $\Gamma = \mathbb{Z}^2 \rtimes_\varphi \mathbb{Z}$ with a similar law. Observe that G is isomorphic to the solvable group $\mathrm{Sol}^3 = \mathbb{R}^2 \rtimes_{\Phi_0} \mathbb{R}$ with group law

$$(x, y, t).(x', y', t') = (x + \lambda^t x', y + \lambda^{-t} y', t + t')$$

(where x and y are the first and second coordinate with respect to the eigenbasis) and T_A^3 is diffeomorphic to the quotient of Sol^3 by a uniform lattice Γ_0. The right action of the image A of the monomorphism

$$(a, b) \in \mathbb{R} \rtimes \mathbb{R}_+^* \mapsto \left(a, 0, \frac{\log b}{\log \lambda}\right) \in \mathrm{Sol}^3$$

defines a foliation $\mathscr{F}$ on T_A^3. The Lebesgue measure on T_A^3 defined by the volume form $\Omega = dx \wedge dy \wedge dt$ is a tangentially smooth measure. Since the Riemannian volume along the right orbits is given by

$$\frac{da \wedge db}{b^2} = (\log \lambda)\, \lambda^{-t} dx \wedge dt,$$

the density function is equal to $\lambda^t/\log \lambda$. In the orbit of the identity element, the sequence $V_n = \{(a, b) \in A : -1 \le a \le 1,\ e^{-n \log \lambda} \le b \le 1\}$ becomes an η-averaging sequence (where η is the modular form of μ). Indeed, on the one hand, we have

$$\text{area}_\eta(V_n) = \int_{V_n} \frac{1}{\log \lambda} \lambda^t (\log \lambda) \lambda^{-t} dx \wedge dt = \int_1^1 dx \int_{-n}^0 dt = 2n.$$

On the other hand, the modified length of a smooth curve $\sigma(t) = (a(t), b(t))$ (with $0 \le t \le L$) is given by $\text{length}_\eta(\sigma) = \int_0^L \sqrt{a'(t)^2 + b'(t)^2}\, dt$, and so we have that

$$\text{length}_\eta(\partial V_n) = 2(2 + (1 - e^{n \log \lambda})) \le 6.$$

By replacing the orbit corresponding to $y = 0$ with another orbit, it is easy to see that all leaves are η-Følner. As in the previous example, all η-averaging sequences define (up to multiplication by a constant) the same harmonic measure, the Lebesgue measure.

6. Final comments

6A. ***Discrete and continuous averaging sequences.*** Comparing the discrete and continuous settings, a natural question arises: what is the relation between δ-averaging and η-averaging sequences? Let us first notice that repeating the same argument as in the classical case (see [Kanai 1985, Theorem 4.1]), the boundedness condition derived from Harnack's inequalities in Remark 4.9(iii) implies that *the leaf L_y is η-Følner if and only if the equivalence class $\mathcal{R}[y]$ is δ-Følner.* But then what is the relation between the harmonic measures defined by δ-averaging and η-averaging sequences? In this case, the answer is more subtle, and we have to use an important result of R. Lyons and Sullivan [1984], completed later by Kaimanovich [1992] and, independently, by W. Ballman and F. Ledrappier [1996], about the discretization of harmonic functions on Riemannian manifolds. First, according to [Lyons and Sullivan 1984, Theorem 6], if μ is a harmonic measure, then the transverse measure ν (well defined up to equivalence) is π-harmonic, where π is a transition kernel defining a random walk on $\mathcal{R}$, different from the simple random walk considered in Definition 3.4. Reciprocally, assuming that T admits a relatively compact neighborhood which meets almost every leaf in a recurrent set, [Ballmann and Ledrappier 1996, Main Theorem] implies that μ is harmonic if ν is π-harmonic.

6B. ***Amenability.*** It is not a coincidence that all the examples in Section 5 are amenable: according to a result of Kaimanovich [1997], amenable foliations admit always averaging sequences. In fact, if $\mathcal{F}$ is an amenable foliation with respect to a tangentially smooth measure μ, then $\mathcal{F}$ is η- Følner, that is, μ-almost every leaf is η-Følner; see [Alcalde Cuesta and Rechtman 2011, Proposition 4.3]. This paper can be viewed as a sequel to [Alcalde Cuesta and Rechtman 2011] where we proved that minimal η-Følner foliations are μ-amenable (assuming that the modified volume of the plaques is bounded). To complete the series, we have to prove that any foliation is amenable with respect to a tangentially smooth measure μ constructed from an averaging sequence using Theorem 4.10.

References

[Alcalde Cuesta and Rechtman 2011] F. Alcalde Cuesta and A. Rechtman, "Minimal Følner foliations are amenable", *Discrete Contin. Dyn. Syst.* **31**:3 (2011), 685–707. MR 2825634 Zbl 05988150

[Alcalde Cuesta et al. 2009] F. Alcalde Cuesta, A. Lozano Rojo, and M. Macho Stadler, "Dynamique transverse de la lamination de Ghys–Kenyon", pp. 1–16 in *Équations différentielles et singularités: en l'honneur de J. M. Aroca*, edited by F. Cano et al., Astérisque **323**, Soc. Math. France, Paris, 2009. MR 2011g:37062 Zbl 1203.37011

[Álvarez López and Candel 2009] J. A. Álvarez López and A. Candel, "Equicontinuous foliated spaces", *Math. Z.* **263**:4 (2009), 725–774. MR 2010i:53040 Zbl 1177.53026

[Ballmann and Ledrappier 1996] W. Ballmann and F. Ledrappier, "Discretization of positive harmonic functions on Riemannian manifolds and Martin boundary", pp. 77–92 in *Actes de la Table Ronde de Géométrie Différentielle: en l'honneur de Marcel Berger* (Luminy, 1992), edited by A. L. Besse, Sémin. Congr. **1**, Soc. Math. France, Paris, 1996. MR 97m:58207 Zbl 0885.53037

[Benedetti and Gambaudo 2003] R. Benedetti and J.-M. Gambaudo, "On the dynamics of $\mathbb{G}$-solenoids: applications to Delone sets", *Ergodic Theory Dynam. Systems* **23**:3 (2003), 673–691. MR 2004f:37019 Zbl 1124.37009

[Bermúdez 2004] M. Bermúdez, *Laminations Boréliennes*, thesis, Université Claude Bernard, Lyon, 2004, Available at http://tinyurl.com/Bermudez-2004.

[Blanc 2001] E. Blanc, *Propriétés génériques des laminations*, thesis, Université Claude Bernard, Lyon, 2001.

[Candel 2003] A. Candel, "The harmonic measures of Lucy Garnett", *Adv. Math.* **176**:2 (2003), 187–247. MR 2004m:58057 Zbl 1031.58003

[Deroin 2003] B. Deroin, *Laminations par variétés complexes*, thesis, École Normale Supérieure de Lyon, 2003.

[Følner 1955] E. Følner, "On groups with full Banach mean value", *Math. Scand.* **3** (1955), 243–254. MR 18,51f Zbl 0067.01203

[Garnett 1983] L. Garnett, "Foliations, the ergodic theorem and Brownian motion", *J. Funct. Anal.* **51**:3 (1983), 285–311. MR 84j:58099 Zbl 0524.58026

[Ghys 1995] É. Ghys, "Topologie des feuilles génériques", *Ann. of Math.* (2) **141**:2 (1995), 387–422. MR 96b:57032 Zbl 0843.57026

[Ghys 1999] É. Ghys, "Laminations par surfaces de Riemann", pp. ix, xi, 49–95 in *Dynamique et géométrie complexes* (Lyon, 1997), Panor. Synthèses **8**, Soc. Math. France, Paris, 1999. In French; translated by L. Kay in *Complex dynamics and geometry*, SMF/AMS Texts and Monographs **10**, Amer. Soc. Math., Providence, RI, 2003, pp. 43–84. MR 2001g:37068 Zbl 1018.37028

[Ghys and Sergiescu 1980] É. Ghys and V. Sergiescu, "Stabilité et conjugaison différentiable pour certains feuilletages", *Topology* **19**:2 (1980), 179–197. MR 81k:57022 Zbl 0478.57017

[Goodman and Plante 1979] S. E. Goodman and J. F. Plante, "Holonomy and averaging in foliated sets", *J. Differential Geom.* **14**:3 (1979), 401–407. MR 81m:57020 Zbl 0475.57007

[Gromov 1993] M. Gromov, "Asymptotic invariants of infinite groups", pp. 1–295 in *Geometric group theory* (Sussex, 1991), vol. 2, edited by G. A. Niblo and M. A. Roller, London Math. Soc. Lecture Note Ser. **182**, Cambridge University Press, Cambridge, 1993. MR 95m:20041 Zbl 0841.20039

[Haefliger 1984] A. Haefliger, "Groupoïdes d'holonomie et classifiants", pp. 70–97 in *Structure transverse des feuilletages* (Toulouse, 1982), edited by J. Pradines, Astérisque **116**, Soc. Math. France, Paris, 1984. MR 86c:57026a Zbl 0562.57012

[Haefliger 2002] A. Haefliger, "Foliations and compactly generated pseudogroups", pp. 275–295 in *Foliations: geometry and dynamics* (Warsaw, 2000), edited by P. Walczak et al., World Scientific, River Edge, NJ, 2002. MR 2003g:58029 Zbl 1002.57059

[Hurder and Katok 1987] S. Hurder and A. Katok, "Ergodic theory and Weil measures for foliations", *Ann. of Math.* (2) **126**:2 (1987), 221–275. MR 89d:57042 Zbl 0645.57021

[Kaimanovich 1992] V. A. Kaimanovich, "Discretization of bounded harmonic functions on Riemannian manifolds and entropy", pp. 213–223 in *Potential theory* (Nagoya, 1990), edited by M. Kishi, de Gruyter, Berlin, 1992. MR 94b:31007 Zbl 0768.58054

[Kaimanovich 1997] V. A. Kaimanovich, "Amenability, hyperfiniteness, and isoperimetric inequalities", *C. R. Acad. Sci. Paris Sér. I Math.* **325**:9 (1997), 999–1004. MR 98j:28014 Zbl 0981.28014

[Kaimanovich 2000] V. A. Kaimanovich, "The Poisson formula for groups with hyperbolic properties", *Ann. of Math.* (2) **152**:3 (2000), 659–692. MR 2002d:60064 Zbl 0984.60088

[Kaimanovich 2001] V. A. Kaimanovich, "Equivalence relations with amenable leaves need not be amenable", pp. 151–166 in *Topology, ergodic theory, real algebraic geometry*, edited by V. Turaev and A. Vershik, Amer. Math. Soc. Transl. Ser. 2 **202**, Amer. Math. Soc., Providence, RI, 2001. MR 2003a:37009 Zbl 0990.28013

[Kanai 1985] M. Kanai, "Rough isometries, and combinatorial approximations of geometries of noncompact Riemannian manifolds", *J. Math. Soc. Japan* **37**:3 (1985), 391–413. MR 87d:53082 Zbl 0554.53030

[Lozano Rojo 2006] Á. Lozano Rojo, "The Cayley foliated space of a graphed pseudogroup", pp. 267–272 in *XIV Fall Workshop on Geometry and Physics* (Bilbao, 2005), edited by L. C. de Andrés et al., Publ. R. Soc. Mat. Esp. **8**, R. Soc. Mat. Esp., Madrid, 2006. MR 2008j:58027 Zbl 1158.58008

[Lozano Rojo 2011] Á. Lozano Rojo, "An example of a non-uniquely ergodic lamination", *Ergodic Theory Dynam. Systems* **31**:2 (2011), 449–457. MR 2776384 Zbl 1221.37049

[Lyons and Sullivan 1984] T. Lyons and D. Sullivan, "Function theory, random paths and covering spaces", *J. Differential Geom.* **19**:2 (1984), 299–323. MR 86b:58130 Zbl 0554.58022

[Moore and Schochet 2006] C. C. Moore and C. L. Schochet, *Global analysis on foliated spaces*, 2nd ed., Mathematical Sciences Research Institute Publications **9**, Cambridge University Press, New York, 2006. MR 2006i:58035 Zbl 1091.58015

[Paulin 1999] F. Paulin, "Propriétés asymptotiques des relations d'équivalences mesurées discrètes", *Markov Process. Related Fields* **5**:2 (1999), 163–200. MR 2001m:37010 Zbl 0937.28015

[Plante 1975] J. F. Plante, "Foliations with measure preserving holonomy", *Ann. of Math.* (2) **102**:2 (1975), 327–361. MR 52 #11947 Zbl 0314.57018

[Renault 1980] J. Renault, *A groupoid approach to C^*-algebras*, Lecture Notes in Mathematics **793**, Springer, Berlin, 1980. MR 82h:46075 Zbl 0433.46049

[Renault 2005] J. Renault, "The Radon–Nikodým problem for approximately proper equivalence relations", *Ergodic Theory Dynam. Systems* **25**:5 (2005), 1643–1672. MR 2006h:46065 Zbl 1093.46035

[Schapira 2003] B. Schapira, "Mesures quasi-invariantes pour un feuilletage et limites de moyennes longitudinales", *C. R. Math. Acad. Sci. Paris* **336**:4 (2003), 349–352. MR 2004j:37043 Zbl 1030.57043

[Sullivan 1976] D. Sullivan, "Cycles for the dynamical study of foliated manifolds and complex manifolds", *Invent. Math.* **36**:1 (1976), 225–255. MR 55 #6440 Zbl 0335.57015

Received January 8, 2011. Revised May 26, 2011.

FERNANDO ALCALDE CUESTA
DEPARTAMENTO DE XEOMETRÍA E TOPOLOXÍA
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
RÚA LOPE GÓMEZ DE MARZOA S/N
15782 SANTIAGO DE COMPOSTELA
SPAIN
fernando.alcalde@usc.es

ANA RECHTMAN
INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE
UNIVERSITÉ DE STRASBOURG
7 RUE RENÉ DESCARTES
67084 STRASBOURG
FRANCE
rechtman@unistra.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

AFFINE GROUP SCHEMES OVER SYMMETRIC MONOIDAL CATEGORIES

ABHISHEK BANERJEE

A well known result of Deligne shows that an affine commutative group scheme of rank r is annihilated by its rank. The purpose of this paper is to extend this result to affine group schemes over symmetric monoidal categories.

1. Introduction

One of the most important results in the study of group schemes is the following, presented in [Tate and Oort 1970].

Theorem 1.1 (Deligne's lemma). *Let $G = \mathrm{Spec}(A)$ be an affine commutative group scheme over a commutative, Noetherian ring k. Assume that A is a flat k-algebra of rank $r \geq 1$. Then, for any k-algebra B, all elements in the group $G(B)$ have an order dividing r.*

The purpose of this paper is to obtain an analogous result for group schemes in the relative algebraic geometry over a symmetric monoidal category. More precisely, we let $(\mathbf{C}, \otimes, 1)$ denote an abelian closed symmetric monoidal category. For instance, $\mathbf{C}$ could be the category of sheaves of abelian groups over a topological space, the category of comodules over a flat Hopf algebroid, the derived category of modules over a commutative ring k as well as chain complexes over all these categories. When $\mathbf{C} = k\text{-}Mod$, the category of modules over a commutative ring k, the algebraic geometry over $\mathbf{C}$ reduces to the usual algebraic geometry over $\mathrm{Spec}(k)$.

Given $(\mathbf{C}, \otimes, 1)$ as above, we refer to commutative and unital monoids in $\mathbf{C}$ as algebras in $\mathbf{C}$. Then, we define an affine commutative group scheme G free of finite rank over $\mathbf{C}$ to be a covariant functor from algebras in $\mathbf{C}$ to the category of abelian groups that satisfies certain conditions (see Definition 3.2 and Definition 3.3). The main result of this article is the following theorem:

The author is happy to acknowledge support from the Max-Planck-Institut für Mathematik, Bonn, where most of this paper was written.
MSC2010: 14L15, 18D10.
Keywords: group schemes, symmetric monoidal categories.

Theorem 1.2. *Let* $(\mathbf{C}, \otimes, 1)$ *be an abelian, closed,* $\mathbb{C}$*-linear symmetric monoidal category and let* G *be an affine commutative group scheme over* $\mathbf{C}$ *free and of finite rank* $r \geq 1$*. Then, for any algebra* B *in* $\mathbf{C}$ *and any element* u *in the group* $G(B)$*, we have* $u^r = 1_B$*, where* 1_B *denotes the identity element of* $G(B)$*.* (For the definition of $G(B)$, see (3-5).)

The relative algebraic geometry over a symmetric monoidal category has been developed in various works, such as [Deligne 1990; Hakim 1972; Toën and Vaquié 2009]. It is therefore natural to ask whether arithmetic geometry can be similarly developed in the general framework of symmetric monoidal categories. In particular, since the theory of finite flat group schemes is closely linked to arithmetic (see [Tate 1997], for instance), they are a natural starting point for such a theory. For more on group schemes, we refer the reader to [Demazure and Gabriel 1970].

2. Notations

In this section, we introduce notation that we will maintain throughout this paper. We let $(\mathbf{C}, \otimes, 1)$ denote an abelian symmetric monoidal category. Further, we suppose that $\mathbf{C}$ is closed, i.e., for any two objects $X, Y \in \mathbf{C}$, there exists an internal Hom object $\underline{\mathrm{Hom}}(X, Y)$ in $\mathbf{C}$ such that the functor

$$Z \mapsto \mathrm{Hom}(Z \otimes X, Y) \tag{2-1}$$

from $\mathbf{C}$ to the category of sets is represented by $\underline{\mathrm{Hom}}(X, Y)$. Here, we also note that, for any objects X, Y, Z and W in $\mathbf{C}$, we have

$$\begin{aligned} \mathrm{Hom}(W, \underline{\mathrm{Hom}}(Z, \underline{\mathrm{Hom}}(X, Y))) &\cong \mathrm{Hom}(W \otimes Z, \underline{\mathrm{Hom}}(X, Y)) \\ &\cong \mathrm{Hom}(W \otimes Z \otimes X, Y) \\ &\cong \mathrm{Hom}(W, \underline{\mathrm{Hom}}(Z \otimes X, Y)). \end{aligned} \tag{2-2}$$

Hence, it follows from Yoneda's lemma that we have a natural isomorphism

$$\underline{\mathrm{Hom}}(Z, \underline{\mathrm{Hom}}(X, Y)) \cong \underline{\mathrm{Hom}}(Z \otimes X, Y) \tag{2-3}$$

for any X, Y, Z and W in $\mathbf{C}$. Further, since $\mathbf{C}$ is an abelian category, $\mathbf{C}$ is additive and hence finite direct sums coincide with finite direct products in $\mathbf{C}$. For any object $X \in \mathbf{C}$ and any integer $r \in \mathbb{Z}$, $r > 0$, we let X^r denote the direct sum (or direct product) of r-copies of X in $\mathbf{C}$.

By an algebra in $\mathbf{C}$, we will always mean a commutative monoid object with unit in $\mathbf{C}$. The category of algebras in $\mathbf{C}$ will be denoted by Alg. More precisely, an algebra in $\mathbf{C}$ is an object A in $\mathbf{C}$ with a multiplication map $m_A : A \otimes A \to A$ and a unit map $u_A : 1 \to A$. satisfying the compatibility conditions for making A a commutative monoid with unit (see [Mac Lane 1998], for instance).

For any algebra A, we let A-*Mod* denote the category of A-modules in $\mathbf{C}$. Then, each $(A\text{-}Mod, \otimes_A, A)$ is also a closed symmetric monoidal category. Given any A-modules M and N, we will denote by $\mathrm{Hom}_A(M, N)$ the set of morphisms from M to N in A-*Mod* and the internal Hom object by $\underline{\mathrm{Hom}}_A(M, N)$. It is clear that $\mathrm{Hom}_A(M, N)$ is an abelian group. Further, the category of unitary commutative monoids in A-*Mod* will be denoted by A-*Alg*. For any two A-algebras B and B', we will denote by $\mathrm{Hom}_{A\text{-}Alg}(B, B')$ the set of A-algebra morphisms from B to B'. If $f : A \to B$ is a morphism of algebras, for any A-module M and B-module N, we have natural isomorphisms

$$T : \mathrm{Hom}_A(M, N) \cong \mathrm{Hom}_B(M \otimes_A B, N) \tag{2-4}$$

described as follows: given $g \in \mathrm{Hom}_A(M, N)$, we define

$$T(g) \in \mathrm{Hom}_B(M \otimes_A B, N)$$

as the composition

$$T(g) : M \otimes_A B \xrightarrow{g\otimes_A 1} N \otimes_A B \longrightarrow N, \tag{2-5}$$

where the morphism $N \otimes_A B \to N$ in (2-5) follows from the B-module structure of N. Conversely, given $h \in \mathrm{Hom}_B(M \otimes_A B, N)$, it is clear that we have $h = T(h')$, where $h' \in \mathrm{Hom}_A(M, N)$ is defined as the composition

$$h' : M \cong M \otimes_A A \xrightarrow{1\otimes_A f} M \otimes_A B \xrightarrow{h} N. \tag{2-6}$$

Furthermore, for any object X in A-*Mod*, we note that

$$\begin{aligned} \mathrm{Hom}_A(X, \underline{\mathrm{Hom}}_A(M, N)) &\cong \mathrm{Hom}_A(X \otimes_A M, N) \\ &\cong \mathrm{Hom}_B(X \otimes_A M \otimes_A B, N) \\ &\cong \mathrm{Hom}_B((X \otimes_A B) \otimes_B (M \otimes_A B), N) \\ &\cong \mathrm{Hom}_B(X \otimes_A B, \underline{\mathrm{Hom}}_B(M \otimes_A B, N)) \\ &\cong \mathrm{Hom}_A(X, \underline{\mathrm{Hom}}_B(M \otimes_A B, N)). \end{aligned} \tag{2-7}$$

Using (2-7), it follows from Yoneda's lemma that we have natural isomorphisms in A-*Mod*:

$$\underline{\mathrm{Hom}}_A(M, N) \cong \underline{\mathrm{Hom}}_B(M \otimes_A B, N). \tag{2-8}$$

3. Affine group schemes

Let $(\mathbf{C}, \otimes, 1)$ be an abelian, closed, symmetric monoidal category as described in Section 2 and let A be an algebra in $\mathbf{C}$. Then, it is well known (see, for instance, [May 2001]) that the collection of endomorphisms $\mathrm{Hom}_A(A, A)$ is an ordinary commutative ring with identity. We start with the following result.

Proposition 3.1. *Let A be an algebra in* **C**. *Then, there is a natural isomorphism*

$$\mathrm{Hom}(1, A) \xrightarrow{\sim} \mathrm{Hom}_A(A, A). \tag{3-1}$$

Proof. Define a map $S : \mathrm{Hom}(1, A) \to \mathrm{Hom}_A(A, A)$ thus: given $f \in \mathrm{Hom}(1, A)$, let $S(f) \in \mathrm{Hom}_A(A, A)$ be the composition

$$A \xrightarrow{\sim} A \otimes 1 \xrightarrow{1 \otimes f} A \otimes A \xrightarrow{m_A} A, \tag{3-2}$$

where $m_A : A \otimes A \to A$ in (3-2) is the multiplication map on the algebra A. Conversely, we define a map $T : \mathrm{Hom}_A(A, A) \to \mathrm{Hom}(1, A)$ as follows: given $g \in \mathrm{Hom}_A(A, A)$, we let $T(g) \in \mathrm{Hom}(1, A)$ denote the composition

$$1 \xrightarrow{u_A} A \xrightarrow{g} A, \tag{3-3}$$

where the map $u_A : 1 \to A$ in (3-3) is the "unit map" for the algebra A. It is easy to check that the associations S and T are inverse to each other and hence we have an isomorphism $\mathrm{Hom}(1, A) \xrightarrow{\sim} \mathrm{Hom}_A(A, A)$. □

Following [Toën and Vaquié 2009], we define $\mathrm{Aff}_{\mathbf{C}} := \mathrm{Alg}^{op}$ to be the category of affine schemes over **C**. For any algebra A in **C**, we let $\mathrm{Spec}(A)$ denote the corresponding object of $\mathrm{Aff}_{\mathbf{C}}$. Further, we denote by $\mathfrak{spec}(A)$ the (contravariant) functor on $\mathrm{Aff}_{\mathbf{C}}$ represented by $\mathrm{Spec}(A)$.

Definition 3.2. Let $(\mathbf{C}, \otimes, 1)$ be as above and let *Set* denote the category of sets. An affine group scheme over **C** is a representable functor

$$G = \mathfrak{spec}(A) : \mathrm{Aff}_{\mathbf{C}} \to Set, \tag{3-4}$$

equipped with a composition map $m_G : G \times G \to G$, an inverse map $i_G : G \to G$ and a unit map $e_G : \mathfrak{spec}(1) \to G$ of functors satisfying the group axioms (see [Waterhouse 1979, § 1.4], for instance).

From Yoneda's lemma it follows that if $G = \mathfrak{spec}(A)$ is an affine group scheme in the sense of Definition 3.2, then A is an algebra in **C** equipped with a comultiplication $\Delta_A : A \to A \otimes A$, an antipode $i_A : A \to A$ and a counit $\epsilon_A : A \to 1$ that gives A the structure of a Hopf algebra in **C**. Further, if Grp denotes the category of groups, we can also express G as a functor from algebras in **C** to groups:

$$G : \mathrm{Alg} \to \mathrm{Grp}, \quad G(B) := \mathrm{Hom}_{\mathrm{Aff}_{\mathbf{C}}}(\mathrm{Spec}(B), \mathrm{Spec}(A)) = \mathrm{Hom}_{\mathrm{Alg}}(A, B). \tag{3-5}$$

Further, since the comultiplication $\Delta_A : A \to A \otimes A$ in Alg corresponds to the composition $m_G : G \times G \to G$, it follows that A is cocommutative if and only if, for all algebras B in **C**, the group $G(B)$ is abelian. In this case, we will say that $G = \mathfrak{spec}(A)$ is an affine commutative group scheme over **C**.

Definition 3.3. Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme over $\mathbf{C}$. Then, we say that G is free of finite rank $r \in \mathbb{Z}$, $r > 0$ if $A \cong 1^r$ as objects of $\mathbf{C}$, where 1^r denotes the direct sum of r-copies of the unit object 1 of $\mathbf{C}$.

Further, suppose that B is an algebra in $\mathbf{C}$ and let B' be a B-algebra. Then, B' is said to be a locally free B-algebra of rank r if $B' \cong B^r$ as B-modules. In case $B = 1$, we will simply say that B' is a locally free algebra of rank r.

From now onwards we will always let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme over $\mathbf{C}$ that is free of finite rank r. We also define $A' := \underline{\mathrm{Hom}}(A, 1)$. Then, it is clear that for any object X in $\mathbf{C}$, we have natural isomorphisms

$$\underline{\mathrm{Hom}}(A, X) \cong \underline{\mathrm{Hom}}(1^r, X) \cong \bigoplus^r \underline{\mathrm{Hom}}(1, X) \cong \underline{\mathrm{Hom}}(A, 1) \otimes X \cong A' \otimes X. \tag{3-6}$$

Proposition 3.4. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme over $\mathbf{C}$ that is free of finite rank r. Then, $A' := \underline{\mathrm{Hom}}(A, 1)$ is a commutative and cocommutative Hopf algebra in $\mathbf{C}$ and is also a locally free algebra of rank r.*

Proof. Since $G = \mathfrak{spec}(A)$ is an affine commutative group scheme, we know that A is a commutative and cocommutative Hopf algebra in $\mathbf{C}$. From (2-3) and (3-6), it follows that

$$A' \otimes A' \cong \underline{\mathrm{Hom}}(A, \underline{\mathrm{Hom}}(A, 1)) \cong \underline{\mathrm{Hom}}(A \otimes A, 1). \tag{3-7}$$

It is clear that the multiplication $m_A : A \otimes A \to A$ induces a map

$$A' = \underline{\mathrm{Hom}}(A, 1) \to \underline{\mathrm{Hom}}(A \otimes A, 1),$$

while the comultiplication $\Delta_A : A \to A \otimes A$ induces

$$\underline{\mathrm{Hom}}(A \otimes A, 1) \to \underline{\mathrm{Hom}}(A, 1) = A'.$$

Combining this with (3-7), we obtain a natural multiplication $m_{A'} : A' \otimes A' \to A'$ and a natural comultiplication $\Delta_{A'} : A' \to A' \otimes A'$ on A'. The unit $u_{A'} : 1 \to A'$, the counit $\epsilon_{A'} : A' \to 1$ and the antipode $i_{A'} : A' \to A'$ on A' are obtained by dualizing $\epsilon_A : A \to 1$, $u_A : 1 \to A$ and $i_A : A \to A$ respectively. It is clear that these maps make A' into a commutative and cocommutative Hopf algebra.

Finally, since $A \cong 1^r$, it follows that $A' \cong \underline{\mathrm{Hom}}(A, 1) \cong 1^r$ and hence A' is also a locally free algebra of rank r. □

Proposition 3.5. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme over $\mathbf{C}$ that is free of finite rank r. Let $A' = \underline{\mathrm{Hom}}(A, 1)$. Then:*

(a) *There are natural isomorphisms*

$$\mathrm{Hom}(A, 1) \cong \mathrm{Hom}(1, A') \cong \mathrm{Hom}_{A'}(A', A'). \tag{3-8}$$

Further, each of the objects in (3-8) *carries a comultiplication structure that is compatible with the isomorphisms in* (3-8).

(b) *There are natural isomorphisms*

$$\mathrm{Hom}(A, A) \cong \mathrm{Hom}(1, A' \otimes A) \cong \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A). \tag{3-9}$$

Further, each of the objects in (3-9) *carries a comultiplication structure that is compatible with the isomorphisms in* (3-9).

Proof. (a) Since $A' = \underline{\mathrm{Hom}}(A, 1)$, it is clear that $\mathrm{Hom}(A, 1) \cong \mathrm{Hom}(1, A')$. Since Proposition 3.4 shows that A' is also an algebra, the isomorphism $\mathrm{Hom}(1, A') \cong \mathrm{Hom}_{A'}(A', A')$ follows from Proposition 3.1.

We now describe the comultiplication structure on $\mathrm{Hom}_{A'}(A', A')$. Given f in $\mathrm{Hom}_{A'}(A', A')$, we can define a morphism

$$\delta_1(f) \in \mathrm{Hom}_{A' \otimes A'}(A' \otimes A', A' \otimes A')$$

as follows:

$$\begin{aligned} \delta_1(f) : A' \otimes A' \xrightarrow{\sim} 1 \otimes A' \otimes A' \longrightarrow A' \otimes A' \otimes A' \xrightarrow{f \otimes 1 \otimes 1} A' \otimes A' \otimes A' \\ \xrightarrow{\Delta_{A'} \otimes 1 \otimes 1} A' \otimes A' \otimes A' \otimes A' \xrightarrow{m'_{13} \otimes m'_{24}} A' \otimes A', \end{aligned} \tag{3-10}$$

where $m'_{ij} : A' \otimes A' \to A'$ in (3-10) denotes the multiplication $m_{A'} : A' \otimes A' \to A'$ on A' applied to the i-th and j-th copy of A' appearing in the term $A' \otimes A' \otimes A' \otimes A'$ in (3-10). Since A' is a locally free algebra of rank r, we have natural isomorphisms

$$\mathrm{Hom}_{A'}(A', A') \otimes \mathrm{Hom}_{A'}(A', A') \cong \mathrm{Hom}_{A' \otimes A'}(A' \otimes A', A' \otimes A'). \tag{3-11}$$

Using (3-10) and (3-11), we have a comultiplication

$$\delta_1 : \mathrm{Hom}_{A'}(A', A') \to \mathrm{Hom}_{A'}(A', A') \otimes \mathrm{Hom}_{A'}(A', A'). \tag{3-12}$$

Considering the comultiplication $\Delta_{A'} : A' \to A' \otimes A'$ on A', we have an induced map

$$\delta_2 : \mathrm{Hom}(1, A') \xrightarrow{\mathrm{Hom}(1, \Delta_{A'})} \mathrm{Hom}(1, A' \otimes A') \cong \mathrm{Hom}(1, A') \otimes \mathrm{Hom}(1, A'), \tag{3-13}$$

where the last isomorphism follows from the fact that A' is a locally free algebra. From (3-10), (3-13), and the construction of the isomorphism

$$\mathrm{Hom}(1, A') \cong \mathrm{Hom}_{A'}(A', A')$$

in Proposition 3.1 applied to A', it follows that the comultiplications δ_1 and δ_2 are compatible with the isomorphism $\mathrm{Hom}(1, A') \cong \mathrm{Hom}_{A'}(A', A')$. Finally, since the comultiplication

$$\delta_3 : \mathrm{Hom}(A, 1) \to \mathrm{Hom}(A \otimes A, 1) \cong \mathrm{Hom}(A, 1) \otimes \mathrm{Hom}(A, 1) \tag{3-14}$$

is induced by the multiplication $m_A : A \otimes A \to A$ on A and m_A induces the comultiplication $\Delta_{A'} : A' \to A' \otimes A'$ on A', the maps δ_2 and δ_3 are compatible with the isomorphism $\mathrm{Hom}(A, 1) \cong \mathrm{Hom}(1, A')$.

(b) From (2-4), it follows that

$$\begin{aligned} \mathrm{Hom}(1, A' \otimes A) \cong \mathrm{Hom}_A(A, A' \otimes A) &\cong \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A), \\ \mathrm{Hom}(A, A) &\cong \mathrm{Hom}_A(A \otimes A, A). \end{aligned} \tag{3-15}$$

We also note that, using (2-8) and (3-6), we have

$$\underline{\mathrm{Hom}}_A(A \otimes A, A) \cong \underline{\mathrm{Hom}}(A, A) \cong A' \otimes A \tag{3-16}$$

From (3-16), it follows that the dual of $A \otimes A$ in the category A-*Mod* is $A' \otimes A$. Further, the comultiplication $\Delta_A : A \to A \otimes A$ induces a comultiplication

$$\Delta_A^A : A \otimes A \to (A \otimes A) \otimes_A (A \otimes A)$$

on the A-algebra $A \otimes A$ as follows:

$$\Delta_A^A := \Delta_A \otimes 1_A : A \otimes A \to A \otimes A \otimes A \cong (A \otimes A) \otimes_A (A \otimes A) \tag{3-17}$$

making $A \otimes A$ into a Hopf algebra in A-*Mod*. Applying the result of part (a) to the object $A \otimes A$ in A-*Mod*, we have compatible comultiplications on each of the following isomorphic objects

$$\mathrm{Hom}_A(A \otimes A, A) \cong \mathrm{Hom}_A(A, A' \otimes A) \cong \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A). \tag{3-18}$$

Using the isomorphisms in (3-15), we have compatible induced comultiplications on each of the following isomorphic objects:

$$\begin{aligned} \delta_1^A &: \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A) \\ &\qquad \to \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A) \otimes \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A), \\ \delta_2^A &: \mathrm{Hom}(1, A' \otimes A) \to \mathrm{Hom}(1, A' \otimes A) \otimes \mathrm{Hom}(1, A' \otimes A), \\ \delta_3^A &: \mathrm{Hom}(A, A) \to \mathrm{Hom}(A, A) \otimes \mathrm{Hom}(A, A). \end{aligned}$$

□

4. Norm map and grouplike elements

From now onwards, we will assume that the closed abelian symmetric monoidal category $(\mathbf{C}, \otimes, 1)$ is $\mathbb{C}$-linear. As before, we let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme that is free of finite rank r. Let $A' = \underline{\mathrm{Hom}}(A, 1)$. Given any algebra B in $\mathbf{C}$, we will construct a map

$$N_B : \mathrm{Hom}_{B \otimes A}(B \otimes A, B \otimes A) \to \mathrm{Hom}_B(B, B), \tag{4-1}$$

which corresponds to the norm map in the context of ordinary $\mathbb{Z}$-algebras. We will refer to the B-algebra $B \otimes A$ as B_A.

Let M be a B-module. Since the category B-*Mod* is $\mathbb{C}$-linear, the notion of exterior product extends to it. For any integer $n \geq 1$, we can consider the tensor product $M^{\otimes_B n} := M \otimes_B M \otimes_B \otimes_B \cdots \otimes_B M$ (n-times). Then, the symmetric group S_n acts on $M^{\otimes_B n}$ by permutations, i.e., for each $\sigma \in S_n$, we have an induced map $\sigma : M^{\otimes_B n} \to M^{\otimes_B n}$ of B-modules. We then consider the morphism

$$q_M^n : M^{\otimes_B n} \to M^{\otimes_B n} \qquad q_M^n := 1 - \frac{1}{n!}\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\sigma. \tag{4-2}$$

It is clear that the morphism $q_M^n \in \mathrm{Hom}_B(M^{\otimes_B n}, M^{\otimes_B n})$ is an idempotent. Since $\mathbf{C}$ is an abelian category, we can form the cokernel of q_M^n, which we denote by $\bigwedge_B^n M$. Further, since q_M^n is an idempotent, the cokernel $\bigwedge_B^n M$ is a direct summand of $M^{\otimes_B n}$. It follows that for any $n \geq 1$, q_M^n induces a morphism

$$\mathrm{Hom}(\textstyle\bigwedge_B^n)(M) : \mathrm{Hom}_B(M^{\otimes_B n}, M^{\otimes_B n}) \to \mathrm{Hom}_B(\bigwedge_B^n M, \bigwedge_B^n M). \tag{4-3}$$

In particular, therefore, taking $M = B_A$ and $n = r$, we have a map

$$\mathrm{Hom}(\textstyle\bigwedge_B^r)(B_A) : \mathrm{Hom}_B(B_A^{\otimes_B r}, B_A^{\otimes_B r}) \to \mathrm{Hom}_B(\bigwedge_B^r B_A, \bigwedge_B^r B_A). \tag{4-4}$$

Also, for any objects X, Y in B-*Mod*, the exterior product satisfies

$$\textstyle\bigwedge_B^n(X \oplus Y) \cong \displaystyle\bigoplus_{k+l=n} \textstyle\bigwedge_B^k X \otimes_B \bigwedge_B^l Y. \tag{4-5}$$

In the situation above, since A is a locally free algebra of rank r, i.e., $A \cong 1^r$, it follows that $B_A = B \otimes A \cong B^r$. Hence, B_A is a locally free B-algebra of rank r.

Lemma 4.1. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Let B be an algebra in* $\mathbf{C}$. *Then, there exists a natural isomorphism $\bigwedge_B^r B_A \cong B$ of B-modules.*

Proof. For any $k \geq 2$, we consider the morphism

$$q_B^k : B^{\otimes_B k} \to B^{\otimes_B k}. \tag{4-6}$$

Since B is a commutative monoid, any morphism $\sigma : B^{\otimes_B k} \cong B \to B^{\otimes_B k} \cong B$ induced by some $\sigma \in S_k$ corresponds to the identity map $1_B : B \to B$. Since $\sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) = 0$, it follows that q_B^k is the identity. Hence

$$\textstyle\bigwedge_B^k B := \mathrm{Coker}(q_B^k) = 0. \tag{4-7}$$

It follows from (4-5) and (4-7) that

$$\begin{aligned} \textstyle\bigwedge_B^r B_A &\cong \textstyle\bigwedge_B^r B^r \\ &\cong \bigoplus_{k_1+k_2+\cdots+k_r=r} \textstyle\bigwedge_B^{k_1} B \otimes_B \cdots \otimes_B \bigwedge_B^{k_r} B \cong B \otimes_B \cdots \otimes_B B \cong B. \end{aligned} \tag{4-8}$$

□

Proposition 4.2. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Let B be an algebra in* **C**. *Then, there exists a norm map*

$$N_B : \mathrm{Hom}_{B\otimes A}(B \otimes A, B \otimes A) \to \mathrm{Hom}_B(B, B) \tag{4-9}$$

that is compatible with composition on $\mathrm{Hom}_{B\otimes A}(B\otimes A, B\otimes A)$ *and* $\mathrm{Hom}_B(B, B)$.

Proof. We set $B_A = B \otimes A$ as above. First, we note that we have a forgetful map

$$\mathrm{Hom}_{B\otimes A}(B \otimes A, B \otimes A) \to \mathrm{Hom}_B(B \otimes A, B \otimes A) = \mathrm{Hom}_B(B_A, B_A). \tag{4-10}$$

Following this, we consider the morphism

$$\mathrm{Hom}_B(B_A, B_A) \to \mathrm{Hom}_B(B_A^{\otimes r}, B_A^{\otimes r}), \qquad f \mapsto f^{\otimes r}. \tag{4-11}$$

From (4-4) and Lemma 4.1, we have

$$\begin{aligned}\mathrm{Hom}({\textstyle\bigwedge}^r_B)(B_A) : \mathrm{Hom}_B(B_A^{\otimes r}, B_A^{\otimes r}) \to \mathrm{Hom}_B({\textstyle\bigwedge}^r_B B_A, {\textstyle\bigwedge}^r_B B_A)\\ \cong \mathrm{Hom}_B(B, B).\end{aligned} \tag{4-12}$$

Composing the morphisms in (4-10), (4-11) and (4-12), we have the map

$$N_B : \mathrm{Hom}_{B\otimes A}(B \otimes A, B \otimes A) \to \mathrm{Hom}_B(B, B). \tag{4-13}$$

Finally, it is clear from the construction that N_B is compatible with composition on $\mathrm{Hom}_{B\otimes A}(B \otimes A, B \otimes A)$ and $\mathrm{Hom}_B(B, B)$. □

By composing the maps in (4-11) and (4-12) in the proof of Proposition 4.2, it follows that we have a norm map $\mathrm{Hom}_B(B \otimes A, B \otimes A) \to \mathrm{Hom}_B(B, B)$ for any algebra B in **C** which we will continue to denote by N_B.

Let $f : B \to C$ be a morphism of algebras in **C**. Then, it follows from base change that f induces maps

$$\begin{aligned}\mathrm{Hom}(f) : \mathrm{Hom}_B(B, B) \to \mathrm{Hom}_C(C, C),\\ \mathrm{Hom}_A(f \otimes 1) : \mathrm{Hom}_{B\otimes A}(B \otimes A, B \otimes A) \to \mathrm{Hom}_{C\otimes A}(C \otimes A, C \otimes A).\end{aligned} \tag{4-14}$$

Further, since the morphisms (4-10), (4-11) and (4-12) are all natural with respect to base change, the following diagram is commutative:

$$\begin{array}{ccc}\mathrm{Hom}_{B\otimes A}(B \otimes A, B \otimes A) & \xrightarrow{\mathrm{Hom}_A(f\otimes 1)} & \mathrm{Hom}_{C\otimes A}(C \otimes A, C \otimes A)\\ {\scriptstyle N_B}\downarrow & & {\scriptstyle N_C}\downarrow\\ \mathrm{Hom}_B(B, B) & \xrightarrow{\mathrm{Hom}(f)} & \mathrm{Hom}_C(C, C)\end{array} \tag{4-15}$$

Lemma 4.3. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Then, for any algebra B in* **C**, *we have natural isomorphisms*

$$\mathrm{Hom}_{\mathrm{Alg}}(A, B) \cong \mathrm{Hom}_{B\text{-}Alg}(A \otimes B, B). \tag{4-16}$$

Proof. We know that we have an isomorphism

$$T : \mathrm{Hom}(A, B) \xrightarrow{\cong} \mathrm{Hom}_B(A \otimes B, B). \tag{4-17}$$

Suppose that $f : A \to B$ is a morphism of algebras. Then, $f \otimes 1 : A \otimes B \to B \otimes B$ is a morphism of B-algebras. Further, the multiplication $m_B : B \otimes B \to B$ is also a map of B-algebras. It follows that

$$T(f) = m_B \circ (f \otimes 1) : A \otimes B \xrightarrow{f \otimes 1} B \otimes B \xrightarrow{m_B} B$$

is a morphism of B-algebras. Hence, T restricts to a morphism

$$T^{\mathrm{alg}} : \mathrm{Hom}_{\mathrm{Alg}}(A, B) \to \mathrm{Hom}_{B\text{-}Alg}(A \otimes B, B). \tag{4-18}$$

Next, we choose some $g \in \mathrm{Hom}_{B\text{-}Alg}(A \otimes B, B) \subseteq \mathrm{Hom}_B(A \otimes B, B)$. Then, it follows that $g = T(f)$, where f is given by the composition

$$f : A \cong A \otimes 1 \xrightarrow{1 \otimes e_B} A \otimes B \xrightarrow{g} B. \tag{4-19}$$

Here $e_B : 1 \to B$ is the unit map of the algebra B. Since both maps in (4-19) are morphisms of algebras, $f \in \mathrm{Hom}(A, B)$ is actually a morphism of algebras. It follows that T^{alg} is a surjection. Further, since T^{alg} is obtained by restricting the isomorphism T, T^{alg} must be injective. Hence, we have an isomorphism

$$T^{\mathrm{alg}} : \mathrm{Hom}_{\mathrm{Alg}}(A, B) \xrightarrow{\cong} \mathrm{Hom}_{B\text{-}Alg}(A \otimes B, B). \qquad \square$$

Proposition 4.4. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Let $A' = \underline{\mathrm{Hom}}(A, 1)$. Then*:

(a) *Let $g \in \mathrm{Hom}(A, 1)$ be a morphism that corresponds to $h \in \mathrm{Hom}_{A'}(A', A')$ under the isomorphism $\mathrm{Hom}_{A'}(A', A') \cong \mathrm{Hom}(A, 1)$ in* (3-8). *Then, $g : A \to 1$ is a morphism of algebras if and only if $\delta_1(h) = h \otimes h$ in the notation of Proposition 3.5.*

(b) *Let $g \in \mathrm{Hom}(A, A)$ be a morphism that corresponds to*

$$h \in \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A) \tag{4-20}$$

under the isomorphism $\mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A) \cong \mathrm{Hom}(A, A)$ in (3-9). *Then, $g : A \to A$ is a morphism of algebras if and only if $\delta_1^A(h) = h \otimes h$ in the notation of Proposition 3.5.*

Proof. We maintain the notation of the proof of Proposition 3.5.

(a) Using Proposition 3.5(a), it suffices to check that $g \in \mathrm{Hom}(A, 1)$ is a morphism of algebras if and only if $\delta_3(g) = g \otimes g$ where δ_3 denotes the comultiplication on $\mathrm{Hom}(A, 1)$.

By definition of δ_3 in (3-14), we know that $\delta_3(g)$ is equal to the composition

$$A \otimes A \xrightarrow{m_A} A \xrightarrow{g} 1 \xrightarrow{\cong} 1 \otimes 1. \tag{4-21}$$

It is immediate from (4-21) that $\delta_3(g) = g \otimes g$ if and only if $g : A \to 1$ is a morphism of algebras.

(b) Using Proposition 3.5(b), we know that the comultiplication δ_3^A on $\mathrm{Hom}(A, A)$ corresponds to the comultiplication δ_1^A on $\mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A)$ via the isomorphism

$$\mathrm{Hom}(A, A) \cong \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A)$$

in (3-9). It therefore suffices to check that $g : A \to A$ is a morphism of algebras if and only if $\delta_3^A(g) = g \otimes g$.

From (2-4), we know that

$$\mathrm{Hom}(A, A) \cong \mathrm{Hom}_A(A \otimes A, A). \tag{4-22}$$

Further, from Lemma 4.3, we know that the isomorphism in (4-22) restricts to an isomorphism

$$\mathrm{Hom}_{\mathrm{Alg}}(A, A) \cong \mathrm{Hom}_{A\text{-}Alg}(A \otimes A, A) \tag{4-23}$$

From the proof of Proposition 3.5, we also know that the comultiplication δ_3^A on $\mathrm{Hom}(A, A)$ is induced by the comultiplication on $\mathrm{Hom}_A(A \otimes A, A)$, also denoted δ_3^A. Hence if $g \in \mathrm{Hom}(A, A)$ corresponds to $g' \in \mathrm{Hom}_A(A \otimes A, A)$, $\delta_3^A(g) = g \otimes g$ if and only if $\delta_3^A(g') = g' \otimes g'$.

Applying the result of part (a) to the A-algebra $A \otimes A$ in A-*Mod*, it follows that $g' : A \otimes A \to A$ is a morphism of A-algebras, i.e., $g' \in \mathrm{Hom}_{A\text{-}Alg}(A \otimes A, A)$ if and only if $\delta_3^A(g') = g' \otimes g'$. Since $\mathrm{Hom}_{\mathrm{Alg}}(A, A) \cong \mathrm{Hom}_{A\text{-}Alg}(A \otimes A, A)$, the result follows. □

5. Analogue of Deligne's lemma

We will now complete the proof of Theorem 1.2 stated in the introduction.

Proposition 5.1. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Let $A' = \underline{\mathrm{Hom}}(A, 1)$. Then, the morphism*

$$N_{A'} : \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A) \to \mathrm{Hom}_{A'}(A', A')$$

restricts to a homomorphism of groups from $G(A)$ to $G(1)$.

Proof. Let $f \in G(A) \subseteq \mathrm{Hom}(A, A) \cong \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A)$, i.e., f is a morphism of algebras. From Proposition 4.4, we know that $\delta_3^A(f) = f \otimes f$.

We consider the morphism $\Delta_{A'} : A' \to A' \otimes A'$ of algebras in $\mathbf{C}$. It follows from (4-15) that we have a commutative diagram

$$\begin{array}{ccc} \mathrm{Hom}_{A'\otimes A}(A'\otimes A,\, A'\otimes A) & \xrightarrow{\mathrm{Hom}_A(\Delta_{A'}\otimes 1)} & \mathrm{Hom}_{A'\otimes A'\otimes A}(A'\otimes A'\otimes A,\, A'\otimes A'\otimes A) \\ {\scriptstyle N_{A'}}\downarrow & & \downarrow{\scriptstyle N_{A'\otimes A'}} \\ \mathrm{Hom}_{A'}(A',\, A') & \xrightarrow{\mathrm{Hom}(\Delta_{A'})} & \mathrm{Hom}_{A'\otimes A'}(A'\otimes A',\, A'\otimes A') \end{array} \tag{5-1}$$

It follows that

$$\mathrm{Hom}(\Delta_{A'})(N_{A'}(f)) = N_{A'\otimes A'}(\mathrm{Hom}_A(\Delta_{A'} \otimes 1)(f)). \tag{5-2}$$

We note that

$$\Delta_{A'} \otimes 1 : A' \otimes A \to A' \otimes A' \otimes A \cong (A' \otimes A) \otimes_A (A' \otimes A) \tag{5-3}$$

is the coproduct $\Delta' : (A' \otimes A) \to (A' \otimes A) \otimes_A (A' \otimes A)$ on the A-algebra $A' \otimes A$ and hence determines the comultiplication on $\mathrm{Hom}_{A'\otimes A}(A' \otimes A,\, A' \otimes A)$. Since $\delta_3^A(f) = f \otimes f$, it follows from Proposition 4.4 that

$$\begin{aligned} \mathrm{Hom}_A(\Delta_{A'} \otimes 1)(f) = f \otimes_A f : (A' \otimes A) \otimes_A (A' \otimes A) = A' \otimes A' \otimes A \\ \longrightarrow A' \otimes A' \otimes A = (A' \otimes A) \otimes_A (A' \otimes A). \end{aligned} \tag{5-4}$$

The morphism $f \otimes_A f$ in (5-4) can be described by the composition

$$(A'\otimes A)\otimes_A (A'\otimes A) \xrightarrow{f\otimes_A 1} (A'\otimes A)\otimes_A (A'\otimes A) \xrightarrow{1\otimes_A f} (A'\otimes A)\otimes_A (A'\otimes A). \tag{5-5}$$

Consider the morphism $e'_{A'} : A' \to A' \otimes A'$ of algebras obtained by base changing the unit morphism $e_{A'} : 1 \to A'$ with A'. Then, we have a commutative diagram

$$\begin{array}{ccc} \mathrm{Hom}_{A'\otimes A}(A'\otimes A,\, A'\otimes A) & \xrightarrow{\mathrm{Hom}_A(e'_{A'}\otimes 1)} & \mathrm{Hom}_{A'\otimes A'\otimes A}(A'\otimes A'\otimes A,\, A'\otimes A'\otimes A) \\ {\scriptstyle N_{A'}}\downarrow & & \downarrow{\scriptstyle N_{A'\otimes A'}} \\ \mathrm{Hom}_{A'}(A',\, A') & \xrightarrow{\mathrm{Hom}(e'_{A'})} & \mathrm{Hom}_{A'\otimes A'}(A'\otimes A',\, A'\otimes A') \end{array} \tag{5-6}$$

From (5-6), it follows that

$$1 \otimes N_{A'}(f) = N_{A'\otimes A'}(1 \otimes f) \tag{5-7}$$

Now, the comultiplication on $\mathrm{Hom}_{A'}(A', A')$ is induced by the morphism $\mathrm{Hom}(\Delta_{A'})$ in (5-1). We also note that $1 \otimes_A f : (A' \otimes A) \otimes_A (A' \otimes A) \to (A' \otimes A) \otimes_A (A' \otimes A)$ is identical to $1 \otimes f : A' \otimes A' \otimes A \to A' \otimes A' \otimes A$. Hence, we have

$$\begin{aligned} \delta_1(N_{A'}(f)) &= \mathrm{Hom}(\Delta_{A'})(N_{A'}(f)) = N_{A'\otimes A'}(\mathrm{Hom}_A(\Delta_{A'} \otimes 1)(f)) \\ &= N_{A'\otimes A'}(f \otimes_A f) = N_{A'\otimes A'}(f \otimes_A 1)N_{A'\otimes A'}(1 \otimes_A f) \\ &= N_{A'\otimes A'}(f \otimes 1)N_{A'\otimes A'}(1 \otimes f) = (N_{A'}(f) \otimes 1)(1 \otimes N_{A'}(f)) \\ &= N_{A'}(f) \otimes N_{A'}(f), \end{aligned} \tag{5-8}$$

and it now follows from Proposition 4.4(a) that $N_{A'}(f) : A' \to A'$ corresponds to a morphism of algebras from A to 1 under the isomorphism

$$\mathrm{Hom}(A, 1) \cong \mathrm{Hom}_{A'}(A', A').$$

Hence, given $f \in G(A)$, it follows that $N_{A'}(f) \in G(1)$. It is also clear that $N_{A'} : G(A) \to G(1)$ is a homomorphism. □

The result of Proposition 5.1 can be restated as follows: the morphism $N_{A'} : \mathrm{Hom}_{A'\otimes A}(A' \otimes A, A' \otimes A) \to \mathrm{Hom}_{A'}(A', A')$ restricts to a homomorphism $N : G(A) \to G(1)$ that fits into a commutative diagram

$$\begin{array}{ccc} G(A) & \longrightarrow & \mathrm{Hom}_{A'\otimes A}(A' \otimes A, A' \otimes A) \\ {\scriptstyle N}\big\downarrow & & \big\downarrow{\scriptstyle N_{A'}} \\ G(1) & \longrightarrow & \mathrm{Hom}_{A'}(A', A') \end{array} \tag{5-9}$$

We choose any $u \in G(1)$, i.e. a morphism $u : A \to 1$ of algebras. For any algebra B, the unit map $e_B : 1 \to B$ induces a morphism $e_{B*} : G(1) \to G(B)$ and hence we can consider the translation map

$$t_{u,B} : G(B) \to G(B) \tag{5-10}$$

obtained by multiplication with the element $e_{B*}(u)$. By Yoneda lemma, the translations $t_{u,B}$ determine an automorphism $e_*(u) : A \to A$ of algebras. We denote the A'-linear automorphism $1 \otimes e_*(u) : A' \otimes A \to A' \otimes A$ by τ. Since

$$u \in \mathrm{Hom}_{\mathrm{Alg}}(A, 1) \subseteq \mathrm{Hom}(A, 1) \cong \mathrm{Hom}_{A'}(A', A'),$$

we will often write u as a morphism $u : A' \to A'$ of A'-modules.

Lemma 5.2. *Let $u : A \to 1$ be a morphism of algebras and let*

$$\tau := 1 \otimes e_*(u) : A' \otimes A \to A' \otimes A$$

be as described above. Then, $N_{A'}(\tau) = u^r$ where u^r denotes the r-th power of u as an element of the group $G(1)$.

Proof. We know that $\tau : A' \otimes A \to A' \otimes A$ is induced by $u \in \mathrm{Hom}_{\mathrm{Alg}}(A, 1)$ and that $A' \otimes A \cong A'^{\oplus r}$ in A'-*Mod*. From the proof of Lemma 4.1, we know that $N_{A'}(\tau) \in \mathrm{Hom}_{A'}(A', A')$ corresponds to the morphism

$$N_{A'}(\tau) : \textstyle\bigwedge^r_{A'} A'^{\oplus r} \cong A' \to \bigwedge^r_{A'} A'^{\oplus r} \cong A'. \tag{5-11}$$

On each individual summand in $A'^{\oplus r}$, the action of the morphism $\tau : A'^{\oplus r} \to A'^{\oplus r}$ is induced by $u : A' \to A'$. Hence, it follows from (4-8) in the proof of Lemma 4.1

that the induced action of τ on the exterior product $\bigwedge^r_{A'} A'^{\oplus r}$ is given by

$$\begin{array}{ccc} \bigwedge_{A'} A'^{\oplus r} & \xrightarrow{N_{A'}(\tau)=\bigwedge^r_{A'}\tau} & \bigwedge_{A'} A'^{\oplus r} \\ \cong\downarrow & & \cong\downarrow \\ A'\otimes_{A'}\cdots\otimes_{A'} A' & \xrightarrow{u\otimes_{A'}u\otimes_{A'}\otimes_{A'}\cdots\otimes_{A'}u} & A'\otimes_{A'}\cdots\otimes_{A'} A' \\ \cong\downarrow & & \cong\downarrow \\ A' & \xrightarrow{u^r} & A' \end{array} \tag{5-12}$$

The morphism $u^r \in \mathrm{Hom}_{A'}(A', A')$ in (5-12) corresponds to the r-th power of

$$u \in \mathrm{Hom}_{\mathrm{Alg}}(A, 1) \subseteq \mathrm{Hom}(A, 1) \cong \mathrm{Hom}_{A'}(A', A')$$

as an element of $G(1) = \mathrm{Hom}_{\mathrm{Alg}}(A, 1)$. □

Proposition 5.3. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Then, every element of the group $G(1)$ can be annihilated by raising to the r-th power.*

Proof. We choose any $u \in G(1)$ and let $\tau : A' \otimes A \to A' \otimes A$ be as above.

Now, suppose that we have a morphism $f \in \mathrm{Hom}_{A'\otimes A}(A' \otimes A, A' \otimes A)$. We set f' to be the composition

$$1 \xrightarrow{e_{A'\otimes A}} A' \otimes A \xrightarrow{f} A' \otimes A.$$

Then, from the proof of Proposition 3.1, we know that f is equal to the composition

$$A' \otimes A \cong A' \otimes A \otimes 1 \xrightarrow{1\otimes f'} A' \otimes A \otimes A' \otimes A \xrightarrow{m_A\otimes m_{A'}} A' \otimes A. \tag{5-13}$$

We set $f_\tau := \tau \circ f : A' \otimes A \to A' \otimes A$ and denote by f'_τ the composition

$$A'\otimes A\otimes 1 \xrightarrow{1\otimes e_{A'\otimes A}} A'\otimes A\otimes A'\otimes A \xrightarrow{1\otimes f_\tau} A'\otimes A\otimes A'\otimes A \xrightarrow{m_A\otimes m_{A'}} A'\otimes A.$$

It follows that $f'_\tau \in \mathrm{Hom}_{A'\otimes A}(A' \otimes A, A' \otimes A)$. We now consider the following commutative diagram in A'-*Mod*:

$$\begin{array}{ccccccc} A'\otimes A\otimes 1 & \xrightarrow{\tau\otimes e_{A'\otimes A}} & A'\otimes A\otimes A'\otimes A & \xrightarrow{1\otimes f_\tau} & A'\otimes A\otimes A'\otimes A & \xrightarrow{m_A\otimes m_{A'}} & A'\otimes A \\ \downarrow= & & & & & & =\downarrow \\ A'\otimes A\otimes 1 & \xrightarrow{1\otimes f'} & A'\otimes A\otimes A'\otimes A & \xrightarrow{\tau\otimes\tau} & A'\otimes A\otimes A'\otimes A & \xrightarrow{m_A\otimes m_{A'}} & A'\otimes A \\ \downarrow= & & & & & & =\downarrow \\ A'\otimes A\otimes 1 & \xrightarrow{1\otimes f'} & A'\otimes A\otimes A'\otimes A & \xrightarrow{m_A\otimes m_{A'}} & A'\otimes A & \xrightarrow{\tau} & A'\otimes A \end{array}$$

The upper rectangle in the figure above is commutative because $f_\tau = \tau \circ f$, while the lower rectangle commutes because τ is a morphism of algebras. Identifying $A' \otimes A$ with $A' \otimes A \otimes 1$, it now follows that

$$f'_\tau \circ \tau = \tau \circ f \in \mathrm{Hom}_{A'}(A' \otimes A, A' \otimes A). \tag{5-14}$$

Since τ is an automorphism, we have $f'_\tau = \tau f \tau^{-1}$. Then, since $\mathrm{Hom}_{A'}(A', A')$ is commutative,

$$N_{A'}(f'_\tau) = N_{A'}(\tau f \tau^{-1}) = N_{A'}(\tau f) N_{A'}(\tau^{-1}) = N_{A'}(\tau^{-1}) N_{A'}(\tau f) = N_{A'}(f).$$

We also note that if $h_1, h_2 \in G(A) = \mathrm{Hom}_{\mathrm{Alg}}(A, A)$ are two morphisms of algebras, the product $h_1 * h_2 \in G(A)$ corresponds to the morphism

$$h_1 * h_2 : A \xrightarrow{\Delta_A} A \otimes A \xrightarrow{h_1 \otimes h_2} A \otimes A \xrightarrow{m_A} A. \tag{5-15}$$

We have an isomorphism

$$H : \mathrm{Hom}(A, A) \xrightarrow{\cong} \mathrm{Hom}_{A' \otimes A}(A' \otimes A, A' \otimes A). \tag{5-16}$$

In particular, let $f = H(1_A)$. Then, we have $f'_\tau = H((1_A \otimes u) \circ \Delta_A)$. Now, since $N_{A'}(f) = N_{A'}(f'_\tau)$, it follows that

$$\begin{aligned} N_{A'}(f) &= N_{A'}(H((1_A \otimes u) \circ \Delta_A)) \\ &= N_{A'}(H(m_A \circ (1_A \otimes e_*(u)) \circ \Delta_A)) \\ &= N(1_A * e_*(u)) \qquad \text{(using (5-9))} \\ &= N(1_A) * N(e_*(u)) = N_{A'}(f) * N_{A'}(\tau), \end{aligned} \tag{5-17}$$

where the products $N(1_A) * N(e_*(u))$ and $N_{A'}(f) * N_{A'}(\tau)$ are taken in $G(1)$. Finally, from Lemma 5.2, we know that $N_{A'}(\tau) = u^r \in G(1)$. Combining with (5-17), it follows that u^r is the identity element of the group $G(1)$. □

Theorem 5.4. *Let $G = \mathfrak{spec}(A)$ be an affine commutative group scheme free of finite rank r. Then, for any algebra B in $\mathbf{C}$ and any element $u \in G(B)$, we have $u^r = 1_B$, where 1_B denotes the identity element of $G(B)$.*

Proof. For any algebra B in $\mathbf{C}$, we consider the symmetric monoidal category $(B\text{-}Mod, \otimes_B, B)$. Then, if we set $B_A := B \otimes A$, the functor $\mathrm{Hom}_{B\text{-}Alg}(B_A, \_\_)$ defines an affine commutative group scheme G_B on B-*Mod* free of finite rank r.

From Proposition 5.3, it now follows that all elements in the group $G_B(B)$ are annihilated by raising to the r-th power. Further, from Lemma 4.3, it follows that $G_B(B) = \mathrm{Hom}_{B\text{-}Alg}(B \otimes A, B) \cong \mathrm{Hom}_{\mathrm{Alg}}(A, B) = G(B)$. This proves the result. □

References

[Deligne 1990] P. Deligne, "Catégories tannakiennes", pp. 111–195 in *The Grothendieck Festschrift*, vol. 2, edited by P. Cartier et al., Progr. Math. **87**, Birkhäuser, Boston, MA, 1990. MR 92d:14002 Zbl 0727.14010

[Demazure and Gabriel 1970] M. Demazure and P. Gabriel, *Groupes algébriques, I: géométrie algébrique, généralités, groupes commutatifs*, Masson, Paris, 1970. MR 46 #1800 Zbl 0203.23401

[Hakim 1972] M. Hakim, *Topos annelés et schémas relatifs*, Ergebnisse der Math. **64**, Springer, Berlin, 1972. MR 51 #500 Zbl 0246.14004

[Mac Lane 1998] S. Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Mathematics **5**, Springer, New York, 1998. MR 2001j:18001 Zbl 0906.18001

[May 2001] J. P. May, "Picard groups, Grothendieck rings, and Burnside rings of categories", *Adv. Math.* 1 (2001), 1–16. MR 2002k:18011 Zbl 0994.18004

[Tate 1997] J. Tate, "Finite flat group schemes", pp. 121–154 in *Modular forms and Fermat's last theorem* (Boston, 1995), edited by G. Cornell et al., Springer, New York, 1997. MR 1638478 Zbl 0924.14024

[Tate and Oort 1970] J. Tate and F. Oort, "Group schemes of prime order", *Ann. Sci. École Norm. Sup.* (4) **3** (1970), 1–21. MR 42 #278 Zbl 0195.50801

[Toën and Vaquié 2009] B. Toën and M. Vaquié, "Au-dessous de Spec $\mathbb{Z}$", *J. K-Theory* **3**:3 (2009), 437–500. MR 2010j:14006 Zbl 1177.14022

[Waterhouse 1979] W. C. Waterhouse, *Introduction to affine group schemes*, Graduate Texts in Mathematics **66**, Springer, New York, 1979. MR 82e:14003 Zbl 0442.14017

Received May 29, 2011. Revised September 25, 2011.

ABHISHEK BANERJEE
DEPARTMENT OF MATHEMATICS
OHIO STATE UNIVERSITY
231 W 18TH AVENUE
COLUMBUS, OH 43210
UNITED STATES

abhishekbanerjee1313@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

EIGENVALUE ESTIMATES ON DOMAINS IN COMPLETE NONCOMPACT RIEMANNIAN MANIFOLDS

DAGUANG CHEN, TAO ZHENG AND MIN LU

In this paper, we obtain universal inequalities for eigenvalues of the Dirichlet eigenvalue problem of the Laplacian and the clamped plate problem on a bounded domain in an n-dimensional ($n \geq 3$) noncompact simply connected complete Riemannian manifold with sectional curvature Sec satisfying $-K^2 \leq \text{Sec} \leq -k^2$, where $K \geq k \geq 0$ are constants. When M is $\mathbb{H}^n(-1)$ ($n \geq 3$), these inequalities become ones previously found by Cheng and Yang.

1. Introduction

Let M be an n-dimensional complete Riemannian manifold and $\Omega \subset M$ a bounded domain in M. The Dirichlet eigenvalue problem of the Laplacian is

$$\begin{cases} \Delta u = -\lambda u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1-1}$$

It is well known that the spectrum of this problem is real and discrete:

$$0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \nearrow \infty,$$

where each λ_i has finite multiplicity which is repeated according to its multiplicity.

A Dirichlet eigenvalue problem of the biharmonic operator or a clamped plate problem that describes the characteristic vibrations of a clamped plate is given by

$$\begin{cases} \Delta^2 u = \Gamma u & \text{in } \Omega, \\ u = \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega, \end{cases} \tag{1-2}$$

where Δ^2 is the biharmonic operator on M and ν denotes the outward normal derivative on $\partial\Omega$. We will denote eigenvalues and the corresponding real eigenfunctions by $\{\Gamma_i\}_{i=1}^{\infty}$ and $\{u_i\}_{i=1}^{\infty}$, respectively. The eigenvalues Γ_i satisfy

$$0 < \Gamma_1 \leq \Gamma_2 \leq \Gamma_3 \leq \cdots \nearrow \infty.$$

D. Chen is partly supported by NSFC grant number 11101234.
MSC2010: 35P15, 58J50, 58G25, 53C42.
Keywords: Laplacian, the Dirichlet problem, the clamped plate problem, eigenvalues, the universal inequality.

When M is a Euclidean space $\mathbb{R}^n$, these are estimates for the eigenvalues (1-1) that do not involve domain dependencies [Protter 1988]; see also [Ashbaugh 1999; 2002]. The main developments were obtained by Payne, Pólya, and Weinberger [Payne et al. 1956], Hile and Protter [1980], and Yang [1991]. More recently, for the Dirichlet eigenvalue problems of the Laplacian on a bounded domain in the n-dimensional unit sphere, complex projective space, and compact homogeneous Riemannian manifolds, Cheng and Yang [2005; 2006b; 2007] obtained the Yang-type inequalities for eigenvalues. For a bounded domain Ω in a complete Riemannian manifold M, the first author and Cheng [Chen and Cheng 2008] proved a Yang-type inequality by using the Nash embedding theorem (compare [El Soufi et al. 2009; Harrell 2007]).

By making use of estimates for eigenvalues of the eigenvalue problem of the Schröinger like operator with a weight, Harrell and Michel [1994], Ashbaugh [2002], and Ashbaugh and Hermi [2007] have obtained several results. In fact, for $n=2$, the Laplacian on $\mathbb{H}^2(-1)$ is like to the Laplacian on $\mathbb{R}^2$ with a weight. However, for $n>2$, this property does not hold. Cheng and Yang [2009] found appropriate trial functions and obtained

$$\sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)^2 \le 4\sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)\Big(\lambda_i-\frac{(n-1)^2}{4}\Big). \tag{1-3}$$

In this paper, we first treat the Dirichlet eigenvalue problem (1-1) of the Laplacian on a bounded domain of a complete noncompact Riemannian manifold M.

Theorem 1.1. *Assume that M^n $(n\ge 3)$ is a noncompact simply connected complete Riemannian manifold with sectional curvature* Sec *satisfying $-K^2\le \mathrm{Sec}\le -k^2$, where $K\ge k\ge 0$ are constants.For a bounded domain Ω in M, let λ_i be the i-th eigenvalue of the eigenvalue problem* (1-1)*. Then we obtain*

$$\sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)^2 \le \sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)(4\lambda_i-(n-1)^2k^2+2(n-1)(K^2-k^2)). \tag{1-4}$$

Remark. If $k=K=1$, that is, M is a hyperbolic space $\mathbb{H}^n(-1)$, the eigenvalue inequality (1-4) agrees with (1-3) obtained by Cheng and Yang [2009].

The other purpose of this paper is to investigate estimates for eigenvalues of the clamped plate problem (1-2) on bounded domains Ω in a complete Riemannian manifold M^n.

For the universal inequalities for eigenvalues of the clamped plate problem in a bounded domain in $\mathbb{R}^n$, Payne et al. [1955; 1956] proved that

$$\Gamma_{k+1}-\Gamma_k \le \frac{8(n+2)}{n^2k}\sum_{i=1}^{k}\Gamma_i, \quad k=1,2,\dots. \tag{1-5}$$

Hile and Yeh [1984] obtained

$$\sum_{i=1}^{k} \frac{\Gamma_i^{\frac{1}{2}}}{\Gamma_{k+1}-\Gamma_i} \geq \frac{n^2 k^{3/2}}{8(n+2)} \left(\sum_{i=1}^{k} \Gamma_i\right)^{-\frac{1}{2}}, \quad k=1,2,\ldots. \tag{1-6}$$

Hook [1990] and Chen and Qian [1990] independently proved

$$\frac{n^2 k^2}{8(n+2)} \leq \left(\sum_{i=1}^{k} \frac{\Gamma_i^{\frac{1}{2}}}{\Gamma_{k+1}-\Gamma_i}\right)\left(\sum_{i=1}^{k} \Gamma_i^{\frac{1}{2}}\right), \quad k=1,2,\ldots. \tag{1-7}$$

Cheng and Yang [2006a] gave an affirmative answer for a problem on universal inequalities for eigenvalues proposed by Ashbaugh [1999]: they proved that

$$\Gamma_{k+1} - \frac{1}{k}\sum_{i=1}^{k} \Gamma_i \leq \left(\frac{8(n+2)}{n^2}\right)^{\frac{1}{2}} \frac{1}{k}\left(\sum_{i=1}^{k} \Gamma_i(\Gamma_{k+1}-\Gamma_i)\right)^{\frac{1}{2}}, \quad k=1,2,\ldots. \tag{1-8}$$

For domains in a unit sphere, Wang and Xia [2007] gave a universal inequality for the clamped plate problem (1-2). They proved

$$\sum_{i=1}^{k} (\Gamma_{k+1}-\Gamma_i)^2 \leq \frac{8(n+2)}{n^2} \sum_{i=1}^{k} (\Gamma_{k+1}-\Gamma_i)\left(\Gamma_i^{\frac{1}{2}} + \frac{n^2}{2n+4}\right)\left(\Gamma_i^{\frac{1}{2}} + \frac{n^2}{4}\right). \tag{1-9}$$

For an n-dimensional complete manifold M, Cheng, Ichikawa, and Mametsuka [Cheng et al. 2010] obtained

$$\begin{aligned}\sum_{i=1}^{k} (\Gamma_{k+1}-\Gamma_i)^2 \\ \leq \frac{8(n+2)}{n^2} \sum_{i=1}^{k} (\Gamma_{k+1}-\Gamma_i)\left(\Gamma_i^{\frac{1}{2}} + \frac{n^2}{2n+4}\sup_{\Omega}|H|^2\right)\left(\Gamma_i^{\frac{1}{2}} + \frac{n^2}{4}\sup_{\Omega}|H|^2\right).\end{aligned} \tag{1-10}$$

For the real hyperbolic space $\mathbb{H}^n(-1)$, Cheng and Yang [2011] proved that

$$\sum_{i=1}^{k} (\Gamma_{k+1}-\Gamma_i)^2 \leq 24 \sum_{i=1}^{k} (\Gamma_{k+1}-\Gamma_i)\left(\Gamma_i^{\frac{1}{2}} - \frac{(n-1)^2}{4}\right)\left(\Gamma_i^{\frac{1}{2}} - \frac{(n-1)^2}{6}\right). \tag{1-11}$$

That paper motivated the present one, where we treat the clamped plate problem on a bounded domain of a noncompact simply connected complete Riemannian manifold M^n.

Theorem 1.2. *Assume that M^n $(n \geq 3)$ is a noncompact simply connected complete Riemannian manifold with sectional curvature* Sec *satisfying $-K^2 \leq \mathrm{Sec} \leq -k^2$, where $K \geq k \geq 0$ are constants. For a bounded domain Ω in M, let Γ_i be the i-th* *eigenvalue of the eigenvalue problem* (1-2). *Then we have*

$$\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2 \le 24\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{4}((n-1)k^2-2(K^2-k^2))\Big) \times \Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\Big). \tag{1-12}$$

Remark. If $k = K = 1$, that is, M^n is a hyperbolic space $\mathbb{H}^n(-1)$, then the eigenvalue inequality (1-12) agrees with (1-11) obtained by Cheng and Yang. Wang and Xia [2011] generalized (1-11) under the assumption that there exists some function whose norm of gradient is 1 and whose Laplacian is a constant.

From Theorem 1.2, we can immediately obtain the following.

Corollary 1.3. *Let* Γ_i *be the i-th eigenvalue of the eigenvalue problem* (1-2). *Then we have*

$$\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2 \le 24\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)\Big(\Gamma_i-\frac{(n-1)^2}{16}((n-1)k^2-2(K^2-k^2))^2\Big).$$

2. Preliminaries

Let B and C be $(n-1)\times(n-1)$ real symmetrical matrixes. If all the eigenvalues of B are equal or greater than all the ones of C, then we write $B \succ C$.

Let (M, g) be an n-dimensional Riemannian manifold and D the Riemannian connection. The curvature tensor is a (1,3)-tensor defined by

$$R(X, Y)Z = D_X D_Y Z - D_Y D_X Z - D_{[X,Y]}Z \tag{2-1}$$

for all $X, Y, Z \in \Gamma(\mathrm{TM})$. Let $\gamma : [0, b) \longrightarrow M$ be the minimal normal geodesic and $\{e_i(t)\}_{i=1}^n$ parallel orthonormal frame fields along $\gamma(t)$ such that $e_n(t) = \gamma(t)$. Let

$$J_i(t) = \sum_{j=1}^{n-1} f_{ij}(t)e_j(t), i = 1, \dots, n-1,$$

be the normal Jacobi fields along the geodesic $\gamma(t)$; that is

$$\ddot{f}_{ij} - f_{il}R_{njnl} = 0, \quad f_{ij}(0) = 0, \quad \dot{f}_{ij}(0) = \delta_{ij}, \tag{2-2}$$

where

$$\dot{f}_{ij} = \frac{\mathrm{d}}{\mathrm{d}t} f_{ij}(t), \quad \ddot{f}_{ij} = \frac{\mathrm{d}^2}{\mathrm{d}t^2} f_{ij}(t), \quad R_{njnl} = g(R(e_n, e_l)e_n, e_j) = R_{nlnj}.$$

Set

$$f(t) = (f_{ij}(t))_{(n-1)\times(n-1)}, \quad K(t) = (R_{nlnj}(\gamma(t)))_{(n-1)\times(n-1)},$$

where $f_{ij}(t)$ is on column j and row i. Then (2-2) can be written as

$$\begin{cases} \ddot{f}(t) - f(t)K(t) = 0, & 0 < t < b, \\ f(0) = 0, \\ \dot{f}(0) = I_{n-1}, \end{cases} \tag{2-3}$$

where I_{n-1} is the $(n-1) \times (n-1)$ unit matrix.

Define the distance function $r(x) = \text{distance}(x, \gamma(0))$. Then

$$\text{Hess}\, r(\gamma(t)) = f(t)^{-1} \dot{f}(t), \quad \Delta r(\gamma(t)) = \text{tr}(f(t)^{-1} \dot{f}(t)). \tag{2-4}$$

Assume that Ω is a bounded domain in an n-dimensional noncompact simply connected complete Riemannian manifold (M, g) with section curvature Sec satisfying $-K^2 \le \text{Sec} \le -k^2$, where $0 \le k \le K$ are constants. For $p \notin \overline{\Omega}$ fixed, define the distance function $r(x) = \text{distance}(x, p)$. Then from the Hessian comparison theorem (cf. [Wu et al. 1989]), we have

$$K \frac{\cosh Kr}{\sinh Kr} I_{n-1} \succ \text{Hess}\, r \succ k \frac{\cosh kr}{\sinh kr} I_{n-1}. \tag{2-5}$$

From (2-4) and (2-5), we have

$$(n-1)k \frac{\cosh kr}{\sinh kr} \le \Delta r \le (n-1)K \frac{\cosh Kr}{\sinh Kr}. \tag{2-6}$$

Since $\partial_r \Delta r = -|\text{Hess}\, r|^2 - \text{Ric}(\partial_r, \partial_r)$ (cf. [Petersen 1998]), we have

$$-\partial_r \Delta r \le (n-1)K^2 \frac{\cosh^2 Kr}{\sinh^2 Kr} - (n-1)k^2. \tag{2-7}$$

3. Proof of Theorem 1.1

Theorem 3.1 [Cheng and Yang 2006b]. *Let λ_i be the i-th eigenvalue of the above eigenvalue problem* (1-1) *and u_i the orthonormal eigenfunction corresponding to λ_i; that is, u_i satisfies*

$$\begin{cases} u_i = -\lambda u_i & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \\ \int_\Omega u_i u_j = \delta_{ij} & \text{for all } i, j = 1, 2, \dots. \end{cases}$$

Then for any $f \in C^3(\Omega) \cap C^2(\partial\Omega)$, we have

$$\sum_{i=1}^{k} (\lambda_{k+1} - \lambda_i)^2 \int_\Omega |\nabla f|^2 u_i^2 \le \sum_{i=1}^{k} (\lambda_{k+1} - \lambda_i) \int_\Omega (2\nabla f \cdot \nabla u_i + u_i \Delta f)^2. \tag{3-1}$$

Proof of Theorem 1.1. Taking $f = r$ in the formula (3-1), we have

$$\sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)^2\int_\Omega |\nabla r|^2 u_i^2 \le \sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)\int_\Omega (2\nabla r\cdot\nabla u_i + u_i\Delta r)^2.$$

Since $|\nabla r| = 1$, we have

$$\sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)^2 \le \sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)\int_\Omega (2\nabla r\cdot\nabla u_i + u_i\Delta r)^2.$$

From (2-6) and (2-7), we obtain

$$\begin{aligned}
\text{(3-2)}\quad & \int_\Omega \Big(2\nabla r\cdot\nabla u_i + u_i\Delta r\Big)^2 \\
&= 4\int_\Omega (\nabla r\cdot\nabla u_i)^2 + 4\int_\Omega u_i\Delta r\nabla r\cdot\nabla u_i + \int_\Omega (u_i\Delta r)^2 \\
&\le 4\int_\Omega |\nabla u_i|^2 - \int_\Omega u_i^2(\Delta r)^2 - 2\int_\Omega u_i^2\nabla r\cdot\nabla\Delta r \\
&= 4\int_\Omega |\nabla u_i|^2 - \int_\Omega u_i^2(\Delta r)^2 - 2\int_\Omega u_i^2\partial_r\Delta r \\
&= 4\int_\Omega |\nabla u_i|^2 - \int_\Omega u_i^2(\Delta r)^2 + 2\int_\Omega u_i^2(\mathrm{Ric}(\partial_r,\partial_r) + |\mathrm{Hess}\, r|^2) \\
&\le 4\lambda_i - (n-1)^2k^2\int_\Omega u_i^2\frac{\cosh^2 kr}{\sinh^2 kr} \\
&\qquad\qquad - 2(n-1)k^2 + 2(n-1)K^2\int_\Omega u_i^2\frac{\cosh^2 Kr}{\sinh^2 Kr} \\
&= 4\lambda_i - (n-1)^2k^2 + 2(n-1)(K^2-k^2) \\
&\qquad\qquad - (n-1)^2\int_\Omega \frac{k^2}{\sinh^2 kr}u_i^2 + 2(n-1)\int_\Omega \frac{K^2}{\sinh^2 Kr}u_i^2.
\end{aligned}$$

Since $K \ge k \ge 0$ and $r > 0$, we have

$$\frac{K}{\sinh Kr} \le \frac{k}{\sinh kr}. \tag{3-3}$$

Since $n \ge 3$, we have

$$(n-1)^2\frac{k^2}{\sinh^2 kr} - 2(n-1)\frac{K^2}{\sinh^2 Kr} \ge (n-1)(n-3)\frac{k^2}{\sinh^2 kr} \ge 0. \tag{3-4}$$

Finally, we have

$$\sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)^2 \le \sum_{i=1}^{k}(\lambda_{k+1}-\lambda_i)\big(4\lambda_i - (n-1)^2k^2 + 2(n-1)(K^2-k^2)\big). \qquad \square$$

4. Proof of Theorem 1.2

Let u_i be the i-th orthonormal eigenfunction corresponding to the eigenvalue Γ_i, $i = 1, \dots, k$; that is,

$$\begin{cases} \Delta^2 u_i = \Gamma_i u_i & \text{in } \Omega, \\ u_i = \dfrac{\partial u_i}{\partial \nu} = 0 & \text{on } \partial\Omega, \\ \displaystyle\int_\Omega u_i u_j = \delta_{ij} & \text{for any } i, j. \end{cases} \tag{4-1}$$

Defining the functions

$$\phi_i = r u_i - \sum_{j=1}^{k} a_{ij} u_j,$$

where

$$a_{ij} = \int_\Omega r u_i u_j,$$

we have

$$\phi_i|_{\partial\Omega} = \frac{\partial \phi_i}{\partial \nu}\Big|_{\partial\Omega} = 0 \quad \text{and} \quad \int_\Omega \phi_i u_j = 0 \qquad \text{for all } i, j = 1, \dots, k. \tag{4-2}$$

Therefore, we know that ϕ_is are trial functions. From the Rayleigh–Ritz inequality [Chavel 1984], we have

$$\Gamma_{k+1} \le \frac{1}{\|\phi_i\|^2} \int_\Omega (\Delta \phi_i)^2, \tag{4-3}$$

where

$$\|\phi_i\|^2 = \int_\Omega \phi_i^2.$$

From (4-1) and (4-2), we have

$$\begin{aligned} \Gamma_{k+1} \int_\Omega \phi_i^2 \le \int_\Omega (\Delta\phi_i)^2 = \int_\Omega \phi_i \Delta^2 \phi_i = \int_\Omega \phi_i \Delta^2 \Big(r u_i - \sum_{j=1}^{k} a_{ij} u_j \Big) \\ = \int_\Omega \phi_i \Delta^2 (r u_i) = \int_\Omega \phi_i (\Delta^2 (r u_i) - \Gamma_i r u_i) + \Gamma_i \int_\Omega \phi_i^2, \end{aligned}$$

that is,

$$(\Gamma_{k+1} - \Gamma_i) \|\phi_i\|^2 \le \int_\Omega \phi_i (\Delta^2 (r u_i) - \Gamma_i r u_i).$$

From the definition of ϕ_i and (4-2), we have

$$(4\text{-}4)\quad (\Gamma_{k+1}-\Gamma_i)\|\phi_i\|^2 \le \int_\Omega \Big(ru_i-\sum_{j=1}^k a_{ij}u_j\Big)(\Delta^2(ru_i)-\Gamma_i ru_i) = \int_\Omega ru_i(\Delta^2(ru_i)-\Gamma_i ru_i)+\sum_{j=1}^k a_{ij}^2(\Gamma_i-\Gamma_j) = \int_\Omega ru_i(\Delta(u_i\Delta r)+2\Delta(\nabla r\cdot\nabla u_i)+2\nabla r\cdot\nabla\Delta u_i+\Delta r\Delta u_i) + \sum_{j=1}^k a_{ij}^2(\Gamma_i-\Gamma_j).$$

From (2-6), (2-7), and Stokes' theorem, by a direct calculation, we have

$$(4\text{-}5)\quad \int_\Omega ru_i(\Delta(u_i\Delta r)+2\Delta(\nabla r\cdot\nabla u_i)+2\nabla r\cdot\nabla\Delta u_i+\Delta r\Delta u_i)$$

$$= \int_\Omega (\Delta(ru_i)(u_i\Delta r+2\nabla r\cdot\nabla u_i)+u_i\nabla r^2\cdot\nabla\Delta u_i+u_i r\Delta r\Delta u_i)$$

$$= \int_\Omega \big((u_i\Delta r+2\nabla r\cdot\nabla u_i+r\Delta u_i)(u_i\Delta r+2\nabla r\cdot\nabla u_i) + u_i\nabla r^2\cdot\nabla\Delta u_i+u_i r\Delta r\Delta u_i\big)$$

$$= \int_\Omega \big((\Delta r)^2u_i^2+2\nabla r\cdot\nabla u_i^2\Delta r+4(\nabla r\cdot\nabla u_i)^2+2ru_i\Delta r\Delta u_i+\nabla r^2\cdot\nabla u_i\Delta u_i\big) + \int_\Omega u_i\nabla r^2\cdot\nabla\Delta u_i$$

$$= \int_\Omega \big((\Delta r)^2u_i^2+2\nabla r\cdot\nabla u_i^2\Delta r+4(\nabla r\cdot\nabla u_i)^2+2ru_i\Delta r\Delta u_i+\nabla r^2\cdot\nabla(u_i\Delta u_i)\big)$$

$$= \int_\Omega ((\Delta r)^2u_i^2+2\nabla r\cdot\nabla u_i^2\Delta r+4(\nabla r\cdot\nabla u_i)^2+(2r\Delta r-\Delta r^2)u_i\Delta u_i)$$

$$= \int_\Omega (-(\Delta r)^2u_i^2-2u_i^2\nabla r\cdot\nabla\Delta r+4(\nabla r\cdot\nabla u_i)^2-2u_i\Delta u_i)$$

$$\le \int_\Omega (4|\nabla u_i|^2-2u_i\Delta u_i)-\int_\Omega u_i^2(2\nabla r\cdot\nabla\Delta r+(\Delta r)^2)$$

$$\le \int_\Omega u_i^2\Big(-(n-1)^2k^2\frac{\cosh^2 kr}{\sinh^2 kr}-2(n-1)k^2+2(n-1)K^2\frac{\cosh^2 Kr}{\sinh^2 Kr}\Big) + \int_\Omega (4|\nabla u_i|^2+2u_i(-\Delta u_i))$$

Since $n \geq 3$, from (3-4), we have

$$\text{(4-6)} \quad \int_\Omega ru_i(\Delta(u_i\Delta r) + 2\Delta(\nabla r\cdot\nabla u_i) + 2\nabla r\cdot\nabla\Delta u_i + \Delta r\Delta u_i)$$
$$\leq 6\int_\Omega u_i(-\Delta u_i) - \int_\Omega u_i^2((n-1)^2k^2 - 2(n-1)(K^2-k^2))$$
$$\leq 6\Big(\int_\Omega (\Delta u_i)^2\Big)^{\frac{1}{2}} - (n-1)((n-1)k^2 - 2(K^2-k^2))$$
$$= 6\Big(\Gamma_i^{\frac{1}{2}} - \frac{n-1}{6}((n-1)k^2 - 2(K^2-k^2))\Big).$$

From (4-4) and (4-6), we deduce

$$\text{(4-7)} \quad (\Gamma_{k+1} - \Gamma_i)\|\phi_i\|^2$$
$$\leq 6\Big(\Gamma_i^{\frac{1}{2}} - \frac{n-1}{6}((n-1)k^2 - 2(K^2-k^2))\Big) + \sum_{j=1}^k a_{ij}^2(\Gamma_i - \Gamma_j).$$

Defining

$$b_{ij} = \int_\Omega \big(\nabla r\cdot\nabla u_i + \tfrac{1}{2}u_i\Delta r\big)u_j,$$

we have

$$b_{ij} = -b_{ji}.$$

From the definitions of b_{ij} and ϕ_i, we obtain

$$\text{(4-8)} \quad -2\int_\Omega \phi_i(\nabla r\cdot\nabla u_i + \tfrac{1}{2}\Delta r u_i)$$
$$= -2\int_\Omega \Big(ru_i - \sum_{j=1}^k a_{ij}u_j\Big)\big(\nabla r\cdot\nabla u_i + \tfrac{1}{2}\Delta r u_i\big)$$
$$= -2\int_\Omega ru_i\big(\nabla r\cdot\nabla u_i + \tfrac{1}{2}\Delta r u_i\big) + 2\sum_{j=1}^k a_{ij}b_{ij}$$
$$= -\int_\Omega \big(\tfrac{1}{2}\nabla r^2\cdot\nabla u_i^2 + r\Delta r u_i^2\big) + 2\sum_{j=1}^k a_{ij}b_{ij}$$
$$= 1 + 2\sum_{j=1}^k a_{ij}b_{ij}.$$

From (4-2), (4-8), and the Cauchy–Schwartz inequality, we have

$$1+2\sum_{j=1}^{k}a_{ij}b_{ij} = -2\int_\Omega \phi_i\big(\nabla r\cdot\nabla u_i+\tfrac{1}{2}u_i\Delta r\big) \tag{4-9}$$
$$= -2\int_\Omega \phi_i\Big(\nabla r\cdot\nabla u_i+\tfrac{1}{2}u_i\Delta r-\sum_{j=1}^{k}b_{ij}u_j\Big)$$
$$\le \alpha_i\|\phi_i\|^2+\frac{1}{\alpha_i}\Big\|\nabla r\cdot\nabla u_i+\tfrac{1}{2}u_i\Delta r-\sum_{j=1}^{k}b_{ij}u_j\Big\|^2$$
$$= \alpha_i\|\phi_i\|^2+\frac{1}{\alpha_i}\Big(\|\nabla r\cdot\nabla u_i+\tfrac{1}{2}u_i\Delta r\|^2-\sum_{j=1}^{k}b_{ij}^2\Big),$$

where $\alpha_i > 0$ is a positive constant.

If $\Gamma_{k+1}-\Gamma_i > 0$, defining

$$\alpha_i = (\Gamma_{k+1}-\Gamma_i)\beta_i \quad \text{for } \beta_i > 0,$$

we infer that

$$(\Gamma_{k+1}-\Gamma_i)^2\Big(1+2\sum_{j=1}^{k}a_{ij}b_{ij}\Big) \tag{4-10}$$
$$\le (\Gamma_{k+1}-\Gamma_i)^3\beta_i\|\phi_i\|^2$$
$$+m\frac{1}{\beta_i}(\Gamma_{k+1}-\Gamma_i)\Big(\|\nabla r\cdot\nabla u_i+\tfrac{1}{2}u_i\Delta r\|^2-\sum_{j=1}^{k}b_{ij}^2\Big).$$

From (2-6), (2-7), and (3-4), we obtain

$$\int_\Omega(2\nabla r\cdot\nabla u_i+u_i\Delta r)^2 \tag{4-11}$$
$$= 4\int_\Omega(\nabla r\cdot\nabla u_i)^2+4\int_\Omega u_i\Delta r\nabla r\cdot\nabla u_i+\int_\Omega(u_i\Delta r)^2$$
$$\le 4\int_\Omega|\nabla u_i|^2-\int_\Omega u_i^2(\Delta r)^2-2\int_\Omega u_i^2\nabla r\cdot\nabla\Delta r$$
$$\le 4\Gamma_i^{\frac{1}{2}}-(n-1)\big((n-1)k^2-2(K^2-k^2)\big)$$
$$\le 4\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{4}((n-1)k^2-2(K^2-k^2))\Big).$$

Therefore, from (4-7), (4-9), (4-10), and (4-11), we obtain

$$
\begin{aligned}
(4\text{-}12)\quad &(\Gamma_{k+1}-\Gamma_i)^2\Big(1+2\sum_{j=1}^{k}a_{ij}b_{ij}\Big)\\
&\le(\Gamma_{k+1}-\Gamma_i)^2\beta_i\Big(6\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\Big)+\sum_{j=1}^{k}a_{ij}^2(\Gamma_i-\Gamma_j)\Big)\\
&\quad+\frac{1}{\beta_i}(\Gamma_{k+1}-\Gamma_i)\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{4}((n-1)k^2-2(K^2-k^2))\Big)\\
&\quad-\frac{1}{\beta_i}(\Gamma_{k+1}-\Gamma_i)\sum_{j=1}^{k}b_{ij}^2.
\end{aligned}
$$

From the antisymmetry of b_{ij} and the Cauchy–Schwartz inequality, we have

$$
\begin{aligned}
&2\sum_{i,j=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2a_{ij}b_{ij}\\
&\qquad-\sum_{i,j=1}^{k}(\Gamma_{k+1}-\Gamma_i)(\Gamma_i-\Gamma_j)^2\beta_i a_{ij}^2-\sum_{i,j=1}^{k}\frac{1}{\beta_i}(\Gamma_{k+1}-\Gamma_i)b_{ij}^2\le 0.
\end{aligned}
$$

From the above inequality and (4-12), we obtain

$$
\begin{aligned}
(4\text{-}13)\quad &\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2\\
&\le 6\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2\beta_i\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\Big)\\
&\quad+\sum_{i=1}^{k}\frac{1}{\beta_i}(\Gamma_{k+1}-\Gamma_i)\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{4}((n-1)k^2-2(K^2-k^2))\Big)\\
&\quad+\sum_{i,j=1}^{k}(\Gamma_{k+1}-\Gamma_i)(\Gamma_{k+1}-\Gamma_j)(\Gamma_i-\Gamma_j)\beta_i a_{ij}^2.
\end{aligned}
$$

From the variational principle, we can prove that

$$\Gamma_i\ge\lambda_i^2,$$

where λ_i denotes the i-th eigenvalue of the Dirichlet eigenvalue problem of the Laplacian on the same domain Ω. Since $4\lambda_1\ge(n-1)^2k^2-2(K^2-k^2)$ from (3-2), setting

$$\beta_i=\beta\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\Big)^{-1}\quad\text{for }\beta>0$$

gives us

$$
\begin{aligned}
\sum_{i,j=1}^{k} &(\Gamma_{k+1}-\Gamma_i)(\Gamma_{k+1}-\Gamma_j)(\Gamma_i-\Gamma_j)\beta_i a_{ij}^2 \\
&= \frac{1}{2}\sum_{i,j=1}^{k}(\Gamma_{k+1}-\Gamma_i)(\Gamma_{k+1}-\Gamma_j)(\Gamma_i-\Gamma_j)(\beta_i-\beta_j)a_{ij}^2 \\
&= -\frac{1}{2}\beta\sum_{i,j=1}^{k}\frac{(\Gamma_{k+1}-\Gamma_i)(\Gamma_{k+1}-\Gamma_j)(\Gamma_i-\Gamma_j)(\Gamma_i^{\frac{1}{2}}-\Gamma_j^{\frac{1}{2}})}{\big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\big)\times\big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\big)}a_{ij}^2 \\
&\le 0.
\end{aligned}
$$

Therefore, we obtain

$$
\begin{aligned}
\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2 \le 6\beta\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)^2 &\\
+\frac{1}{\beta}\sum_{i=1}^{k}(\Gamma_{k+1}-\Gamma_i)\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{4}((n-1)k^2-2(K^2-k^2))\Big)&\cdot \\
\times\Big(\Gamma_i^{\frac{1}{2}}-\frac{n-1}{6}((n-1)k^2-2(K^2-k^2))\Big)&.
\end{aligned}
$$

Finally, taking $\beta=\frac{1}{12}$, we deduce (1-12). This completes the proof of Theorem 1.2.

Acknowledgements

We thank Professor Hongcang Yang for his encouragement, suggestions, and support. We would also like to thank the referee for helpful comments and for bringing [Wang and Xia 2011] to our attention.

References

[Ashbaugh 1999] M. S. Ashbaugh, "Isoperimetric and universal inequalities for eigenvalues", pp. 95–139 in *Spectral theory and geometry* (Edinburgh, 1998), edited by E. B. Davies and Y. Safarov, London Math. Soc. Lecture Note Ser. **273**, Cambridge University Press, Cambridge, 1999. MR 2001a:35131 Zbl 0937.35114

[Ashbaugh 2002] M. S. Ashbaugh, "The universal eigenvalue bounds of Payne–Pólya–Weinberger, Hile–Protter, and H. C. Yang: spectral and inverse spectral theory", *Proc. Indian Acad. Sci.* (*Math. Sci.*) **112**:1 (2002), 3–30. MR 2004c:35302 Zbl 1199.35261

[Ashbaugh and Hermi 2007] M. S. Ashbaugh and L. Hermi, "On Harrell–Stubbe type inequalities for the discrete spectrum of a self-adjoint operator", preprint, 2007. arXiv 0712.4396

[Chavel 1984] I. Chavel, *Eigenvalues in Riemannian geometry*, Pure and Applied Mathematics **115**, Academic Press, Orlando, FL, 1984. MR 86g:58140 Zbl 0551.53001

[Chen and Cheng 2008] D. Chen and Q.-M. Cheng, "Extrinsic estimates for eigenvalues of the Laplace operator", *J. Math. Soc. Japan* **60**:2 (2008), 325–339. MR 2010b:35323 Zbl 1147.35060

[Chen and Qian 1990] Z. C. Chen and C. L. Qian, "Estimates for discrete spectrum of Laplacian operator with any order", *J. China Univ. Sci. Tech.* **20**:3 (1990), 259–266. MR 92c:35087 Zbl 0748.35022

[Cheng and Yang 2005] Q.-M. Cheng and H. C. Yang, "Estimates on eigenvalues of Laplacian", *Math. Ann.* **331**:2 (2005), 445–460. MR 2005i:58038 Zbl 1122.35086

[Cheng and Yang 2006a] Q.-M. Cheng and H. C. Yang, "Inequalities for eigenvalues of a clamped plate problem", *Trans. Amer. Math. Soc.* **358**:6 (2006), 2625–2635. MR 2006m:35263 Zbl 1096.35095

[Cheng and Yang 2006b] Q.-M. Cheng and H. C. Yang, "Inequalities for eigenvalues of Laplacian on domains and compact complex hypersurfaces in complex projective spaces", *J. Math. Soc. Japan* **58**:2 (2006), 545–561. MR 2007k:58051 Zbl 1127.35026

[Cheng and Yang 2007] Q.-M. Cheng and H. C. Yang, "Bounds on eigenvalues of Dirichlet Laplacian", *Math. Ann.* **337**:1 (2007), 159–175. MR 2007k:35064 Zbl 1110.35052

[Cheng and Yang 2009] Q.-M. Cheng and H. C. Yang, "Estimates for eigenvalues on Riemannian manifolds", *J. Differential Equations* **247**:8 (2009), 2270–2281. MR 2010j:58066 Zbl 1180.35390

[Cheng and Yang 2011] Q.-M. Cheng and H. C. Yang, "Universal inequalities for eigenvalues of a clamped plate problem on a hyperbolic space", *Proc. Amer. Math. Soc.* **139**:2 (2011), 461–471. MR 2012b:35234 Zbl 1209.35089

[Cheng et al. 2010] Q.-M. Cheng, T. Ichikawa, and S. Mametsuka, "Estimates for eigenvalues of a clamped plate problem on Riemannian manifolds", *J. Math. Soc. Japan* **62**:2 (2010), 673–686. MR 2011e:58039 Zbl 1191.35192

[El Soufi et al. 2009] A. El Soufi, E. M. Harrell, II, and S. Ilias, "Universal inequalities for the eigenvalues of Laplace and Schrödinger operators on submanifolds", *Trans. Amer. Math. Soc.* **361**:5 (2009), 2337–2350. MR 2010e:58032 Zbl 1162.58009

[Harrell 2007] E. M. Harrell, II, "Commutators, eigenvalue gaps, and mean curvature in the theory of Schrödinger operators", *Comm. Partial Differential Equations* **32**:3 (2007), 401–413. MR 2008i:35041 Zbl 05150097

[Harrell and Michel 1994] E. M. Harrell, II and P. L. Michel, "Commutator bounds for eigenvalues, with applications to spectral geometry", *Comm. Partial Differential Equations* **19**:11-12 (1994), 2037–2055. MR 95i:58182 Zbl 0815.35078

[Hile and Protter 1980] G. N. Hile and M. H. Protter, "Inequalities for eigenvalues of the Laplacian", *Indiana Univ. Math. J.* **29**:4 (1980), 523–538. MR 82c:35052 Zbl 0454.35064

[Hile and Yeh 1984] G. N. Hile and R. Z. Yeh, "Inequalities for eigenvalues of the biharmonic operator", *Pacific J. Math.* **112**:1 (1984), 115–133. MR 85k:35170 Zbl 0541.35059

[Hook 1990] S. M. Hook, "Domain-independent upper bounds for eigenvalues of elliptic operators", *Trans. Amer. Math. Soc.* **318**:2 (1990), 615–642. MR 90h:35075 Zbl 0727.35096

[Payne et al. 1955] L. E. Payne, G. Pólya, and H. F. Weinberger, "Sur le quotient de deux fréquences propres consécutives", *C. R. Acad. Sci. Paris* **241** (1955), 917–919. MR 17,372d Zbl 0065.08801

[Payne et al. 1956] L. E. Payne, G. Pólya, and H. F. Weinberger, "On the ratio of consecutive eigenvalues", *J. Math. and Phys.* **35** (1956), 289–298. MR 18,905c Zbl 0073.08203

[Petersen 1998] P. Petersen, *Riemannian geometry*, vol. 171, Grad. Texts in Math., Springer, New York, 1998. 2nd ed. published in 2006. MR 98m:53001 Zbl 0914.53001

[Protter 1988] M. H. Protter, "Universal inequalities for eigenvalues", pp. 111–120 in *Maximum principles and eigenvalue problems in partial differential equations* (Knoxville, TN, 1987), edited by P. W. Schaefer, Pitman Res. Notes Math. Ser. **175**, Longman Scientific & Technical, Harlow, 1988. MR 89k:35167 Zbl 0663.35052

[Wang and Xia 2007] Q. Wang and C. Xia, "Universal bounds for eigenvalues of the biharmonic operator on Riemannian manifolds", *J. Funct. Anal.* **245**:1 (2007), 334–352. MR 2008e:58033 Zbl 1113.58013

[Wang and Xia 2011] Q. Wang and C. Xia, "Inequalities for eigenvalues of a clamped plate problem", *Calc. Var. Partial Differential Equations* **40**:1-2 (2011), 273–289. MR 2012a:35215 Zbl 1205.35175

[Wu et al. 1989] H. Wu, C. L. Shen, and Y. L. Yu, *Introduction to Riemannian Geometry*, Peking University Press, Peking, 1989. In Chinese.

[Yang 1991] H. C. Yang, "An estimate of the difference between consecutive eigenvalues", preprint IC/91/60, International Centre for Theoretical Physics (ICTP), Trieste, 1991.

Received January 18, 2011. Revised October 26, 2011.

DAGUANG CHEN
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING, 100084
CHINA
dgchen@math.tsinghua.edu.cn

TAO ZHENG
HUA LOO-KENG KEY LABORATORY OF MATHEMATICS
CHINESE ACADEMY OF SCIENCES
BEIJING, 100190
CHINA
zhengtao08@mails.gucas.ac.cn

MIN LU
DEPARTMENT OF APPLIED MATHEMATICS
NANJING AUDIT UNIVERSITY
NANJING, 210029
CHINA
luminm@foxmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

REALIZING THE LOCAL WEIL REPRESENTATION OVER A NUMBER FIELD

GERALD CLIFF AND DAVID MCNEILLY

Let F be a non-Archimedean local field whose residue field has order q and characteristic $p \neq 2$. We show that the Weil representations of the symplectic group $\mathrm{Sp}(2n, F)$ can be realized over the field

$$E_0 = \begin{cases} \mathbb{Q}(\sqrt{p}, \sqrt{-p}), & \text{if } q \text{ is not a square;} \\ \mathbb{Q}(\sqrt{-p}), & \text{if } q \text{ is a square and } p \equiv 1 \bmod 4; \\ \mathbb{Q}(\sqrt{-1}), & \text{if } q \text{ is a square and } p \equiv 3 \bmod 4. \end{cases}$$

Furthermore, the field E_0 is shown to be optimal if $q \equiv 1 \bmod 4$.

1. Introduction

Let F be a non-Archimedean local field whose residue field has order q and characteristic $p \neq 2$. Our main result is that the Weil representations of the symplectic group $\mathrm{Sp}(2n, F)$, can be realized over the number field

$$E_0 = \begin{cases} \mathbb{Q}(\sqrt{p}, \sqrt{-p}), & \text{if } q \text{ is not a square;} \\ \mathbb{Q}(\sqrt{-p}), & \text{if } q \text{ is a square and } p \equiv 1 \bmod 4; \\ \mathbb{Q}(\sqrt{-1}), & \text{if } q \text{ is a square and } p \equiv 3 \bmod 4. \end{cases}$$

This answers a question raised by D. Prasad [1998]. A consequence of this, also pointed out by Prasad, is that the local theta correspondence can be defined for representations which are realized over E_0.

Let λ be a nontrivial, continuous, complex, unitary character of the additive group of the field F. We shall use $\mathbb{Q}(\lambda)$ to denote the field obtained by adjoining all of the character values of λ to $\mathbb{Q}$, and set $E = \mathbb{Q}(\lambda)(\sqrt{-1})$. We observe that E is an algebraic extension of $\mathbb{Q}$. Indeed, if F has characteristic 0, E is the field obtained from $\mathbb{Q}$ by adjoining $\sqrt{-1}$ and all p-power roots of unity. On the other hand, if $\operatorname{char} F = p$ then E is the number field obtained by adjoining a primitive $4p$-th root of unity to $\mathbb{Q}$.

MSC2000: 11F70, 22E50.
Keywords: Weil representation, local fields.

Let $F^{2n} = X \oplus Y$ be a decomposition of F^{2n} as a direct sum of totally isotropic F-subspaces, with respect to the alternating form on F^{2n} used to define the symplectic group $Sp(2n, F)$. Ranga Rao [1993] provided an explicit realization of the Weil representation W_λ of $\mathrm{Sp}(2n, F)$ associated with λ as integral operators acting on the Bruhat-Schwartz space $\mathcal{S}(X)$ of complex valued, locally constant functions on X of compact support. For a subfield L of $\mathbb{C}$, define $\mathcal{S}(X, L)$ to be the space of locally constant functions on X of compact support having values in L. Observing that the Haar measure $\mu_{\lambda,g}$ used to define the operators $W_\lambda(g)$ is $\mathbb{Q}(\sqrt{q})$-rational, we are able to show that the space $\mathcal{S}(X, E)$ is invariant under the Weil representation W_λ, hence provides a realization of the Weil representation over the algebraic extension E. In particular, this provides an affirmative answer to Prasad's question in the case char $F = p$.

The latter half of the paper is devoted to the construction a 1-cocycle δ on $\mathrm{Gal}\big(E/E_0\big)$ with values in $\mathrm{GL}\big(\mathcal{S}(X, E)\big)$ such that

$$ {}^{\sigma}W_\lambda(g) = \delta(\sigma)^{-1} W_\lambda(g)\delta(\sigma), \quad g \in \mathrm{Sp}(V). \tag{I}$$

Using Galois descent, we show that there exists $\alpha \in \mathrm{GL}(\mathcal{S}(X, E))$ such that $\delta(\sigma) = \alpha^{-1}\, {}^{\sigma}\alpha$ for $\sigma \in \mathrm{Gal}\big(E/E_0\big)$.

Main theorem. *The operators $\alpha W_\lambda(g)\alpha^{-1}$ leave $\mathcal{S}(X, E_0)$ invariant, and provide a form of the Weil representation realized over E_0.*

We should remark that we fail to provide an explicit description of the operator α. As such, the problem of finding an explicit realization of the Weil representation over E_0 remains open.

To indicate how we find the 1-cocycle satisfying (I), for the rest of the introduction we assume that F has characteristic 0. The Galois group of $\mathbb{Q}(\lambda)/\mathbb{Q}$ is isomorphic to the units $\mathbb{Z}_p^*$ of the p-adic integers. For an element s of $\mathbb{Z}_p^*$, we let σ_s denote the corresponding element of $\mathrm{Gal}(\mathbb{Q}(\lambda)/\mathbb{Q})$. For an element $t \in F^*$, we define the character $\lambda[t]$ of F by $\lambda[t](r) = \lambda(tr)$, $r \in F$.

For $t \in F^*$, let $g_t \in \mathrm{Sp}(2n, F)$ and $f_t \in \mathrm{GL}(2n, F)$ be defined by

$$(x + y)g_t = t^{-1}x + ty,$$
$$(x + y)f_t = x + ty,$$

where $x \in X$, $y \in Y$. In general, f_t is not in $\mathrm{Sp}(2n, F)$, but conjugation by f_t leaves $\mathrm{Sp}(2n, F)$ invariant. We have

$$W_\lambda(g^{f_t}) = W_{\lambda[t]}(g), \quad g \in \mathrm{Sp}(V). \tag{II}$$

Furthermore, observing f_{t^2} is the composite $tI \circ g_t$, we show

$$W_\lambda(g^{f_{t^2}}) = W_\lambda(g_t)^{-1} W_\lambda(g) W_\lambda(g_t). \tag{III}$$

On the other hand, restriction to $\mathbb{Q}(\lambda)$ identifies $\mathrm{Gal}(E/E_0)$ with $(F^*)^2 \cap \mathbb{Z}_p^*$. If $\sigma \in \mathrm{Gal}(E/E_0)$, we can write

$$\sigma|_{\mathbb{Q}(\lambda)} = \sigma_{t^2}$$

for some $t \in F^*$. We note

$$\text{(IV)} \qquad {}^{\sigma}W_\lambda(g) = W_{\lambda[t^2]}(g).$$

In light of (II) and (III), we deduce the fundamental identity

$${}^{\sigma}W_\lambda(g) = W_\lambda(g_t)^{-1} W_\lambda(g) W_\lambda(g_t).$$

The last equation is used to show that $\delta(\sigma) = W_\lambda(g_t)$ satisfies (I) and almost satisfies the one-cocycle condition. An actual one-cocycle is obtained by slightly modifying the operators $W_\lambda(g_t)$.

The paper concludes with an investigation of the optimality of the field E_0. Our main tool is the K-types associated with the compact subgroup $\mathrm{Sp}(\mathscr{L})$ of elements preserving a lattice $\mathscr{L}$ on which the symplectic form $\langle\ ,\ \rangle$ is nondegenerate. If $q \equiv 1 \mod 4$ then it is impossible to realize the K-types in a proper subfield of E_0, which allows us to deduce that E_0 is optimal for realizing W_λ. If $q \equiv 3 \mod 4$, the K-types can be realized over the proper subfield $\mathbb{Q}(\sqrt{-p})$ of E_0. In this case, the possibility of realizing the Weil representation over the smaller field is left open.

2. Preliminary remarks on local fields, characters and measures

We fix some notation and recall some elementary facts about the characters of the additive group of a local field. Further details can be found in the first two chapters of [Weil 1974].

Let F be a non-Archimedean local field, $\mathbb{O}$ its ring of integers, and $\mathfrak{m}$ the maximal ideal of $\mathbb{O}$. The order of the residue class field $\kappa = \mathbb{O}/\mathfrak{m}$ shall be denoted q; we note that q is power of $p = \mathrm{char}\,\kappa$. We assume throughout that p is different from 2; in particular, 2 is a unit of $\mathbb{O}$.

Given a fractional $\mathbb{O}$-ideal $\mathfrak{a}$, there exists an unique integer $v(\mathfrak{a})$, the *valuation of* $\mathfrak{a}$, such that

$$\mathfrak{a} = \mathfrak{m}^{v(\mathfrak{a})}.$$

If $s \in F$ is nonzero, the valuation of the ideal $s\mathbb{O}$ is referred to as the valuation of s, denoted $v(s)$. The absolute value on F is related to the valuation v on F by

$$|s| = q^{-v(s)}, \quad s \in F, s \neq 0.$$

Let λ be a nontrivial, continuous, unitary, complex linear character of F^+. The continuity of λ ensures that its kernel contains a fractional $\mathbb{O}$-ideal. The fact that λ is nontrivial allows one to deduce that the set of all such fractional $\mathbb{O}$-ideals has a unique maximal element $\mathfrak{i} = \mathfrak{i}_\lambda$, *the conductor of* λ. The *level of* λ is defined to be the valuation of $\mathfrak{i}_\lambda$.

Given $n \geq 1$, let

$$\nu_{p^n} = \{ z \in \mathbb{C} : z^{p^n} = 1 \}, \quad \nu_{p^\infty} = \bigcup_{n=1}^{\infty} \nu_{p^n}.$$

(The more customary symbol μ will be used to denote a measure.)

Lemma 1. *We have*

$$\operatorname{im} \lambda = \begin{cases} \nu_p & \text{if } \operatorname{char} F = p, \\ \nu_{p^\infty} & \text{if } \operatorname{char} F = 0. \end{cases}$$

Proof. Take $x \in F$. If $\operatorname{char} F = p$ then

$$1 = \lambda(0) = \lambda(px) = \lambda(x)^p.$$

This shows $\operatorname{im} \lambda \subseteq \nu_p$. Equality follows from the fact $\operatorname{im} \lambda$ is a nontrivial subgroup of the simple abelian group ν_p.

If $\operatorname{char} F = 0$ then, since $p \in \mathfrak{m}$, there exists an $n \geq 0$ such that $p^n x \in \mathfrak{i}_\lambda$. For such n,

$$1 = \lambda(p^n x) = \lambda(x)^{p^n}.$$

Then $\operatorname{im} \lambda \subseteq \nu_{p^\infty}$. If the inclusion were proper then there would exist $m \geq 0$ such that $\operatorname{im} \lambda = \nu_{p^m}$. In this case, if $x \in F$ then

$$\lambda(x) = \lambda\left(p^m \cdot \frac{x}{p^m}\right) = \lambda\left(\frac{x}{p^m}\right)^{p^m} = 1$$

since $\lambda(x/p^m)$ is a p^m-th root of unity. As this would contradict the nontriviality of λ, $\operatorname{im} \lambda = \nu_{p^\infty}$. □

Define $\mathbb{Q}(\lambda)$ to be the field obtained by adjoining to $\mathbb{Q}$ all the character values $\lambda(x)$, $x \in F$. Define

$$\mathcal{P} \simeq \begin{cases} \mathbb{Z}/p\mathbb{Z} & \text{if } \operatorname{char} F = p, \\ \mathbb{Z}_p & \text{if } \operatorname{char} F = 0. \end{cases}$$

Note that $\mathcal{P}$ is the topological closure of the prime ring of F.

Lemma 2. *There is a canonical topological isomorphism*

$$\operatorname{Gal}(\mathbb{Q}(\lambda)/\mathbb{Q}) \simeq \mathcal{P}^*.$$

Proof. The preceding lemma ensures that $\operatorname{im}\lambda$ is invariant under the action of Galois, hence restriction yields a homomorphism

$$\operatorname{Gal}(\mathbb{Q}(\lambda)/\mathbb{Q}) \to \operatorname{Aut}(\operatorname{im}\lambda) \simeq \begin{cases} (\mathbb{Z}/p\mathbb{Z})^* & \text{if char } F = p, \\ \mathbb{Z}_p^* & \text{if char } F = 0. \end{cases}$$

It is readily checked that this map is an isomorphism of topological groups. The proof is completed by appealing to the description of $\mathcal{P}$ given above. □

The pairing

$$(s, t) \to \lambda(st), \qquad s, t \in F,$$

is nondegenerate and leads to an identification of F^+ with its Pontryagin dual [Weil 1974, II.5]. The image of $s \in F$ in the dual shall be denoted $\lambda[s]$:

$$\lambda[s](t) = \lambda(st), \qquad t \in F.$$

Let $\mu = dt$ be a Haar measure on F^+. If ϕ is a locally constant, complex valued function on F of compact support, the Fourier transform $\mathcal{F}_\lambda\phi$ is the complex valued function on F defined by

$$\mathcal{F}_\lambda\phi(s) = \int_F \lambda[s](t)\phi(t)\,dt, \qquad s \in F.$$

It can be shown that $\mathcal{F}_\lambda\phi$ is locally constant and has compact support. Furthermore, the general theory of Fourier transforms asserts the existence of a positive constant c, depending only on the Haar measure dt, such that

$$(\mathcal{F}_\lambda\mathcal{F}_\lambda\phi)\,(t) = c\phi(-t), \qquad t \in F.$$

There is a unique Haar measure on F^+ for which $c = 1$; it shall be denoted $d_\lambda t$ and will be referred to as the *self-dual Haar measure associated with* λ [Weil 1974, VII.2].

Lemma 3. *If λ has level l then the associated self-dual Haar measure is characterized by the condition*

$$\int_{\mathcal{O}} d_\lambda t = q^{l/2}. \tag{1}$$

Proof. This follows from [Weil 1974, Corollary 3, VII.2]. □

Corollary. *If $s \in F^*$ then*

$$d_{\lambda[s]}t = |s|^{1/2} d_\lambda t.$$

Proof. Since $\mathfrak{i}_\lambda = s\mathfrak{i}_{\lambda[s]}$, the levels l_1 of λ and l_2 of $\lambda[s]$ satisfy the relation $l_1 = v(s) + l_2$. Therefore, Lemma 3 yields

$$\int_{\mathcal{O}} d_{\lambda[s]}t = q^{l_2/2} = q^{-v(s)/2} q^{l_1/2} = |s|^{1/2} \int_{\mathcal{O}} d_\lambda t.$$ □

3. The Schrödinger and Weil representations

Let $\langle\ ,\ \rangle$ be a nondegenerate, alternating, F-bilinear form on a finite dimensional F-vector space V. The *Heisenberg group* H is the group on $V \times F$ having multiplication

$$(v, t)(v', t') = \big(v + v', t + t' + \langle v, v'\rangle/2\big), \qquad t,\ t' \in F,\ v,\ v' \in V.$$

Let λ be a nontrivial, continuous, unitary, complex linear character of F^+. Since $Z(H) = 0 \times F \simeq F^+$, it may be viewed as a character of the center of the Heisenberg group H.

Theorem (Stone, von Neumann). *There exists a smooth, irreducible representation of H having central character* λ. *Such a representation is necessarily admissible, and is unique up to isomorphism.*

A proof of the Stone-von Neumann Theorem can be found in [Mœglin et al. 1987, 2.I]. The representation provided by the Stone-von Neumann Theorem is referred to as *the Schrödinger representation of type* λ.

The symplectic group

$$\mathrm{Sp}(V) = \{g \in \mathrm{GL}(V) : \langle vg, wg\rangle = \langle v, w\rangle, v, w \in V\}$$

acts on the Heisenberg group H as a group of automorphisms as follows: if $g \in \mathrm{Sp}(V)$ and $(t, v) \in H$ then

$$(t, v)g = (t, vg).$$

Given a Schrödinger representation S_λ of type λ and $g \in \mathrm{Sp}(V)$, consider the representation S_λ^g of H defined by

$$S_\lambda^g(h) = S_\lambda(hg), \quad h \in H.$$

It is readily verified that S_λ^g is a smooth, irreducible representation of H. Furthermore, observing that g acts trivially on $Z(H)$, S_λ^g has central character λ. The Stone-von Neumann Theorem allows us to conclude that the representation S_λ and S_λ^g are equivalent, hence the ambient space affording S_λ admits an operator $W_\lambda(g)$ for which

$$S_\lambda^g(h) = W_\lambda(g)^{-1} S_\lambda(h) W_\lambda(g), \quad h \in H.$$

In light of Schur's Lemma, the operator $W_\lambda(g)$ is uniquely defined up to multiplication by a nonzero constant. As a result, the map

$$g \mapsto W_\lambda(g), \quad g \in \mathrm{Sp}(V),$$

is a projective representation of $\mathrm{Sp}(V)$, called a *Weil representation of type* λ.

In this paper we consider the Schrödinger models of S_λ and W_λ [Kudla 1996, Lemma 2.2, Proposition 2.3; Mœglin et al. 1987, 2.I.4(a) and 2.II.6; Ranga Rao 1993, §3]. Let

$$V = X + Y$$

where X and Y are maximal, totally isotropic subspaces. The Schrödinger model is realized in the Bruhat-Schwartz space $\mathcal{S}(X)$ of locally constant functions $f : X \to \mathbb{C}$ of compact support: if $x \in X$, $y \in Y$ and $t \in F$ then $S_\lambda((x + y, t))$ is the operator defined by

$$[S_\lambda((x+y,t))\phi]\,(x') = \lambda\left(t + \frac{\langle x, y\rangle}{2} + \langle x', y\rangle\right)\phi(x+x'), \quad \phi \in \mathcal{S}(X), x' \in X.$$

The description of the Weil representation requires some additional notation. Viewing $x + y \in V$ as a row vector (x, y), each $g \in \mathrm{Sp}(V)$ can be expressed in the matrix form

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2}$$

where $a : X \to X$, $b : X \to Y$, $c : Y \to X$, and $d : Y \to Y$. With this notation, set

$$Y_g = Y/\ker c.$$

If μ_g is a Haar measure on Y_g then the action of $W_\lambda(g)$ on $\mathcal{S}(X)$ is given by

$$[W_\lambda(g)\phi]\,(x) = \int_{Y_g} \lambda\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa+yc)\,d\mu_g y, \tag{3}$$

for $\phi \in \mathcal{S}(X)$ and $x \in X$. Note that the integral appearing in (3) is well-defined, for the integrand is constant on the cosets of $\ker c$, hence can be viewed as a function on Y_g. The fact $\phi \in \mathcal{S}(X)$ can be used to show that the integrand belongs to $\mathcal{S}(Y_g)$, hence the integral converges, and that the resulting function $W_\lambda(g)\phi$ belongs to $\mathcal{S}(X)$.

We now recall a particular choice of Haar measures $\mu_{\lambda,g}$ on Y_g, $g \in \mathrm{Sp}(V)$ [Ranga Rao 1993, §3.3]. Fix a basis $x_1, \ldots, x_n$ of X and let $y_1, \ldots, y_n$ be the dual basis of Y defined by the conditions

$$\langle x_i, y_j\rangle = \delta_{ij}, \quad 1 \le i, j \le n.$$

Let τ_i, $0 \le i \le n$, be the element of $\mathrm{Sp}(V)$ defined by

$$x_j\tau_i = \begin{cases} -y_j & \text{if } j \le i, \\ x_j & \text{if } i < j, \end{cases}$$

$$y_j\tau_i = \begin{cases} x_j & \text{if } j \le i, \\ y_j & \text{if } i < j. \end{cases}$$

We note that Y_{τ_i} can be identified with the subspace of Y spanned by the elements $y_1, \dots, y_i$. We define

$$d\mu_{\lambda,\tau_i} y = \prod_{k=1}^{i} d_\lambda y_k, \tag{4}$$

where $d_\lambda y_k$ is the self-dual Haar measure associated with λ.

Let

$$P = \{g \in \mathrm{Sp}(V) : Yg = g\},$$

the parabolic subgroup that leaves Y invariant. If $\dim Y_g = i$ then [Ranga Rao 1993, Theorem 2.14] ensures the existence of elements p_1 and p_2 of P such that

$$g = p_1 \tau_i p_2.$$

Observing that the operator p_1 induces an isomorphism $\overline{p}_1 : Y_g \to Y_{\tau_i}$, we set

$$\mu_{\lambda,g} = |\det(p_1 p_2|_Y)|^{-1/2}\, \overline{p}_1 \cdot \mu_{\lambda,\tau_i}. \tag{5}$$

Here, $\overline{p}_1 \cdot \mu_{\lambda,\tau_i}$ denotes the pullback of the Haar measure μ_{λ,τ_i} to Y_g via $\overline{p}_1$: if E is a measurable subset of Y_g then

$$\overline{p}_1 \cdot \mu_{\lambda,\tau_i}(O) = \mu_{\lambda,\tau_i}(O\overline{p}_1).$$

Theorem 4. *The measures $\mu_{\lambda,g}$, $g \in Sp(V)$, are well-defined. The projective representation W_λ of* $\mathrm{Sp}(V)$ *defined by* (3) *with the Haar measures $\mu_g = \mu_{\lambda,g}$ has the following properties.*

(i) *If $g \in \mathrm{Sp}(V)$ and $p_1,\ p_2 \in P$ then $W_\lambda(p_1 g p_2) = W_\lambda(p_1)W_\lambda(g)W_\lambda(p_2)$; in particular W_λ restricts to an ordinary representation of P.*

(ii) *If $\phi \in \mathcal{S}(X)$ and $p = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in P$ then*

$$[W_\lambda(p)\phi](x) = |\det a|^{1/2} \lambda\left(\frac{\langle xa, xb\rangle}{2}\right)\phi(xa), \quad x \in X.$$

Proof. This follows from [Ranga Rao 1993, Theorem 3.5]. □

Lemma 5. *If $s \in F^*$ and $g \in \mathrm{Sp}(V)$ then $\mu_{\lambda[s],g} = |s|_{Y_g}^{1/2} \mu_{\lambda,g}$.*

Proof. In light of the Corollary to Lemma 3, (4) yields

$$d\mu_{\lambda[s],\tau_i} y = \prod_{k=1}^{i} d_{\lambda[s]} y_k = \prod_{k=1}^{i} \left[|s|^{1/2} d_\lambda y_k\right] = |s|^{i/2} \prod_{k=1}^{i} d_\lambda y_k = |s|^{i/2} d\mu_{\lambda,\tau_i} y.$$

Therefore, we obtain from (5) and the fact that Y_g has dimension i over F that

$$\begin{aligned}\mu_{\lambda[s],g} &= |\det(p_1 p_2|_Y)|^{-1/2}\ \overline{p}_1 \cdot \mu_{\lambda[s],\tau_i} = |s|^{i/2}\, |\det(p_1 p_2|_Y)|^{-1/2}\ \overline{p}_1 \cdot \mu_{\lambda,\tau_i} \\ &= |s|^{i/2} \mu_{\lambda,g} = |s|_{Y_g}^{1/2} \mu_{\lambda,g}\end{aligned}$$

. □

Let μ be a Haar measure on a totally disconnected topological group A. If O_1 and O_2 are nonempty compact open sets in A then the ratio

$$(O_1 : O_2) = \frac{\mu(O_1)}{\mu(O_2)}$$

is a rational number [Cartier 1979, I.1.1]. Hence, if $\mu(O)$ lies in a subfield L of $\mathbb{C}$ for some nonempty compact open set O then the same is true for all nonempty compact open sets. The measure μ is said to *L-rational* if this is the case.

Lemma 6. *The measures $\mu_{\lambda,g}$, $g \in \mathrm{Sp}(V)$, are $\mathbb{Q}(\sqrt{q})$-rational.*

Proof. If $t \in F^*$ then $|t|$ is a power of q. Therefore, (5) shows that it is sufficient to verify that the measures μ_{λ,τ_i} are $\mathbb{Q}(\sqrt{q})$-rational. Formulas (1) and (4) ensure that this is indeed the case: if $\mathcal{Y}_i = \sum_{k=1}^{i} \mathcal{O} y_k$ then

$$\int_{\mathcal{Y}_i} d\mu_{\lambda,\tau_i} y = q^{il/2}. \qquad \square$$

If L is a subfield of $\mathbb{C}$, let $\mathcal{S}(A, L)$ denote the space of locally constant, L-valued functions on A of compact support.

Lemma 7. *Let A be a totally disconnected topological group, $L \subseteq K$ an extension of fields, and μ a L-rational Haar measure on A. If $\phi \in \mathcal{S}(A, K)$ then $\int_A \phi \, d\mu$ belongs to K.*

Proof. Since $\phi \in \mathcal{S}(A, K)$, there exists compact open subsets $A_1, \dots, A_k$ of A and scalars $c_1, \dots, c_k$ in K such that

$$\phi = \sum_{i=1}^{k} c_i \chi_{A_i}.$$

Here, χ_{A_i} denotes the characteristic function of A_i. Since $\mu(A_i) \in L \subseteq K$, it follows that

$$\int_A \phi \, d\mu = \sum_{i=1}^{k} c_i \mu(A_i)$$

lies in K. $\square$

Let $\mathbb{Q}(\lambda)$ be the character field of λ and set

$$E = \mathbb{Q}(\lambda)(\sqrt{-1}).$$

Observe that Lemma 1 ensures that $\mathbb{Q}(\sqrt{q})$ is a subfield of E.

Proposition 8. *The operators $W_\lambda(g)$, $g \in \mathrm{Sp}(V)$, leave the subspace $\mathcal{S}(X, E)$ invariant.*

Proof. If $\phi \in \mathcal{S}(X, E)$ then the integrand in (3) lies in $\mathcal{S}(Y_g, E)$, since $\mathbb{Q}(\lambda) \subseteq E$. In light of Lemma 6, Lemma 7 applied in the case $A = Y_g$, $K = E$, $L = \mathbb{Q}(\sqrt{q})$, and $\mu = \mu_{\lambda,g}$ allows us to deduce that the integral (3) lies in E. It follows immediately that $W_\lambda(g)\phi \in \mathcal{S}(X, E)$. □

In particular, if F has odd characteristic p, the preceding result allows one to conclude that the Weil representation W_λ can be realized over the number field $\mathbb{Q}(\nu_{4p})$.

4. Galois action

By Lemma 1, E is a Galois extension of $\mathbb{Q}$. Its Galois group acts on $\mathcal{S}(X, E)$: if $\sigma \in \mathrm{Gal}(E/\mathbb{Q})$ and $\phi \in \mathcal{S}(X, E)$ then

$$(\sigma(\phi))(x) = \sigma(\phi(x)), \quad x \in X. \tag{6}$$

There is an associated Galois action on $\mathrm{End}\,\mathcal{S}(X, E)$: if $\sigma \in G$ and $T \in \mathrm{End}\,\mathcal{S}(X, E)$ then

$${}^{\sigma}T(\phi) = \sigma\big[T\left(\sigma^{-1}(\phi)\right)\big], \quad \phi \in \mathcal{S}(X, E). \tag{7}$$

The Galois group also permutes the unitary characters of F^+: if $\sigma \in \mathrm{Gal}(E/\mathbb{Q})$ and λ is a unitary character of F^+ then ${}^{\sigma}\lambda$ is the character defined by

$${}^{\sigma}\lambda(t) = \sigma(\lambda(t)), \quad t \in F^+.$$

Let $\mathcal{P}$ be the topological closure of the prime ring of F. The image of $s \in \mathcal{P}^*$ in $\mathrm{Gal}(\mathbb{Q}(\lambda)/\mathbb{Q})$ under the canonical isomorphism of Lemma 2 will be denoted σ_s.

Lemma 9. *Let $\sigma \in \mathrm{Gal}(E/\mathbb{Q})$. If $\sigma|_{\mathbb{Q}(\lambda)} = \sigma_s$ then ${}^{\sigma}\lambda = \lambda[s]$.*

Proof. (char $F = 0$) Let $\mathfrak{i}$ be the conductor of λ. Given $t \in F$, fix $n \geq 1$ such that $t \in p^{-n}\mathfrak{i}$. Since $p^n t \in \mathfrak{i}$,

$$1 = \lambda(p^n t) = \lambda(t)^{p^n},$$

thus $\lambda(t) \in \nu_{p^n}$. Fixing $r \in \mathbb{Z}$ such that $s \equiv r \mod p^n\mathcal{P}$,

$$({}^{\sigma}\lambda)(t) = \sigma(\lambda(t)) = \lambda(t)^r = \lambda(rt) = \lambda(st),$$

the last equality following from the fact $rt \equiv st \mod \mathfrak{i}$. □

Given $\sigma \in \mathrm{Gal}(E/\mathbb{Q})$, let ${}^{\sigma}W_\lambda$ be the projective representation defined by

$$({}^{\sigma}W_\lambda)(g) = {}^{\sigma}(W_\lambda(g)), \quad g \in \mathrm{Sp}(V).$$

Proposition 10. *Let $\sigma \in \mathrm{Gal}(E/\mathbb{Q}(\sqrt{q}))$. If $\sigma|_{\mathbb{Q}(\lambda)} = \sigma_s$ then ${}^{\sigma}W_\lambda(g) = W_{\lambda[s]}(g)$.*

The proof of Proposition 10 is based on the integral formula (3) and the following result:

Lemma 11. *Let A be a totally disconnected topological group, $L \subseteq K$ an extension of fields, and μ a L-rational Haar measure on A. If σ is an L-automorphism of K then, for all $\phi \in \mathcal{S}(A, K)$,*

$$\int_A \sigma(\phi)\, d\mu = \sigma\left(\int_A \phi\, d\mu\right).$$

Proof. Using the notation introduced in the proof of Lemma 7, if $\phi = \sum_{i=1}^{k} c_i \chi_{A_i}$ then

$$\sigma(\phi) = \sum_{i=1}^{k} \sigma(c_i)\chi_{A_i}.$$

Therefore, since $\mu(A_i) \in L$ is fixed by σ,

$$\begin{aligned}\int \sigma(\phi)\, d\mu &= \sum_{i=1}^{k} \sigma(c_i)\mu(A_i) = \sum_{i=1}^{k} \sigma(c_i)\sigma(\mu(A_i)) \\ &= \sigma\left(\sum_{i=1}^{k} c_i \mu(A_i)\right) = \sigma\left(\int_A \phi\, d\mu\right).\end{aligned}$$

□

Proof of Proposition 10. Let $g \in \mathrm{Sp}(V)$, $\phi \in \mathcal{S}(X, E)$, and $x \in X$. We assume g has the matrix representation (2). Lemma 6 asserts that the measure $\mu_{\lambda,g}$ is $\mathbb{Q}(\sqrt{q})$-rational. Applying Lemma 11 to the case $A = Y_g$, $L = \mathbb{Q}(\sqrt{q})$, $K = E$, and $\mu = \mu_{\lambda,g}$, the definition of ${}^{\sigma}W_\lambda$, the formula (3), and Lemma 9 yield

$$\begin{aligned}&\left[{}^{\sigma}W_\lambda(g)\phi\right](x) \\ &\quad= \sigma\left[W_\lambda(g)(\sigma^{-1}\phi)(x)\right] \\ &\quad= \sigma\left[\int_{Y_g} \lambda\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)(\sigma^{-1}\phi)(xa+yc)\, d\mu_{\lambda,g}y\right] \\ &\quad= \int_{Y_g} {}^{\sigma}\lambda\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa+yc)\, d\mu_{\lambda,g}y \\ &\quad= \int_{Y_g} \lambda[s]\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa+yc)\, d\mu_{\lambda,g}y.\end{aligned}$$

Observing $s \in \mathcal{P}^* \subseteq \mathcal{O}^*$, Lemma 5 implies that $\mu_{\lambda[s],g} = \mu_{\lambda,g}$. The preceding calculation thus gives

$$\begin{aligned}&\left[{}^{\sigma}W_\lambda(g)\phi\right](x) \\ &\quad= \int_{Y_g}\left[\lambda[s]\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa+yc)\right] d\mu_{\lambda[s],g}y \\ &\quad= \left[W_{\lambda[s]}(g)\phi\right](x).\end{aligned}$$

□

5. Action of symplectic similitudes

In the previous section, we described the action of Galois on the projective representations W_λ. Here, we discuss an action of the group of symplectic similitudes on the Weil representations.

Given $s \in F^*$, let f_s be the element of $\mathrm{GL}(V)$ defined by

$$(x+y)f_s = x + sy, \quad x \in X,\, y \in Y.$$

Conjugation by f_s leaves the symplectic group $\mathrm{Sp}(V)$ invariant. In fact, if $g \in \mathrm{Sp}(V)$ is expressed in the matrix form (2) then

$$g^{f_s} = \begin{pmatrix} a & sb \\ s^{-1}c & d \end{pmatrix}. \tag{8}$$

In particular, we note that the spaces Y_g and $Y_{g^{f_s}}$ are equal, since $\ker c = \ker s^{-1}c$.

Lemma 12. *If $s \in F^*$ then $\mu_{\lambda, g^{f_s}} = |s|_{Y_g}^{-1/2} \mu_{\lambda,g}$.*

Proof. Let $p_{i,s}$, $0 \le i \le n$, be the elements of $\mathrm{Sp}(V)$ defined by

$$x_j p_{i,s} = \begin{cases} s^{-1}x_j & \text{if } j \le i, \\ x_j & \text{if } i < j, \end{cases}$$
$$y_j p_{i,s} = \begin{cases} sy_j & \text{if } j \le i, \\ y_j & \text{if } i < j. \end{cases}$$

Note that $p_{i,s} \in P$ and

$$\det(p_{i,s}|_Y) = s^i.$$

Moreover, one readily verifies that

$$\tau_i^{f_s} = \tau_i p_{i,s}.$$

Let $g \in G$. If $g = p_1 \tau_i p_2$, p_1, $p_2 \in P$, then

$$g^{f_s} = (p_1\tau_i p_2)^{f_s} = p_1^{f_s}\tau_i^{f_s} p_2^{f_s} = p_1^{f_s}\tau_i(p_{i,s}p_2^{f_s}).$$

Observing that both $p_1^{f_s}$ and $p_{i,s}p_2^{f_s}$ belong to P, (5) yields

$$\mu_{\lambda, g^{f_s}} = |\det(p_1^{f_s} p_{i,s} p_2^{f_s}|_Y)|^{-1/2}\, \overline{p_1^{f_s}} \cdot \mu_{\lambda,\tau_i}.$$

Using (8), if $p \in P$ then $p^{f_s}|_Y = p|_Y$. As a consequence,

$$\overline{p_1^{f_s}} = \overline{p}_1 : Y_g \to Y_{\tau_i}.$$

In light of these observations,

$$\det(p_1^{f_s} p_{i,s} p_2^{f_s}|_Y) = \det(p_1 p_{i,s} p_2|_Y) = \det(p_{i,s}|_Y)\cdot \det(p_1p_2|_Y) = s^i \det(p_1p_2|_Y);$$

hence

$$\mu_{\lambda,g^{f_s}} = |s^i \det(p_1 p_2|_Y)|^{-1/2}\overline{p}_1 \cdot \mu_{\lambda,\tau_i} = |s|^{-i/2}\mu_{\lambda,g} = |s|_{Y_g}^{-1/2}\mu_{\lambda,g},$$

since Y_g has dimension i over F. □

Let $W_\lambda^{f_s}$ be the projective representation of $\mathrm{Sp}(V)$ defined by

$$W_\lambda^{f_s}(g) = W_\lambda\left(g^{f_s}\right).$$

For the proof of the next result, let $|\alpha|_V$ denote the module of an automorphism α of an F-vector space V [Weil 1974, I.2]. We have

$$|\alpha|_V = |\det \alpha|.$$

In particular, the module of left multiplication by $s \in F^*$ on V satisfies

$$|s|_V = |s|^{\dim V}.$$

Proposition 13. *If $s \in F^*$ then $W_\lambda^{f_s} = W_{\lambda[s]}$.*

Proof. Let $g \in \mathrm{Sp}(V)$. We assume that g has the matrix representation (2), hence that of g^{f_s} is given by (8). If $\phi \in \mathcal{S}(X)$ and $x \in X$ then the integral formula (3) and Lemma 12 yield

$$\begin{aligned}
&\left[W_\lambda(g^{f_s})\phi\right](x)\\
&= \int_{Y_{g^{f_s}}} \lambda\left(\frac{\langle xa, sxb\rangle - 2\langle sxb, s^{-1}yc\rangle + \langle s^{-1}yc, yd\rangle}{2}\right)\phi(xa + s^{-1}yc)\, d\mu_{\lambda,g^{f_s}}\, y\\
&= |s|_{Y_g}^{-1/2}\int_{Y_g} \lambda\left(\frac{\langle xa, sxb\rangle - 2\langle sxb, s^{-1}yc\rangle + \langle s^{-1}yc, yd\rangle}{2}\right)\phi(xa+s^{-1}yc)\, d\mu_{\lambda,g}\, y.
\end{aligned}$$

Replacing y by sy, the definition of $|s|_{Y_g}$ and Lemma 5 yield

$$\begin{aligned}
&\left[W_\lambda(g^{f_s})\phi\right](x)\\
&\quad= |s|_{Y_g}^{-1/2}|s|_{Y_g}\int_{Y_g} \lambda\left(\frac{\langle xa, sxb\rangle - 2\langle sxb, yc\rangle + \langle yc, syd\rangle}{2}\right)\phi(xa + yc)\, d\mu_{\lambda g}\, y\\
&\quad= |s|_{Y_g}^{1/2}\int_{Y_g} \lambda\left(s \cdot \frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa + yc)\, d\mu_{\lambda,g}\, y\\
&\quad= |s|_{Y_g}^{1/2}\int_{Y_g} \lambda[s]\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa + yc)\, d\mu_{\lambda,g}\, y\\
&\quad= \int_{Y_g} \lambda[s]\left(\frac{\langle xa, xb\rangle - 2\langle xb, yc\rangle + \langle yc, yd\rangle}{2}\right)\phi(xa + yc)\, d\mu_{\lambda[s],g}\, y\\
&\quad= [W_{\lambda[s]}(g)\phi](x).
\end{aligned}$$

This completes the proof of the proposition. □

6. The fundamental identity

Let

$$\mathfrak{G} = \{\sigma \in \mathrm{Gal}(E/\mathbb{Q}(\sqrt{q})) : \exists s \in \mathbb{O}^* \text{ such that } \sigma|_{\mathbb{Q}(\lambda)} = \sigma_{s^2}\}.$$

Note that $\mathfrak{G}$ is a subgroup of $\mathrm{Gal}(E/\mathbb{Q}(\sqrt{q}))$. Given $s \in F^*$, let $g_s \in \mathrm{Sp}(V)$ be the map defined by

$$(x+y)g_s = s^{-1}x + sy, \quad x \in X, y \in Y.$$

We observe that g_s lies in the parabolic subgroup P that leaves Y invariant and is related to the operator f_{s^2} defined earlier by the identity

$$f_{s^2} = sI \circ g_s.$$

Proposition 14. *Let $\sigma \in \mathfrak{G}$ and $g \in \mathrm{Sp}(V)$. If $\sigma|_{\mathbb{Q}(\lambda)} = \sigma_{s^2}$, $s \in \mathbb{O}^*$, then*

$$^{\sigma}W_\lambda(g) = W_\lambda(g_s)^{-1} W_\lambda(g) W_\lambda(g_s).$$

Proof. In light of Propositions 10 and 13,

$$^{\sigma}W_\lambda(g) = W_{\lambda[s^2]}(g) = W_\lambda^{f_{s^2}}(g) = W_\lambda(g^{f_{s^2}}) = W_\lambda(g^{g_s}).$$

Applying Theorem 4(i) with $p_1^{-1} = p_2 = g_s$,

$$W_\lambda(g^{g_s}) = W_\lambda(g_s^{-1}) W_\lambda(g) W_\lambda(g_s) = W_\lambda(g_s)^{-1} W_\lambda(g) W_\lambda(g_s).$$

This completes the proof of the proposition. □

Corollary. *If $t \in F^*$ and $\sigma \in \mathfrak{G}$ then $^{\sigma}W_\lambda(g_t) = W_\lambda(g_t)$.*

Proof. Fix $s \in \mathbb{O}^*$ such that $\sigma|_{\mathbb{Q}(\lambda)} = \sigma_{s^2}$. Observing that g_s and g_t are commuting elements of P, the preceding proposition combines with Theorem 4(i) to yield

$$^{\sigma}W_\lambda(g_t) = W_\lambda(g_s)^{-1} W_\lambda(g_t) W_\lambda(g_s) = W_\lambda(g_s^{-1} g_t g_s) = W_\lambda(g_t),$$

as required. □

7. The cocycle

Our aim in this section is the construction of a 1-cocycle δ on

$$\mathfrak{H} = \mathrm{Gal}(E/E_0)$$

with values in $\mathrm{GL}(\mathcal{S}(X, E))$ satisfying the identity (I):

$$^{\sigma}W_\lambda(g) = \delta(\sigma)^{-1} W_\lambda(g) \delta(\sigma), \quad g \in \mathrm{Sp}(V),\ \sigma \in \mathfrak{H}.$$

When combined with restriction to $\mathbb{Q}(\lambda)$, the canonical isomorphism of Lemma 2 yields

$$\mathfrak{H} \simeq \mathrm{Gal}(\mathbb{Q}(\lambda)/E_0 \cap \mathbb{Q}(\lambda))) \simeq (F^*)^2 \cap \mathcal{P}^*. \tag{9}$$

Let

$$o = \begin{cases} 2(p-1) & \text{if } q \text{ is a square,} \\ p-1 & \text{if } q \text{ is not a square,} \end{cases}$$

and fix a primitive o-th root of unity $\epsilon \in F^*$. Furthermore, let

$$U_1 = \begin{cases} \{1\} & \text{if char } F = p, \\ \{r \in \mathcal{P} : r \equiv 1 \bmod p\} & \text{if char } F = 0. \end{cases}$$

Since p is odd, the map $r \mapsto r^2$ is an automorphism of the pro-p group U_1. This allows us to conclude that

$$(F^*)^2 \cap \mathcal{P}^* = \langle \epsilon^2 \rangle \times U_1.$$

The isomorphism (9) identifies U_1 with $\mathrm{Gal}\big(E/\mathbb{Q}(\nu_p, \sqrt{-1})\big)$, where ν_p is the group of complex p-th roots of unity. This in turn leads to an identification of $\langle \epsilon^2 \rangle$ with

$$\mathfrak{H}/\,\mathrm{Gal}\big(E/\mathbb{Q}(\nu_p, \sqrt{-1})\big) \simeq \mathrm{Gal}\big(\mathbb{Q}(\nu_p, \sqrt{-1})/E_0\big).$$

In particular, the element η of $\mathfrak{H}$ characterized by

$$\eta|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^2} \tag{10}$$

has order $o/2$ and restricts to a generator of $\mathrm{Gal}\big(\mathbb{Q}(\nu_p, \sqrt{-1})/E_0\big)$.

Given $\sigma \in \mathfrak{H}$, there is a unique integer i, $1 \le i \le o/2$, and a unique element $s \in U_1$, such that

$$\sigma|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^{2i}s^2}.$$

If τ is a second element of $\mathfrak{H}$, say

$$\tau|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^{2j}t^2}, \quad 1 \le j \le o/2, \quad t \in U_1,$$

then

$$\sigma\tau|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^{2k}(st)^2},$$

where $st \in U_1$ and

$$k = \begin{cases} i+j & \text{if } i+j \le o/2, \\ i+j-o/2 & \text{if } i+j > o/2. \end{cases}$$

Our initial attempt at the construction of the cocycle is to define

$$D(\sigma) = W_\lambda(g_{\epsilon^i s}), \quad \sigma|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^{2i}s^2}, \quad 1 \le i \le o/2, \quad s \in U_1.$$

Proposition 14 ensures that

$${}^\sigma W_\lambda(g) = D(\sigma)^{-1} W_\lambda(g) D(\sigma), \quad g \in \mathrm{Sp}(V), \sigma \in \mathfrak{H}. \tag{11}$$

Assuming σ and τ are as above, the definition of D yields

$$D(\sigma\tau) = W_\lambda(g_{\epsilon^k st}).$$

On the other hand, the Corollary to Proposition 14 gives

$${}^\sigma D(\tau) = {}^\sigma W_\lambda(g_{\epsilon^j t}) = W_\lambda(g_{\epsilon^j t}),$$

hence Theorem 4(i) yields

$$D(\sigma)^\sigma D(\tau) = W_\lambda(g_{\epsilon^i s}) W_\lambda(g_{\epsilon^j t}) = W_\lambda(g_{\epsilon^{i+j} st}).$$

If $i + j \leq o/2$ then

$$W_\lambda(g_{\epsilon^{i+j} st}) = W_\lambda(g_{\epsilon^k st}).$$

If $i + j > o/2$ then, observing $\epsilon^{o/2} = -1$, Theorem 4(i) yields

$$W_\lambda(g_{\epsilon^{i+j} st}) = W_\lambda(g_{-\epsilon^k st}) = W_\lambda(\iota) W_\lambda(g_{\epsilon^k st}),$$

where $\iota = g_{-1}$ is the central involution of $\mathrm{Sp}(V)$ that maps $v \in V$ to $-v$. In summary,

$$D(\sigma)^\sigma D(\tau) = \begin{cases} D(\sigma\tau) & \text{if } i + j \leq o/2, \\ W_\lambda(\iota) D(\sigma\tau) & \text{if } i + j > o/2. \end{cases} \tag{12}$$

In particular, D is not a 1-cocycle; to get one we must account for the factor $W_\lambda(\iota)$.

Since $\iota \in P$, Theorem 4(ii) implies that if ϕ belongs to $\mathcal{S}(X, E)$ then

$$[W_\lambda(\iota)\phi](x) = \phi(-x), \quad x \in X.$$

In particular, $W_\lambda(\iota)$ is an involution, hence the operators

$$\rho_e = \frac{1}{2}(I + W_\lambda(\iota)) \quad \text{and} \quad \rho_o = \frac{1}{2}(I - W_\lambda(\iota))$$

are orthogonal idempotents. Furthermore, recalling $\iota = g_{-1}$, the Corollary to Proposition 14 shows that both ρ_e and ρ_o are fixed by the action of Galois. Finally, since $I = \rho_e + \rho_o$, it is easily verified that the operators

$$\rho_e + c\rho_o, \quad c \in E, c \neq 0,$$

are invertible.

Lemma 15. *The norm equation*

$$N(u) = -1, \quad N : \mathbb{Q}(\nu_p, \sqrt{-1}) \to E_0$$

has a solution.

Proof. The case $p \equiv 1 \bmod 4$ is covered by [Cliff et al. 2004, Lemma 24], an application of the Hasse Norm Theorem. Suppose $p \equiv 3 \bmod 4$. If q is not a square then the extension $\mathbb{Q}(\nu_p, \sqrt{-1})/E_0$ has odd degree $(p-1)/2$, hence -1 is a solution of the norm equation. If q is square then the extension has degree $p-1 \equiv 2 \bmod 4$. In this case, $\sqrt{-1} \in E_0$ is a solution. □

Let u be a solution of the norm equation of the preceding lemma. Given $\sigma \in \mathfrak{H}$, set

$$A(\sigma) = \rho_e + \left(\prod_{l=0}^{i-1} \eta^l(u)\right)\rho_0, \quad \text{where } \sigma|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^{2i}s^2}, \quad 1 \le i \le o/2, \quad s \in U_1$$

where η satisfies (10). The remarks preceding Lemma 15 ensure that $A(\sigma) \in \mathrm{GL}(\mathscr{S}(X, E))$. With the notation introduced earlier, if σ and τ belong to $\mathfrak{H}$ then

$$A(\sigma\tau) = \rho_e + \left(\prod_{l=0}^{k-1} \eta^l(u)\right)\rho_0.$$

On the other hand, observing

$$\sigma\eta^{-i}|_{\mathbb{Q}(\lambda)} = \sigma_{\epsilon^{2i}s^2}\sigma_{\epsilon^2}^{-i} = \sigma_{\epsilon^{2i}s^2}\sigma_{\epsilon^{-2i}} = \sigma_{s^2},$$

the fact (9) identifies U_1 with $\mathrm{Gal}\big(E/\mathbb{Q}(\nu_p, \sqrt{-1})\big)$ allows us to deduce that the restrictions of σ and η^i to $\mathbb{Q}(\nu_p, \sqrt{-1})$ coincide. Therefore,

$$\begin{aligned} {}^{\sigma}A(\tau) &= {}^{\sigma}\!\left[\rho_e + \left(\prod_{l=0}^{j-1} \eta^l(u)\right)\rho_0\right] = \rho_e + {}^{\sigma}\!\left(\prod_{l=0}^{j-1} \eta^l(u)\right)\rho_0 \\ &= \rho_e + {}^{\eta^i}\!\left(\prod_{l=0}^{j-1} \eta^l(u)\right)\rho_0 = \rho_e + \left(\prod_{l=i}^{i+j-1} \eta^l(u)\right)\rho_0; \end{aligned}$$

hence

$$\begin{aligned} A(\sigma)^{\sigma}A(\tau) &= \left[\rho_e + \left(\prod_{l=0}^{i-1} \eta^l(u)\right)\rho_0\right]\left[\rho_e + \left(\prod_{l=i}^{i+j-1} \eta^l(u)\right)\rho_0\right] \\ &= \left[\rho_e + \left(\prod_{l=0}^{i+j-1} \eta^l(u)\right)\rho_0\right]. \end{aligned}$$

If $i + j \le o/2$ then

$$\prod_{l=0}^{i+j-1} \eta^l(u) = \prod_{l=0}^{k-1} \eta^l(u),$$

hence

$$A(\sigma)^{\sigma}A(\tau) = A(\sigma\tau).$$

If $i+j > o/2$ then the choice of η and u yield

$$\prod_{l=0}^{i+j-1}\eta^l(u) = \left(\prod_{l=0}^{(o-2)/2}\eta^l(u)\right)\left(\prod_{l=o/2}^{i+j-1}\eta^l(u)\right) = N(u)\prod_{l=0}^{k-1}\eta^l(u) = -\prod_{l=0}^{k-1}\eta^l(u).$$

Observing that $\rho_e = \rho_e W_\lambda(\iota)$ and $-\rho_o = \rho_o W_\lambda(\iota)$,

$$A(\sigma)^{\sigma}A(\tau) = \rho_e - \left(\prod_{l=0}^{k-1}\eta^l(u)\right)\rho_0 = \left[\rho_e + \left(\prod_{l=0}^{k-1}\eta^l(u)\right)\rho_0\right]W_\lambda(\iota) = A(\sigma\tau)W_\lambda(\iota).$$

In summary,

$$A(\sigma)^{\sigma}A(\tau) = \begin{cases} A(\sigma\tau) & \text{if } i+j \le o/2, \\ A(\sigma\tau)W_\lambda(\iota) & \text{if } i+j > o/2. \end{cases} \tag{13}$$

Consider the map $\delta : \mathfrak{H} \to \mathrm{GL}(\mathscr{S}(X, E))$ given by

$$\delta(\sigma) = A(\sigma)D(\sigma).$$

If $\sigma, \tau \in \mathfrak{H}$ are as above

$$\delta(\sigma)^{\sigma}\delta(\tau) = (A(\sigma)D(\sigma))^{\sigma}(A(\tau)D(\tau)) = A(\sigma)D(\sigma)^{\sigma}A(\tau)^{\sigma}D(\tau).$$

By Theorem 4(ii), ${}^{\sigma}A(\tau) \in E[W_\lambda(i)]$ commutes with $D(\sigma) = W_\lambda(g_{\epsilon^i s})$, hence

$$A(\sigma)D(\sigma)^{\sigma}A(\tau)^{\sigma}D(\tau) = A(\sigma)^{\sigma}A(\tau)D(\sigma)^{\sigma}D(\tau).$$

If $i+j > o/2$ then (12) and (13) yield

$$A(\sigma)^{\sigma}A(\tau)D(\sigma)^{\sigma}D(\tau) = A(\sigma\tau)W_\lambda(\iota)W_\lambda(\iota)D(\sigma\tau) = A(\sigma\tau)D(\sigma\tau).$$

Since this is trivially true if $i+j \le o/2$, we conclude

$$\delta(\sigma)^{\sigma}\delta(\tau) = A(\sigma\tau)D(\sigma\tau) = \delta(\sigma\tau).$$

This shows that δ is a 1-cocycle. Furthermore, if $g \in \mathrm{Sp}(V)$ then Theorem 4(i) shows that $A(\sigma) \in E[W_\lambda(\iota)]$ commutes with $W_\lambda(g)$, hence (11) yields

$$\begin{aligned}\delta(\sigma)^{-1}W_\lambda(g)\delta(\sigma) &= (A(\sigma)D(\sigma))^{-1}W_\lambda(g)A(\sigma)D(\sigma)\\ &= D(\sigma)^{-1}A(\sigma)^{-1}W_\lambda(g)A(\sigma)D(\sigma)\\ &= D(\sigma)^{-1}W_\lambda(g)D(\sigma)\\ &= {}^{\sigma}W_\lambda(g),\end{aligned}$$

which verifies that (I) is satisfied.

8. The triviality of the cocycle

Let $\delta : \mathfrak{H} \to \mathrm{GL}(\mathscr{S}(X, E))$ be the 1-cocycle satisfying (I) constructed above.

Lemma 16. *If $\phi \in \mathscr{S}(X, E)$ then there exists an open subgroup $\mathfrak{K}$ of $\mathfrak{H}$ such that*

$$\delta(\sigma)\phi = \phi, \quad \sigma \in \mathfrak{K}.$$

Proof. If char $F = p$ then $\mathfrak{H}$ is a finite discrete group, so one may take $\mathfrak{K}$ to be the trivial subgroup.

Assume char $F = 0$. If $\mathfrak{X}$ is a lattice in X then the subgroups

$$p^k\mathfrak{X}, \quad k \in \mathbb{Z},$$

form a local base at the origin. Therefore, given $x \in X$, there exist $i_x \in \mathbb{Z}$ such that ϕ is constant on the coset $x + p^{i_x}\mathfrak{X}$. As the family $\left\{x + p^{i_x}\mathfrak{X} : x \in X\right\}$ is an open cover of X, there exists $x_1, \ldots, x_m$ in X such that

$$\operatorname{supp}\phi \subseteq \bigcup_{j=1}^{m} x_j + p^{i_{x_j}}\mathfrak{X}.$$

Set

$$i = \max\left\{i_{x_1}, \ldots, i_{x_m}\right\}$$

and consider $x + p^i\mathfrak{X} \cap \operatorname{supp}\phi$, $x \in X$. If it is empty then the restriction of ϕ to the coset $x + p^i\mathfrak{X}$ is identically 0. If not, there exists j such that $x + p^i\mathfrak{X} \cap x_j + p^{i_{x_j}}\mathfrak{X}$ is nonempty, hence

$$x + p^i\mathfrak{X} \subseteq x_j + p^{i_{x_j}}\mathfrak{X}$$

by choice of i. The choice of i_{x_j} thus ensures that the restriction of ϕ to $x + p^i\mathfrak{X}$ is the constant function with value $\phi(x_j)$. We conclude that ϕ is constant on the $p^i\mathfrak{X}$-cosets of X.

Let $\sigma \in \mathfrak{H}$. If $\sigma|_{\mathbb{Q}(\lambda)} = \sigma_{r^2}$, $r \in U_1$, then by construction $\delta(\sigma) = W_\lambda(g_r)$. Observing

$$g_r = \begin{pmatrix} r^{-1} \cdot 1_X & 0 \\ 0 & r \cdot 1_Y \end{pmatrix} \in P,$$

if $x \in X$ then Theorem 4(i) yields

$$(\delta(\sigma)\phi)(x) = (W_\lambda(g_r)\phi)(x) = |r|^{-\dim X/2}\lambda\left(\frac{\langle r^{-1}x, rx\rangle}{2}\right)\phi(r^{-1}x) = \phi(r^{-1}x),$$

since r is a unit and $\langle\ ,\ \rangle$ is F-bilinear and alternating. Fix $j \in \mathbb{Z}$ such that $i > j$ and

$$\operatorname{supp}\phi \subseteq p^j\mathfrak{X}.$$

If $x \notin p^j\mathfrak{X}$ then neither is $r^{-1}x$, so the choice of j ensures that

$$(\delta(\sigma)\phi)(x) = \phi(r^{-1}x) = 0 = \phi(x).$$

On the other hand, suppose $x \in p^j\mathfrak{X}$. In this case, if $r \equiv 1 \mod p^{i-j}$ then

$$r^{-1}x + p^i\mathfrak{X} = x + p^{i-j}p^j\mathfrak{X} + p^i\mathfrak{X} = x + p^i\mathfrak{X},$$

hence the choice of i ensures that

$$(\delta(\sigma)\phi)(x) = \phi(r^{-1}x) = \phi(x).$$

In light of the preceding discussion,

$$\mathfrak{K} = \{\sigma \in \mathfrak{H} : \sigma|_{\mathbb{Q}(\lambda)} = \sigma_{r^2}, r \equiv 1 \mod p^{i-j}\} = \mathrm{Gal}\big(E/\mathbb{Q}\big(\nu_{p^{i-j}}, \sqrt{-1}\big)\big)$$

has the required properties. □

Let K/k be a Galois extension and M a K-vector space equipped with an semilinear action of the Galois group $\mathrm{Gal}(K/k)$: if $\sigma \in \mathrm{Gal}(K/k)$, $m \in M$ and $e \in K$ then

$$\sigma(em) = \sigma(e)\sigma(m).$$

For such an action, the fixed-point set

$$M^{\mathrm{Gal}(K/k)} = \{m \in M : m = \sigma(m) \text{ for all } \sigma \in \mathrm{Gal}(K/k)\}$$

is a k-vector space. The canonical action of $\mathrm{Gal}(K/k)$ on K yields a semilinear action on the tensor product $K \otimes_k M^{\mathrm{Gal}(K/k)}$:

$$\sigma(e \otimes m) = \sigma(e) \otimes m, \quad \sigma \in \mathrm{Gal}(K/k),\ e \in E,\ m \in M^{\mathrm{Gal}(K/k)}.$$

The action of Galois on M is said to be *smooth* if the stabilizer of each $m \in M$ is open in $\mathrm{Gal}(K/k)$.

Proposition 17. *[Galois Descent] If M is a K-vector space equipped with a semilinear, smooth action of* $\mathrm{Gal}(K/k)$ *then the canonical map*

$$\psi : K \otimes_k M_k \to M$$

is a K-linear isomorphism of $\mathrm{Gal}(K/k)$*-modules.*

Proof. The case $K = k_s$, the separable closure of k, is proved in [Borel 1991, AG.14.2]. The general case is proved using the same argument, *mutatis mutandis*. □

Proposition 18. *There exists $\alpha \in \mathrm{GL}\,(\mathcal{S}(X, E))$ such that*

$$\delta(\sigma) = \alpha^{-1\sigma}\alpha, \quad \sigma \in \mathfrak{H}. \tag{14}$$

Proof. The canonical action (7) of $\mathfrak{H}$ on $\mathcal{S}(X, E)$ is clearly semilinear. It is furthermore smooth, since each element of $\mathcal{S}(X, E)$ takes only finitely many values in E.

On the other hand, since δ is a 1-cocycle, then

$$(\sigma, \phi) \mapsto \delta(\sigma)\sigma(\phi), \quad \sigma \in \mathfrak{H}, \phi \in \mathcal{S}(X, E),$$

is also an action of $\mathfrak{H}$ on $\mathcal{S}(X, E)$, referred to as the twisted action by δ. It is semi-linear, since δ takes values in $\mathrm{GL}\,(\mathcal{S}(X, E))$. Since the original action is smooth, if $\phi \in \mathcal{S}(X, E)$ then there exists an open subgroup $\mathfrak{H}_1$ such that

$$\sigma(\phi) = \phi, \quad \sigma \in \mathfrak{H}_1.$$

Furthermore, Lemma 16 asserts that there is an open subgroup $\mathfrak{K}$ of $\mathfrak{H}$ such that

$$\delta(\sigma)\phi = \phi, \quad \sigma \in \mathfrak{K}.$$

Therefore, if $\sigma \in \mathfrak{H}_1 \cap \mathfrak{K}$ then

$$\delta(\sigma)\sigma(\phi) = \delta(\sigma)\phi = \phi.$$

This shows that the stabilizer of ϕ under the twisted action contains the open subgroup $\mathfrak{H}_1 \cap \mathfrak{K}$. Since it is the union of its $\mathfrak{H}_1 \cap \mathfrak{K}$-cosets, it follows that the stabilizer of ϕ under the twisted action is open. We conclude that the twisted action is smooth.

Using $\mathcal{S}(X, E)$ and ${}_{\delta}\mathcal{S}(X, E)$ to denote the $\mathfrak{H}$-modules defined by the natural and twisted actions, respectively, Galois Descent asserts the existence of E-linear, $\mathfrak{H}$-equivariant isomorphisms

$${}_{\delta}\mathcal{S}(X, E) \simeq E \otimes_{E_0} {}_{\delta}\mathcal{S}(X, E)^{\mathfrak{H}} \quad \text{and} \quad E \otimes_{E_0} \mathcal{S}(X, E)^{\mathfrak{H}} \simeq \mathcal{S}(X, E).$$

In particular,

$$\dim_{E_0} {}_{\delta}\mathcal{S}(X, E)^{\mathfrak{H}} = \dim_E \mathcal{S}(X, E) = \dim_{E_0} \mathcal{S}(X, E)^{\mathfrak{H}},$$

so ${}_{\delta}\mathcal{S}(X, E)^{\mathfrak{H}}$ and $\mathcal{S}(X, E)^{\mathfrak{H}}$ are E_0-isomorphic. As any such isomorphism extends by scalars to a E-linear, $\mathfrak{H}$-equivariant isomorphism

$$E \otimes_{E_0} {}_{\delta}\mathcal{S}(X, E)^{\mathfrak{H}} \simeq E \otimes_{E_0} \mathcal{S}(X, E)^{\mathfrak{H}},$$

we conclude that

$${}_{\delta}\mathcal{S}(X, E) \simeq \mathcal{S}(X, E).$$

Let $\alpha \in \mathrm{GL}\,(\mathcal{S}(X, E))$ be a $\mathfrak{H}$-equivariant isomorphism ${}_{\delta}\mathcal{S}(X, E) \to \mathcal{S}(X, E)$. If $\sigma \in \mathfrak{H}$ and $\phi \in \mathfrak{H}$ then the definition of the twisted action ensures that

$$\alpha\delta(\sigma)\sigma(\phi) = \sigma(\alpha\phi);$$

hence

$$\delta(\sigma)\phi = \alpha^{-1}\alpha\delta(\sigma)\sigma(\sigma^{-1}(\phi)) = \alpha^{-1}\sigma\big(\alpha(\sigma^{-1}(\phi))\big) = \alpha^{-1\sigma}\alpha(\phi). \qquad \square$$

9. Proof of the main theorem

Fix $\alpha \in \mathrm{GL}(\mathcal{S}(X, E))$ satisfying the conclusion of Proposition 18. In light of (9) and (14), if $\sigma \in \mathfrak{H}$ and $g \in \mathrm{Sp}(V)$ then

$$^{\sigma}\left(\alpha W_\lambda(g)\alpha^{-1}\right) = {}^{\sigma}\alpha\, {}^{\sigma}W_\lambda(g)({}^{\sigma}\alpha)^{-1} = {}^{\sigma}\alpha\delta(\sigma)^{-1}W_\lambda(g)\delta(\sigma)({}^{\sigma}\alpha)^{-1} = \alpha W_\lambda(g)\alpha^{-1}.$$

The compatibility of the Galois actions (6) and (7) allows us to deduce that the operators

$$\alpha W_\lambda(g)\alpha^{-1}, \quad g \in \mathrm{Sp}(V),$$

leave

$$\mathcal{S}(X, E)^{\mathfrak{H}} = \mathcal{S}(X, E^{\mathfrak{H}}) = \mathcal{S}(X, E_0)$$

invariant, hence provide a projective Weil representation realized over E_0.

10. Optimality of the field E_0

It is natural to ask if the field E_0 is optimal in the sense that the Weil representation W_λ may not be realized over a proper subfield. To investigate this, fix a lattice $\mathcal{L}$ of V on which the symplectic form $\langle\ ,\ \rangle$ is nondegenerate and consider the K-types of the Weil representation W_λ obtained by restricting to the compact subgroup $\mathrm{Sp}(\mathcal{L})$ [Prasad 1998].

Given a natural number k, let Γ_k denote that normal subgroup of $\mathrm{Sp}(\mathcal{L})$ consisting of those elements g for which

$$vg \equiv v \bmod \mathfrak{m}^k\mathcal{L}, \quad v \in \mathcal{L},$$

and let Fix_k be the space of Γ_k-fixed points in the Weil representation. The nontrivial K-types of W_λ associated with $\mathrm{Sp}(\mathcal{L})$ can be realized as the ±1-eigenspaces of ι, the central involution of $\mathrm{Sp}(V)$, acting on the quotients $\mathrm{Fix}_{2i+2}/\mathrm{Fix}_{2i}$, $i = 0, 1, \dots$. Indeed, in light of Proposition 13 and the remarks preceding Proposition 14, it is sufficient to verify this when λ has level 0 and -1. The first case is an immediate consequence of the description of the K-types provided by [Prasad 1998, Theorem 2], while the second case follows from the analogous result for representations arising from characters of odd level. In particular, if W_λ can be realized over a field L then its K-types can also be realized over L.

The nontrivial K-types of W_λ can be shown to coincide with the irreducible representations $\mathrm{Top}^\pm$ studied in [Cliff et al. 2004]. If $q \equiv 1 \bmod 4$ then Top^- has Schur index 2, by Theorem 26 of that reference. Since Theorem 17 of the same work asserts that its character field is $\mathbb{Q}$ (respectively, $\mathbb{Q}(\sqrt{p})$) if q is square (respectively, not square), Top^- may not be realized over a proper subfield of E_0. The remarks made above allow us to conclude that E_0 is an optimal field for realizing W_λ.

In the case $q \equiv 3 \bmod 4$, the representations $\mathrm{Top}^{\pm}$ all have Schur index 1 and character fields $\mathbb{Q}(\sqrt{-p})$ [Cliff et al. 2004, Theorems 17 and 26]. As a result, the restriction of W_λ to the compact group $\mathrm{Sp}(\mathscr{L})$ can be realized over the subfield $\mathbb{Q}(\sqrt{-p})$ of E_0. The possibility of realizing the entire Weil representation over the field $\mathbb{Q}(\sqrt{-p})$ is left open.

References

[Borel 1991] A. Borel, *Linear algebraic groups*, 2nd ed., Graduate Texts in Mathematics **126**, Springer, New York, 1991. MR 92d:20001 Zbl 0726.20030

[Cartier 1979] P. Cartier, "Representations of p-adic groups: a survey", pp. 111–155 in *Automorphic forms, representations and L-functions* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure. Math. **33**, Amer. Math. Soc., Providence, R.I., 1979. MR 81e:22029 Zbl 0421.22010

[Cliff et al. 2004] G. Cliff, D. McNeilly, and F. Szechtman, "Character fields and Schur indices of irreducible Weil characters", *J. Group Theory* **7**:1 (2004), 39–64. MR 2004m:20086 Zbl 1041.20030

[Kudla 1996] S. S. Kudla, "Notes on the local theta correspondence", lecture notes, 1996, available at http://www.math.toronto.edu/~skudla/castle.pdf.

[Mœglin et al. 1987] C. Mœglin, M.-F. Vignéras, and J.-L. Waldspurger, *Correspondances de Howe sur un corps p-adique*, Lecture Notes in Mathematics **1291**, Springer, Berlin, 1987. MR 91f:11040 Zbl 0642.22002

[Prasad 1998] D. Prasad, "A brief survey on the theta correspondence", pp. 171–193 in *Number theory* (Tiruchirapalli, 1996), edited by V. K. Murty and M. Waldschmidt, Contemp. Math. **210**, Amer. Math. Soc., Providence, RI, 1998. MR 99e:11063 Zbl 0922.11041

[Ranga Rao 1993] R. Ranga Rao, "On some explicit formulas in the theory of Weil representation", *Pacific J. Math.* **157**:2 (1993), 335–371. MR 94a:22037 Zbl 0794.58017

[Weil 1974] A. Weil, *Basic number theory*, 2nd ed., Springer, New York, 1974. MR 55 #302 Zbl 0326.12001

Received April 9, 2009. Revised August 18, 2011.

GERALD CLIFF
DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES
UNIVERSITY OF ALBERTA
EDMONTON, AB T6G 2G1
CANADA
gcliff@math.ualberta.ca

DAVID MCNEILLY
DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES
UNIVERSITY OF ALBERTA
EDMONTON, AB T6G 2G1
CANADA
dam@math.ualberta.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

LAGRANGIAN SUBMANIFOLDS IN COMPLEX PROJECTIVE SPACE WITH PARALLEL SECOND FUNDAMENTAL FORM

FRANKI DILLEN, HAIZHONG LI, LUC VRANCKEN AND XIANFENG WANG

From the Riemannian geometric point of view, one of the most fundamental problems in the study of Lagrangian submanifolds is the classification of Lagrangian submanifolds with parallel second fundamental form. In 1980's, H. Naitoh completely classified the Lagrangian submanifolds with parallel second fundamental form and without Euclidean factor in complex projective space, by using the theory of Lie groups and symmetric spaces. He showed that such a submanifold is always locally symmetric and is one of the symmetric spaces: $\mathrm{SO}(k+1)/\mathrm{SO}(k)$ $(k \geq 2)$, $\mathrm{SU}(k)/\mathrm{SO}(k)$ $(k \geq 3)$, $\mathrm{SU}(k)$ $(k \geq 3)$, $\mathrm{SU}(2k)/\mathrm{Sp}(k)$ $(k \geq 3)$, $\mathrm{E}_6/\mathrm{F}_4$.

In this paper, we completely classify the Lagrangian submanifolds in complex projective space with parallel second fundamental form by an elementary geometrical method. We prove that such a Lagrangian submanifold is either totally geodesic, or the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, or the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form, or one of the standard symmetric spaces: $\mathrm{SU}(k)/\mathrm{SO}(k)$, $\mathrm{SU}(k)$, $\mathrm{SU}(2k)/\mathrm{Sp}(k)$ $(k \geq 3)$, $\mathrm{E}_6/\mathrm{F}_4$.

As the arguments are of a local nature, at the same time, due to the correspondence between C-parallel Lagrangian submanifolds in Sasakian space forms and parallel Lagrangian submanifolds in complex space forms, we can also give a complete classification of all C-parallel submanifolds of S^{2n+1} equipped with its standard Sasakian structure.

1. Introduction

One of the first studies of Lagrangian submanifolds of complex space forms was done by Chen and Ogiue [1974]. Since then such submanifolds have been studied

This research was supported by Tsinghua University–Katholieke Universiteit Leuven Bilateral Scientific Cooperation Fund. Li was supported by NSFC grant number 10971110. Wang was supported by NSFC grant number 11171175 and the Fundamental Research Funds for the Central Universities. Wang is the corresponding author.

MSC2010: primary 53B25; secondary 53C42.

Keywords: Lagrangian submanifolds, complex space forms, complex projective space, parallel second fundamental form.

by many authors and a lot of progress has been made in order to understand them properly. Notwithstanding, several open problems remain.

One of the first questions asked and solved by Naitoh in a series of papers [1980; 1981a; 1981b; 1982; 1983a] was the classification of the parallel Lagrangian submanifolds of the complex projective space. The classification relies heavily on the study of symmetric spaces (and Lie groups), and whereas in the irreducible case the classification is clear, little information is given on how to construct all reducible examples. In this paper, we use the techniques developed in [Hu et al. 2009; 2011] in order to obtain a complete and explicit classification of the Lagrangian submanifolds in complex projective space with parallel second fundamental form by an elementary geometric method. The advantage of this approach is that it also allows the study of related problems in this area, such as:

(i) Which are the biharmonic parallel submanifolds of the complex projective space?

(ii) Which are the second order parallel submanifolds (in the sense of Lumiste [2009])?

(iii) Which are the semiparallel submanifolds?

The main result we show is the following:

Classification theorem. *Let M be a Lagrangian submanifold in $\mathbb{CP}^n(4)$ with constant holomorphic sectional curvature* 4. *Suppose that M has parallel second fundamental form, then either M is totally geodesic, or*

(i) *M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, or*

(ii) *M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form, or*

(iii) $n = \frac{1}{2}k(k+1) - 1$, $k \geq 3$, *and M is congruent with* $\mathrm{SU}(k)/\mathrm{SO}(k)$, *or*

(iv) $n = k^2 - 1$, $k \geq 3$, *and M is congruent with* $\mathrm{SU}(k)$, *or*

(v) $n = 2k^2 - k - 1$, $k \geq 3$, *and M is congruent with* $\mathrm{SU}(2k)/\mathrm{Sp}(k)$, *or*

(vi) $n = 26$ *and M is congruent with* $\mathrm{E}_6/\mathrm{F}_4$.

The Calabi product is a standard technique [Bolton et al. 2009; Castro et al. 2006; Hu et al. 2008; Li and Wang 2011; Rodriguez Montealegre and Vrancken 2009]. It allows one to construct with one (or two) Lagrangian immersions a new Lagrangian immersion. It is recalled in detail in Section 4 of the paper.

The paper is organized as follows. In Section 2, we recall the basic formulas for Lagrangian submanifolds of complex space forms. In Section 3, we give a construction of an appropriate basis and hence decompose the tangent space into 3 orthogonal distributions $\mathfrak{D}_1$, which is 1-dimensional, $\mathfrak{D}_2$ and $\mathfrak{D}_3$. According to the dimension of $\mathfrak{D}_2$, we have n cases $\{\mathfrak{C}_m\}_{1\le m\le n}$ to consider, where $m = \dim \mathfrak{D}_2 + 1$. We show that the case $\{\mathfrak{C}_n\}$ does not occur. In order to get the components of the second fundamental form, we define a bilinear map L from $\mathfrak{D}_2 \times \mathfrak{D}_2$ to $\mathfrak{D}_3$ and give some properties of L. In Section 4, we introduce for any unit vector $v \in \mathfrak{D}_2$ a linear map $P_v : \mathfrak{D}_2 \to \mathfrak{D}_2$ and study its properties. We use the previous results to obtain a direct sum decomposition for $\mathfrak{D}_2$. We prove that there exists an integer k_0 and unit vectors $v_1, \dots, v_{k_0} \in \mathfrak{D}_2$ such that

$$\mathfrak{D}_2 = \{v_1\} \oplus V_{v_1}(0) \oplus \cdots \oplus \{v_{k_0}\} \oplus V_{v_{k_0}}(0),$$

where $V_{v_j}(0)$ is the eigenspace of P_{v_j} with eigenvalue 0. We remark that we always mean an orthogonal sum of vector spaces when we speak of a direct sum. We also find that $\dim V_{v_1}(0) = \cdots = \dim V_{v_{k_0}}(0)$ and the dimension which we denote by $\mathfrak{p}$ can only be equal to 0, 1, 3 or 7 when $k_0 \ge 2$. Note that up to this point all results remain valid assuming only that M is semiparallel. We also recall some characterizations of the Calabi product Lagrangian immersions in $\mathbb{CP}^n(4)$, whose application gives that M is the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form for case $\{\mathfrak{C}_1\}$. In Section 5, we discuss case $\{\mathfrak{C}_m\}_{2\le m\le n-1}$ with $k_0 = 1$. In Sections 6–9, we consider each of the four cases: case $\{\mathfrak{C}_m\}_{2\le m\le n-1}$ with $k_0 \ge 2$ and $\mathfrak{p} = 0, 1, 3, 7$ separately and in each case we obtain a complete classification of the Lagrangian submanifolds in $\mathbb{CP}^n(4)$ with parallel second fundamental form. In Section 10, we complete the proof of the Classification theorem.

2. Preliminaries

In this section, M will always denote an n-dimensional Lagrangian submanifold of $\bar{M}^n(4\varepsilon)$, an n-dimensional complex space form with constant holomorphic sectional curvature 4ε. We denote the Levi-Civita connections on M, $\bar{M}^n(4\varepsilon)$ and the normal bundle by ∇, D and $\nabla^{\perp}_X$ respectively. The formulas of Gauss and Weingarten are given by (see [Chen 1973; 1997a; 1997b; Castro et al. 2006])

$$D_X Y = \nabla_X Y + h(X, Y) \quad \text{and} \quad D_X \xi = -A_\xi X + \nabla^{\perp}_X \xi,$$

where X and Y are tangent vector fields and ξ is a normal vector field on M.

As M is Lagrangian, we have (see [Chen 2001; 2005; Li and Vrancken 2005])

$$\nabla^{\perp}_X JY = J\nabla_X Y \quad \text{and} \quad A_{JX} Y = -Jh(X, Y) = A_{JY} X, \tag{2-1}$$

where h and A denote respectively the second fundamental form and the shape operator.

We denote the curvature tensors of ∇ and $\nabla^{\perp}_X$ by R and $R^{\perp}$, respectively. The first and second covariant derivatives of h are defined by

$$(\nabla h)(X, Y, Z) = \nabla^{\perp}_X h(Y, Z) - h(\nabla_X Y, Z) - h(\nabla_X Z, Y),$$
$$(\nabla^2 h)(X, Y, Z, W) = \nabla^{\perp}_X((\nabla h)(Y, Z, W)) - (\nabla h)(\nabla_X Y, Z, W) - (\nabla h)(Y, \nabla_X Z, W) - (\nabla h)(Y, Z, \nabla_X W),$$

where X, Y, Z and W are tangent vector fields.

The equations of Gauss, Codazzi and Ricci for a Lagrangian submanifold of $\bar{M}^n(4\varepsilon)$ are given by (see [Chen and Ogiue 1974; Chen 1997a; 1997b; 2001])

$$R(X, Y)Z = \varepsilon(\langle Y, Z\rangle X - \langle X, Z\rangle Y) + [A_{JX}, A_{JY}]Z, \tag{2-2}$$
$$(\nabla h)(X, Y, Z) = (\nabla h)(Y, X, Z),$$
$$R^{\perp}(X, Y)JZ = \varepsilon(\langle Y, Z\rangle JX - \langle X, Z\rangle JY) + J[A_{JX}, A_{JY}]Z,$$

where X, Y and Z are tangent vector fields. Note that for a Lagrangian submanifold the equations of Gauss and Ricci are mutually equivalent.

We have the following Ricci identity (see [Montiel and Urbano 1988]):

$$(\nabla^2 h)(X, Y, Z, W) = (\nabla^2 h)(Y, X, Z, W) + JR(X, Y)A_{JZ}W - h(R(X, Y)Z, W) - h(R(X, Y)W, Z), \tag{2-3}$$

where X, Y, Z and W are tangent vector fields.

The Lagrangian condition implies that

$$\langle R^{\perp}(X, Y)JZ, JW\rangle = \langle R(X, Y)Z, W\rangle,$$
$$\langle h(X, Y), JZ\rangle = \langle h(X, Z), JY\rangle,$$

for tangent vector fields X, Y, Z and W.

From now on, we will also assume that M has parallel fundamental form, that is, in each point p of M, $\nabla h = 0$.

Note that the vanishing of ∇h together with the Ricci identity (2-3) imply that

$$(R(X, Y)h)(Z, W) = R^{\perp}(X, Y)\, h(Z, W) - h(R(X, Y)Z, W) - h(Z, R(X, Y)W) \equiv 0, \tag{2-4}$$

for tangent vector fields X, Y, Z and W. Lagrangian submanifolds satisfying the above property are called semiparallel. Using this property, following an idea first introduced by Ejiri [1981], and since then widely applied and very useful for solving various problems in submanifold theory, a special orthonormal basis can be constructed.

3. The construction of an appropriate orthonormal basis

In this section, we will always assume that M is a Lagrangian submanifold of $\bar{M}^n(4\varepsilon)$ with semiparallel second fundamental form, where $\bar{M}^n(4\varepsilon)$ is an n-dimensional complex space form with constant holomorphic sectional curvature 4ε.

Throughout this section, we fix $p \in M$ and let $UM_p = \{u \in T_pM \mid \|u\| = 1\}$. Note that totally geodesic submanifolds in symmetric spaces have been classified completely by Chen and Nagano [1977; 1978], we will assume that p is a non-totally geodesic point and we define $f(u) = \langle h(u,u), Ju\rangle$ for $u \in UM_p$ and take e_1 as a vector in which f attains its maximum. The following lemma can be found in [Li and Vrancken 2005], [Li and Wang 2009] and [Montiel and Urbano 1988].

Lemma 3.1. *There exists an orthonormal basis* $\{e_1, \dots, e_n\}$ *of* T_pM *satisfying*:

(i) $h(e_1, e_i) = \lambda_i Je_i$ *for* $i = 1, \dots, n$, *where* λ_1 *is the maximum of* f.

(ii) $\lambda_i \le \frac{1}{2}\lambda_1$ *for* $i = 2, \dots, n$, *and if* $\lambda_j = \frac{1}{2}\lambda_1$ *for some* j, *then* $f(e_j) = 0$.

Furthermore, by taking $X = Z = W = e_1$, $Y = e_j$ for $j \ge 2$ in (2-4), by Lemma 3.1.(i) there exists a unique m with $1 \le m \le n$ such that

$$\lambda_2 = \lambda_3 = \cdots = \lambda_m = \tfrac{1}{2}\lambda_1 \quad \text{and} \quad \lambda_{m+1} = \cdots = \lambda_n = \mu, \tag{3-1}$$

where

$$\mu := \frac{\lambda_1 - \sqrt{\lambda_1^2 + 4\varepsilon}}{2}.$$

We define $\mathfrak{D}_2 := \operatorname{span}\{e_2, \dots, e_m\}$ and $\mathfrak{D}_3 := \operatorname{span}\{e_{m+1}, \dots, e_n\}$.

Lemma 3.2. *The tangent space* T_pM *can be decomposed as a direct sum of* 3 *orthogonal vector spaces*, *that is*, $T_pM = \mathfrak{D}_1 \oplus \mathfrak{D}_2 \oplus \mathfrak{D}_3$, *where*

(i) $\mathfrak{D}_1$ *is a* 1*-dimensional vector space spanned by the unit tangent vector* e_1,

(ii) $h(e_1, v) = \frac{1}{2}\lambda_1 v$, *for any* $v \in \mathfrak{D}_2$,

(iii) $h(e_1, w) = \mu w$, *for any* $w \in \mathfrak{D}_3$,

(iv) $h(v_1, v_2) - \frac{1}{2}\lambda_1\langle v_1, v_2\rangle Je_1 \in J\mathfrak{D}_3$, *for any* $v_1, v_2 \in \mathfrak{D}_2$.

We have n cases $\{\mathfrak{C}_m\}_{1\le m\le n}$ as follows:

Case $\mathfrak{C}_1 : \lambda_2 = \lambda_3 = \cdots = \lambda_n = \mu$.

Case $\mathfrak{C}_n : \lambda_2 = \lambda_3 = \cdots = \lambda_n = \frac{1}{2}\lambda_1$.

Case $\mathfrak{C}_m : \lambda_2 = \cdots = \lambda_m = \frac{1}{2}\lambda_1$ and $\lambda_{m+1} = \cdots = \lambda_n = \mu$ for $2 \le m \le n-1$.

Our aim in the next sections is to describe explicitly the second fundamental form h when restricted to vectors belonging to $\mathfrak{D}_2$. In view of Lemma 3.2 this is trivial in case that $m = 1$ or $m = n$. We first state:

Theorem 3.3. *Case* $\{\mathfrak{C}_n\}$ *does not occur.*

Proof. Suppose that it did. We use (2-4), and we choose $X = e_1$, $Y = v$, $Z = v$ and $W = v$, with v a unit vector belonging to $\mathfrak{D}_2$. Taking also into account, from the previous lemma, that

$$h(e_1, e_1) = \lambda_1 J e_1, \quad h(e_1, v) = \tfrac{1}{2}\lambda_1 J v \quad \text{and} \quad h(v, v) = \tfrac{1}{2}\lambda_1 J e_1,$$

we find that $\lambda_1 = 0$. This is a contradiction. □

By applying Theorem 4.12 (see also [Li and Wang 2011, Theorem 1.6]), we conclude that M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form for case $\{\mathfrak{C}_1\}$. We will now restrict ourselves in the remainder of this section, as well as in the next sections, to the case $\{\mathfrak{C}_m\}$ when $1 < m < n$. Surprisingly enough it is the form of the second fundamental form restricted to $\mathfrak{D}_2$ which will play a crucial role and in some sense completely describe the immersion. For convenience we write

$$\eta = \tfrac{1}{2}\sqrt{\lambda_1^2 + 4\varepsilon}$$

and without loss of generality we may assume that $\eta \neq 0$.

By Lemma 3.2 we can introduce a bilinear map $L : \mathfrak{D}_2 \times \mathfrak{D}_2 \to \mathfrak{D}_3$ by

$$L(v_1, v_2) := -J\big(h(v_1, v_2) - \tfrac{1}{2}\lambda_1 \langle v_1, v_2\rangle J e_1\big), \quad v_1, v_2 \in \mathfrak{D}_2. \tag{3-2}$$

We will now distinguish vectors belonging to the different vector spaces and so we use the notations $v, v_j \in \mathfrak{D}_2$, $w, w_r \in \mathfrak{D}_3$.

Lemma 3.4. *We have* $\langle h(\mathfrak{D}_3, \mathfrak{D}_3), J\mathfrak{D}_2\rangle = 0$. *The tensor* L *is an isotropic tensor in the sense of O'Neill [1965], that is,*

$$\langle L(v, v), L(v, v)\rangle = \tfrac{1}{2}\lambda_1 \eta \|v\|^2, \quad v \in \mathfrak{D}_2. \tag{3-3}$$

Linearizing this expression, it follows for arbitrary vectors $v_1, v_2, v_3, v_4 \in \mathfrak{D}_2$ *that*

$$\begin{aligned}\langle L(v_1, v_2), L(v_3, v_4)\rangle + \langle L(v_1, v_3), L(v_2, v_4)\rangle + \langle L(v_1, v_4), L(v_2, v_3)\rangle \\ = \tfrac{1}{2}\lambda_1 \eta\big(\langle v_1, v_2\rangle\langle v_3, v_4\rangle + \langle v_1, v_3\rangle\langle v_2, v_4\rangle + \langle v_1, v_4\rangle\langle v_2, v_3\rangle\big).\end{aligned} \tag{3-4}$$

Proof. By taking $Z = W = e_1$ in (2-4) we immediately obtain that for arbitrary vectors x and y, $R(x, y)e_1$ is an eigenvector of A_{Je_1} with eigenvalue $\frac{1}{2}\lambda_1$. So $R(x, y)e_1 \in \mathfrak{D}_2$. Moreover taking $v \in \mathfrak{D}_2$ and $w \in \mathfrak{D}_3$, by the Gauss equation (2-2) we have

$$R(v, w)e_1 = (\mu - \tfrac{1}{2}\lambda_1) A_{Jv} w = -\eta A_{Jv} w,$$

so we have

$$A_{Jv} w \in \mathfrak{D}_2, \quad \text{for all} \quad v \in \mathfrak{D}_2,\ w \in \mathfrak{D}_3, \tag{3-5}$$

which gives the first claim of the lemma.

In order to prove the second claim, we use again (2-4), and we choose $X = e_1$, $Y = v_1$, $Z = v_2$ and $W = v_3$, all belonging to $\mathcal{D}_2$. By using (2-2) and the definition of L, it follows immediately that

$$h(v_1, L(v_2, v_3)) + h(v_2, L(v_1, v_3)) + h(v_3, L(v_1, v_2)) = \tfrac{1}{2}\lambda_1 \eta(\langle v_2, v_3\rangle Jv_1 + \langle v_1, v_3\rangle Jv_2 + \langle v_1, v_2\rangle Jv_3). \tag{3-6}$$

Taking the inner product with v_4 and using the complete symmetry of the cubic form completes the proof. □

We now decompose $\mathcal{D}_3$ as a direct sum of two orthogonal vector spaces. We define $\mathcal{D}_{31}$ to be the vector space $\mathrm{vect}\{L(\mathcal{D}_2, \mathcal{D}_2)\}$ generated by vectors $L(X, Y)$ where $X, Y \in \mathcal{D}_2$, and $\mathcal{D}_{32}$ as its orthogonal complement in $\mathcal{D}_3$. Then by taking $X = e_1$, $Y = v_1$, $Z = v_2$ and $W = w$ in (2-4) where $v_1, v_2 \in \mathcal{D}_2$ and $w \in \mathcal{D}_{32}$ and using the fact that $h(v_2, w) = 0$ we get:

Lemma 3.5. *Let $v_1, v_2 \in \mathcal{D}_2$ and $w \in \mathcal{D}_{32}$. Then*

$$h(L(v_1, v_2), w) = \mu\eta\langle v_1, v_2\rangle Jw. \tag{3-7}$$

Similarly, we also have:

Lemma 3.6. *Let $v_1, v_2, v_3, v_4 \in \mathcal{D}_2$ and let $\{u_1, \ldots, u_{m-1}\}$ be an orthonormal basis of $\mathcal{D}_2$, then we have*

$$\begin{aligned} h(L(v_1, v_2), L(v_3, v_4)) &= \mu\langle L(v_1, v_2), L(v_3, v_4)\rangle Je_1 + \mu\eta\langle v_1, v_2\rangle JL(v_3, v_4) \\ &+ \sum_{i=1}^{m-1}\langle L(v_1, u_i), L(v_3, v_4)\rangle JL(v_2, u_i) + \sum_{i=1}^{m-1}\langle L(v_2, u_i), L(v_3, v_4)\rangle JL(v_1, u_i). \end{aligned} \tag{3-8}$$

Proof. By (2-2), we have for $v, \tilde{v} \in \mathcal{D}_2$ that

$$R(e_1, v)\tilde{v} = (\varepsilon + \tfrac{1}{4}\lambda_1^2)\,\langle v, \tilde{v}\rangle e_1 - \eta L(v, \tilde{v}) = \eta^2\langle v, \tilde{v}\rangle e_1 - \eta L(v, \tilde{v}). \tag{3-9}$$

Similarly, we have for $v \in \mathcal{D}_2$ and $w \in \mathcal{D}_3$ that $R(e_1, v)w = \eta A_{Jv}w$.

As M is semiparallel, we have from (2-4) that

$$R^{\perp}(e_1, v_1)h(v_2, L(v_3, v_4)) = h(R(e_1, v_1)v_2, L(v_3, v_4)) + h(v_2, R(e_1, v_1)L(v_3, v_4)). \tag{3-10}$$

We now compute each of the terms in the above equation separately. Since, by

Lemma 3.4, $h(v_j, L(v_k, v_l)) \in J\mathfrak{D}_2$, we can write

$$h(v_2, L(v_3, v_4)) = \sum_{i=1}^{m-1} \langle h(v_2, L(v_3, v_4)), Ju_i\rangle Ju_i = \sum_{i=1}^{m-1} \langle L(v_2, u_i), L(v_3, v_4)\rangle Ju_i.$$

Therefore, we get

$$R^{\perp}(e_1, v_1)\, h(v_2, L(v_3, v_4)) = \sum_{i=1}^{m-1} \langle L(v_2, u_i), L(v_3, v_4)\rangle R^{\perp}(e_1, v_1) Ju_i$$
$$= \eta^2 \langle L(v_1, v_2), L(v_3, v_4)\rangle Je_1 - \eta \sum_{i=1}^{m-1} \langle L(v_2, u_i), L(v_3, v_4)\rangle JL(v_1, u_i).$$

Next, as $L(v_3, v_4) \in \mathfrak{D}_3$, we have

$$R(e_1, v_1)L(v_3, v_4) = \eta A_{Jv_1} L(v_3, v_4) = \eta \sum_{i=1}^{m-1} \langle L(v_1, u_i), L(v_3, v_4)\rangle u_i.$$

Hence

$$h(v_2, R(e_1, v_1)L(v_3, v_4)) = \eta \sum_{i=1}^{m-1} \langle L(v_1, u_i), L(v_3, v_4)\rangle h(v_2, u_i)$$
$$= \frac{\lambda_1}{2} \eta \langle L(v_1, v_2), L(v_3, v_4)\rangle Je_1 + \eta \sum_{i=1}^{m-1} \langle L(v_1, u_i), L(v_3, v_4)\rangle JL(v_2, u_i).$$

Finally the last term of (3-10) can be computed as follows:

$$h(R(e_1, v_1)v_2, L(v_3, v_4)) = \eta^2 \mu \langle v_1, v_2\rangle JL(v_3, v_4) - \eta h(L(v_1, v_2), L(v_3, v_4)).$$

Combining all three terms now completes the proof of the lemma. □

We note that Equation (3-8) has very important consequences which will be used in sequel sections. For example:

Lemma 3.7. *Assume that $m \geq 3$ and let $\{u_1, \dots, u_{m-1}\}$ be an orthonormal basis of $\mathfrak{D}_2$, then for $p \neq j$, we have*

$$\text{(3-11)} \quad 0 = \big(\eta(\eta + \tfrac{1}{2}\lambda_1) - 4\langle L(u_j, u_p), L(u_j, u_p)\rangle\big) L(u_p, u_j) + \sum_{i \neq p} \big(\langle L(u_p, u_i), L(u_j, u_j)\rangle - 2\langle L(u_j, u_i), L(u_p, u_j)\rangle\big) L(u_j, u_i).$$

In particular, if $L(u_1, u_2) \neq 0$ and $L(u_1, u_i)$ is orthogonal to $L(u_1, u_2)$ for all $i \neq 2$, then

$$\text{(3-12)} \quad \langle L(u_1, u_2), L(u_1, u_2)\rangle = \tfrac{1}{4}\eta(\eta + \tfrac{1}{2}\lambda_1) =: \tau.$$

Proof. We use (3-8). Interchanging the couples of indices {1, 2} and {3, 4} we find the following condition:

$$
\begin{aligned}
\text{(3-13)}\quad 0 = \eta\mu\big(\langle v_1, v_2\rangle L(v_3, v_4) - \langle v_3, v_4\rangle L(v_1, v_2)\big)& \\
+\sum_{i=1}^{m-1}\langle L(v_1, u_i), L(v_3, v_4)\rangle L(v_2, u_i) + \sum_{i=1}^{m-1}\langle L(v_2, u_i), L(v_3, v_4)\rangle L(v_1, u_i)& \\
-\sum_{i=1}^{m-1}\langle L(v_3, u_i), L(v_1, v_2)\rangle L(v_4, u_i) - \sum_{i=1}^{m-1}\langle L(v_4, u_i), L(v_1, v_2)\rangle L(v_3, u_i).&
\end{aligned}
$$

If we take $v_2 = v_3 = v_4 = u_j$ and $v_1 = u_p$ with j and p different, then by using also the isotropy condition, (3-13) reduces to

$$
\begin{aligned}
0 = \big(\eta(\eta + \tfrac{1}{2}\lambda_1) - 4\langle L(u_j, u_p), L(u_j, u_p)\rangle\big)L(u_p, u_j)& \\
+\sum_{i\neq p}\big(\langle L(u_p, u_i), L(u_j, u_j)\rangle - 2\langle L(u_j, u_i), L(u_p, u_j)\rangle\rangle\big)L(u_j, u_i).&
\end{aligned}
$$

Finally (3-12) follows by taking $j = 1$ and $p = 2$ in the (3-11), and by using Lemma 3.4. □

4. A map $P_v : \mathfrak{D}_2 \to \mathfrak{D}_2$ for unit vector $v \in \mathfrak{D}_2$ and a decomposition of $\mathfrak{D}_2$

We now define for any given unit vector $v \in \mathfrak{D}_2$ a linear map $P_v : \mathfrak{D}_2 \to \mathfrak{D}_2$ by

$$
P_v\tilde{v} = A_{Jv}L(v, \tilde{v}) \quad \text{for } \tilde{v} \in \mathfrak{D}_2. \tag{4-1}
$$

It is easily seen that P_v is well defined and a symmetric linear operator satisfying

$$
\langle P_v\tilde{v}, v^*\rangle = \langle A_{Jv}L(v, \tilde{v}), v^*\rangle = \langle L(v, \tilde{v}), L(v, v^*)\rangle = \langle P_v v^*, \tilde{v}\rangle \tag{4-2}
$$

for all $\tilde{v}, v^* \in \mathfrak{D}_2$. Moreover, we have:

Lemma 4.1. *For all unit $v \in \mathfrak{D}_2$, the operator $P_v : \mathfrak{D}_2 \to \mathfrak{D}_2$ has $\sigma = \frac{1}{2}\lambda_1\eta$ as an eigenvalue with eigenvector v. In the orthogonal complement of* $\{v\}$ *the operator has two eigenvalues, namely τ and* 0, *where τ is defined in* (3-12).

Proof. According to (3-2) and (3-3), we have

$$
\langle v, P_v v\rangle = \langle L(v, v), L(v, v)\rangle = \tfrac{1}{2}\lambda_1\eta,
$$

and if $v^* \perp v$, then

$$
\langle v^*, P_v v\rangle = \langle L(v, v^*), L(v, v)\rangle = 0.
$$

This implies that $P_v v = \frac{1}{2}\lambda_1\eta\, v$.

Next, we take an orthonormal basis $\{u_i\}_{i=1}^{m-1}$ of $\mathfrak{D}_2$ consisting of eigenvectors of P_v such that $P_v u_i = \sigma_i u_i$ for $1 \le i \le m-1$, with $u_1 = v$ and $\sigma_1 = \frac{1}{2}\lambda_1\eta$. We take the inner product of formula (3-11) for $j = 1$ and any $p \geq 2$ with $L(v, u_p)$. We have

$$\langle L(u_1, u_p), L(u_1, u_p)\rangle \big(\tau - \langle L(u_1, u_p), L(u_1, u_p)\rangle\big) = 0. \tag{4-3}$$

Here we have used that

$$\langle L(u_1, u_p), L(u_1, u_i)\rangle = \langle u_p, P_{u_1} u_i\rangle = \langle u_p, \sigma_i u_i\rangle = 0 \quad \text{for all } i \neq p.$$

By (4-3), we get either

$$\sigma_p = \langle L(v, u_p), L(v, u_p)\rangle = 0 \quad \text{or} \quad \sigma_p = \langle L(v, u_p), L(v, u_p)\rangle = \tau. \qquad \square$$

In the following we denote by $V_v(0)$ and $V_v(\tau)$ the eigenspaces of P_v (in the orthogonal complement of $\{v\}$) with respect to the eigenvalues 0 and τ, respectively. Note that in exceptional cases it can happen that $\tau = \sigma$.

Lemma 4.2. *Let $u, v \in \mathfrak{D}_2$ be two unit orthogonal vectors. The following statements are equivalent*:

(i) $u \in V_v(0)$.

(ii) $L(u, v) = 0$.

(iii) $L(u, u) = L(v, v)$.

(iv) $v \in V_u(0)$.

Moreover any of the previous statements implies that

(v) $P_u = P_v$ *on* $\{u, v\}^\perp$.

Proof. As $\langle v_1, P_v v_2\rangle = \langle L(v, v_1), L(v, v_2)\rangle$, the equivalence of (i), (ii) and (iv) follows immediately. As u and v are orthogonal, the isotropy condition implies that

$$\langle L(u, u), L(v, v)\rangle + 2\langle L(u, v), L(u, v)\rangle = \tfrac{1}{2}\lambda_1\eta.$$

Because $\langle L(u, u), L(u, u)\rangle = \langle L(v, v), L(v, v)\rangle = \frac{1}{2}\lambda_1\eta$, the equivalence of (ii) and (iii) now follows from the Cauchy–Schwarz inequality.

Now in order to prove (v), we may assume that (i), (ii), (iii) and (iv) are satisfied. As the space spanned by $\{u, v\}$ is invariant by P_u and P_v, also its orthogonal complement is invariant. By taking v_1, v_2 in this orthogonal complement and using the isotropy condition, we find

$$\begin{aligned}\langle v_1, P_v v_2\rangle &= \langle L(v, v_1), L(v, v_2)\rangle\\ &= -\tfrac{1}{2}\langle L(v, v), L(v_1, v_2)\rangle + \tfrac{1}{4}\lambda_1\eta\langle v_1, v_2\rangle\\ &= -\tfrac{1}{2}\langle L(u, u), L(v_1, v_2)\rangle + \tfrac{1}{4}\lambda_1\eta\langle v_1, v_2\rangle\\ &= \langle L(u, v_1), L(u, v_2)\rangle = \langle v_1, P_u v_2\rangle. \qquad \square\end{aligned}$$

Lemma 4.3. *Let $v, \tilde{v} \in \mathcal{D}_2$ be two unit orthogonal vectors. Then the equality*

$$\langle L(v, \tilde{v}), L(v, \tilde{v})\rangle = \tau$$

holds if and only if $\tilde{v} \in V_v(\tau)$.

Moreover, if we assume $u \in V_v(0)$ and the equality holds, then $u \in V_{\tilde{v}}(\tau)$.

Proof. If $\tilde{v} \in V_v(\tau)$, then $\langle L(v, \tilde{v}), L(v, \tilde{v})\rangle = \langle \tilde{v}, P_v\tilde{v}\rangle = \tau$.

Conversely, if $\langle L(v, \tilde{v}), L(v, \tilde{v})\rangle = \tau$, we can write

$$\tilde{v} = \cos\theta v_0 + \sin\theta v_1, \quad |v_0| = |v_1| = 1,$$

where $v_0 \in V_v(0)$ and $v_1 \in V_v(\tau)$. Then we get

$$\tau = \langle L(v, \tilde{v}), L(v, \tilde{v})\rangle = \langle P_v\tilde{v}, \tilde{v}\rangle = \cos^2\theta\tau,$$

which means that $\sin\theta = 0$ and $\tilde{v} = \cos\theta v_1 \in V_v(\tau)$.

Now assume the equality holds. If $u \in V_v(0)$, then as $v \in V_v(\sigma)$ and $\tilde{v} \in V_v(\tau)$, we see that $u, v, \tilde{v}$ are orthonormal vectors. Therefore $P_u\tilde{v} = P_v\tilde{v} = \tau\tilde{v}$ by Lemma 4.2, which means that $\tilde{v} \in V_u(\tau)$. Applying the first part of the lemma now shows that we have $u \in V_{\tilde{v}}(\tau)$. □

Lemma 4.4. *Let $v_1, v_2, v_3 \in \mathcal{D}_2$ be orthonormal vectors satisfying $v_1, v_2 \in V_{v_3}(\tau)$. Then for any vector $v \in \mathcal{D}_2$, we have $\langle L(v_1, v_2), L(v_3, v)\rangle = 0$.*

Proof. Using the linearity of the assertion, we may assume that v is an eigenvector of P_{v_3}. By Lemma 4.2 we only have to consider the case $v \notin V_{v_3}(0)$.

We choose an orthonormal basis $\{u_i\}_{i=1}^{m-1}$ of $\mathcal{D}_2$ consisting of eigenvectors of P_{v_3} such that $u_1 = v_1$, $u_2 = v_2$ and $u_3 = v_3$. We now use (3-13) for $p = 1$, $j = 2$, $k = l = 3$ to obtain

$$\begin{aligned}(4\text{-}4)\quad 0 = -\mu\eta L(v_1, v_2) &+ \sum_{i=1}^{m-1}\langle L(v_1, u_i), L(v_3, v_3)\rangle L(v_2, u_i)\\ &+\sum_{i=1}^{m-1}\langle L(v_2, u_i), L(v_3, v_3)\rangle L(v_1, u_i) - 2\sum_{i=1}^{m-1}\langle L(v_3, u_i), L(v_1, v_2)\rangle L(v_3, u_i).\end{aligned}$$

Remark that if $i = 3$ and $k = 1, 2$, it follows that $\langle L(v_k, u_i), L(v_3, v_3)\rangle = 0$, and if $k = 1, 2$ and $i \neq k, 3$, we have that $\langle L(v_k, u_i), L(v_3, v_3)\rangle = -2\langle v_k, P_{v_3}u_i\rangle = 0$. Using this, together with (3-4) and the assumption we see that (4-4) reduces to

$$(4\text{-}5)\qquad \sum_{i=1}^{m-1}\langle L(v_3, u_i), L(v_1, v_2)\rangle L(v_3, u_i) = 0.$$

Note that we have

$$\langle L(v_3, u_p), L(v_3, u_q)\rangle = \langle u_p, P_{v_3}u_q\rangle = 0 \quad \text{if } p \neq q.$$

Thus (4-5) implies that $\langle L(v_1, v_2), L(v_3, u_i)\rangle = 0$ for all u_i, which immediately implies that for any vector $v \in \mathfrak{D}_2$, we have $\langle L(v_1, v_2), L(v_3, v)\rangle = 0$. □

Using the above lemmas, we can introduce a direct sum decomposition for $\mathfrak{D}_2$, which turns out crucial for our purpose.

Pick any unit vector $v_1 \in \mathfrak{D}_2$ and recall that $\tau = \frac{1}{4}\eta(\eta + \frac{1}{2}\lambda_1)$, then by Lemma 4.1, we have a direct sum decomposition for $\mathfrak{D}_2$

$$\mathfrak{D}_2 = \{v_1\} \oplus V_{v_1}(0) \oplus V_{v_1}(\tau),$$

where here and later on, we denote also by $\{\cdot\}$ the vector space spanned by its elements. If $V_{v_1}(\tau) \neq \{0\}$, we take an arbitrary unit vector $v_2 \in V_{v_1}(\tau)$. Then by Lemma 4.3 we have:

$$v_1 \in V_{v_2}(\tau), \quad V_{v_1}(0) \subset V_{v_2}(\tau) \quad \text{and} \quad V_{v_2}(0) \subset V_{v_1}(\tau).$$

From this we deduce that

$$\mathfrak{D}_2 = \{v_1\} \oplus V_{v_1}(0) \oplus \{v_2\} \oplus V_{v_2}(0) \oplus \big(V_{v_1}(\tau) \cap V_{v_2}(\tau)\big).$$

If $V_{v_1}(\tau) \cap V_{v_2}(\tau) \neq \{0\}$, we further pick a unit vector $v_3 \in V_{v_1}(\tau) \cap V_{v_2}(\tau)$. Then

$$\mathfrak{D}_2 = \{v_3\} \oplus V_{v_3}(0) \oplus V_{v_3}(\tau),$$

and by Lemma 4.3 we have

$$v_1, v_2 \in V_{v_3}(\tau) \quad \text{and} \quad V_{v_1}(0), V_{v_2}(0) \subset V_{v_3}(\tau).$$

It follows that

$$\mathfrak{D}_2 = \{v_1\} \oplus V_{v_1}(0) \oplus \{v_2\} \oplus V_{v_2}(0) \oplus \{v_3\} \oplus V_{v_3}(0) \\ \oplus \big(V_{v_1}(\tau) \cap V_{v_2}(\tau) \cap V_{v_3}(\tau)\big).$$

Considering that $\dim \mathfrak{D}_2 = m - 1$ is finite, we easily obtain by induction:

Proposition 4.5. *There exists an integer k_0 and unit vectors $v_1, \ldots, v_{k_0} \in \mathfrak{D}_2$ so*

$$\mathfrak{D}_2 = \{v_1\} \oplus V_{v_1}(0) \oplus \cdots \oplus \{v_{k_0}\} \oplus V_{v_{k_0}}(0). \tag{4-6}$$

We denote $\{v_k\} \oplus V_{v_k}(0)$ by V_k. In what follows, we will now study the decomposition (4-6) in more detail.

Lemma 4.6. (i) *For any unit vector $u_1 \in \{v_1\} \oplus V_{v_1}(0)$, we have*

$$\{v_1\} \oplus V_{v_1}(0) = \{u_1\} \oplus V_{u_1}(0).$$

(ii) *For any unit vectors $u_1, \tilde{u}_1 \in \{v_1\} \oplus V_{v_1}(0)$ and $u_1 \perp \tilde{u}_1$, we have $L(u_1, \tilde{u}_1) = 0$.*

Proof. (i) We first assume that u_1 is orthogonal to v_1. As then $u_1 \in V_{v_1}(0)$, we have that $L(u_1, v_1) = 0$ and $v_1 \in V_{u_1}(0)$. Also on $\{u_1, v_1\}^\perp$ we have that $P_{u_1} = P_{v_1}$, which implies that the orthogonal complement of u_1 in $V_{v_1}(0)$ coincides with the orthogonal complement of v_1 in $V_{u_1}(0)$. This completes the proof in this case.

Now we consider the general case. If $\dim(V_{v_1}(0)) = 0$, there is nothing to prove. If $\dim(V_{v_1}(0)) \geq 2$, we can take a vector $\tilde{u}$ in that space which is orthogonal to both u_1 and v_1. Applying twice the previous result then completes the proof. If $\dim(V_{v_1}(0)) = 1$, there exists $v_0 \in V_{v_1}(0)$ such that $V_{v_1}(0) = \{v_0\}$. Denote $u_1 = \cos\theta v_1 + \sin\theta v_0$. By Lemma 4.2, we see that

$$L(\cos\theta v_1 + \sin\theta v_0, -\sin\theta v_1 + \cos\theta v_0) = 0$$

and hence $-\sin\theta v_1 + \cos\theta v_0 \in V_{u_1}(0)$. Therefore $\{v_1\} \oplus V_{v_1}(0) \subset \{u_1\} \oplus V_{u_1}(0)$. If we do not have the equality, we can find a vector in the second space which is orthogonal to both v_1 and u_1. As $\{v_1\} \oplus V_{v_1}(0) = \{x\} \oplus V_x(0) = \{u_1\} \oplus V_{u_1}(0)$, we get a contradiction.

In order to prove (ii), we have by (i) that

$$\{v_1\} \oplus V_{v_1}(0) = \{u_1\} \oplus V_{u_1}(0).$$

As u_1 and $\tilde{u}_1$ are orthogonal this implies that $\tilde{u}_1 \in V_{u_1}(0)$. Consequently we see that $L(u_1, \tilde{u}_1) = 0$. □

Lemma 4.7. *In the decomposition* (4-6), *if we pick a unit vector $u_2 \in V_{v_2}(0)$, then there exists a unique vector $u_1 \in v_1 \oplus V_{v_1}(0)$ such that $L(v_1, u_2) = L(v_2, u_1)$. Moreover u_1 is a unit vector belonging to $V_{v_1}(0)$ and $L(v_1, v_2) = -L(u_2, u_1)$.*

Proof. Let $\{u_1^l, \dots, u_{p_l}^l\}$ be an orthonormal basis of $V_{v_l}(0)$, $1 \leq l \leq k_0$, such that $u_1^2 = u_2$. Then

$$\left\{v_1, \dots, v_{k_0}, u_1^1, \dots, u_{p_1}^1, \dots, u_1^{k_0}, \dots, u_{p_{k_0}}^{k_0}\right\} =: \{\tilde{u}_i\}_{1 \leq i \leq m-1}$$

forms an orthonormal basis of $\mathfrak{D}_2$. Now we use (3-8) with the vectors v_2, u_2, v_1, u_1. As by Lemma 4.2 $L(v_2, u_2) = 0$, and by our decomposition $v_1 \in V_{v_2}(\tau)$, we obtain

$$\begin{aligned}
0 &= h(L(v_2, u_2), L(v_1, v_2)) \\
&= \mu \langle L(u_2, v_2), L(v_1, v_2)\rangle Je_1 + \sum_{i=1}^{m-1} \langle L(v_2, \tilde{u}_i), L(v_1, v_2)\rangle JL(u_2, \tilde{u}_i) \\
&\qquad\qquad + \sum_{i=1}^{m-1} \langle L(u_2, \tilde{u}_i), L(v_1, v_2)\rangle JL(v_2, \tilde{u}_i) \\
&= \sum_{i=1}^{m-1} \langle P_{v_2} v_1, \tilde{u}_i\rangle JL(u_2, \tilde{u}_i) + \sum_{i=1}^{m-1} \langle L(u_2, \tilde{u}_i), L(v_1, v_2)\rangle JL(v_2, \tilde{u}_i) \\
&= \tau JL(u_2, v_1) + \sum_{i=1}^{m-1} \langle L(u_2, \tilde{u}_i), L(v_1, v_2)\rangle JL(v_2, \tilde{u}_i).
\end{aligned}$$

From this we see that we can put

$$u_1 = -\frac{1}{\tau}\sum_{i=1}^{m-1}\langle L(u_2, \tilde{u}_i), L(v_1, v_2)\rangle \tilde{u}_i. \tag{4-7}$$

Noting that $u_2 \in V_{v_1}(\tau)$, and applying Lemma 4.4 and Lemma 4.6, we see that the above sum is nonzero only if $\tilde{u}_i = u_2$ and $\tilde{u}_i = v_1$ or if $\tilde{u}_i \in V_{v_1}(0)$.

If $\tilde{u}_i = u_2$, using Lemma 4.2, we get that

$$\langle L(u_2, u_2), L(v_1, v_2)\rangle = \langle L(v_2, v_2), L(v_1, v_2)\rangle = 0,$$

whereas if $\tilde{u}_i = v_1$, we have that

$$\langle L(u_2, v_1), L(v_1, v_2)\rangle = \langle u_2, P_{v_1}v_2\rangle = \tau\langle u_2, v_2\rangle = 0.$$

Consequently $u_1 \in V_{v_1}(0)$.

In order to prove the uniqueness in $v_1 \oplus V_{v_1}(0)$, suppose that $\tilde{u}_1 \in v_1 \oplus V_{v_1}(0)$ such that $L(v_1, u_2) = L(v_2, \tilde{u}_1)$. Then we have $L(v_2, u_1 - \tilde{u}_1) = 0$, hence by Lemma 4.2 we have $u_1 - \tilde{u}_1 \in V_{v_2}(0)$. But we also have $u_1 - \tilde{u}_1 \in v_1 \oplus V_{v_1}(0) \subset V_{v_2}(\tau)$, so we must have $u_1 - \tilde{u}_1 = 0$.

To show that vector $u_1 \in V_{v_1}(0)$ satisfying $L(v_1, u_2) = L(v_2, u_1)$ must be of unit length, we note that as $u_2 \in V_{v_2}(0) \subset V_{v_1}(\tau)$ and $u_1 \in V_{v_1}(0) \subset V_{v_2}(\tau)$, then

$$\langle L(v_1, u_2), L(v_1, u_2)\rangle = \tau \quad \text{and} \quad \langle L(v_2, u_1), L(v_2, u_1)\rangle = \|u_1\|^2\tau.$$

Hence $\|u_1\|^2 = 1$ and u_1 is a unit vector.

In order to prove $L(v_1, u_2) = L(v_2, u_1)$ and $L(v_1, v_2) = -L(u_2, u_1)$ are equivalent, we use the isotropic condition (3-4) and the Cauchy–Schwarz inequality.

Suppose now that $L(v_1, u_2) = L(v_2, u_1)$. We have $v_1,\ u_1 \in V_{v_2}(\tau) = V_{u_2}(\tau)$ by Lemma 4.6, so we get $\langle L(v_1, v_2), L(v_1, v_2)\rangle = \tau$, $\langle L(u_1, u_2), L(u_1, u_2)\rangle = \tau$. As the isotropy condition gives

$$\langle L(v_1, v_2), -L(u_1, u_2)\rangle = \langle L(v_1, u_2), L(v_2, u_1)\rangle = \langle L(v_2, u_1), L(v_2, u_1)\rangle = \tau,$$

then by using the Cauchy–Schwarz inequality we get $L(v_1, v_2) = -L(u_2, u_1)$. The other hand side can be proved in a similar way. □

Lemma 4.8. *In the decomposition* (4-6), *we write* $V_l = \{v_l\} \oplus V_{v_l}(0),\ 1 \le l \le k_0$.

(1) *For any unit vector* $a \in V_j$,

$$h(L(a, a), L(a, a)) = \tfrac{1}{2}\lambda_1\mu\eta Je_1 + \eta(\mu + \lambda_1)JL(a, a). \tag{4-8}$$

(2) *For any unit vectors* $a \in V_j$, $b \in V_l$, $j \neq l$,

$$h(L(a,a), L(a,b)) = \tfrac{1}{2}\eta(\mu+\lambda_1)JL(a,b), \tag{4-9}$$

$$h(L(a,a), L(b,b)) = \tfrac{1}{2}\eta\mu^2 Je_1 + \eta\mu J(L(a,a)+L(b,b)), \tag{4-10}$$

$$h(L(a,b), L(a,b)) = \mu\tau Je_1 + \tau J(L(a,a)+L(b,b)). \tag{4-11}$$

(3) *For unit vectors* $a \in V_j$, $b, b' \in V_l$, $c \in V_q$, $d \in V_s$ *and* j, l, q, s *being distinct*, b *and* b' *being orthogonal*,

$$h(L(a,b), L(a,c)) = \tau JL(b,c), \tag{4-12}$$

$$h(L(a,a), L(b,c)) = \eta\mu JL(b,c), \tag{4-13}$$

$$h(L(a,b), L(a,b')) = 0, \tag{4-14}$$

$$h(L(a,b), L(c,d)) = 0. \tag{4-15}$$

(4) *For orthogonal unit vectors* a_1, $a_2 \in V_j$ *and unit vectors* $b \in V_l$, $c \in V_q$ *with* j, l, q *being distinct*, *we have*

$$h(L(a_1,b), L(a_2,c)) = \tau JL(b,c'), \tag{4-16}$$

where $c' \in V_q$ *is the unique unit vector satisfying* $L(a_2,c) = L(a_1,c')$.

Proof. We take an orthonormal basis of $\mathfrak{D}_2$ in such a way so that it consists of all the orthonormal basis of V_j, $1 \le j \le k_0$. Then the conclusions are direct consequences of Lemma 3.6. For example, to prove (4-12) we combine Lemma 3.6 with the fact $\langle L(a,b), L(a,c)\rangle = \langle b, P_a c\rangle = \tau\langle b,c\rangle = 0$ and the isotropic properties of L. From (4-12) and Lemmas 4.6 and 4.7 we get (4-16). □

Proposition 4.9. *In the decomposition* (4-6), *if* $k_0 = 1$, *then* $\dim(\operatorname{Im} L) = 1$. *If* $k_0 \ge 2$, *then* $\dim V_{v_1}(0) = \cdots = \dim V_{v_{k_0}}(0)$. *We denote the dimension by* $\mathfrak{p}$, *then* $\dim \mathfrak{D}_2 = m-1 = k_0(\mathfrak{p}+1)$. *Moreover*, $\mathfrak{p}$ *can only be equal to* 0, 1, 3 *or* 7.

Proof. When $k_0 = 1$, from Lemma 4.2 and Lemma 4.6 we get that $L(v_1, v_1)$ is a basis of $\operatorname{Im} L$, hence $\dim(\operatorname{Im} L) = 1$. As a direct consequence of Lemma 4.7, for any $j \neq l$, we can define a one to one linear map from $V_{v_j}(0)$ to $V_{v_l}(0)$, which preserves the length of vectors. Hence $V_{v_j}(0)$ and $V_{v_l}(0)$ are isomorphic and have the same dimension which we denote by $\mathfrak{p}$. To make the following discussion meaningful, we now assume $\mathfrak{p} \ge 1$.

Set $V_l = \{v_l\} \oplus V_{v_l}(0)$, $1 \le l \le k_0$. Let $\{v_l, u_1^l, \dots, u_{\mathfrak{p}}^l\}$ be an orthonormal basis of V_l. For each $j = 1, \dots, \mathfrak{p}$, Lemmas 4.6 and 4.7 show that we can define a linear map $\mathfrak{T}_j : V_1 \to V_1$ such that the image $\mathfrak{T}_j(v)$ of any unit vector $v \in V_1$ satisfies

$$L(v, u_j^2) = L(v_2, \mathfrak{T}_j(v)). \tag{4-17}$$

The linear map $\mathfrak{T}_j : V_1 \to V_1$ has these fundamental properties:

(P1) $\langle \mathfrak{T}_j(v), \mathfrak{T}_j(v)\rangle = \langle v, v\rangle$, that is, $\mathfrak{T}_j$ preserves the length of vectors.

(P2) For all $v \in V_1$, we have $\mathfrak{T}_j(v) \perp v$.

(P3) $\mathfrak{T}_j^2 = -\,\mathrm{id}$.

(P4) For all $j \neq l$ and $v \in V_1$, we have $\langle \mathfrak{T}_j(v), \mathfrak{T}_l(v)\rangle = 0$.

Since (P1) and (P2) can be easily seen from Lemma 4.7 and the definition of $\mathfrak{T}_j$, we need only to verify explicitly (P3) and (P4).

For any unit vector $v \in V_1$, we have

$$L(v_2, \mathfrak{T}_j^2(v)) = L(\mathfrak{T}_j(v), u_j^2). \tag{4-18}$$

Using the fact $\{\mathfrak{T}_j(v)\} \oplus V_{\mathfrak{T}_j(v)}(0) = V_1$ and $u_j^2 \in V_{v_2}(0) \subset V_{\mathfrak{T}_j(v)}(\tau)$, we have

$$\begin{aligned}\langle L(\mathfrak{T}_j(v), u_j^2), L(\mathfrak{T}_j(v), u_j^2)\rangle &= \langle L(v_2, \mathfrak{T}_j(v)), L(v_2, \mathfrak{T}_j(v))\rangle \\ &= \langle L(v, v_2), L(v, v_2)\rangle = \tau.\end{aligned}$$

Since $v, \mathfrak{T}_j(v), v_2, u_j^2$ are orthonormal vectors, by (3-4), (4-17) and $L(v_2, u_j^2) = 0$ we see that

$$\begin{aligned}0 &= \langle L(v, v_2), L(\mathfrak{T}_j(v), u_j^2)\rangle + \langle L(v, \mathfrak{T}_j(v)), L(v_2, u_j^2)\rangle + \langle L(v, u_j^2), L(v_2, \mathfrak{T}_j(v))\rangle \\ &= \langle L(v, v_2), L(\mathfrak{T}_j(v), u_j^2)\rangle + \langle L(v_2, \mathfrak{T}_j(v)), L(v_2, \mathfrak{T}_j(v))\rangle.\end{aligned}$$

Applying (4-12) and the Cauchy–Schwarz inequality we deduce

$$L(\mathfrak{T}_j(v), u_j^2) = -L(v, v_2). \tag{4-19}$$

Combining (4-18) and (4-19), we get $L(\mathfrak{T}_j^2(v) + v, v_2) = 0$, which implies that $\mathfrak{T}_j^2(v) + v \in V_{v_2}(0)$. As $\mathfrak{T}_j^2(v) + v \in V_1 \subset V_{v_2}(\tau)$, it follows that $\mathfrak{T}_j^2(v) = -v$ for a unit vector v and then by linearity for all $v \in V_1$, as claimed by (P3).

To verify (P4), we note that, if $j \neq l$ and $\mathfrak{T}_j(v), \mathfrak{T}_l(v) \in V_v(0)$, then by definition

$$L(v_2, \mathfrak{T}_j(v)) = L(v, u_j^2) \perp L(v, u_l^2) = L(v_2, \mathfrak{T}_l(v)).$$

If we assume $\mathfrak{T}_l(v) = a\mathfrak{T}_j(v) + x$, where $x \perp \mathfrak{T}_j(v)$ and $x \in V_v(0)$, then

$$\begin{aligned}0 &= \langle L(v_2, \mathfrak{T}_j(v)), L(v_2, \mathfrak{T}_l(v))\rangle \\ &= \langle L(v_2, \mathfrak{T}_j(v)), aL(v_2, \mathfrak{T}_j(v)) + L(v_2, x)\rangle \\ &= a\langle L(v_2, \mathfrak{T}_j(v)), L(v_2, \mathfrak{T}_j(v))\rangle = a\tau.\end{aligned}$$

Thus $a = 0$ and therefore $\mathfrak{T}_j(v) \perp \mathfrak{T}_l(v)$, as claimed.

We now look at the unit hypersphere $S^{\mathfrak{p}}(1) \subset V_1$, properties (P1)–(P4) above show that at $v \in S^{\mathfrak{p}}(1)$ one has

$$T_v S^{\mathfrak{p}}(1) = \big\{\mathfrak{T}_1(v), \ldots, \mathfrak{T}_{\mathfrak{p}}(v)\big\}.$$

Hence, by the properties (P1)–(P4), $S^{\mathfrak{p}}(1)$ is parallelizable. Then, according to Bott and Milnor [1958] and Kervaire [1958], the dimension $\mathfrak{p}$ can only be equal to 1, 3 or 7. □

From now on we will restrict ourselves to the complex projective case, that is, we will assume that $\epsilon = 1$. From Proposition 4.9 we see that, in order to complete the proof of the Classification theorem, it is sufficient to deal with case $\{\mathfrak{C}_m\}_{2\le m\le n-1}$ with either $k_0 = 1$ or $k_0 \ge 2$ and $\mathfrak{p} = 0,\ 1,\ 3,\ 7$. In most cases the classification will reduce to a Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, or a Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form. These are respectively constructed in the following way, see [Bolton et al. 2009; Castro et al. 2006; Hu et al. 2008; Li and Wang 2011; Rodriguez Montealegre and Vrancken 2009].

Definition 4.10 [Bolton et al. 2009]. Let $\psi_i : (M_i, g_i) \to \mathbb{CP}^{n_i}(4)$, $i = 1, 2$, be two Lagrangian immersions and let $\tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2) : I \to \mathbb{S}^3(1) \subset \mathbb{C}^2$ be a Legendre curve. Then

$$\psi = \Pi(\tilde{\gamma}_1\tilde{\psi}_1; \tilde{\gamma}_2\tilde{\psi}_2) : I \times M_1 \times M_2 \to \mathbb{CP}^n(4)$$

is a Lagrangian immersion, where $n = n_1 + n_2 + 1$, $\tilde{\psi}_i : M_i \to \mathbb{S}^{2n_i+1}(1)$ are horizontal lifts of ψ_i, $i = 1, 2$, respectively and Π is the Hopf fibration. We call ψ a *warped product* Lagrangian immersion of ψ_1 and ψ_2. When n_1 (or n_2) is zero, we call ψ a *warped product* Lagrangian immersion of ψ_2 (or ψ_1) and a point.

Definition 4.11 [Li and Wang 2011]. In Definition 4.10, when

$$\tilde{\gamma}(t) = \left(r_1 e^{i\frac{r_2}{r_1}at}, r_2 e^{-i\frac{r_1}{r_2}at}\right),$$

where r_1, r_2, and a are positive constants with $r_1^2 + r_2^2 = 1$, we call ψ a *Calabi product* Lagrangian immersion of ψ_1 and ψ_2. When n_1 (or n_2) is zero, we call ψ a *Calabi product* Lagrangian immersion of ψ_2 (or ψ_1) and a point.

Using the arguments of Bolton et al. [2009], Calabi products were characterized in Li and Wang [2011]. In particular we recall:

Theorem 4.12 [Li and Wang 2011, Theorem 1.6]. *Let $\psi : M \to \mathbb{CP}^n(4)$ be a Lagrangian immersion. Suppose that M admits orthogonal distributions $\mathfrak{D}_1$ (of dimension 1, spanned by a unit vector E_1) and $\mathfrak{D}_2$ (of dimension $n-1$, spanned by $\{E_2, \dots, E_n\}$), and that there exist local functions λ_1, λ_2 such that $\lambda_1 \ne 2\lambda_2$ and*

$$h(E_1, E_1) = \lambda_1 J E_1 \quad \text{and} \quad h(E_1, E_i) = \lambda_2 J E_i \quad \text{for } i = 2, \dots, n. \tag{4-20}$$

Then M has parallel second fundamental form if and only if ψ is locally a Calabi product Lagrangian immersion of a point and an $(n-1)$-dimensional Lagrangian immersion $\psi_1 : M_1 \to \mathbb{CP}^{n-1}(4)$ which has parallel second fundamental form.

Theorem 4.13 [Li and Wang 2011, Theorem 4.6]. *Let $\psi : M \to \mathbb{CP}^n(4)$ be a Lagrangian immersion. Suppose that M admits three mutually orthogonal distributions $\mathfrak{D}_1$ (spanned by a unit vector E_1), $\mathfrak{D}_2$, and $\mathfrak{D}_3$ of dimension* 1, n_1 *and* n_2 *respectively, with $1+n_1+n_2=n$, and that there are three real constants λ_1, λ_2 and λ_3 that satisfy $2\lambda_3 \neq \lambda_1 \neq 2\lambda_2 \neq 2\lambda_3$ such that for all $E_i \in \mathfrak{D}_2$, $E_\alpha \in \mathfrak{D}_3$,*

$$h(E_1,E_1)=\lambda_1 JE_1, \qquad h(E_1,E_i)=\lambda_2 JE_i, \tag{4-21}$$
$$h(E_1,E_\alpha)=\lambda_3 JE_\alpha, \qquad h(E_i,E_\alpha)=0.$$

Then M has parallel second fundamental form if and only if ψ is locally a Calabi product Lagrangian immersion of two lower-dimensional Lagrangian submanifolds ψ_i $(i=1,2)$ with parallel second fundamental form.

5. Case $\{\mathfrak{C}_m\}_{2\le m\le n-1}$ with $k_0 = 1$

In this section, we consider the case $\mathfrak{C}_m$ for $2 \le m \le n-1$ with $k_0 = 1$. In view of Proposition 4.9 this implies that $\dim(\operatorname{Im} L) = 1$.

Theorem 5.1. *Let $M \subset \mathbb{CP}^n(4)$ be a Lagrangian submanifold in a complex space form with constant holomorphic sectional curvature* 4. *Suppose that M is not totally geodesic and has parallel second fundamental form, that $k_0 = 1$ and that $1 \le \dim \mathfrak{D}_2 = m-1 \le n-2$. Then M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form or the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form.*

Proof. In view of Lemma 4.2 and Lemma 4.6 we see that there exists a unit vector $w_1 \in \operatorname{Im} L \subset \mathfrak{D}_3$ such that

$$L(v_1,v_2)=\sqrt{\frac{\lambda_1\eta}{2}}\langle v_1,v_2\rangle w_1 =: \rho\langle v_1,v_2\rangle w_1, \tag{5-1}$$

for all $v_1,\ v_2 \in \mathfrak{D}_2$.

By (4-8) we get

$$h(w_1,w_1)=\mu Je_1+(2\rho+\mu\eta/\rho)Jw_1. \tag{5-2}$$

By (3-5) we get the operator $A_{Jw_1} : \mathfrak{D}_2 \to \mathfrak{D}_2$ is well defined and self adjoint. From the definition of L, we get for orthonormal vectors $\{v_1,\dots,v_{m-1}\}$ belonging to $\mathfrak{D}_2$ that

$$h(e_1,v_j)=\tfrac{1}{2}\lambda_1 Jv_j, \quad h(w_1,v_j)=\rho Jv_j \quad \text{and} \quad h(v_j,v_k)=\big(\tfrac{1}{2}\lambda_1 Je_1+\rho Jw_1\big)\delta_{jk}$$

for $1 \le j,k \le m-1$.

From $\dim(\operatorname{Im} L) = 1$, we have $\mathfrak{D}_{31} = \{w_1\}$. Denote $\tilde{n} = n-m-1$, then $\dim(\mathfrak{D}_{32}) = \tilde{n}$. We choose $\{\tilde{w}_1,\dots,\tilde{w}_{\tilde{n}}\}$ to be an orthonormal basis of $\mathfrak{D}_{32}$. Then by Lemma 3.4 and Lemma 3.5 we have

$$h(w_1, \tilde{w}_r) = \frac{\mu\eta}{\rho} J\tilde{w}_r, \quad 1 \le r \le \tilde{n}. \tag{5-3}$$

Now we define $T = \alpha e_1 + \beta w_1$ and $T^* = -\beta e_1 + \alpha w_1$, where

$$\alpha = \frac{\rho}{\sqrt{\rho^2+\eta^2}} \quad \text{and} \quad \beta = \frac{\eta}{\sqrt{\rho^2+\eta^2}}. \tag{5-4}$$

Then $\{T,\ T^*,\ v_{i|1\le i\le m-1},\ \tilde{w}_{r|1\le r\le \tilde{n}}\}$ forms an orthonormal basis of T_pM. By (5-2), we easily obtain

$$h(T,T) = \eta_1 JT, \quad h(T,u) = \eta_2 Ju \quad \text{and} \quad h(T,\tilde{w}_r) = \eta_3 J\tilde{w}_r \tag{5-5}$$

for $1 \le r \le \tilde{n}$, where η_1, η_2 and η_3 are defined by

$$\eta_1 = \alpha\big(\tfrac{1}{2}\lambda_1 + \eta\big) + \mu/\alpha, \quad \eta_2 = \alpha\big(\tfrac{1}{2}\lambda_1 + \eta\big) \quad \text{and} \quad \eta_3 = \mu/\alpha, \tag{5-6}$$

which satisfy the relations $\eta_2 \ne \eta_3$, $2\eta_2 \ne \eta_1 \ne 2\eta_3$ and

$$\eta_1 = \eta_2 + \eta_3 \quad \text{and} \quad \eta_2\eta_3 = \mu\big(\eta + \tfrac{1}{2}\lambda_1\big) = -1, \tag{5-7}$$

and $u \in \{T^*, v_1, \dots, v_{m-1}\}$.

From (5-5), we have

$$\begin{cases} T(\eta_1) = \langle (\nabla h)(T,T,T), JT \rangle, \\ u(\eta_1) = \langle (\nabla h)(u,T,T), JT \rangle \quad \text{for } u \in \{T^*, v_1, \dots, v_{m-1}\}, \\ \tilde{w}_r(\eta_1) = \langle (\nabla h)(\tilde{w}_r,T,T), JT \rangle \quad \text{for } 1 \le r \le \tilde{n}, \end{cases} \tag{5-8}$$

Since M has parallel second fundamental form, (5-8) implies that η_1 is constant on M. By a similar argument, we can prove that η_2 and η_3 are also constant on M.

By the Gauss equation (2-2) and Equation (5-5), we have

$$R^\perp(u, \tilde{w}_r)\, h(T,T) = \eta_1(\eta_3 - \eta_2) JA_{Ju}\tilde{w}_r, \tag{5-9}$$

while on the other hand, from (2-4), we have

$$R^\perp(u, \tilde{w}_r)\, h(T,T) = 2(\eta_3 - \eta_2)\, h(T, A_{Ju}\tilde{w}_r). \tag{5-10}$$

Since $\eta_3 - \eta_2 \ne 0$, (5-9) and (5-10) imply that

$$h(T, A_{Ju}\tilde{w}_r) = \tfrac{1}{2}\eta_1 JA_{Ju}\tilde{w}_r, \tag{5-11}$$

so from (2-1), (5-5) and (5-11) we deduce that $h(u,\tilde{w}_r) = JA_{Ju}\tilde{w}_r = 0$.

Now we apply Theorem 4.13 (see also [Li and Wang 2011, Theorem 4.6]) — or, if $\tilde{n} = 0$, Theorem 4.12 (see also [ibid., Theorem 1.6]) — to conclude that M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form or the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form. □

6. Case $\{\mathfrak{C}_m\}_{2\leq m\leq n-1}$ with $k_0 \geq 2$ and $\mathfrak{p} = 0$

Theorem 6.1. *Let $M \subset \mathbb{CP}^n(4)$ be a Lagrangian submanifold in a complex space form with constant holomorphic sectional curvature 4. Suppose that M is not totally geodesic and that M has parallel second fundamental form. Suppose also that $1 \leq \dim \mathfrak{D}_2 = m-1 \leq n-2$, and that k_0 and $\mathfrak{p}$ defined in Section 4 satisfy $k_0 \geq 2$ and $\mathfrak{p} = 0$. Then $n \geq \frac{1}{2}m(m+1)-1$. Moreover, if $n = \frac{1}{2}m(m+1)$, then M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, and if $n \geq \frac{1}{2}m(m+1)+1$, then M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form.*

For the proof we need some observations and a lemma. Suppose M is not totally geodesic. In the present situation, the decomposition (4-6) reduces to $\mathfrak{D}_2 = \{v_1\}\oplus\cdots\oplus\{v_{k_0}\}$. Then $\dim \mathfrak{D}_2 = k_0 = m-1$ and $\{v_1, \ldots, v_{k_0}\}$ forms an orthonormal basis of $\mathfrak{D}_2$.

According to Lemma 3.7 and the fact that for $j \neq l$, $v_j \in V_{v_l}(\tau)$, we have

(6-1) $$\langle L(v_j, v_l), L(v_j, v_l)\rangle = \tau, \quad j \neq l,$$

(6-2) $$\langle L(v_j, v_{l_1}), L(v_j, v_{l_2})\rangle = 0, \quad j, l_1, l_2 \text{ distinct},$$

(6-3) $$\langle L(v_{j_1}, v_{j_2}), L(v_{j_3}, v_{j_4})\rangle = 0, \quad j_1, j_2, j_3, j_4 \text{ distinct}.$$

It follows that $\left\{\frac{1}{\sqrt{\tau}}L(v_j, v_l)\right\}_{1\leq j<l\leq k_0}$ consists of $\frac{1}{2}k_0(k_0-1) = \frac{1}{2}(m-1)(m-2)$ orthonormal vectors. For $\{L(v_j, v_j)\}_{1\leq j\leq k_0}$, we note that

(6-4) $$\langle L(v_j, v_j), L(v_j, v_j)\rangle = \tfrac{1}{2}\lambda_1\eta, \quad 1 \leq j \leq k_0,$$

(6-5) $$\langle L(v_j, v_j), L(v_l, v_l)\rangle = \tfrac{1}{2}\mu\eta, \quad 1 \leq j \neq l \leq k_0,$$

(6-6) $$\langle L(v_j, v_j), L(v_j, v_l)\rangle = 0, \quad 1 \leq j \neq l \leq k_0,$$

(6-7) $$\langle L(v_j, v_j), L(v_{l_1}, v_{l_2})\rangle = 0, \quad 1 \leq j, l_1, l_2 \text{ distinct and } \leq k_0.$$

Then $\{L_j := L(v_1, v_1)+\cdots+L(v_j, v_j) - jL(v_{j+1}, v_{j+1}) \mid 1 \leq j \leq k_0-1\}$ is a set of k_0-1 mutually orthogonal vectors which are all orthogonal to $L(v_j, v_l)$, $j \neq l$. Moreover, we easily have $\langle L_j, L_j\rangle = 2j(j+1)\tau \neq 0$. Hence

(6-8) $$\begin{aligned} w_j &= \frac{1}{\sqrt{2j(j+1)\tau}}L_j, \quad 1 \leq j \leq k_0-1 = m-2,\\ w_{kl} &= \frac{1}{\sqrt{\tau}}L(v_k, v_l), \quad 1 \leq k < l \leq k_0 = m-1, \end{aligned}$$

are $\frac{1}{2}(m-1)(m-2)+(m-2)$ orthonormal vectors in $\mathrm{Im}(L) \subset \mathfrak{D}_3$.

Finally, it is easily known that $\operatorname{Tr} L = L(v_1, v_1) + \cdots + L(v_{k_0}, v_{k_0})$ is orthogonal to the above $\frac{1}{2}(m-1)(m-2) + (m-2)$ vectors and satisfies

(6-9) $$\langle \operatorname{Tr} L, \operatorname{Tr} L\rangle = \tfrac{1}{2}k_0\eta(\lambda_1 + (k_0-1)\mu) =: \rho^2,$$

where $\rho \geq 0$. These results imply that

(6-10) $$\begin{aligned} n &= 1 + \dim \mathfrak{D}_2 + \dim \mathfrak{D}_3 \\ &\geq 1 + (m-1) + \tfrac{1}{2}(m-1)(m-2) + (m-2) = \tfrac{1}{2}m(m+1) - 1. \end{aligned}$$

Lemma 6.2. *We have that* $\operatorname{Tr} L = 0$ *if and only if* $n = \frac{1}{2}m(m+1) - 1$.

Proof. Suppose $\operatorname{Tr} L = 0$, we can first prove that $\mathfrak{D}_3 = \operatorname{Im}(L)$. If not, we can choose a vector $w \in \mathfrak{D}_3$ which is orthogonal to $\operatorname{Im}(L)$, then by (3-7) we get

$$0 = h(\operatorname{Tr} L, w) = (m-1)\mu\eta J w,$$

hence we get $w = 0$ which is a contraction. So we have

$$n = 1 + \dim \mathfrak{D}_2 + \dim \mathfrak{D}_3 = 1 + (m-1) + \tfrac{1}{2}(m+1)(m-2) = \tfrac{1}{2}m(m+1) - 1.$$

On the other hand, suppose that $n = \frac{1}{2}m(m+1) - 1$. By Equation (6-10) we get $\dim \mathfrak{D}_3 = \frac{1}{2}(m-1)(m-2) + (m-2)$ hence $\operatorname{Tr} L = 0$. □

Proof of Theorem 6.1. We need to consider two cases:

(i) $n = \frac{1}{2}m(m+1)$.

(ii) $n \geq \frac{1}{2}m(m+1) + 1$.

We define a unit vector $t = \frac{1}{\rho}\operatorname{Tr} L$.

In case (i), the previous results and particularly (6-9) show that

$$\{t,\ w_{kl|1\leq k<l\leq m-1},\ w_{j|1\leq j\leq m-2}\}$$

is an orthonormal basis of $\operatorname{Im}(L) = \mathfrak{D}_3$. By direct calculations with application of Lemma 3.6, Lemma 4.8 and (6-1)–(6-8), we have:

Lemma 6.3. *Under the above assumptions*, *we have*

(6-11) $$\begin{aligned} &h(t, e_1) = \mu J t, \quad h(t, u) = \frac{\rho}{k_0} J u, \quad h(t, w) = \frac{2\rho}{k_0} J w, \\ &h(t, t) = \mu J e_1 + \Big(\frac{2\rho}{k_0} + \frac{k_0\mu\eta}{\rho}\Big) J t, \end{aligned}$$

where $u = v_i$ *for* $1 \leq i \leq k_0 = m-1$, *and* w *stands for either* w_j *or* w_{kl}, *with* $1 \leq j \leq k_0 - 1 = m-2$ *and* $1 \leq k < l \leq k_0 = m-1$.

Put $T = \alpha e_1 + \beta t$ and $T^* = -\beta e_1 + \alpha t$, where

$$\alpha = \frac{\rho}{\sqrt{\rho^2 + k_0^2 \eta^2}} \quad \text{and} \quad \beta = \frac{k_0 \eta}{\sqrt{\rho^2 + k_0^2 \eta^2}}. \tag{6-12}$$

Then $\{T,\ T^*,\ v_{i|1\le i\le m-1},\ w_{j|1\le j\le m-2},\ w_{kl|1\le k<l\le m-1}\}$ is an orthonormal basis of T_pM. By Lemma 6.3 we easily obtain:

Lemma 6.4. *Under the above assumptions, we have*

$$h(T,T) = \eta_1 JT \quad \textit{and} \quad h(T,u) = \eta_2 Ju, \tag{6-13}$$

where η_1 and η_2 are defined by

$$\eta_1 = \alpha\big(\tfrac{1}{2}\lambda_1 + \eta\big) + \mu/\alpha \quad \textit{and} \quad \eta_2 = \alpha\big(\tfrac{1}{2}\lambda_1 + \eta\big), \tag{6-14}$$

which satisfy the relation

$$\eta_1\eta_2 - \eta_2^2 = \mu(\tfrac{1}{2}\lambda_1 + \eta) = -1, \tag{6-15}$$

where u stands for one of T^*, v_i, w_j, w_{kl} and $1 \le i, k, l \le m-1,\ 1 \le j \le m-2$.

We note that $\eta_1 \ne 2\eta_2$. Otherwise, by definition we have $\mu/\alpha = \alpha(\frac{1}{2}\lambda_1 + \eta)$, then by using the definition of α, ρ and the fact that $\eta \ne 0$ for case $\{\mathfrak{C}_m\}$ we get

$$\lambda_1 + 2\eta = \lambda_1 + \sqrt{\lambda_1^2 + 4} = 0,$$

which cannot happen.

Based on the conclusions of Lemma 6.4, we can apply Theorem 4.12 (see also Theorem 1.6 in [Li and Wang 2011]) to conclude that in case (i) M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form.

In case (ii), we proceed in the same way. We still have that

$$\{t,\ w_{kl|1\le k<l\le m-1},\ w_{j|1\le j\le m-2}\}$$

is an orthonormal basis of $\mathrm{Im}(L)$. But now we no longer have that $\mathrm{Im}(L)$ coincides with $\mathfrak{D}_3$. Denote $\tilde{n} = n - \frac{1}{2}m(m+1)$ and choose $\tilde{w}_1, \dots, \tilde{w}_{\tilde{n}}$ in the orthogonal complement of $\mathrm{Im}(L)$ in $\mathfrak{D}_3$ such that

$$\{t,\ w_{kl|1\le k<l\le m-1}, w_{j|1\le j\le m-2},\ \tilde{w}_{r|1\le r\le \tilde{n}}\}$$

is an orthonormal basis of $\mathfrak{D}_3$. Then, besides (6-11), we further use (3-7) to get

$$h(t, \tilde{w}_r) = \frac{k_0 \mu \eta}{\rho} J\tilde{w}_r, \quad 1 \le r \le \tilde{n}. \tag{6-16}$$

Now we define T and T^* as in case (i). Similarly to Lemma 6.4, we can easily show:

Lemma 6.5. *For case* (ii), *we have*

$$h(T,T)=\eta_1 JT,\quad h(T,u)=\eta_2 Ju \quad \text{and} \quad h(T,\tilde{w}_r)=\eta_3 J\tilde{w}_r, \tag{6-17}$$

for $1\le r\le \tilde{n}$. *Here* η_1 *and* η_2 *are defined by* (6-14) *and* $\eta_3=\mu/\alpha$. *These satisfy the relations* $\eta_2\neq\eta_3,\ 2\eta_2\neq\eta_1\neq 2\eta_3$,

$$\eta_1=\eta_2+\eta_3 \quad \text{and} \quad \eta_2\eta_3=\mu(\eta+\tfrac{1}{2}\lambda_1)=-1, \tag{6-18}$$

where u *is one of* T^*, v_i, w_j, w_{kl} *and* $1\le i,k,l\le m-1,\ 1\le j\le m-2$.

Based on the conclusions of Lemma 6.5, after a similar argument as in the proof of Theorem 5.1, we can apply Theorem 4.13 (see also [Li and Wang 2011, Theorem 4.6]) to conclude that in case (ii) M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form. This completes the proof of Theorem 6.1. □

7. Case $\{\mathfrak{C}_m\}_{2\le m\le n-1}$ with $k_0\ge 2$ and $\mathfrak{p}=1$

Theorem 7.1. *Let* $M\subset\mathbb{CP}^n(4)$ *be a Lagrangian submanifold in a complex space form with constant holomorphic sectional curvature* 4. *Suppose that* M *is not totally geodesic and has parallel second fundamental form. Suppose also that* $1\le\dim\mathfrak{D}_2=m-1\le n-2$, *and* k_0 *and* $\mathfrak{p}$ *defined in Section 4 satisfy* $k_0\ge 2$ *and* $\mathfrak{p}=1$. *Then* $n\ge\frac{1}{4}(m+1)^2-1$. *Moreover, if* $n=\frac{1}{4}(m+1)^2$, *then* M *is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, and if* $n\ge\frac{1}{4}(m+1)^2+1$, *then* M *is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form.*

Lemma 7.2. *Suppose* $\dim\mathfrak{D}_2=m-1\ge 1$, $k_0\ge 2$ *and* $\mathfrak{p}=1$. *Then from the decomposition* (4-6) *there exist unit vectors* $u_j\in V_{v_j}(0),\ 1\le j\le k_0=\frac{1}{2}(m-1)$, *such that the orthonormal basis* $\{v_1,u_1,\dots,v_{k_0},u_{k_0}\}$ *of* $\mathfrak{D}_2$ *satisfies the relations*

$$L(v_j,u_l)=-L(u_j,v_l) \quad \text{and} \quad L(v_j,v_l)=L(u_j,u_l) \tag{7-1}$$

for $1\le j,l\le k_0$.

Proof. We have the decomposition (4-6) with $\dim V_{v_j}(0)=1,\ 1\le j\le k_0$. Let $V_{v_2}(0)=\{u_2\}$, here u_2 is a unit vector.

According to Lemma 4.7, for each $j\neq 2$, we have a unique unit vector u_j in $V_{v_j}(0)$ satisfying

$$L(v_j,-u_2)=L(u_j,v_2) \quad \text{and} \quad L(u_j,u_2)=L(v_j,v_2) \tag{7-2}$$

for $1\le j\le k_0,\ j\neq 2$. The lemma now follows from the following claim. □

Claim 7.3. $L(v_j,u_l)=-L(u_j,v_l)$ and $L(v_j,v_l)=L(u_j,u_l)$ for $1\le j,\ l\le k_0$, $j,l\neq 2$.

Proof. For $j = l$, the fact that $u_j \in V_{v_j}(0)$ implies $L(v_j, u_j) = 0$. It follows that $L(u_j, u_j) = L(v_j, v_j)$.

Next, for $k_0 \geq 3$, we fix $j, l \neq 2$ such that $j \neq l$. By Lemma 4.7, there exists a unique unit vector in $V_{v_j}(0)$, denoted $u_j(l)$, such that

$$L(v_j, u_l) = -L(u_j(l), v_l). \tag{7-3}$$

Since both unit vectors u_j and $u_j(l)$ are in $V_{v_j}(0)$ and $\dim V_{v_j}(0) = 1$, we have $u_j(l) = \epsilon u_j$ with $\epsilon = \pm 1$, which implies that $u_j(l) - \epsilon u_j = 0$ and

$$L(v_j, u_l) = -\epsilon L(u_j, v_l) \quad \text{and} \quad L(v_j, v_l) = \epsilon L(u_j, u_l). \tag{7-4}$$

By using (7-2) and Lemma 4.8, we find that

$$\begin{aligned} h(L(u_j, u_l), L(v_2, u_j)) &= \tau J L(u_l, v_2) = -\tau J L(v_l, u_2) \quad \text{and} \\ h(L(v_j, v_l), L(v_2, u_j)) &= h(L(v_j, v_l), -L(v_j, u_2)) = -\tau J L(v_l, u_2), \end{aligned}$$

which imply

$$0 = h(L(v_j, v_l) - \epsilon L(u_j, u_l), L(v_2, u_j)) = -\tau(1-\epsilon) J L(v_l, u_2). \tag{7-5}$$

Combining equations (7-4) and (7-5) we get $\epsilon = 1$, which completes the proof of the claim. □

Remark 7.4. For $\mathfrak{p} = 1$ we have $\dim \mathfrak{D}_2 = 2k_0$. Denote

$$V_j = \{v_j\} \oplus V_{v_j}(0) = \{v_j\} \oplus \{u_j\}, \quad 1 \leq j \leq k_0.$$

For each $1 \leq j \leq k_0$, we define a linear map $J_0 : V_j \to V_j$ by setting

$$J_0 v_j = u_j \quad \text{and} \quad J_0 u_j = -v_j.$$

Then $J_0 : \mathfrak{D}_2 \to \mathfrak{D}_2$ is an almost complex structure and Lemma 7.2 shows that it satisfies the relations

$$L(J_0 u, v) = -L(u, J_0 v) \quad \text{and} \quad L(J_0 u, J_0 v) = L(u, v) \tag{7-6}$$

for all $u, v \in \mathfrak{D}_2$.

Let $\{v_1, u_1, \dots, v_{k_0}, u_{k_0}\}$ be the orthonormal basis of $\mathfrak{D}_2$ from Lemma 7.2. Combining Lemma 4.4 with the fact that $u_j, v_j \in V_{v_l}(\tau) = V_{u_l}(\tau)$ for $j \neq l$, we have

$$\langle L(v_j, u_l), L(v_j, u_l)\rangle = \langle L(v_j, v_l), L(v_j, v_l)\rangle = \tau, \tag{7-7}$$

for $j \neq l$. Next we get

$$\begin{aligned} \langle L(u_j, v_{l_1}), L(u_j, v_{l_2})\rangle &= \langle L(v_j, u_{l_1}), L(v_j, u_{l_2})\rangle \\ &= \langle L(v_j, v_{l_1}), L(v_j, v_{l_2})\rangle = 0, \end{aligned} \tag{7-8}$$

for j, l_1, l_2 distinct. Then

$$\langle L(v_{j_1}, v_{j_2}), L(v_{j_3}, v_{j_4})\rangle = 0, \qquad j_1, j_2, j_3, j_4 \text{ distinct,} \tag{7-9}$$

$$\langle L(v_j, v_l), L(v_{j_1}, u_{l_1})\rangle = 0, \qquad j \neq l \text{ and } j_1 \neq l_1. \tag{7-10}$$

Thus

$$\left\{\tfrac{1}{\sqrt{\tau}} L(v_j, v_l)\right\}_{1\le j<l\le k_0} \cup \left\{\tfrac{1}{\sqrt{\tau}} L(v_j, u_l)\right\}_{1\le j<l\le k_0}$$

consists of $k_0(k_0 - 1) = \frac{1}{4}(m-1)(m-3)$ orthonormal vectors. For the subset $\{L(v_j, v_j) = L(u_j, u_j)\}_{1\le j\le k_0}$, we note that

$$\langle L(v_j, v_j), L(v_j, v_j)\rangle = \lambda_1\eta/2, \tag{7-11}$$

$$\langle L(v_j, v_j), L(v_l, v_l)\rangle = \mu\eta/2, \tag{7-12}$$

$$\langle L(v_j, v_j), L(v_j, v_l)\rangle = \langle L(v_j, v_j), L(v_j, u_l)\rangle = 0, \tag{7-13}$$

$$\langle L(v_j, v_j), L(v_{l_1}, v_{l_2})\rangle = \langle L(v_j, v_j), L(v_{l_1}, u_{l_2})\rangle = 0, \tag{7-14}$$

where $1 \le j \neq l \le k_0$ and $1 \le j, l_1, l_2$ distinct $\le k_0$.

As in the previous section, we see that

$$\{L_j := L(v_1, v_1) + \cdots + L(v_j, v_j) - jL(v_{j+1}, v_{j+1}) \mid 1 \le j \le k_0 - 1\}$$

are $k_0 - 1 = \frac{1}{2}(m-3)$ mutually orthogonal vectors which are orthogonal to all $L(v_j, v_l)$ and $L(v_j, u_l)$, $j \neq l$. We also easily have $\langle L_j, L_j\rangle = 2j(j+1)\tau \neq 0$. Hence

$$\begin{cases} w_j = \frac{1}{\sqrt{2j(j+1)\tau}} L_j, & 1 \le j \le k_0 - 1 = \frac{1}{2}(m-3), \\ w_{kl} = \frac{1}{\sqrt{\tau}} L(v_k, v_l), & 1 \le k < l \le k_0 = \frac{1}{2}(m-1), \\ w'_{kl} = \frac{1}{\sqrt{\tau}} L(v_k, u_l), & 1 \le k < l \le k_0 = \frac{1}{2}(m-1), \end{cases} \tag{7-15}$$

are $\frac{1}{4}(m+1)(m-3)$ orthonormal vectors in $\mathrm{Im}(L) \subset \mathfrak{D}_3$.

Finally, it is easily verified that $\frac{1}{2}\mathrm{Tr}\, L = L(v_1, v_1) + \cdots + L(v_{k_0}, v_{k_0})$ is orthogonal to the above $(m+1)(m-3)/4$ vectors and satisfies

$$\tfrac{1}{4}\langle \mathrm{Tr}\, L, \mathrm{Tr}\, L\rangle = \tfrac{1}{2}k_0\eta(\lambda_1 + (k_0 - 1)\mu) =: \rho^2, \qquad \rho \ge 0. \tag{7-16}$$

Similarly as in the previous section we get that

Lemma 7.5. *We have* $\mathrm{Tr}\, L = 0$ *if and only if* $n = \frac{1}{4}(m+1)^2 - 1$.

Proof of Theorem 7.1. We define a unit vector $t = \frac{1}{2\rho}\mathrm{Tr}\, L$. Again we need to consider two cases.

(i) $n = \frac{1}{4}(m+1)^2$. The previous results show that the set $\{t, w_{kl}, w'_{kl}, w_j\}$, where we have $1 \le k < l \le \frac{1}{2}(m-1)$ and $1 \le j \le \frac{1}{2}(m-3)$, is an orthonormal basis of $\mathrm{Im}(L) = \mathfrak{D}_3$. By direct calculations applying Lemma 3.6, Lemma 4.8 and (7-7)-(7-14) we obtain again the expressions of (6-11) for $u = v_i, u_i$ and $w = w_j, w_{kl}, w'_{kl}$ with $1 \le i, k, l \le \frac{1}{2}(m-1)$ and $1 \le j \le \frac{1}{2}(m-3)$. Proceeding then in the same way as before, we can again apply Theorem 4.12 (see also [Li and Wang 2011, Theorem 1.6]) to conclude that in this case, M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form.

(ii) $n \ge \frac{1}{4}(m+1)^2 + 1$. Here we see that $\{t, w_{kl}, w'_{kl}, w_j\}$, where j, k, l are as before, is still an orthonormal basis of $\mathrm{Im}(L)$. But now $\mathrm{Im}(L) \subsetneq \mathfrak{D}_3$. Introduce the notation

$$\tilde{n} = n - \tfrac{1}{4}(m+1)^2 \ge 1$$

and choose $w'_1, \dots, w'_{\tilde{n}}$ in the orthogonal complement of $\mathrm{Im}(L)$ in $\mathfrak{D}_3$, such that

$$\{t,\ w_{kl}, w'_{kl}, w_j, w'_r\}$$

where j, k, l are as before and $1 \le r \le \tilde{n}$, is an orthonormal basis of $\mathfrak{D}_3$. Then (3-7) gives that

$$h(t, w'_r) = \frac{k_0 \mu \eta}{2\rho} J w'_r, \quad 1 \le r \le \tilde{n}, \tag{7-17}$$

and we can again proceed exactly as in the previous section to conclude that in this case, M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form. □

8. Case $\mathfrak{C}_m$ $(2 \le m \le n-1)$ with $k_0 \ge 2$ and $\mathfrak{p} = 3$

Theorem 8.1. *Let $M \subset \mathbb{CP}^n(4)$ be a Lagrangian submanifold in a complex space form with constant holomorphic sectional curvature* 4. *Suppose that M is not totally geodesic and that it has parallel second fundamental form. Suppose also that* $1 \le \dim \mathfrak{D}_2 = m - 1 \le n - 2$, *and k_0 and $\mathfrak{p}$ defined in Section 4 satisfy $k_0 \ge 2$ and* $\mathfrak{p} = 3$. *Then $n \ge \frac{1}{8}(m-1)(m+5)$. If $n = \frac{1}{8}(m-1)(m+5)+1$, then M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, and if $n \ge \frac{1}{8}(m-1)(m+5)+2$, then M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form.*

Lemma 8.2. *Suppose* $\dim \mathfrak{D}_2 = m - 1 \ge 1$, $k_0 \ge 2$ *and* $\mathfrak{p} = 3$. *Then from the decomposition* (4-6) *there exist unit orthogonal vectors*

$$x_j, y_j, z_j \in V_{v_j}(0), \quad 1 \le j \le k_0 = \tfrac{1}{4}(m-1),$$

such that the orthonormal basis $\{v_1, x_1, y_1, z_1, \ldots, v_{k_0}, x_{k_0}, y_{k_0}, z_{k_0}\}$ *of* $\mathfrak{D}_2$ *satisfies*

$$\begin{aligned} L(x_j, x_l) &= L(y_j, y_l) = L(z_j, z_l) = L(v_j, v_l), \\ L(v_j, x_l) &= -L(x_j, v_l) = -L(y_j, z_l) = L(y_l, z_j), \\ L(v_j, y_l) &= -L(y_j, v_l) = -L(z_j, x_l) = L(x_j, z_l), \\ L(v_j, z_l) &= -L(z_j, v_l) = -L(x_j, y_l) = L(x_l, y_j), \end{aligned} \tag{8-1}$$

for all $1 \le j, l \le k_0$.

Proof. We use the decomposition (4-6) with $\dim V_{v_j}(0) = 3$ for $1 \le j \le k_0$.

Denote $V_j = \{v_j\} \oplus V_{v_j}(0)$. First we choose arbitrary orthonormal vectors $x_1, y_1 \in V_{v_1}(0)$, next by using Lemma 4.6 and Lemma 4.7 we can first find unit vectors $x_j, y_j \in V_{v_j}(0)$, $j \ge 2$ such that

$$\begin{aligned} &L(x_j, x_1) = L(y_j, y_1) = L(v_j, v_1), && L(v_j, x_1) = -L(x_j, v_1), \\ &L(v_j, y_1) = -L(y_j, v_1), && L(x_1, y_j) = -L(x_j, y_1). \end{aligned} \tag{8-2}$$

Next we choose z_j, z_1^j such that $L(v_j, z_1^j) = -L(z_j, v_1) = -L(x_j, y_1)$. By using the Cauchy–Schwarz inequality, we have

$$\begin{aligned} L(x_j, x_1) &= L(y_j, y_1) = L(z_j, z_1^j) = L(v_j, v_1), \\ L(v_j, x_1) &= -L(x_j, v_1) = -L(y_j, z_1^j) = L(y_1, z_j), \\ L(v_j, y_1) &= -L(y_j, v_1) = -L(z_j, x_1) = L(x_j, z_1^j), \\ L(v_j, z_1^j) &= -L(z_j, v_1) = -L(x_j, y_1) = L(x_1, y_j). \end{aligned} \tag{8-3}$$

Claim 8.3. For all $j \ge 2$, the families $\{x_1, y_1, z_1^j\}$ and $\{x_j, y_j, z_j\}$ of (8-3) are orthonormal bases of $V_{v_1}(0)$ and $V_{v_j}(0)$, respectively.

Proof of claim. In fact, from (8-3) we have

$$\begin{aligned} \tau\langle z_1^j, x_1\rangle &= \langle L(v_j, z_1^j), L(v_j, x_1)\rangle = \langle L(x_j, -y_1), L(x_j, -v_1)\rangle = \tau\langle y_1, v_1\rangle = 0, \\ \tau\langle z_1^j, y_1\rangle &= \langle L(v_j, z_1^j), L(v_j, y_1)\rangle = \langle L(y_j, -x_1), L(y_j, -v_1)\rangle = \tau\langle x_1, v_1\rangle = 0, \end{aligned}$$

hence we get $\{x_1, y_1, z_1^j\}$ is an orthonormal basis of $V_{v_1}(0)$.

For $j \ge 2$, from (8-3) we have

$$\tau\langle x_j, y_j\rangle = \langle L(v_1, x_j), L(v_1, y_j)\rangle = \langle L(v_j, -x_1), L(v_j, -y_1)\rangle = \tau\langle x_1, y_1\rangle = 0,$$

similarly, we get $\langle x_j, z_j\rangle = \langle x_1, z_1^j\rangle = 0$ and $\langle y_j, z_j\rangle = \langle y_1, z_1^j\rangle = 0$. This completes the proof. □

Claim 8.4. The vectors z_1^j and z_1^l of (8-3) are equal for all $2 \le j, l \le k_0$. If we denote this common value by z_1, then (8-1) holds.

Proof of claim. By Claim 8.3, we know that for $j \neq l$, $j, l \geq 2$ we have $z_1^j = \varepsilon_{jl} z_1^l$ with $\varepsilon_{jl} = \pm 1$. From Lemma 4.8 and (8-3) we get

$$\begin{aligned} \varepsilon_{jl} \tau JL(v_j, v_l) &= h(L(v_j, z_1^j), L(v_l, z_1^l)) \\ &= h(L(x_j, y_1), L(x_l, y_1)) = \tau JL(x_j, x_l). \end{aligned} \tag{8-4}$$

Similarly, we get

$$\begin{aligned} &\varepsilon_{jl} L(v_j, v_l) = L(y_j, y_l) = L(z_j, z_l), \\ &\varepsilon_{jl} L(x_j, x_l) = L(y_j, y_l) = L(z_j, z_l) = L(v_j, v_l). \end{aligned} \tag{8-5}$$

From (8-4) and (8-5) we get $\varepsilon_{jl} = 1$ and

$$L(v_j, v_l) = L(x_j, x_l) = L(y_j, y_l) = L(z_j, z_l), \quad j \neq l,\ j, l \geq 2. \tag{8-6}$$

Let $z_1 = z_1^2 = \cdots = z_1^{k_0}$, then by (8-3) and Lemma 4.8 we have

$$\begin{aligned} \tau JL(x_j, y_l) &= h(L(y_1, x_j), L(y_1, y_l)) \\ &= h(L(v_1, z_j), L(v_1, v_l)) = \tau JL(z_j, v_l). \end{aligned} \tag{8-7}$$

From (8-6) and (8-7), and by using Lemma 4.6 and Lemma 4.7 we get that (8-1) holds. □

Combining the above claims completes the proof of the lemma. □

Remark 8.5. Having fixed the orthonormal basis of $\mathfrak{D}_2$ satisfying (8-1), we can now define three almost complex structures $J_1, J_2, J_3 : \mathfrak{D}_2 \to \mathfrak{D}_2$ such that for all $1 \leq j \leq k_0$,

$$\begin{aligned} J_1 v_j &= x_j, & J_2 v_j &= y_j, & J_3 v_j &= z_j, \\ J_1 x_j &= -v_j, & J_2 y_j &= -v_j, & J_3 z_j &= -v_j, \end{aligned} \tag{8-8}$$

and furthermore J_1, J_2 and J_3 satisfy

$$J_1 \circ J_1 = J_2 \circ J_2 = J_3 \circ J_3 = -\operatorname{id} \quad \text{and} \quad J_1 J_2 = -J_2 J_1 = J_3. \tag{8-9}$$

Then we define a quaternionic structure $\{J_1, J_2, J_3\}$ on $\mathfrak{D}_2$. It is important to remark that (8-1) is equivalent to the relations

$$L(J_s u, v) = -L(u, J_s v) \quad \text{and} \quad L(J_s u, J_s v) = L(u, v) \tag{8-10}$$

for all $s = 1, 2, 3$ and $u, v \in \mathfrak{D}_2$.

We have $m - 1 = 4k_0$ and $k_0 \geq 2$. Let $\{v_1, x_1, y_1, z_1, \ldots, v_{k_0}, x_{k_0}, y_{k_0}, z_{k_0}\}$ be an orthonormal basis of $\mathfrak{D}_2$ as constructed in Lemma 8.2. Applying Lemma 4.4 and the fact that for $j \neq l$, $v_j, x_j, y_j, z_j \in V_{v_l}(\tau) = V_{x_l}(\tau) = V_{y_l}(\tau) = V_{z_l}(\tau)$, we easily show that

$$\begin{aligned}(8\text{-}11)\quad \langle L(v_j,x_l),L(v_j,x_l)\rangle &= \langle L(v_j,y_l),L(v_j,y_l)\rangle\\ &= \langle L(v_j,z_l),L(v_j,z_l)\rangle = \langle L(v_j,v_l),L(v_j,v_l)\rangle = \tau,\end{aligned}$$

for $j \neq l$. We also get

$$\begin{aligned}(8\text{-}12)\quad \langle L(x_j,v_{l_1}),L(x_j,v_{l_2})\rangle &= \langle L(v_j,x_{l_1}),L(v_j,x_{l_2})\rangle = \langle L(y_j,v_{l_1}),L(y_j,v_{l_2})\rangle\\ &= \langle L(v_j,y_{l_1}),L(v_j,y_{l_2})\rangle = \langle L(z_j,v_{l_1}),L(z_j,v_{l_2})\rangle\\ &= \langle L(v_j,z_{l_1}),L(v_j,z_{l_2})\rangle = \langle L(v_j,v_{l_1}),L(v_j,v_{l_2})\rangle\\ &= 0,\end{aligned}$$

for j, l_1, l_2 distinct. Next we get

$$\begin{aligned}(8\text{-}13)\quad \langle L(v_{j_1},v_{j_2}),L(v_{j_3},v_{j_4})\rangle &= \langle L(v_{j_1},x_{j_2}),L(v_{j_3},x_{j_4})\rangle = \langle L(v_{j_1},y_{j_2}),L(v_{j_3},y_{j_4})\rangle\\ &= \langle L(v_{j_1},z_{j_2}),L(v_{j_3},z_{j_4})\rangle = 0,\end{aligned}$$

for j_1, j_2, j_3, j_4 distinct, and then

$$\begin{aligned}(8\text{-}14)\qquad \langle L(v_j, v_l), L(v_{j_1}, x_{l_1})\rangle &= \langle L(v_j, v_l), L(v_{j_1}, y_{l_1})\rangle\\ &= \langle L(v_j, v_l), L(v_{j_1}, z_{l_1})\rangle = 0,\end{aligned}$$

for $j \neq l$ and $j_1 \neq l_1$.

For $\{L(v_j, v_j) = L(x_j, x_j) = L(y_j, y_j) = L(z_j, z_j)\}_{1\leq j\leq k_0}$, we note that

$$(8\text{-}15)\qquad \langle L(v_j, v_j), L(v_j, v_j)\rangle = \tfrac{1}{2}\lambda_1\eta,$$

$$(8\text{-}16)\qquad \langle L(v_j, v_j), L(v_l, v_l)\rangle = \frac{n+1}{4(n-i)}\lambda_1^2 - 2\tau = \tfrac{1}{2}\mu\eta,$$

$$(8\text{-}17)\qquad \langle L(v_j, v_j), L(v_j, v_l)\rangle = \langle L(v_j, v_j), L(v_j, u_l)\rangle = 0,$$

$$(8\text{-}18)\qquad \langle L(v_j, v_j), L(v_{l_1}, v_{l_2})\rangle = \langle L(v_j, v_j), L(v_{l_1}, u_{l_2})\rangle = 0,$$

for $1 \leq j, l, l_1, l_2 \leq k_0$ distinct. Similarly to the previous section, we deduce that

$$\{L_j := L(v_1, v_1) + \cdots + L(v_j, v_j) - jL(v_{j+1}, v_{j+1}) \mid 1 \leq j \leq k_0 - 1\}$$

are $k_0 - 1 = \frac{1}{4}(m-5)$ mutually orthogonal vectors which are orthogonal to all of the vectors $L(v_j, v_l)$, $L(v_j, x_l)$, $L(v_j, y_l)$, and $L(v_j, z_l)$, where $j \neq l$. Also, we have $\langle L_j, L_j\rangle = 2j(j+1)\tau \neq 0$. Hence the vectors

$$w_j = \frac{1}{\sqrt{2j(j+1)\tau}}L_j, \quad w_{kl} = \frac{1}{\sqrt{\tau}}L(v_k, v_l),$$

$$w^1_{kl} = \frac{1}{\sqrt{\tau}}L(v_k, x_l), \quad w^2_{kl} = \frac{1}{\sqrt{\tau}}L(v_k, y_l), \quad w^3_{kl} = \frac{1}{\sqrt{\tau}}L(v_k, z_l),$$

where $1 \leq j \leq k_0 = \frac{1}{4}(m-1)$ and $1 \leq k < l \leq k_0$, comprise $2k_0(k_0-1) + k_0 - 1 = \frac{1}{8}(m+1)(m-5)$ orthonormal vectors in $\mathrm{Im}(L) \subset \mathfrak{D}_3$.

Finally, from Lemma 8.2, (8-15) and (8-16) it is easily known that the vector

$$\operatorname{Tr} L = 4\big(L(v_1, v_1) + \cdots + L(v_{k_0}, v_{k_0})\big)$$

is orthogonal to the above $\frac{1}{8}(m+1)(m-5)$ vectors and satisfies

$$\frac{1}{16}\langle \operatorname{Tr} L, \operatorname{Tr} L\rangle = \tfrac{1}{2}k_0\eta(\lambda_1 + (k_0-1)\mu) =: \rho^2, \quad \rho \geq 0. \tag{8-19}$$

The above results imply that

$$n = 1 + \dim \mathcal{D}_2 + \dim \mathcal{D}_3 \geq 1 + (m-1) + \tfrac{1}{8}(m+1)(m-5) = \tfrac{1}{8}(m-1)(m+5).$$

Lemma 8.6. *We have* $\operatorname{Tr} L = 0$ *if and only if* $n = \frac{1}{8}(m-1)(m+5)$.

Proof of Theorem 8.1. We need to consider two cases:

(i) $n = \frac{1}{8}(m-1)(m+5) + 1$.

(ii) $n \geq \frac{1}{8}(m-1)(m+5) + 2$.

In case (i), we have that

$$\{t,\ w_j\,|_{1\leq j\leq (i-5)/4},\ w_{kl},\ w^1_{kl},\ w^2_{kl},\ w^3_{kl}\,|_{1\leq k<l\leq (i-1)/4}\}$$

is an orthonormal basis of $\operatorname{Im}(L) = \mathcal{D}_3$. In case (ii), in order to have an orthonormal basis we still need to add an orthonormal basis of $\mathcal{D}_{32}$.

As in the previous sections, we get that (6-13) in case (i) and (6-17) in case (ii) are satisfied. Consequently we deduce that in case (i), M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, and in case (ii), we deduce that M is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form. This completes the proof. □

9. Case $\{\mathfrak{C}_m\}_{2\leq m\leq n-1}$ with $k_0 \geq 2$ and $\mathfrak{p} = 7$

Theorem 9.1. *Let* $M \subset \mathbb{CP}^n(4)$ *be a Lagrangian submanifold in a complex space form with constant holomorphic sectional curvature* 4. *Suppose that* M *is not totally geodesic and has parallel second fundamental form. Suppose also that* $1 \leq \dim \mathcal{D}_2 = m-1 \leq n-2$ *and* k_0 *and* $\mathfrak{p}$ *defined in Section 4 satisfy* $k_0 \geq 2$ *and* $\mathfrak{p} = 7$. *Then* $k_0 = 2$ *and* $m = 17$, *which implies that* $n \geq 26$. *Moreover, if* $n = 27$ *we have that* M *is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, and if* $n \geq 28$, *then* M *is locally the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form.*

Lemma 9.2. *Suppose* $\dim \mathcal{D}_2 = m-1 \geq 1, k_0 \geq 2$ *and* $\mathfrak{p} = 7$. *Then from the decomposition* (4-6), *if* $k_0 \geq 2$, *we can choose an orthonormal basis* $\{x_j\}_{1\leq j\leq 7}$

for $V_{v_1}(0)$ *and an orthonormal basis* $\{y_j\}_{1\le j\le 7}$ *for* $V_{v_2}(0)$ *so that by identifying* $e_j(v_1)=x_j$ *and* $e_j(v_2)=y_j$, *we have the relations*

$$L(e_j(v_1), e_l(v_2)) = -L(v_1, e_j e_l(v_2)) = -L(e_l e_j(v_1), v_2), \tag{9-1}$$

for $1\le j, l\le 7$, *where the product is defined by the following multiplication table*:

$\cdot$	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_1	$-\mathrm{id}$	e_3	$-e_2$	e_5	$-e_4$	$-e_7$	e_6
e_2	$-e_3$	$-\mathrm{id}$	e_1	e_6	e_7	$-e_4$	$-e_5$
e_3	e_2	$-e_1$	$-\mathrm{id}$	e_7	$-e_6$	e_5	$-e_4$
e_4	$-e_5$	$-e_6$	$-e_7$	$-\mathrm{id}$	e_1	e_2	e_3
e_5	e_4	$-e_7$	e_6	$-e_1$	$-\mathrm{id}$	$-e_3$	e_2
e_6	e_7	e_4	$-e_5$	$-e_2$	e_3	$-\mathrm{id}$	$-e_1$
e_7	$-e_6$	e_5	e_4	$-e_3$	$-e_2$	e_1	$-\mathrm{id}$

Proof. Let $k_0\ge 2$ and suppose we have the decomposition (4-6) with $\dim V_{v_j}(0)=7$ $(1\le j\le k_0)$.

Denote $V_j=\{v_j\}\oplus V_{v_j}(0)$. First we choose arbitrary orthonormal vectors $x_1, x_2\in V_{v_1}(0)$. Next we can use Lemma 4.6 and Lemma 4.7 to consecutively find unit vectors $y_1, y_2\in V_{v_2}(0)$, $x_3\in V_{v_1}(0)$ and $y_3\in V_{v_2}(0)$ satisfying

$$L(y_1, v_1) = -L(x_1, v_2), \qquad L(y_2, v_1) = -L(x_2, v_2), \tag{9-2}$$

$$L(y_1, x_2) = -L(v_2, x_3), \qquad L(y_3, v_1) = -L(x_3, v_2). \tag{9-3}$$

Now we pick an arbitrary unit vector $x_4\in V_{v_1}(0)$ so that it is orthogonal to all x_1, x_2 and x_3. Then we can take unit vectors $x_5, x_6, x_7\in V_{v_1}(0)$ and unit vectors $y_4, y_5, y_6, y_7\in V_{v_2}(0)$ inductively such that the following hold:

$$L(x_4,y_1)=-L(y_4,x_1)=-L(v_2,x_5)=L(v_1,y_5), \tag{9-4}$$

$$L(x_4,y_2)=-L(v_2,x_6)=L(v_1,y_6), \quad L(x_4,y_3)=-L(v_2,x_7)=L(v_1,y_7). \tag{9-5}$$

From the previous equations, together with the isotropy conditions and the Cauchy–Schwarz inequality, it immediately follows that $L(x_i, y_i)=L(v_1, v_2)$, for $i=1,\dots 7$. Applying once more the same properties it also follows that $L(x_i, y_j)=-L(x_j, y_i)$ and $L(x_i, v_2)=-L(y_i, v_1)$.

From (9-3) and (9-4) it additionally follows that

$$\begin{aligned} L(y_1, x_3) &= L(x_2, v_2), & L(x_4, v_2) &= L(x_5, y_1),\\ L(x_4, y_5) &= -L(v_1, y_1), & L(x_4, v_2) &= L(x_6, y_2),\\ L(x_4, y_6) &= -L(v_1, y_2), & L(x_4, v_2) &= L(x_7, y_3),\\ L(x_4, y_7) &= -L(v_1, y_3). \end{aligned}$$

Hence $L(x_4, v_2) = L(x_5, y_1) = L(x_6, y_2) = L(x_7, y_3)$. Repeating now the same procedure on the newly found identities shows that L has the desired form.

Finally note that the fact that $\{v_1, x_1, \dots, x_7\}$ and $\{v_2, y_1, \dots, y_7\}$ are orthonormal can be seen as follows. First, we have

$$\tau\langle x_1, x_3\rangle = \langle L(v_2, x_3), L(v_2, x_1)\rangle = \langle L(x_1, y_2), L(v_2, x_1)\rangle = \tau\langle v_2, y_2\rangle = 0.$$

The other equations are obtained similarly. □

Lemma 9.3. *Suppose* $\dim \mathcal{D}_2 = m-1 \geq 1$ *and* $\mathfrak{p} = 7$. *If* $k_0 \geq 2$ *in the decomposition* (4-6), *then in fact* $k_0 = 2$.

Proof. Suppose on the contrary that $k_0 \geq 3$. To choose a basis for $V_{v_3}(0)$, we follow the same ideas as in Lemma 9.2 for $V_{v_1}(0)$ and $V_{v_2}(0)$. Let x_1, x_2, x_3 be given as in Lemma 9.2, then we have unique unit vectors $z_1, z_2 \in V_{v_3}(0)$ and $\tilde{x}_3 \in V_{v_1}(0)$ that satisfy

$$L(z_1, v_1) = -L(x_1, v_3), \quad L(z_2, v_1) = -L(x_2, v_3) \quad \text{and} \quad L(z_1, x_2) = -L(v_3, \tilde{x}_3).$$

Now we pick an arbitrary unit vector $x_4 \in V_{v_1}(0)$ so that it is orthogonal to x_1, x_2, x_3 and $\tilde{x}_3$. Then we can choose unit vectors $\tilde{x}_5, \tilde{x}_6, \tilde{x}_7 \in V_{v_1}(0)$ and vectors $z_3, z_4, z_5, z_6, z_7 \in V_{v_3}(0)$ inductively by the following conditions:

$$\begin{aligned} &L(z_3, v_1) = -L(\tilde{x}_3, v_3),\\ &L(x_4, z_2) = -L(v_3, \tilde{x}_6) = L(v_1, z_6),\\ &L(x_4, z_3) = -L(v_3, \tilde{x}_7) = L(v_1, z_7),\\ &L(x_4, z_1) = -L(z_4, x_1) = -L(v_3, \tilde{x}_5) = L(v_1, z_5). \end{aligned}$$

Then, similarly to the proof of Lemma 9.2, we get that $\{z_1, z_2, z_3, z_4, z_5, z_6, z_7\}$ forms an orthonormal basis of $V_{v_3}(0)$ together with the relations between inner products of L:

$$\text{(9-6)} \quad L(e_j(v_1), e_l(v_3)) = -L(v_1, e_j e_l(v_3)) = -L(e_l e_j(v_1), v_3), \quad 1 \leq j, l \leq 7,$$

where $e_j e_l$ denotes a product defined by the multiplication table in Lemma 9.2.

We have two orthonormal bases of $V_{v_1}(0)$, namely $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ and $\{x_1, x_2, \tilde{x}_3, x_4, \tilde{x}_5, \tilde{x}_6, \tilde{x}_7\}$. We first show that $\tilde{x}_i = x_i$ for $i = 3, 5, 6, 7$:

By Lemma 4.8 and the relations between the inner products of L, we get

$$\tau L(y_1, z_1) = h(L(y_1, x_2), L(z_1, x_2)) = h(L(v_1, y_3), L(v_1, z_3)) = \tau L(y_3, z_3).$$

Similarly, we get $L(v_2, v_3) = L(y_j, z_j)$ for $j = 1, \dots, 7$.

Since $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ and $\{x_1, x_2, \tilde{x}_3, x_4, \tilde{x}_5, \tilde{x}_6, \tilde{x}_7\}$ are two orthonormal bases for $V_{v_1}(0)$, we may assume that $x_3 = b_3\tilde{x}_3 + b_5\tilde{x}_5 + b_6\tilde{x}_6 + b_7\tilde{x}_7$. Then by Lemma 4.8 and the relations between the inner products of L, we get

$$\begin{aligned}\tau L(y_2, z_2) &= h(L(v_1, y_2), L(v_1, z_2)) \\ &= -h(L(x_3, y_1), L(v_1, z_2)) \\ &= b_3 h(L(\tilde{x}_3, y_1), L(\tilde{x}_3, z_1)) + b_5 h(L(\tilde{x}_5, y_1), L(\tilde{x}_5, z_7)) \\ &\qquad - b_6 h(L(\tilde{x}_6, y_1), L(\tilde{x}_6, z_4)) - b_7 h(L(\tilde{x}_7, y_1), L(\tilde{x}_7, z_5)) \\ &= b_3 \tau L(y_1, z_1) + b_5 \tau L(y_1, z_7) - b_6 \tau L(y_1, z_4) - b_7 \tau L(y_1, z_5).\end{aligned}$$

By the relations between the inner products of L, we get $L(y_1, z_1) = L(y_2, z_2)$, and that $L(y_1, z_4)$, $L(y_1, z_5)$ and $L(y_1, z_7)$ are orthogonal to each other. Hence we get $b_3 = 1$, $b_5 = b_6 = b_7 = 0$ and $x_3 = \tilde{x}_3$. By a similar argument, we can prove that $\tilde{x}_i = x_i$ for $i = 5, 6, 7$.

In order to complete the proof of Lemma 9.3, we will first use (9-1) and (9-6) to show that we have also similar relations between the spaces $V_2 = \{v_2\} \oplus V_{v_2}(0)$ and $V_3 = \{v_3\} \oplus V_{v_3}(0)$, that is,

$$\text{(9-7)} \quad L(e_j(v_2), e_l(v_3)) = -L(v_2, e_j e_l(v_3)) = -L(e_l e_j(v_2), v_3), \quad 1 \le j, l \le 7,$$

where $e_j e_l$ denotes a product defined by the multiplication table in Lemma 9.2.

For $j = l$, by Lemma 4.8, (9-1) and (9-6) we have

$$\begin{aligned}\tau JL(e_j(v_2), e_j(v_3)) &= h(L(e_j(v_2), e_k(v_1)), L(e_j(v_3), e_k(v_1))) \\ &= h(-L(e_j e_k(v_1), v_2), -L(e_j e_k(v_1), v_3)) = \tau JL(v_2, v_3).\end{aligned}$$

For $j \ne l$, from the table in Lemma 9.2 we have that there exists a unique k such that $e_l e_j = \epsilon e_k$, $e_j e_k = \epsilon e_l$, $e_k e_l = \epsilon e_j$, where ϵ is 1 or -1. Then by Lemma 4.8, (9-1) and (9-6) we have

$$\begin{aligned}\tau JL(e_j(v_2), e_l(v_3)) &= h(L(e_j(v_2), v_1), L(e_l(v_3), v_1)) \\ &= h(L(-\epsilon e_l e_k(v_2), v_1), L(e_l(v_3), v_1)) \\ &= \epsilon h(L(e_l(v_1), e_k(v_2)), -L(e_l(v_1), v_3)) \\ &= -\epsilon \tau JL(e_k(v_2), v_3) = -\tau L(e_l e_j(v_2), v_3)\end{aligned}$$

and

$$\begin{aligned}\tau JL(v_2, e_j e_l(v_3)) &= h\big(L(v_2, e_k(v_1)), L(e_j e_l(v_3), e_k(v_1))\big) \\ &= h\big(L(v_2, \epsilon e_l e_j(v_1)), L(-\epsilon e_k(v_3), e_k(v_1)))\big) \\ &= h\big(L(v_1, -\epsilon e_l e_j(v_2)), L(-\epsilon v_3, v_1))\big) = \tau JL(e_l e_j(v_2), v_3).\end{aligned}$$

From (9-1), (9-6), (9-7) and Lemma 4.8 we have

$$\text{(9-8)} \qquad h(L(v_1, y_6) + L(x_1, y_7), L(x_2, v_3)) = 0.$$

On the other hand, we have

$$h(L(v_1, y_6), L(x_2, v_3)) = h(L(v_1, y_6), -L(v_1, z_2)) = -\tau JL(y_6, z_2),$$
$$h(L(x_1, y_7), L(x_2, v_3)) = h(L(x_1, y_7), -L(x_1, z_3)) = -\tau JL(y_7, z_3).$$

These together with (9-8) give that

$$L(y_6, z_2) + L(y_7, z_3) = 0. \tag{9-9}$$

From (9-7) we have $L(y_6, z_2) = L(y_7, z_3)$. We also have that

$$\langle L(y_6, z_2), L(y_6, z_2)\rangle = \tau,$$

so we get a contradiction with (9-9). This completes the proof. □

Proof of Theorem 9.1. By Lemma 9.3, we have $k_0 = 2$, $m = 8k_0 + 1 = 17$ and $\dim \mathfrak{D}_2 = m - 1 = 16$.

Let $\{v_1, v_2, x_j, y_j \mid 1 \le j \le 7\}$ be the orthonormal basis of $\mathfrak{D}_2$ as constructed in Lemma 9.2, whose elements satisfy (9-1). Define $L_1 = L(v_1, v_1) - L(v_2, v_2)$, then direct calculation shows that

$$\langle L_1, L_1\rangle = 4\tau \neq 0. \tag{9-10}$$

We now easily see that the nine vectors

$$w_0 = \frac{1}{2\sqrt{\tau}} L_1, \quad w_1 = \frac{1}{\sqrt{\tau}} L(v_1, v_2) \quad \text{and} \quad w_{j+1} = \frac{1}{\sqrt{\tau}} L(v_1, y_j), \quad 1 \le j \le 7,$$

in $\operatorname{Im}(L) \subset \mathfrak{D}_3$ are orthonormal one to another.

Note that $\operatorname{Tr} L = 8(L(v_1, v_1) + L(v_2, v_2))$ is orthogonal to the above nine vectors. Using (3-3) and (3-4), the vector $\operatorname{Tr} L$ obviously satisfies

$$\tfrac{1}{64}\langle \operatorname{Tr} L, \operatorname{Tr} L\rangle = \tfrac{1}{2} k_0 \eta(\lambda_1 + (k_0 - 1)\mu) = \eta(\lambda_1 + \mu) =: \rho^2, \quad \rho \ge 0.$$

Then we have the conclusion

$$n = 1 + \dim \mathfrak{D}_2 + \dim \mathfrak{D}_3 \ge 1 + 16 + 9 = 26,$$

and as proved in previous sections we see that $n = 26$ if and only if $\operatorname{Tr} L = 0$.

When $n = 27$ or $n \ge 28$, we can still define a unit vector $t = \frac{1}{8\rho} \operatorname{Tr} L$. As before we get the same expressions as in Lemma 6.3, 6.4 and 6.5 which allows us to conclude that M is locally the Calabi product of a point with a lower-dimensional Lagrangian submanifold with parallel second fundamental form, or the Calabi product of two lower-dimensional Lagrangian submanifolds with parallel second fundamental form. □

10. The remaining cases

In this section we will complete the proof of the classification theorem. Let $k = k_0 + 1$, we will show that if M is neither totally geodesic nor can be decomposed as a Calabi product then one of the following applies:

(i) $n = \frac{1}{2}k(k+1) - 1,\ k \geq 3$, and M is congruent with $\mathrm{SU}(k)/\mathrm{SO}(k)$.

(ii) $n = k^2 - 1$, $k \geq 3$, and M is congruent with $\mathrm{SU}(k)$.

(iii) $n = 2k^2 - k - 1,\ k \geq 3$, and M is congruent with $\mathrm{SU}\big(2k\big)/\mathrm{Sp}(k)$.

(iv) $n = 26$ and M is congruent with $\mathrm{E}_6/\mathrm{F}_4$.

From Naitoh [1981b; 1983a; 1983b] we see that there indeed exist parallel immersions of the above spaces of the previously mentioned dimensions into the complex projective space.

From the previous remaining sections, each of the resulting cases corresponds to one of the cases $\mathfrak{p} = 0, 1, 3, 7$ with $\mathcal{D}_{32} = \{0\}$ (from Lemma 6.2, Lemma 7.5, Lemma 8.6 and the arguments in Section 9) and $\operatorname{Tr} L$ vanishing. Note that in each of the above cases, the vanishing of $\operatorname{Tr} L$ allows to determine λ_1 explicitly. We also have in each of the cases a basis and we can compute the components of the second fundamental form from Lemmas 3.2, 3.4 and 4.8. For example in the case of $\mathfrak{p} = 0$, this basis is spanned by

$$\{e_1,\ v_1, \dots, v_{k_0},\ L(v_j, v_j)|_{1 \leq j \leq k_0 - 1},\ L(v_j, v_k) \mid 1 \leq j < k \leq k_0\}.$$

As M is parallel we can extend this basis using parallel translation thus obtaining the same expression of the second fundamental form at every point. Applying then the lemma of Cartan, as the previously mentioned spaces are also parallel and therefore must admit a similar basis, shows that M is isometric with one of the previously mentioned spaces. Finally applying the uniqueness result for Lagrangian immersions shows also that the immersion of M is congruent to one of Naitoh's examples.

References

[Bolton et al. 2009] J. Bolton, C. Rodriguez Montealegre, and L. Vrancken, "Characterizing warped-product Lagrangian immersions in complex projective space", *Proc. Edinb. Math. Soc.* (2) **52**:2 (2009), 273–286. MR 2010d:53063 Zbl 1166.53010

[Bott and Milnor 1958] R. Bott and J. Milnor, "On the parallelizability of the spheres", *Bull. Amer. Math. Soc.* **64** (1958), 87–89. MR 21 #1590 Zbl 0082.16602

[Castro et al. 2006] I. Castro, H. Li, and F. Urbano, "Hamiltonian-minimal Lagrangian submanifolds in complex space forms", *Pacific J. Math.* **227**:1 (2006), 43–63. MR 2007k:53092 Zbl 1129.53039

[Chen 1973] B.-Y. Chen, *Geometry of submanifolds*, Pure and Applied Mathematics **22**, Marcel Dekker, New York, 1973. MR 50 #5697 Zbl 0262.53036

[Chen 1997a] B.-Y. Chen, "Interaction of Legendre curves and Lagrangian submanifolds", *Israel J. Math.* **99**:1 (1997), 69–108. MR 98i:53086 Zbl 0884.53014

[Chen 1997b] B.-Y. Chen, "Complex extensors and Lagrangian submanifolds in complex Euclidean spaces", *Tohoku Math. J.* (2) **49**:2 (1997), 277–297. MR 98g:53096 Zbl 0877.53041

[Chen 2001] B.-Y. Chen, "Riemannian geometry of Lagrangian submanifolds", *Taiwanese J. Math.* **5**:4 (2001), 681–723. MR 2002k:53154 Zbl 1002.53053

[Chen 2005] B.-Y. Chen, "Classification of Lagrangian surfaces of constant curvature in complex projective plane", *J. Geom. Phys.* **53**:4 (2005), 428–460. MR 2005k:53145 Zbl 1072.53027

[Chen and Nagano 1977] B.-Y. Chen and T. Nagano, "Totally geodesic submanifolds of symmetric spaces, I", *Duke Math. J.* **44**:4 (1977), 745–755. MR 56 #16543 Zbl 0368.53038

[Chen and Nagano 1978] B.-Y. Chen and T. Nagano, "Totally geodesic submanifolds of symmetric spaces, II", *Duke Math. J.* **45**:2 (1978), 405–425. MR 58 #7494 Zbl 0384.53024

[Chen and Ogiue 1974] B.-Y. Chen and K. Ogiue, "On totally real submanifolds", *Trans. Amer. Math. Soc.* **193** (1974), 257–266. MR 49 #11433 Zbl 0286.53019

[Ejiri 1981] N. Ejiri, "Totally real submanifolds in a 6-sphere", *Proc. Amer. Math. Soc.* **83**:4 (1981), 759–763. MR 83a:53033 Zbl 0474.53051

[Hu et al. 2008] Z. Hu, H. Li, and L. Vrancken, "Characterizations of the Calabi product of hyperbolic affine hyperspheres", *Results Math.* **52** (2008), 299–314. MR 2009i:53006 Zbl 1161.53013

[Hu et al. 2009] Z. Hu, H. Li, U. Simon, and L. Vrancken, "On locally strongly convex affine hypersurfaces with parallel cubic form, I", *Diff. Geom. Appl.* **27**:2 (2009), 188–205. MR 2010b:53015 Zbl 05544179

[Hu et al. 2011] Z. Hu, H. Li, and L. Vrancken, "Locally strongly convex affine hypersurfaces with parallel cubic form", *J. Differential Geom.* **87**:2 (2011), 239–308. MR 2788657 Zbl 1220.53015

[Kervaire 1958] M. A. Kervaire, "Non-parallelizability of the n-sphere for $n > 7$", *Proc. Nat. Acad. Sci. USA* **44**:3 (1958), 280–283. Zbl 0093.37303

[Li and Vrancken 2005] H. Li and L. Vrancken, "A basic inequality and new characterization of Whitney spheres in a complex space form", *Israel J. Math.* **146** (2005), 223–242. MR 2006d:53071 Zbl 1076.53072

[Li and Wang 2009] H. Li and X. Wang, "Isotropic Lagrangian submanifolds in complex Euclidean space and complex hyperbolic space", *Results Math.* **56**:1-4 (2009), 387–403. MR 2011a:53106 Zbl 1185.53038

[Li and Wang 2011] H. Li and X. Wang, "Calabi product Lagrangian immersions in complex projective space and complex hyperbolic space", *Results Math.* **59**:3-4 (2011), 453–470. MR 2793467 Zbl 05911457

[Lumiste 2009] Ü. Lumiste, *Semiparallel submanifolds in space forms*, Springer, New York, 2009. MR 2010h:53080 Zbl 1156.53002

[Montiel and Urbano 1988] S. Montiel and F. Urbano, "Isotropic totally real submanifolds", *Math. Z.* **199**:1 (1988), 55–60. MR 89f:53086 Zbl 0677.53064

[Naitoh 1980] H. Naitoh, "Isotropic submanifolds with parallel second fundamental forms in symmetric spaces", *Osaka J. Math.* **17**:1 (1980), 95–110. MR 80m:53043 Zbl 0427.53022

[Naitoh 1981a] H. Naitoh, "Isotropic submanifolds with parallel second fundamental form in $P^m(c)$", *Osaka J. Math.* **18**:2 (1981), 427–464. MR 83b:53051 Zbl 0471.53036

[Naitoh 1981b] H. Naitoh, "Totally real parallel submanifolds in $P^n(c)$", *Tokyo J. Math.* **4**:2 (1981), 279–306. MR 83h:53072 Zbl 0485.53044

[Naitoh 1983a] H. Naitoh, "Parallel submanifolds of complex space forms, I", *Nagoya Math. J.* **90** (1983), 85–117. MR 85d:53026a Zbl 0509.53046

[Naitoh 1983b] H. Naitoh, "Parallel submanifolds of complex space forms, II", *Nagoya Math. J.* **91** (1983), 119–149. MR 85d:53026b Zbl 0502.53045

[Naitoh and Takeuchi 1982] H. Naitoh and M. Takeuchi, "Totally real submanifolds and symmetric bounded domains", *Osaka J. Math.* **19**:4 (1982), 717–731. MR 84d:53058 Zbl 0547.53028

[O'Neill 1965] B. O'Neill, "Isotropic and Kähler immersions", *Canad. J. Math.* **17** (1965), 907–915. MR 32 #1654 Zbl 0171.20503

[Rodriguez Montealegre and Vrancken 2009] C. Rodriguez Montealegre and L. Vrancken, "Warped product minimal Lagrangian immersions in complex projective space", *Results Math.* **56**:1-4 (2009), 405–420. MR 2011b:53148 Zbl 1190.53079

Received January 10, 2011. Revised May 31, 2011.

FRANKI DILLEN
KATHOLIEKE UNIVERSITEIT LEUVEN
DEPARTEMENT WISKUNDE
CELESTIJNENLAAN 200B, BOX 2400
BE-3001 LEUVEN
BELGIUM
franki.dillen@wis.kuleuven.be

HAIZHONG LI
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING 100084
CHINA
hli@math.tsinghua.edu.cn

LUC VRANCKEN
UNIVERSITÉ DE LILLE NORD DE FRANCE
F-59000 LILLE
UVHC, LAMAV
F-59313 VALENCIENNES
FRANCE

and

KATHOLIEKE UNIVERSITEIT LEUVEN
DEPARTEMENT WISKUNDE
CELESTIJNENLAAN 200B, BOX 2400
BE-3001 LEUVEN
BELGIUM
luc.vrancken@univ-valenciennes.fr

XIANFENG WANG
SCHOOL OF MATHEMATICAL SCIENCES AND LPMC
NANKAI UNIVERSITY
TIANJIN 300071 CHINA
wangxianfeng@nankai.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

ULTRA-DISCRETIZATION OF THE $D_4^{(3)}$-GEOMETRIC CRYSTAL TO THE $G_2^{(1)}$-PERFECT CRYSTALS

MANA IGARASHI, KAILASH C. MISRA AND TOSHIKI NAKASHIMA

Let $\mathfrak{g}$ be an affine Lie algebra and $\mathfrak{g}^L$ its Langlands dual. It was conjectured by Kashiwara, Nakashima, and Okado that $\mathfrak{g}$ has a positive geometric crystal whose ultra-discretization is isomorphic to the limit of certain coherent family of perfect crystals for $\mathfrak{g}^L$. We prove that the ultra-discretization of the positive geometric crystal for $\mathfrak{g} = D_4^{(3)}$ given by Igarashi and Nakashima is isomorphic to the limit of the coherent family of perfect crystals for $\mathfrak{g}^L = G_2^{(1)}$ constructed by Misra, Mohamad, and Okado.

1. Introduction

Let $A = (a_{ij})_{i,j \in I}$, where $I = \{0, 1, \dots, n\}$, be an affine Cartan matrix and let $(A, \{\alpha_i\}_{i \in I}, \{\alpha_i^\vee\}_{i \in I})$ be a given Cartan datum. Let $\mathfrak{g} = \mathfrak{g}(A)$ denote the associated affine Lie algebra [Kac 1990] and $U_q(\mathfrak{g})$ denote the corresponding quantum affine algebra. Let $P = \mathbb{Z}\Lambda_0 \oplus \mathbb{Z}\Lambda_1 \oplus \cdots \oplus \mathbb{Z}\Lambda_n \oplus \mathbb{Z}\delta$ denote the affine weight lattice and $P^\vee = \mathbb{Z}\alpha_0^\vee \oplus \mathbb{Z}\alpha_1^\vee \oplus \cdots \oplus \mathbb{Z}\alpha_n^\vee \oplus \mathbb{Z}d$ the dual affine weight lattice. For a dominant weight $\lambda \in P^+ = \{\mu \in P \mid \mu(h_i) \geq 0 \text{ for all } i \in I\}$ of level $l = \lambda(\mathbf{c})$ (where $\mathbf{c}$ is the canonical central element), Kashiwara [1990] defined the crystal base $(L(\lambda), B(\lambda))$ for the integrable highest weight $U_q(\mathfrak{g})$-module $V(\lambda)$. The crystal $B(\lambda)$ is the $q = 0$ limit of the canonical basis [Lusztig 1990] or the global crystal basis [Kashiwara 1991]. It has many interesting combinatorial properties. To give an explicit realization of $B(\lambda)$, the notions of affine crystal and perfect crystal were introduced in [Kang et al. 1992a]. It is shown there that the affine crystal $B(\lambda)$ for the level $l \in \mathbb{Z}_{>0}$ integrable highest weight $U_q(\mathfrak{g})$-module $V(\lambda)$ can be realized as the semi-infinite tensor product $\cdots \otimes B_l \otimes B_l \otimes B_l$, where B_l is a perfect crystal of level l. This is known as the path realization.

Kang et al. [1994] remarked that one needs a coherent family of perfect crystals $\{B_l\}_{l \geq 1}$ in order to give a path realization of the Verma module $M(\lambda)$ (or $U_q^-(\mathfrak{g})$). In particular, the crystal $B(\infty)$ of $U_q^-(\mathfrak{g})$ can be realized as the semi-infinite tensor

Misra was supported in part by NSA Grant H98230-08-1-0080. Nakashima was supported in part by JSPS Grants in Aid for Scientific Research #22540031.
MSC2010: primary 17B37, 17B67; secondary 22E65, 14M15.
Keywords: geometric crystals, perfect crystals, ultra-discretization.

product $\cdots \otimes B_\infty \otimes B_\infty \otimes B_\infty$ where B_∞, is the limit of the coherent family of perfect crystals $\{B_l\}_{l\geq 1}$.

At least one coherent family $\{B_l\}_{l\geq 1}$ of perfect crystals and its limit is known for $\mathfrak{g} = A_n^{(1)}, B_n^{(1)}, C_n^{(1)}, D_n^{(1)}, A_{2n-1}^{(2)}, A_{2n}^{(2)}, D_{n+1}^{(2)}, D_4^{(3)}, G_2^{(1)}$. (See [Kang et al. 1992b; 1994; Yamane 1998; Kashiwara et al. 2007; Misra et al. 2010].)

A perfect crystal is indeed a crystal for certain finite-dimensional modules of the quantum affine algebra $U_q(\mathfrak{g})$ named after Kirillov and Reshetikhin [1987], and known as KR-modules for short. KR-modules are parametrized by two integers, $i \in I \setminus \{0\}$ and $l > 0$. Let $\{\varpi_i\}_{i\in I\setminus\{0\}}$ be the set of level 0 fundamental weights [Kashiwara 2002]. Hatayama et al. [1999; 2002] conjectured that any KR-module $W(l\varpi_i)$ admits a crystal base $B^{i,l}$ in the sense of Kashiwara and that $B^{i,l}$ is perfect if l is a multiple of $c_i^\vee := \max(1, 2/(\alpha_i, \alpha_i))$. This conjecture has been proved for quantum affine algebras $U_q(\mathfrak{g})$ of classical types [Okado and Schilling 2008; Fourier et al. 2009; 2010]. When $\{B^{i,l}\}_{l\geq 1}$ is a coherent family of perfect crystals we denote its limit by $B_\infty(\varpi_i)$, or just B_∞ if there is no confusion.

The notion of geometric crystals is a geometric analog to Kashiwara's crystal [Kashiwara 1990]. It was defined in [Berenstein and Kazhdan 2000] for reductive algebraic groups and extended to general Kac–Moody groups in [Nakashima 2005a]. For a given Cartan datum $(A, \{\alpha_i\}_{i\in I}, \{\alpha_i^\vee\}_{i\in I})$, a geometric crystal is defined as a quadruple $\mathcal{V}(\mathfrak{g}) = (X, \{e_i\}_{i\in I}, \{\gamma_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$, where X is an algebraic variety, $e_i : \mathbb{C}^\times \times X \to X$ are rational $\mathbb{C}^\times$-actions and $\gamma_i, \varepsilon_i : X \to \mathbb{C}$ $(i \in I)$ are rational functions satisfying certain conditions (see Definition 2.1). Geometric crystals have many properties similar to algebraic crystals. For instance, the product of two geometric crystals admits the structure of a geometric crystal if they are induced from unipotent crystals [Berenstein and Kazhdan 2000]. A geometric crystal is said to be a positive geometric crystal if it admits a positive structure (see Definition 2.5). A remarkable relation between positive geometric crystals and algebraic crystals is the ultra-discretization functor $\mathcal{UD}$ between them (page 123). Applying this functor, positive rational functions are transferred to piecewise linear functions by the simple correspondence:

$$x \times y \mapsto x + y, \qquad \frac{x}{y} \mapsto x - y, \qquad x + y \mapsto \max(x, y).$$

Let G denote the affine Kac–Moody group associated with the affine Lie algebra $\mathfrak{g}$. Let $B^\pm$ be fixed Borel subgroups and T the maximal torus of G such that $B^+ \cap B^- = T$. Set $y_i(c) := \exp(cf_i)$, and let $\alpha_i^\vee(c) \in T$ be the image of $c \in \mathbb{C}^\times$ under the group morphism $\mathbb{C}^\times \to T$ induced by the simple coroot $\alpha_i^\vee$. We set $Y_i(c) := y_i(c^{-1})\, \alpha_i^\vee(c) = \alpha_i^\vee(c)\, y_i(c)$. Let W and $\widetilde{W}$ be the Weyl group and extended Weyl group associated with $\mathfrak{g}$. The Schubert cell

$$X_w := BwB/B,$$

where $w = s_{i_1} \cdots s_{i_k} \in W$, is birationally isomorphic to the variety

$$B_\iota^- := \{Y_{i_1}(x_1) \cdots Y_{i_k}(x_k) \mid x_1, \dots, x_k \in \mathbb{C}^\times\} \subset B^-,$$

and X_w has a natural geometric crystal structure, where $\iota = i_1, \dots, i_k$ is a reduced word for w. [Berenstein and Kazhdan 2000; Nakashima 2005a].

Let $W(\varpi_i)$ be the KR-module (also called the fundamental representation) of $U_q(\mathfrak{g})$ with ϖ_i as an extremal weight (see [Kashiwara 2002]). Denote its specialization at $q = 1$ by the same symbol, $W(\varpi_i)$. It is a finite-dimensional $\mathfrak{g}$-module (not necessarily irreducible). Let $\mathbb{P}(\varpi_i)$ be the projective space $(W(\varpi_i) \setminus \{0\})/\mathbb{C}^\times$. For any $i \in I$ the translation $t(c_i^\vee \varpi_i)$ belongs to $\widetilde{W}$ (see [Kashiwara et al. 2008]). For a subset J of I, let us denote by $\mathfrak{g}_J$ the subalgebra of $\mathfrak{g}$ generated by $\{e_i, f_i\}_{i \in J}$. For an integral weight μ, define $I(\mu) := \{j \in I \mid \langle \alpha_j^\vee, \mu \rangle \geq 0\}$.

Conjecture 1.1 [Kashiwara et al. 2008]. *For any $i \in I \setminus \{0\}$ there exist a unique variety X endowed with a positive $\mathfrak{g}$-geometric crystal structure and a rational mapping $\pi : X \to \mathbb{P}(\varpi_i)$ satisfying the following properties*:

(i) *For an arbitrary extremal vector $u \in W(\varpi_i)_\mu$, writing the translation $t(c_i^\vee \mu)$ as $\tau w \in \widetilde{W}$ with a Dynkin diagram automorphism τ and $w = s_{i_1} \cdots s_{i_k}$, there exists a birational mapping $\xi : B^-_{i_1,\dots,i_k} \to X$ such that ξ is a morphism of $\mathfrak{g}_{I(\mu)}$-geometric crystals and that the composition $\pi \circ \xi : B^-_{i_1,\dots,i_k} \to \mathbb{P}(\varpi_i)$ coincides with $Y_{i_1}(x_1) \cdots Y_{i_k}(x_k) \mapsto Y_{i_1}(x_1) \cdots Y_{i_k}(x_k)\bar{u}$, where $\bar{u}$ is the line including u.*

(ii) *The ultra-discretization* (*Section 2*) *of X is isomorphic to the crystal $B_\infty = B_\infty(\varpi_i)$ of the Langlands dual $\mathfrak{g}^L$.*

In [Kashiwara et al. 2008], it was shown that this conjecture is true for $i = 1$ and $\mathfrak{g} = A_n^{(1)}, B_n^{(1)}, C_n^{(1)}, D_n^{(1)}, A_{2n-1}^{(2)}, A_{2n}^{(2)}, D_{n+1}^{(2)}$. In [Nakashima 2007], a positive geometric crystal for $\mathfrak{g} = G_2^{(1)}$ and $i = 1$ was constructed and it was shown in [Nakashima 2010] that the ultra-discretization of this positive geometric crystal is isomorphic to the limit of the coherent family of perfect crystals for $\mathfrak{g}^L = D_4^{(3)}$ given in [Kashiwara et al. 2007].

More recently, two of the authors have constructed a positive geometric crystal for $\mathfrak{g} = D_4^{(3)}$, $i = 1$ in [Igarashi and Nakashima 2010]. In this paper we describe the structure of the crystal obtained by the ultra-discretization of the geometric crystal $\mathcal{V}(\mathfrak{g})$ constructed in [Igarashi and Nakashima 2010] and then prove that it is isomorphic to the limit B_∞ of the coherent family of perfect crystals for its Langlands dual $\mathfrak{g}^L = G_2^{(1)}$ constructed in [Misra et al. 2010]. This proves Conjecture 4.5 in [Igarashi and Nakashima 2010].

This paper is organized as follows. In Section 2, we recall necessary definitions and facts about geometric crystals. In Section 3, we review needed facts about affine crystals and perfect crystals. We recall from [Misra et al. 2010] the coherent family of perfect crystals for $\mathfrak{g} = G_2^{(1)}$ and its limit in Section 4. In Section 5, we review the positive geometric crystal $\mathcal{V}(\mathfrak{g})$ for $\mathfrak{g} = D_4^{(3)}$ constructed in [Igarashi and Nakashima 2010]. In Section 6, we state and prove our main result, Theorem 6.1.

2. Geometric crystals

In this section, we review Kac–Moody groups and geometric crystals following [Peterson and Kac 1983; Kumar 2002; Berenstein and Kazhdan 2000].

Kac–Moody algebras and Kac–Moody groups. Fix a symmetrizable generalized Cartan matrix $A = (a_{ij})_{i,j\in I}$ with a finite index set I. Let $(\mathfrak{t}, \{\alpha_i\}_{i\in I}, \{\alpha_i^\vee\}_{i\in I})$ be the associated root data, where $\mathfrak{t}$ is a vector space over $\mathbb{C}$ and $\{\alpha_i\}_{i\in I} \subset \mathfrak{t}^*$ and $\{\alpha_i^\vee\}_{i\in I} \subset \mathfrak{t}$ are linearly independent satisfying $\alpha_j(\alpha_i^\vee) = a_{ij}$.

The Kac–Moody Lie algebra $\mathfrak{g} = \mathfrak{g}(A)$ associated with A is the Lie algebra over $\mathbb{C}$ generated by $\mathfrak{t}$, the Chevalley generators e_i and f_i $(i \in I)$ with the usual defining relations [Kac and Peterson 1983; Peterson and Kac 1983]. There is the root space decomposition $\mathfrak{g} = \bigoplus_{\alpha\in\mathfrak{t}^*} \mathfrak{g}_\alpha$. Denote the set of roots by

$$\Delta := \{\alpha \in \mathfrak{t}^* \mid \alpha \neq 0,\ \mathfrak{g}_\alpha \neq (0)\}.$$

Set $Q = \sum_i \mathbb{Z}\alpha_i$, $Q_+ = \sum_i \mathbb{Z}_{\geq 0}\alpha_i$, $Q^\vee := \sum_i \mathbb{Z}\alpha_i^\vee$ and $\Delta_+ := \Delta \cap Q_+$. An element of Δ_+ is called a *positive root*. Let $P \subset \mathfrak{t}^*$ be a weight lattice such that $\mathbb{C} \otimes P = \mathfrak{t}^*$, whose element is called a weight.

Define simple reflections $s_i \in \mathrm{Aut}(\mathfrak{t})$ $(i \in I)$ by $s_i(h) := h - \alpha_i(h)\alpha_i^\vee$; they generate the Weyl group W, which acts on $\mathfrak{t}^*$ by

$$s_i(\lambda) := \lambda - \lambda(\alpha_i^\vee)\alpha_i.$$

Set $\Delta^{\mathrm{re}} := \{w(\alpha_i) \mid w \in W,\ i \in I\}$, whose elements are called real roots.

Let $\mathfrak{g}'$ be the derived Lie algebra of $\mathfrak{g}$ and G the Kac–Moody group associated with $\mathfrak{g}'$ [Peterson and Kac 1983]. Let $U_\alpha := \exp \mathfrak{g}_\alpha$ $(\alpha \in \Delta^{\mathrm{re}})$ be a one-parameter subgroup of G. The group G is generated by U_α $(\alpha \in \Delta^{\mathrm{re}})$. Let $U^\pm$ be the subgroup generated by $U_{\pm\alpha}$ $(\alpha \in \Delta_+^{\mathrm{re}} = \Delta^{\mathrm{re}} \cap Q_+)$, *i.e.*, $U^\pm := \langle U_{\pm\alpha} | \alpha \in \Delta_+^{\mathrm{re}}\rangle$.

For any $i \in I$, there exists a unique homomorphism; $\phi_i : SL_2(\mathbb{C}) \to G$ such that

$$\phi_i\left(\begin{pmatrix} c & 0 \\ 0 & c^{-1} \end{pmatrix}\right) = c^{\alpha_i^\vee}, \quad \phi_i\left(\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\right) = \exp(te_i), \quad \phi_i\left(\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}\right) = \exp(tf_i),$$

where $c \in \mathbb{C}^\times$ and $t \in \mathbb{C}$. Set $\alpha_i^\vee(c) := c^{\alpha_i^\vee}$, $x_i(t) := \exp(te_i)$, $y_i(t) := \exp(tf_i)$, $G_i := \phi_i(SL_2(\mathbb{C}))$, $T_i := \phi_i(\{\mathrm{diag}(c, c^{-1}) | c \in \mathbb{C}^\vee\})$ and $N_i := N_{G_i}(T_i)$. Let T (resp. N) be the subgroup of G with the Lie algebra $\mathfrak{t}$ (resp. generated by the N_i's), which is called a *maximal torus* in G, and let $B^\pm = U^\pm T$ be the Borel subgroup of G. We have the isomorphism $\phi : W \xrightarrow{\sim} N/T$ defined by $\phi(s_i) = N_i T/T$. An element

$\bar{s}_i := x_i(-1)y_i(1)x_i(-1) = \phi_i\left(\begin{pmatrix} 0 & \pm1 \\ \mp1 & 0 \end{pmatrix}\right)$ is in $N_G(T)$, which is a representative of $s_i \in W = N_G(T)/T$.

Geometric crystals. Let X be an ind-variety, $\gamma_i : X \to \mathbb{C}$ and $\varepsilon_i : X \to \mathbb{C}$ $(i \in I)$ rational functions on X, and $e_i : \mathbb{C}^\times \times X \to X$ $((c, x) \mapsto e_i^c(x))$ a rational $\mathbb{C}^\times$-action.

Definition 2.1. A quadruple $(X, \{e_i\}_{i\in I}, \{\gamma_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$ is a G (or $\mathfrak{g}$)-*geometric crystal* if it satisfies these conditions:

(i) $\{1\} \times X \subset \mathrm{dom}(e_i)$ for any $i \in I$.

(ii) $\gamma_j(e_i^c(x)) = c^{a_{ij}}\gamma_j(x)$.

(iii) The e_i satisfy

$$\begin{aligned}
&e_i^{c_1}e_j^{c_2} = e_j^{c_2}e_i^{c_1} && \text{if } a_{ij} = a_{ji} = 0,\\
&e_i^{c_1}e_j^{c_1c_2}e_i^{c_2} = e_j^{c_2}e_i^{c_1c_2}e_j^{c_1} && \text{if } a_{ij} = a_{ji} = -1,\\
&e_i^{c_1}e_j^{c_1^2c_2}e_i^{c_1c_2}e_j^{c_2} = e_j^{c_2}e_i^{c_1c_2}e_j^{c_1^2c_2}e_i^{c_1} && \text{if } a_{ij} = -2,\ a_{ji} = -1,\\
&e_i^{c_1}e_j^{c_1^3c_2}e_i^{c_1^2c_2}e_j^{c_1^3c_2^2}e_i^{c_1c_2}e_j^{c_2} = e_j^{c_2}e_i^{c_1c_2}e_j^{c_1^3c_2^2}e_i^{c_1^2c_2}e_j^{c_1^3c_2}e_i^{c_1} && \text{if } a_{ij} = -3,\ a_{ji} = -1.
\end{aligned}$$

(iv) $\varepsilon_i(e_i^c(x)) = c^{-1}\varepsilon_i(x)$ and $\varepsilon_i(e_j^c(x)) = \varepsilon_i(x)$ if $a_{i,j} = a_{j,i} = 0$.

Condition (iv) is slightly modified from the one in [Igarashi and Nakashima 2010; Nakashima 2007; 2010].

Let W be the Weyl group associated with $\mathfrak{g}$. Define $R(w)$ for $w \in W$ by

$$R(w) := \{(i_1, i_2, \dots, i_l) \in I^l \mid w = s_{i_1}s_{i_2}\cdots s_{i_l}\},$$

where l is the length of w. Then $R(w)$ is the set of reduced words of w. For a word $\mathbf{i} = (i_1, \dots, i_l) \in R(w)$ $(w \in W)$, set $\alpha^{(j)} := s_{i_l}\cdots s_{i_{j+1}}(\alpha_{i_j})$ $(1 \le j \le l)$ and

$$e_{\mathbf{i}} : T \times X \to X, \quad (t, x) \mapsto e_{\mathbf{i}}^t(x) := e_{i_1}^{\alpha^{(1)}(t)}e_{i_2}^{\alpha^{(2)}(t)}\cdots e_{i_l}^{\alpha^{(l)}(t)}(x).$$

Condition (iii) above amounts to saying that $e_{\mathbf{i}} = e_{\mathbf{i}'}$ for any $w \in W$ and $\mathbf{i}, \mathbf{i}' \in R(w)$.

Geometric crystal on Schubert cell. Let $w \in W$ be a Weyl group element and take a reduced expression $w = s_{i_1}\cdots s_{i_l}$. Let $X := G/B$ be the flag variety, which is an ind-variety and $X_w \subset X$ the Schubert cell associated with w, which has a natural geometric crystal structure [Berenstein and Kazhdan 2000; Nakashima 2005a]. For $\mathbf{i} := (i_1, \dots, i_k)$, set

$$B_{\mathbf{i}}^- := \{Y_{\mathbf{i}}(c_1, \dots, c_k) := Y_{i_1}(c_1)\cdots Y_{i_l}(c_k) \mid c_1 \cdots, c_k \in \mathbb{C}^\times\} \subset B^-, \tag{2-1}$$

where $Y_i(c) := y_i(\frac{1}{c})\alpha_i^\vee(c)$. This has a geometric crystal structure [Nakashima 2005a] isomorphic to X_w. The explicit forms of the action e_i^c, the rational function ε_i and γ_i on $B_{\mathbf{i}}^-$ are given by

$$e_i^c(Y_{\mathbf{i}}(c_1, \dots, c_k)) = Y_{\mathbf{i}}(\mathscr{C}_1, \dots, \mathscr{C}_k)),$$

where

$$(2\text{-}2)\quad \mathscr{C}_j := c_j \cdot \frac{\displaystyle\sum_{1\le m\le j, i_m=i} \frac{c}{c_1^{a_{i_1,i}}\cdots c_{m-1}^{a_{i_{m-1},i}} c_m} + \sum_{j<m\le k, i_m=i} \frac{1}{c_1^{a_{i_1,i}}\cdots c_{m-1}^{a_{i_{m-1},i}} c_m}}{\displaystyle\sum_{1\le m<j, i_m=i} \frac{c}{c_1^{a_{i_1,i}}\cdots c_{m-1}^{a_{i_{m-1},i}} c_m} + \sum_{j\le m\le k, i_m=i} \frac{1}{c_1^{a_{i_1,i}}\cdots c_{m-1}^{a_{i_{m-1},i}} c_m}},$$

$$(2\text{-}3)\quad \varepsilon_i(Y_{\mathbf{i}}(c_1,\dots,c_k)) = \sum_{1\le m\le k, i_m=i} \frac{1}{c_1^{a_{i_1,i}}\cdots c_{m-1}^{a_{i_{m-1},i}} c_m},$$

$$(2\text{-}4)\quad \gamma_i(Y_{\mathbf{i}}(c_1,\dots,c_k)) = c_1^{a_{i_1,i}}\cdots c_k^{a_{i_k,i}}.$$

Positive structure, ultra-discretizations and tropicalizations. The setting is the same as in [Kashiwara et al. 2008]. Let $T = (\mathbb{C}^\times)^l$ be an algebraic torus over $\mathbb{C}$, with character lattice $X^*(T) := \mathrm{Hom}(T, \mathbb{C}^\times) \cong \mathbb{Z}^l$ and cocharacter lattice $X_*(T) := \mathrm{Hom}(\mathbb{C}^\times, T) \cong \mathbb{Z}^l$. Set $R := \mathbb{C}(c)$ and define

$$v : R\setminus\{0\} \to \mathbb{Z}, \quad f(c) \mapsto \deg f(c),$$

where deg is the degree of poles at $c = \infty$. Note that for $f_1, f_2 \in R\setminus\{0\}$, we have

$$(2\text{-}5)\qquad v(f_1 f_2) = v(f_1) + v(f_2), \quad v\left(\frac{f_1}{f_2}\right) = v(f_1) - v(f_2).$$

A nonzero rational function on an algebraic torus T is called *positive* if it can be written as g/h where g and h are a positive linear combination of characters of T.

Definition 2.2. Let $f : T \to T'$ be a rational morphism between two algebraic tori T and T'. We say that f is *positive* if $\eta\circ f$ is positive for any character $\eta : T' \to \mathbb{C}$.

Denote by $\mathrm{Mor}^+(T, T')$ the set of positive rational morphisms from T to T'.

Lemma 2.3 [Berenstein and Kazhdan 2000]. *For any* $f \in \mathrm{Mor}^+(T_1, T_2)$ *and any* $g \in \mathrm{Mor}^+(T_2, T_3)$, *the composition* $g\circ f$ *is well-defined and lies in* $\mathrm{Mor}^+(T_1, T_3)$.

By Lemma 2.3, we can define a category $\mathscr{T}_+$ whose objects are algebraic tori over $\mathbb{C}$ and arrows are positive rational morphisms.

Let $f : T \to T'$ be a positive rational morphism of algebraic tori T and T'. We define a map $\widehat{f} : X_*(T) \to X_*(T')$ by

$$\langle \eta, \widehat{f}(\xi)\rangle = v(\eta\circ f\circ \xi),$$

where $\eta \in X^*(T')$ and $\xi \in X_*(T)$.

Lemma 2.4 [Berenstein and Kazhdan 2000]. *For any algebraic tori* T_1, T_2, T_3, *and positive rational morphisms* $f \in \mathrm{Mor}^+(T_1, T_2)$, $g \in \mathrm{Mor}^+(T_2, T_3)$, *we have* $\widehat{g\circ f} = \widehat{g}\circ\widehat{f}$.

Let $\mathfrak{Set}$ denote the category of sets and set maps. By the lemma, we obtain a functor

$$\begin{array}{cccc} \mathcal{UD}: & \mathcal{T}_+ & \to & \mathfrak{Set} \\ & T & \mapsto & X_*(T) \\ & (f : T \to T') & \mapsto & (\widehat{f} : X_*(T) \to X_*(T'))). \end{array}$$

Definition 2.5 [Berenstein and Kazhdan 2000]. Let

$$\chi = (X, \{e_i\}_{i\in I}, \{\mathrm{wt}_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$$

be a geometric crystal, T' an algebraic torus and $\theta : T' \to X$ a birational isomorphism. The isomorphism θ is called *positive structure* on χ if

(i) for any $i \in I$ the rational functions $\gamma_i \circ \theta : T' \to \mathbb{C}$ and $\varepsilon_i \circ \theta : T' \to \mathbb{C}$ are positive, and

(ii) for any $i \in I$, the rational morphism $e_{i,\theta} : \mathbb{C}^\times \times T' \to T'$ defined by $e_{i,\theta}(c, t) := \theta^{-1} \circ e_i^c \circ \theta(t)$ is positive.

Let $\theta : T \to X$ be a positive structure on a geometric crystal $\chi = (X, \{e_i\}_{i\in I}, \{\mathrm{wt}_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$. Applying the functor $\mathcal{UD}$ to positive rational morphisms $e_{i,\theta} : \mathbb{C}^\times \times T' \to T'$ and $\gamma_i, \varepsilon_i \circ \theta : T' \to \mathbb{C}$ (the notations are as above), we obtain

$$\begin{aligned} \tilde{e}_i &:= \mathcal{UD}(e_{i,\theta}) : \mathbb{Z} \times X_*(T) \to X_*(T), \\ \mathrm{wt}_i &:= \mathcal{UD}(\gamma_i \circ \theta) : X_*(T') \to \mathbb{Z}, \\ \varepsilon_i &:= \mathcal{UD}(\varepsilon_i \circ \theta) : X_*(T') \to \mathbb{Z}. \end{aligned}$$

Now, for given positive structure $\theta : T' \to X$ on a geometric crystal $\chi = (X, \{e_i\}_{i\in I}, \{\mathrm{wt}_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$, we associate the quadruple $(X_*(T'), \{\tilde{e}_i\}_{i\in I}, \{\mathrm{wt}_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$ with a free pre-crystal structure (see [Berenstein and Kazhdan 2000, 2.2]) and denote it by $\mathcal{UD}_{\theta,T'}(\chi)$.

Theorem 2.6 [Berenstein and Kazhdan 2000; Nakashima 2005a]. *For any geometric crystal $\chi = (X, \{e_i\}_{i\in I}, \{\gamma_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$ and positive structure $\theta : T' \to X$, the associated pre-crystal $\mathcal{UD}_{\theta,T'}(\chi) = (X_*(T'), \{\tilde{e}_i\}_{i\in I}, \{\mathrm{wt}_i\}_{i\in I}, \{\varepsilon_i\}_{i\in I})$ is a crystal* (see [Berenstein and Kazhdan 2000, 2.2]).

Now, let $\mathcal{GC}^+$ be the category whose objects are triplets (χ, T', θ), where $\chi = (X, \{e_i\}, \{\gamma_i\}, \{\varepsilon_i\})$ is a geometric crystal and $\theta : T' \to X$ is a positive structure on χ, and whose morphisms $f : (\chi_1, T_1', \theta_1) \to (\chi_2, T_2', \theta_2)$ are given by morphisms $\varphi : X_1 \to X_2$ ($\chi_i = (X_i, \dots)$) such that

$$f := \theta_2^{-1} \circ \varphi \circ \theta_1 : T_1' \to T_2',$$

is a positive rational morphism. Let $\mathcal{CR}$ be the category of crystals. Theorem 2.6 yields:

Corollary 2.7. *The map* $\mathcal{UD} = \mathcal{UD}_{\theta,T'}$ *defined above is a functor*

$$\begin{array}{cccc} \mathcal{UD}: & \mathcal{GC}^+ & \to & \mathcal{CR} \\ & (\chi, T', \theta) & \mapsto & X_*(T'), \\ & (f:(\chi_1, T'_1, \theta_1) \to (\chi_2, T'_2, \theta_2)) & \mapsto & (\widehat{f}: X_*(T'_1) \to X_*(T'_2)). \end{array}$$

We call the functor $\mathcal{UD}$ *"ultra-discretization"* as in [Nakashima 2005a; 2005b] instead of "tropicalization" as in [Berenstein and Kazhdan 2000]. And for a crystal B, if there exists a geometric crystal χ and a positive structure $\theta : T' \to X$ on χ such that $\mathcal{UD}(\chi, T', \theta) \cong B$ as crystals, we call an object (χ, T', θ) in $\mathcal{GC}^+$ a *tropicalization* of B, where it is not known that this correspondence is a functor.

3. Limit of perfect crystals

We review limit of perfect crystals following [Kang et al. 1994]. (See also [Kang et al. 1992a; 1992b].)

Crystals. First we review the theory of crystals, which is the notion obtained by abstracting the combinatorial properties of crystal bases. Let $(A, \{\alpha_i\}_{i\in I}, \{\alpha_i^\vee\}_{i\in I})$ be a Cartan data.

Definition 3.1. A *crystal* B is a set endowed with maps

$$\begin{aligned} &\mathrm{wt}: B \to P, \\ &\varepsilon_i : B \to \mathbb{Z} \sqcup \{-\infty\}, \qquad \varphi_i : B \to \mathbb{Z} \sqcup \{-\infty\} \qquad \text{for } i \in I, \\ &\tilde{e}_i : B \sqcup \{0\} \to B \sqcup \{0\}, \quad \tilde{f}_i : B \sqcup \{0\} \to B \sqcup \{0\} \quad \text{for } i \in I, \\ &\tilde{e}_i(0) = \tilde{f}_i(0) = 0. \end{aligned}$$

satisfying the following axioms, for all $b, b_1, b_2 \in B$:

$$\begin{aligned} &\varphi_i(b) = \varepsilon_i(b) + \langle \alpha_i^\vee, \mathrm{wt}(b)\rangle, \\ &\mathrm{wt}(\tilde{e}_i b) = \mathrm{wt}(b) + \alpha_i \quad \text{if } \tilde{e}_i b \in B, \\ &\mathrm{wt}(\tilde{f}_i b) = \mathrm{wt}(b) - \alpha_i \quad \text{if } \tilde{f}_i b \in B, \\ &\tilde{e}_i b_2 = b_1 \iff \tilde{f}_i b_1 = b_2, \\ &\varepsilon_i(b) = -\infty \implies \tilde{e}_i b = \tilde{f}_i b = 0. \end{aligned}$$

The following tensor product structure is one of the most crucial properties of crystals.

Theorem 3.2. *Let* B_1 *and* B_2 *be crystals, and set*

$$B_1 \otimes B_2 := \{b_1 \otimes b_2;\ b_j \in B_j\ (j = 1, 2)\}.$$

(i) $B_1 \otimes B_2$ *is a crystal.*

(ii) *For* $b_1 \in B_1$ *and* $b_2 \in B_2$, *we have*

$$\tilde{f}_i(b_1 \otimes b_2) = \begin{cases} \tilde{f}_i b_1 \otimes b_2 & \text{if } \varphi_i(b_1) > \varepsilon_i(b_2), \\ b_1 \otimes \tilde{f}_i b_2 & \text{if } \varphi_i(b_1) \le \varepsilon_i(b_2), \end{cases}$$
$$\tilde{e}_i(b_1 \otimes b_2) = \begin{cases} b_1 \otimes \tilde{e}_i b_2 & \text{if } \varphi_i(b_1) < \varepsilon_i(b_2), \\ \tilde{e}_i b_1 \otimes b_2 & \text{if } \varphi_i(b_1) \ge \varepsilon_i(b_2). \end{cases}$$

Definition 3.3. Let B_1 and B_2 be crystals. A *strict morphism* of crystals

$$\psi : B_1 \to B_2$$

is a map $\psi : B_1 \sqcup \{0\} \to B_2 \sqcup \{0\}$ such that $\psi(0) = 0$, $\psi(B_1) \subset B_2$, ψ commutes with all $\tilde{e}_i$ and $\tilde{f}_i$, and

$$\mathrm{wt}(\psi(b)) = \mathrm{wt}(b), \quad \varepsilon_i(\psi(b)) = \varepsilon_i(b), \quad \varphi_i(\psi(b)) = \varphi_i(b) \text{ for any } b \in B_1.$$

A bijective strict morphism is called an *isomorphism of crystals*.

Example 3.4. If (L, B) is a crystal base, then B is a crystal. Hence, for the crystal base $(L(\infty), B(\infty))$ of the nilpotent subalgebra $U_q^-(\mathfrak{g})$ of the quantum algebra $U_q(\mathfrak{g})$, $B(\infty)$ is a crystal.

Example 3.5. For $\lambda \in P$, set $T_\lambda := \{t_\lambda\}$. We define a crystal structure on T_λ by

$$\tilde{e}_i(t_\lambda) = \tilde{f}_i(t_\lambda) = 0, \quad \varepsilon_i(t_\lambda) = \varphi_i(t_\lambda) = -\infty, \quad \mathrm{wt}(t_\lambda) = \lambda.$$

Definition 3.6. For a crystal B, a colored oriented graph structure is associated with B by

$$b_1 \xrightarrow{i} b_2 \Longleftrightarrow \tilde{f}_i b_1 = b_2.$$

We call this graph the *crystal graph* of B.

Affine weights. Let $\mathfrak{g}$ be an affine Lie algebra. The sets $\mathfrak{t}$, $\{\alpha_i\}_{i\in I}$ and $\{\alpha_i^\vee\}_{i\in I}$ be as in Section 2. We take $\dim \mathfrak{t} = \#I + 1$. Let $\delta \in Q_+$ be the unique element satisfying $\{\lambda \in Q \mid \langle\alpha_i^\vee, \lambda\rangle = 0 \text{ for any } i \in I\} = \mathbb{Z}\delta$ and $\mathbf{c} \in \mathfrak{g}$ be the canonical central element satisfying $\{h \in Q^\vee \mid \langle h, \alpha_i\rangle = 0 \text{ for any } i \in I\} = \mathbb{Z}\mathbf{c}$. We write, as in [Kac 1990, 6.1],

$$\mathbf{c} = \sum_i a_i^\vee \alpha_i^\vee, \qquad \delta = \sum_i a_i \alpha_i.$$

Let $(\ ,\ )$ be the nondegenerate W-invariant symmetric bilinear form on $\mathfrak{t}^*$ normalized by $(\delta, \lambda) = \langle \mathbf{c}, \lambda\rangle$ for $\lambda \in \mathfrak{t}^*$. Let us set $\mathfrak{t}_{\mathrm{cl}}^* := \mathfrak{t}^*/\mathbb{C}\delta$ and let $\mathrm{cl} : \mathfrak{t}^* \to \mathfrak{t}_{\mathrm{cl}}^*$ be the canonical projection. Here we have $\mathfrak{t}_{\mathrm{cl}}^* \cong \bigoplus_i (\mathbb{C}\alpha_i^\vee)^*$. Set $\mathfrak{t}_0^* := \{\lambda \in \mathfrak{t}^* \mid \langle \mathbf{c}, \lambda\rangle = 0\}$, $(\mathfrak{t}_{\mathrm{cl}}^*)_0 := \mathrm{cl}(\mathfrak{t}_0^*)$. Since $(\delta, \delta) = 0$, we have a positive definite symmetric form on $\mathfrak{t}_{\mathrm{cl}}^*$ induced by the one on $\mathfrak{t}^*$. Let $\Lambda_i \in \mathfrak{t}_{\mathrm{cl}}^*$ $(i \in I)$ be a classical weight such that $\langle \alpha_i^\vee, \Lambda_j\rangle = \delta_{i,j}$, which is called a fundamental weight. We choose P so that $P_{\mathrm{cl}} := \mathrm{cl}(P)$ coincides with $\bigoplus_{i\in I} \mathbb{Z}\Lambda_i$ and we call P_{cl} a *classical weight lattice*.

Perfect crystals and their limits. Let $\mathfrak{g}$ be an affine Lie algebra, let P_{cl} be a classical weight lattice as above and set $(P_{\mathrm{cl}})_l^+ := \{\lambda \in P_{\mathrm{cl}} \mid \langle \mathbf{c}, \lambda\rangle = l,\ \langle \alpha_i^\vee, \lambda\rangle \geq 0\}$ $(l \in \mathbb{Z}_{>0})$.

Definition 3.7. A crystal B is a *perfect crystal* of level l if the following conditions are satisfied:

(i) $B \otimes B$ is connected as a crystal graph.

(ii) There exists $\lambda_0 \in P_{\mathrm{cl}}$ such that

$$\mathrm{wt}(B) \subset \lambda_0 + \sum_{i \neq 0} \mathbb{Z}_{\leq 0}\mathrm{cl}(\alpha_i), \qquad \#B_{\lambda_0} = 1.$$

(iii) There exists a finite-dimensional $U_q'(\mathfrak{g})$-module V with a crystal pseudobase B_{ps} such that $B \cong B_{\mathrm{ps}}/\pm 1$.

(iv) For any $b \in B$, we have $\langle \mathbf{c}, \varepsilon(b)\rangle \geq l$.

(v) The maps $\varepsilon, \varphi : B^{\min} := \{b \in B \mid \langle \mathbf{c}, \varepsilon(b)\rangle = l\} \longrightarrow (P_{\mathrm{cl}}^+)_l$ are bijective, where $\varepsilon(b) := \sum_i \varepsilon_i(b)\Lambda_i$ and $\varphi(b) := \sum_i \varphi_i(b)\Lambda_i$.

Let $\{B_l\}_{l\geq 1}$ be a family of perfect crystals of level l and set $J := \{(l, b) \mid l > 0,\ b \in B_l^{\min}\}$.

Definition 3.8. A crystal B_∞ with an element b_∞ is called a *limit of* $\{B_l\}_{l\geq 1}$ if

(i) $\mathrm{wt}(b_\infty) = \varepsilon(b_\infty) = \varphi(b_\infty) = 0$;

(ii) for any $(l, b) \in J$, there exists an embedding of crystals

$$f_{(l,b)} : T_{\varepsilon(b)} \otimes B_l \otimes T_{-\varphi(b)} \hookrightarrow B_\infty, \quad t_{\varepsilon(b)} \otimes b \otimes t_{-\varphi(b)} \mapsto b_\infty;$$

(iii) $B_\infty = \bigcup_{(l,b)\in J} \mathrm{Im} f_{(l,b)}$.

As for the crystal T_λ, see Example 3.5. If a limit exists for a family $\{B_l\}$, we say that $\{B_l\}$ is a *coherent family* of perfect crystals.

Here is one of the most important properties of limit of perfect crystals.

Proposition 3.9. *For the crystal $B(\infty)$ as in Example 3.4, we have an isomorphism of crystals*

$$B(\infty) \otimes B_\infty \xrightarrow{\sim} B(\infty).$$

4. Perfect crystals of type $G_2^{(1)}$

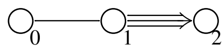
In this section, we review the family of perfect crystals of type $G_2^{(1)}$ and its limit [Misra et al. 2010].

We fix the data for $G_2^{(1)}$. Let $\{\alpha_0, \alpha_1, \alpha_2\}$, $\{\alpha_0^\vee, \alpha_1^\vee, \alpha_2^\vee\}$ and $\{\Lambda_0, \Lambda_1, \Lambda_2\}$ be the set of simple roots, simple coroots and fundamental weights, respectively. The

Cartan matrix $A=(a_{ij})_{i,j=0,1,2}$ is given by

$$A=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix},$$

and its Dynkin diagram is as follows:



The standard null root δ and the canonical central element $\mathbf{c}$ are given by

$$\delta=\alpha_0+2\alpha_1+3\alpha_2 \quad \text{and} \quad \mathbf{c}=\alpha_0^\vee+2\alpha_1^\vee+\alpha_2^\vee,$$

where $\alpha_0=2\Lambda_0-\Lambda_1+\delta$, $\alpha_1=-\Lambda_0+2\Lambda_1-3\Lambda_2$, and $\alpha_2=-\Lambda_1+2\Lambda_2$.

For a positive integer l we introduce $G_2^{(1)}$-crystals B_l and B_∞ as

$$B_l=\left\{ b=(b_1,b_2,b_3,\bar b_3,\bar b_2,\bar b_1)\in(\mathbb{Z}_{\ge 0}/3)^6 \;\middle|\; \begin{array}{l} 3b_3\equiv 3\bar b_3 \ (\mathrm{mod}\ 2), \\ \sum_{i=1,2}(b_i+\bar b_i)+\frac{1}{2}(b_3+\bar b_3)\le l \\ b_1,\bar b_1, b_2-b_3, \bar b_3-\bar b_2\in\mathbb{Z} \end{array} \right\},$$

$$B_\infty=\left\{ b=(b_1,b_2,b_3,\bar b_3,\bar b_2,\bar b_1)\in(\mathbb{Z}/3)^6 \;\middle|\; \begin{array}{l} 3b_3\equiv 3\bar b_3 \ (\mathrm{mod}\ 2), \\ b_1,\bar b_1, b_2-b_3, \bar b_3-\bar b_2\in\mathbb{Z} \end{array} \right\}.$$

Now we describe the explicit crystal structures of B_l and B_∞. Indeed, most of them coincide with each other except for ε_0 and φ_0. In the rest of this section, we use the following convention: $(x)_+=\max(x,0)$. For $b=(b_1,b_2,b_3,\bar b_3,\bar b_2,\bar b_1)$ we define

$$s(b)=b_1+b_2+\tfrac{1}{2}(b_3+\bar b_3)+\bar b_2+\bar b_1, \tag{4-1}$$

and

$$z_1=\bar b_1-b_1,\quad z_2=\bar b_2-\bar b_3,\quad z_3=b_3-b_2,\quad z_4=\tfrac{1}{2}(\bar b_3-b_3). \tag{4-2}$$

Now we define conditions and (F_1)–(F_6) as follows:

$$\begin{cases} (F_1) & z_1+z_2+z_3+3z_4\le 0,\ z_1+z_2+3z_4\le 0,\ z_1+z_2\le 0,\ z_1\le 0, \\ (F_2) & z_1+z_2+z_3+3z_4\le 0,\ z_2+3z_4\le 0,\ z_2\le 0,\ z_1>0, \\ (F_3) & z_1+z_3+3z_4\le 0,\ z_3+3z_4\le 0,\ z_4\le 0,\ z_2>0,\ z_1+z_2>0, \\ (F_4) & z_1+z_2+3z_4>0,\ z_2+3z_4>0,\ z_4>0,\ z_3\le 0,\ z_1+z_3\le 0, \\ (F_5) & z_1+z_2+z_3+3z_4>0,\ z_3+3z_4>0,\ z_3>0,\ z_1\le 0, \\ (F_6) & z_1+z_2+z_3+3z_4>0,\ z_1+z_3+3z_4>0,\ z_1+z_3>0,\ z_1>0. \end{cases} \tag{4-3}$$

Conditions (E_i), for $1\le i\le 6$, are defined from (F_i) by replacing $>$ with $\ge$ and $\le$ with $<$.

We also define

(4-4) $A = (0, z_1, z_1 + z_2, z_1 + z_2 + 3z_4, z_1 + z_2 + z_3 + 3z_4, 2z_1 + z_2 + z_3 + 3z_4)$.

Then for $b = (b_1, b_2, b_3, \bar{b}_3, \bar{b}_2, \bar{b}_1) \in B_l$ or B_∞, the values of $\tilde{e}_i b$, $\tilde{f}_i b$, $\varepsilon_i(b)$, and $\varphi_i(b)$, for $i = 0, 1, 2$, are as follows:

$$\tilde{e}_0 b = \begin{cases} (b_1 - 1, \dots) & \text{if } (E_1), \\ (\dots, b_3 - 1, \bar{b}_3 - 1, \dots, \bar{b}_1 + 1) & \text{if } (E_2), \\ (\dots, b_2 - \frac{2}{3}, b_3 - \frac{2}{3}, \bar{b}_3 + \frac{4}{3}, \bar{b}_2 + \frac{1}{3}, \dots) & \text{if } (E_3) \text{ and } z_4 = -\frac{1}{3}, \\ (\dots, b_2 - \frac{1}{3}, b_3 - \frac{4}{3}, \bar{b}_3 + \frac{2}{3}, \bar{b}_2 + \frac{2}{3}, \dots) & \text{if } (E_3) \text{ and } z_4 = -\frac{2}{3}, \\ (\dots, b_3 - 2, \dots, \bar{b}_2 + 1, \dots) & \text{if } (E_3) \text{ and } z_4 \neq -\frac{1}{3}, -\frac{2}{3}, \\ (\dots, b_2 - 1, \dots, \bar{b}_3 + 2, \dots) & \text{if } (E_4), \\ (b_1 - 1, \dots, b_3 + 1, \bar{b}_3 + 1, \dots) & \text{if } (E_5), \\ (\dots, \bar{b}_1 + 1) & \text{if } (E_6), \end{cases}$$

$$\tilde{f}_0 b = \begin{cases} (b_1 + 1, \dots) & \text{if } (F_1), \\ (\dots, b_3 + 1, \bar{b}_3 + 1, \dots, \bar{b}_1 - 1) & \text{if } (F_2), \\ (\dots, b_3 + 2, \dots, \bar{b}_2 - 1, \dots) & \text{if } (F_3), \\ (\dots, b_2 + \frac{1}{3}, b_3 + \frac{4}{3}, \bar{b}_3 - \frac{2}{3}, \bar{b}_2 - \frac{2}{3}, \dots) & \text{if } (F_4) \text{ and } z_4 = \frac{1}{3}, \\ (\dots, b_2 + \frac{2}{3}, b_3 + \frac{2}{3}, \bar{b}_3 - \frac{4}{3}, \bar{b}_2 - \frac{1}{3}, \dots) & \text{if } (F_4) \text{ and } z_4 = \frac{2}{3}, \\ (\dots, b_2 + 1, \dots, \bar{b}_3 - 2, \dots) & \text{if } (F_4) \text{ and } z_4 \neq \frac{1}{3}, \frac{2}{3}, \\ (b_1 + 1, \dots, b_3 - 1, \bar{b}_3 - 1, \dots) & \text{if } (F_5), \\ (\dots, \bar{b}_1 - 1) & \text{if } (F_6), \end{cases}$$

$$\tilde{e}_1 b = \begin{cases} (\dots, \bar{b}_2 + 1, \bar{b}_1 - 1) & \text{if } \bar{b}_2 - \bar{b}_3 \geq (b_2 - b_3)_+, \\ (\dots, b_3 + 1, \bar{b}_3 - 1, \dots) & \text{if } \bar{b}_2 - \bar{b}_3 < 0 \leq b_3 - b_2, \\ (b_1 + 1, b_2 - 1, \dots) & \text{if } (\bar{b}_2 - \bar{b}_3)_+ < b_2 - b_3, \end{cases}$$

$$\tilde{f}_1 b = \begin{cases} (b_1 - 1, b_2 + 1, \dots) & \text{if } (\bar{b}_2 - \bar{b}_3)_+ \leq b_2 - b_3, \\ (\dots, b_3 - 1, \bar{b}_3 + 1, \dots) & \text{if } \bar{b}_2 - \bar{b}_3 \leq 0 < b_3 - b_2, \\ (\dots, \bar{b}_2 - 1, \bar{b}_1 + 1) & \text{if } \bar{b}_2 - \bar{b}_3 > (b_2 - b_3)_+, \end{cases}$$

$$\tilde{e}_2 b = \begin{cases} (\dots, \bar{b}_3 + \frac{2}{3}, \bar{b}_2 - \frac{1}{3}, \dots) & \text{if } \bar{b}_3 \geq b_3, \\ (\dots, b_2 + \frac{1}{3}, b_3 - \frac{2}{3}, \dots) & \text{if } \bar{b}_3 < b_3, \end{cases}$$

$$\tilde{f}_2 b = \begin{cases} (\dots, b_2 - \frac{1}{3}, b_3 + \frac{2}{3}, \dots) & \text{if } \bar{b}_3 \leq b_3, \\ (\dots, \bar{b}_3 - \frac{2}{3}, \bar{b}_2 + \frac{1}{3}, \dots) & \text{if } \bar{b}_3 > b_3, \end{cases}$$

$$\varepsilon_1(b)=\bar{b}_1+(\bar{b}_3-\bar{b}_2+(b_2-b_3)_+)_+,\quad \varphi_1(b)=b_1+(b_3-b_2+(\bar{b}_2-\bar{b}_3)_+)_+,$$
$$\varepsilon_2(b)=3\bar{b}_2+\tfrac{3}{2}(b_3-\bar{b}_3)_+,\qquad \varphi_2(b)=3b_2+\tfrac{3}{2}(\bar{b}_3-b_3)_+,$$
$$\varepsilon_0(b)=\begin{cases} l-s(b)+\max A-(2z_1+z_2+z_3+3z_4) & b\in B_l,\\ -s(b)+\max A-(2z_1+z_2+z_3+3z_4) & b\in B_\infty,\end{cases}$$
$$\varphi_0(b)=\begin{cases} l-s(b)+\max A & b\in B_l,\\ -s(b)+\max A & b\in B_\infty.\end{cases}$$

For $b\in B_l$ if $\tilde{e}_i b$ or $\tilde{f}_i b$ does not belong to B_l, namely, if b_j or $\bar{b}_j$ for some j becomes negative or $s(b)$ exceeds l, we understand it to be 0.

Theorem 4.1 [Misra et al. 2010]. (i) *The $G_2^{(1)}$-crystal B_l is a perfect crystal of level l.*

(ii) *The family of the perfect crystals $\{B_l\}_{l\geq 1}$ forms a coherent family and the crystal B_∞ is its limit with the vector $b_\infty=(0,0,0,0,0,0)$.*

As was shown in [Misra et al. 2010], the minimal elements are given by

$$(B_l)_{\min}=\{(\alpha,\beta,\beta,\beta,\beta,\alpha)\mid \alpha\in\mathbb{Z}_{\geq 0},\ \beta\in(\mathbb{Z}_{\geq 0})/3,\ 2\alpha+3\beta\leq l\}.$$

Set $J=\{(l,b)\mid l\in\mathbb{Z}_{\geq 1}, b\in(B_l)_{\min}\}$ and let the maps $\varepsilon,\ \varphi:(B_l)_{\min}\to(P_{\rm cl}^+)_l$ be as in Definition 3.7. Then we have $\mathrm{wt}\, b_\infty=0$ and

$$\varepsilon_i(b_\infty)=\varphi_i(b_\infty)=0\quad\text{for } i=0,1,2.$$

For $(l,b_0)\in J$, since $\varepsilon(b_0)=\varphi(b_0)$, one can set $\lambda=\varepsilon(b_0)=\varphi(b_0)$. For

$$b=(b_1,b_2,b_3,\bar{b}_3,\bar{b}_2,\bar{b}_1)\in B_l$$

we define a map

$$f_{(l,b_0)}:\ T_\lambda\otimes B_l\otimes B_{-\lambda}\to B_\infty$$

by

$$f_{(l,b_0)}(t_\lambda\otimes b\otimes t_{-\lambda})=b'=(\nu_1,\nu_2,\nu_3,\bar{\nu}_3,\bar{\nu}_2,\bar{\nu}_1)$$

where $b_0=(\alpha,\beta,\beta,\beta,\beta,\alpha)$, and

$$\nu_1=b_1-\alpha,\quad \bar{\nu}_1=\bar{b}_1-\alpha,$$
$$\nu_2=b_2-\beta,\quad \bar{\nu}_2=\bar{b}_2-\beta,$$
$$\nu_3=b_3-\beta,\quad \bar{\nu}_3=\bar{b}_3-\beta.$$

Finally, we obtain

$$B_\infty=\bigcup_{(l,b)\in J}\mathrm{Im}\, f_{(l,b)}.$$

5. Affine geometric crystal $\mathcal{V}_1(D_4^{(3)})$

Fundamental representation $W(\varpi_1)$ for $D_4^{(3)}$. Let $\mathbf{c} = \sum_i a_i^\vee \alpha_i^\vee$ be the canonical central element in an affine Lie algebra $\mathfrak{g}$ (see [Kac 1990, 6.1]), $\{\Lambda_i \mid i \in I\}$ the set of fundamental weights as in the previous section and $\varpi_1 := \Lambda_1 - a_1^\vee \Lambda_0$ the fundamental weight (of level 0). Let $W(\varpi_1)$ be the fundamental representation of $U_q'(\mathfrak{g})$ associated with ϖ_1 [Kashiwara 2002].

By [Kashiwara 2002, Theorem 5.17], $W(\varpi_1)$ is a finite-dimensional irreducible integrable $U_q'(\mathfrak{g})$-module and has a global basis with a simple crystal. Thus, we can consider the specialization $q = 1$ and obtain the finite-dimensional $\mathfrak{g}$-module $W(\varpi_1)$, which we call a fundamental representation of $\mathfrak{g}$ and use the same notation as above.

We shall present the explicit form of $W(\varpi_1)$ for $\mathfrak{g} = D_4^{(3)}$.

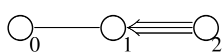
$W(\varpi_1)$ for $D_4^{(3)}$. The Cartan matrix $A = (a_{i,j})_{i,j=0,1,2}$ of type $D_4^{(3)}$ is given by

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -3 \\ 0 & -1 & 2 \end{pmatrix}.$$

Then the simple roots are

$$\alpha_0 = 2\Lambda_0 - \Lambda_1 + \delta, \quad \alpha_1 = -\Lambda_0 + 2\Lambda_1 - \Lambda_2, \quad \alpha_2 = -3\Lambda_1 + 2\Lambda_2,$$

and the Dynkin diagram is this:



The $D_4^{(3)}$-module $W(\varpi_1)$ is an 8-dimensional module with the basis

$$\{v_1, v_2, v_3, v_0, \varnothing, v_{\bar{3}}, v_{\bar{2}}, v_{\bar{1}}\}.$$

The explicit form of $W(\varpi_1)$ is given in [Kashiwara et al. 2007].

$$\mathrm{wt}(v_1) = \Lambda_1 - 2\Lambda_0, \quad \mathrm{wt}(v_2) = -\Lambda_0 - \Lambda_1 + \Lambda_2, \quad \mathrm{wt}(v_3) = -\Lambda_0 + 2\Lambda_1 - \Lambda_2,$$
$$\mathrm{wt}(v_{\bar{i}}) = -\mathrm{wt}(v_i) \ (i = 1, 2, 3), \quad \mathrm{wt}(v_0) = \mathrm{wt}(\varnothing) = 0.$$

The actions of e_i and f_i on these basis vectors are given as follows:

$$f_0(v_0, v_{\bar{3}}, v_{\bar{2}}, v_{\bar{1}}, \varnothing) = \left(v_1, v_2, v_3, \varnothing + \tfrac{1}{2}v_0, \tfrac{3}{2}v_1\right),$$
$$f_1(v_1, v_3, v_0, v_{\bar{2}}) = (v_2, v_0, 2v_{\bar{3}}, v_{\bar{1}}), \quad f_2(v_2, v_{\bar{3}}) = (v_3, v_{\bar{2}}),$$
$$e_0(v_1, v_2, v_3, v_0, \varnothing) = \left(\varnothing + \tfrac{1}{2}v_0, v_{\bar{3}}, v_{\bar{2}}, v_{\bar{1}}, \tfrac{3}{2}v_{\bar{1}}\right),$$
$$e_1(v_2, v_0, v_{\bar{3}}, v_{\bar{1}}) = (v_1, 2v_3, v_0, v_{\bar{2}}), \quad e_2(v_3, v_{\bar{2}}) = (v_2, v_{\bar{3}}),$$

where we give nontrivial actions only.

Construction of the affine geometric crystal $\mathcal{V}_1(D_4^{(3)})$ in $W(\varpi_1)$. In this section, we follow [Igarashi and Nakashima 2010]. For $\xi \in (\mathfrak{t}^*_{\mathrm{cl}})_0$, let $t(\xi)$ be the translation as in [Kashiwara 2002, Section 4] and $\widetilde{\varpi}_i$ as in [Kashiwara 2005]; indeed, $\widetilde{\varpi}_i := \max(1, 2/(\alpha_i, \alpha_i))\varpi_i$. Then we have

$$t(\widetilde{\varpi}_1) = s_0 s_1 s_2 s_1 s_2 s_1 =: w_1,$$
$$t(\mathrm{wt}(v_{\bar{2}})) = s_2 s_1 s_2 s_1 s_0 s_1 =: w_2.$$

Associated with these Weyl group elements w_1 and w_2, we define algebraic varieties $\mathcal{V}_1 = \mathcal{V}_1(D_4^{(3)})$ and $\mathcal{V}_2 = \mathcal{V}_2(D_4^{(3)}) \subset W(\varpi_1)$ respectively:

$$\mathcal{V}_1 := \{V_1(x) := Y_0(x_0)Y_1(x_1)Y_2(x_2)Y_1(x_3)Y_2(x_4)Y_1(x_5)v_1 \mid x_i \in \mathbb{C}^\times,\ 0 \le i \le 5\},$$
$$\mathcal{V}_2 := \{V_2(y) := Y_2(y_2)Y_1(y_1)Y_2(y_4)Y_1(y_3)Y_0(y_0)Y_1(y_5)v_{\bar{2}} \mid y_i \in \mathbb{C}^\times,\ 0 \le i \le 5\}.$$

Owing to the explicit forms of f_i's on $W(\varpi_1)$ as above, we have $f_0^3 = 0$, $f_1^3 = 0$ and $f_2^2 = 0$ and then

$$Y_i(c) = \Big(1 + \frac{f_i}{c} + \frac{f_i^2}{2c^2}\Big)\alpha_i^\vee(c) \ \ (i = 0, 1), \quad Y_2(c) = \Big(1 + \frac{f_2}{c}\Big)\alpha_2^\vee(c).$$

We get explicit forms of $V_1(x) \in \mathcal{V}_1$ and $V_2(y) \in \mathcal{V}_2$ as in [Nakashima 2007]:

$$V_1(x) = \sum_{1 \le i \le 3} \big(X_i v_i + X_{\bar{i}} v_{\bar{i}}\big) + X_0 v_0 + X_\varnothing \varnothing,$$
$$V_2(y) = \sum_{1 \le i \le 3} \big(Y_i v_i + Y_{\bar{i}} v_{\bar{i}}\big) + Y_0 v_0 + Y_\varnothing \varnothing.$$

where the rational functions X_i's and Y_i's are all positive in $(x_0, \dots, x_5)$ and $(y_0, \dots, y_5)$ respectively (as for their explicit forms, see [Igarashi and Nakashima 2010]) and for any x there exist a unique rational function $a(x)$ and y such that $V_2(y) = a(x)V_1(x)$. Using this result, we get the positive birational isomorphism $\overline{\sigma} : \mathcal{V}_1 \to \mathcal{V}_2$ $(V_1(x) \mapsto V_2(y))$ and we know that its inverse $\overline{\sigma}^{-1}$ is also positive. The actions of $\bar{e}_0^c$ on $V_2(y)$ (respectively $\overline{\gamma}_0(V_2(y))$ and $\overline{\varepsilon}_0(V_2(y))$) are induced from the ones on $Y_2(y_2)Y_1(y_1)Y_2(y_4)Y_1(y_3)Y_0(y_0)Y_1(y_5)$ as an element of the geometric crystal $\mathcal{V}_2$. We define the action e_0^c on $V_1(x)$ by

$$e_0^c(V_1(x)) := \overline{\sigma}^{-1} \circ \bar{e}_0^c \circ \overline{\sigma}(V_1(x)). \tag{5-1}$$

We also define $\gamma_0(V_1(x))$ and $\varepsilon_0(V_1(x))$ by

$$\gamma_0(V_1(x)) := \overline{\gamma}_0(\overline{\sigma}(V_1(x))), \qquad \varepsilon_0(V_1(x)) := \overline{\varepsilon}_0(\overline{\sigma}(V_1(x))). \tag{5-2}$$

Theorem 5.1 [Igarashi and Nakashima 2010]. *Together with* (5-1), (5-2) *on* $\mathcal{V}_1$, *we obtain a positive affine geometric crystal* $\chi := (\mathcal{V}_1, \{e_i\}_{i \in I}, \{\gamma_i\}_{i \in I}, \{\varepsilon_i\}_{i \in I})$

($I=\{0,1,2\}$), *whose explicit form is as follows*: *first we have* e_i^c, γ_i, *and* ε_i, *for* $i=1,2$, *from* (2-2), (2-3), *and* (2-4):

$$e_1^c(V_1(x)) = V_1(x_0, \mathscr{C}_1 x_1, x_2, \mathscr{C}_3 x_3, x_4, \mathscr{C}_5 x_5),$$
$$e_2^c(V_1(x)) = V_1(x_0, x_1, \mathscr{C}_2 x_2, x_3, \mathscr{C}_4 x_4, x_5),$$

where

$$\mathscr{C}_1 = \frac{\dfrac{c\,x_0}{x_1}+\dfrac{x_0\,x_2}{x_1{}^2\,x_3}+\dfrac{x_0\,x_2\,x_4}{x_1{}^2\,x_3{}^2\,x_5}}{\dfrac{x_0}{x_1}+\dfrac{x_0\,x_2}{x_1{}^2\,x_3}+\dfrac{x_0\,x_2\,x_4}{x_1{}^2\,x_3{}^2\,x_5}},\quad \mathscr{C}_3 = \frac{\dfrac{c\,x_0}{x_1}+\dfrac{c\,x_0\,x_2}{x_1{}^2\,x_3}+\dfrac{x_0\,x_2\,x_4}{x_1{}^2\,x_3{}^2\,x_5}}{\dfrac{c\,x_0}{x_1}+\dfrac{x_0\,x_2}{x_1{}^2\,x_3}+\dfrac{x_0\,x_2\,x_4}{x_1{}^2\,x_3{}^2\,x_5}},$$

$$\mathscr{C}_5 = \frac{c\left(\dfrac{x_0}{x_1}+\dfrac{x_0x_2}{x_1^2x_3}+\dfrac{x_0x_2x_4}{x_1^2x_3^2x_5}\right)}{\dfrac{cx_0}{x_1}+\dfrac{cx_0x_2}{x_1^2x_3}+\dfrac{x_0x_2x_4}{x_1^2x_3^2x_5}},\quad \mathscr{C}_2 = \frac{\dfrac{cx_1{}^3}{x_2}+\dfrac{x_1{}^3x_3{}^3}{x_2{}^2x_4}}{\dfrac{x_1{}^3}{x_2}+\dfrac{x_1{}^3\,x_3{}^3}{x_2{}^2\,x_4}},\quad \mathscr{C}_4 = \frac{c\left(\dfrac{x_1{}^3}{x_2}+\dfrac{x_1{}^3x_3{}^3}{x_2{}^2x_4}\right)}{\dfrac{c\,x_1{}^3}{x_2}+\dfrac{x_1{}^3\,x_3{}^3}{x_2{}^2\,x_4}},$$

$$\varepsilon_1(V_1(x)) = \frac{x_0}{x_1}+\frac{x_0\,x_2}{x_1{}^2\,x_3}+\frac{x_0\,x_2\,x_4}{x_1{}^2\,x_3{}^2\,x_5},\quad \varepsilon_2(V_1(x)) = \frac{x_1{}^3}{x_2}+\frac{x_1{}^3\,x_3{}^3}{x_2{}^2\,x_4},$$
$$\gamma_1(V_1(x)) = \frac{x_1^2x_3^2x_5^2}{x_0x_2x_4},\qquad \gamma_2(V_1(x)) = \frac{x_2^2x_4^2}{x_1^3x_3^3x_5^3}.$$

We also have e_0^c, ε_0 *and* γ_0 *on* $V_1(x)$:

$$e_0^c(V_1(x)) = V_1\left(\frac{D}{c\cdot E}x_0, \frac{F}{c\cdot E}x_1, \frac{G}{c^3\cdot E^3}x_2, \frac{D\cdot H}{c^2\cdot E\cdot F}x_3, \frac{D^3}{c^3\cdot G}x_4, \frac{D}{c\cdot H}x_5\right),$$
$$\varepsilon_0(V_1(x)) = \frac{E}{x_0{}^3\,x_2\,x_3},\qquad \gamma_0(V_1(x)) = \frac{x_0^2}{x_1x_3x_5},$$

where

$$\begin{aligned}
D &= c^2x_0{}^2x_2x_3 + x_1x_2x_3{}^2x_5 + c\,x_0\big(x_1x_3{}^3 + x_2(x_3{}^2 + x_1x_4 + x_1x_3x_5)\big),\\
E &= x_0{}^2x_2x_3 + x_1x_2x_3{}^2x_5 + x_0\big(x_1x_3{}^3 + x_2(x_3{}^2 + x_1x_4 + x_1x_3x_5)\big),\\
F &= x_2x_3{}^2(x_0 + x_1x_5) + c\,x_0\big(x_0x_2x_3 + x_1(x_3{}^3 + x_2x_4 + x_2x_3x_5)\big),\\
G &= c^3x_0{}^6x_2{}^3x_3{}^3 + 3c^2x_0{}^5x_2{}^3x_3{}^4 + 3c^2x_0{}^5x_1x_2{}^2x_3{}^5 + 3c\,x_0{}^4x_2{}^3x_3{}^5\\
&\quad + 6c\,x_0{}^4x_1x_2{}^2x_3{}^6 + x_0{}^3x_2{}^3x_3{}^6 + 3c\,x_0{}^4x_1{}^2x_2x_3{}^7 + 3x_0{}^3x_1x_2{}^2x_3{}^7\\
&\quad + 3x_0{}^3x_1{}^2x_2x_3{}^8 + x_0{}^3x_1{}^3x_3{}^9 + 3c^3x_0{}^5x_1x_2{}^3x_3{}^2x_4 + 6c^2x_0{}^4x_1x_2{}^3x_3{}^3x_4\\
&\quad + 3c\,x_0{}^4x_1{}^2x_2{}^2x_3{}^4x_4 + 3c^3x_0{}^4x_1{}^2x_2{}^2x_3{}^4x_4 + 3c\,x_0{}^3x_1x_2{}^3x_3{}^4x_4\\
&\quad + 3x_0{}^3x_1{}^2x_2{}^2x_3{}^5x_4 + 3c^2x_0{}^3x_1{}^2x_2{}^2x_3{}^5x_4 + 2x_0{}^3x_1{}^3x_2x_3{}^6x_4\\
&\quad + c^3x_0{}^3x_1{}^3x_2x_3{}^6x_4 + 3c^3x_0{}^4x_1{}^2x_2{}^3x_3x_4{}^2 + 3c^2x_0{}^3x_1{}^2x_2{}^3x_3{}^2x_4{}^2\\
&\quad + x_0{}^3x_1{}^3x_2{}^2x_3{}^3x_4{}^2 + 2c^3x_0{}^3x_1{}^3x_2{}^2x_3{}^3x_4{}^2 + c^3x_0{}^3x_1{}^3x_2{}^3x_4{}^3
\end{aligned}$$

$$
\begin{aligned}
&+3c^3x_0{}^5x_1x_2{}^3x_3{}^3x_5+9c^2x_0{}^4x_1x_2{}^3x_3{}^4x_5+6c^2x_0{}^4x_1{}^2x_2{}^2x_3{}^5x_5\\
&+9cx_0{}^3x_1x_2{}^3x_3{}^5x_5+12cx_0{}^3x_1{}^2x_2{}^2x_3{}^6x_5+3x_0{}^2x_1x_2{}^3x_3{}^6x_5\\
&+3cx_0{}^3x_1{}^3x_2x_3{}^7x_5+6x_0{}^2x_1{}^2x_2{}^2x_3{}^7x_5+3x_0{}^2x_1{}^3x_2x_3{}^8x_5\\
&+6c^3x_0{}^4x_1{}^2x_2{}^3x_3{}^2x_4x_5+12c^2x_0{}^3x_1{}^2x_2{}^3x_3{}^3x_4x_5+3cx_0{}^3x_1{}^3x_2{}^2x_3{}^4x_4x_5\\
&+3c^3x_0{}^3x_1{}^3x_2{}^2x_3{}^4x_4x_5+6cx_0{}^2x_1{}^2x_2{}^3x_3{}^4x_4x_5+3x_0{}^2x_1{}^3x_2{}^2x_3{}^5x_4x_5\\
&+3c^2x_0{}^2x_1{}^3x_2{}^2x_3{}^5x_4x_5+3c^3x_0{}^3x_1{}^3x_2{}^3x_3x_4{}^2x_5+3c^2x_0{}^2x_1{}^3x_2{}^3x_3{}^2x_4{}^2x_5\\
&+3c^3x_0{}^4x_1{}^2x_2{}^3x_3{}^3x_5{}^2+9c^2x_0{}^3x_1{}^2x_2{}^3x_3{}^4x_5{}^2+3c^2x_0{}^3x_1{}^3x_2{}^2x_3{}^5x_5{}^2\\
&+9cx_0{}^2x_1{}^2x_2{}^3x_3{}^5x_5{}^2+6cx_0{}^2x_1{}^3x_2{}^2x_3{}^6x_5{}^2+3x_0x_1{}^2x_2{}^3x_3{}^6x_5{}^2\\
&+3x_0x_1{}^3x_2{}^2x_3{}^7x_5{}^2+3c^3x_0{}^3x_1{}^3x_2{}^3x_3{}^2x_4x_5{}^2+6c^2x_0{}^2x_1{}^3x_2{}^3x_3{}^3x_4x_5{}^2\\
&+3cx_0x_1{}^3x_2{}^3x_3{}^4x_4x_5{}^2+c^3x_0{}^3x_1{}^3x_2{}^3x_3{}^3x_5{}^3+3c^2x_0{}^2x_1{}^3x_2{}^3x_3{}^4x_5{}^3\\
&+3cx_0x_1{}^3x_2{}^3x_3{}^5x_5{}^3+x_1{}^3x_2{}^3x_3{}^6x_5{}^3,
\end{aligned}
$$

$$H=cx_0{}^2x_2x_3+x_0x_2x_3{}^2+x_0x_1x_3{}^3+x_0x_1x_2x_4+cx_0x_1x_2x_3x_5+x_1x_2x_3{}^2x_5.$$

6. Ultra-discretization

We denote the positive structure on χ as in the previous section by $\theta : T' := (\mathbb{C}^\times)^6 \to \mathcal{V}_1$ $(x \mapsto V_1(x))$. Then by Corollary 2.7 we obtain the ultra-discretization $\mathcal{UD}(\chi, T', \theta)$, which is a Kashiwara's crystal. Now we show that the conjecture in [Igarashi and Nakashima 2010] is correct.

Theorem 6.1. *The crystal $\mathcal{UD}(\chi, T', \theta)$ as above is isomorphic to the crystal B_∞ of type $G_2^{(1)}$ as in Section 4.*

To show this, we display the explicit crystal structure on $\mathcal{X} := \mathcal{UD}(\chi, T', \theta)$. Note that $\mathcal{UD}(\chi) = \mathbb{Z}^6$ as a set. Here as for variables in $\mathcal{X}$, we use the same notations $c, x_0, x_1, \dots, x_5$ as for χ.

For $x = (x_0, x_1, \dots, x_5) \in \mathcal{X}$, it follows from the results in the previous section that the functions wt_i and ε_i $(i = 0, 1, 2)$ are given as

$$\mathrm{wt}_0(x) = 2x_0 - x_1 - x_3 - x_5,\ \ \mathrm{wt}_1(x) = 2(x_1 + x_3 + x_5) - x_0 - x_2 - x_4,$$
$$\mathrm{wt}_2(x) = 2(x_2 + x_4) - 3(x_1 - x_3 - x_5).$$

Set

$$
\begin{aligned}
&\alpha{:=}2x_0{+}x_2{+}x_3, \quad \beta{:=}x_1{+}x_2{+}2x_3{+}x_5, \quad \gamma{:=}x_0{+}x_1{+}3x_3,\\
&\delta{:=}x_0{+}x_2{+}2x_3, \quad \epsilon{:=}x_0{+}x_1{+}x_2{+}x_4, \quad \phi{:=}x_0{+}x_1{+}x_2{+}x_3{+}x_5.
\end{aligned}
\tag{6-1}
$$

Then we have

$$
\begin{aligned}
&\varepsilon_0(x) = \max(\alpha, \beta, \gamma, \delta, \epsilon, \phi) - (3x_0{+}x_2{+}x_3),\\
&\varepsilon_1(x) = \max(x_0{-}x_1,\, x_0{+}x_2{-}2x_1{-}x_3,\, x_0{+}x_2{+}x_4{-}2x_1{-}2x_3{-}x_5),\\
&\varepsilon_2(x) = \max(3x_1{-}x_2,\, 3x_1{+}3x_3{-}2x_2{-}x_4).
\end{aligned}
\tag{6-2}
$$

Indeed, from the explicit form of G in the previous section we have $\mathcal{UD}(G)|_{c=-1}$ $= \max(-3+3\alpha,\ -2+2\alpha+\delta,\ -2+2\alpha+\gamma,\ -1+\alpha+2\delta,\ -1+\alpha+\gamma+\delta,\ 3\delta,$ $-1+\alpha+2\gamma,\ \gamma+2\delta,\ 2\gamma+\delta,\ 3\gamma,\ -3+2\alpha+\epsilon,\ -2+\alpha+\delta+\epsilon,\ -1+\alpha+\gamma+\epsilon,$ $-1+2\delta+\epsilon, \gamma+\delta+\epsilon, 2\gamma+\epsilon, -3+\alpha+2\epsilon, -2+\delta+2\epsilon, \gamma+2\epsilon, -3+3\epsilon, -3+2\alpha+\phi,$ $-2+\alpha+\delta+\phi,\ -2+\alpha+\gamma+\phi,\ -1+2\delta+\phi,\ -1+\gamma+\delta+\phi,\ \beta+2\delta,\ -1+2\gamma+\phi,$ $\beta+\gamma+\delta,\ \beta+2\gamma,\ -3+\alpha+\epsilon+\phi,\ -2+\delta+\epsilon+\phi,\ -1+\gamma+\epsilon+\phi,\ -1+\beta+\delta+\epsilon,$ $\beta+\gamma+\epsilon,\ -3+2\epsilon+\phi,\ -2+\beta+2\epsilon,\ -3+\alpha+2\phi,\ -2+\delta+2\phi,\ -2+\gamma+2\phi,$ $-1+\alpha+2\beta, -1+\beta+\gamma+\phi, 2\beta+\delta, 2\beta+\gamma, -3+\epsilon+2\phi, -2+\beta+\epsilon+\phi, -1+2\beta+\epsilon,$ $-3+3\phi,\ -2+\beta+2\phi,\ -1+2\beta+\phi,\ 3\beta)$.

We simplify this by using the following lemma:

Lemma 6.2. *For $m_1, \dots, m_k \in \mathbb{R}$ and $t_1, \dots, t_k \in \mathbb{R}_{\geq 0}$ such that $t_1 + \cdots + t_k = 1$, we have*

$$\max\left(m_1, \dots, m_k, \sum_{i=1}^{k} t_i m_i\right) = \max(m_1, \dots, m_k).$$

Since we have

$$-2+2\alpha+\delta = \frac{2(-3+3\alpha)+3\delta}{3}, \qquad -2+2\alpha+\gamma = \frac{2(-3+3\alpha)+3\gamma}{3},$$
$$-1+\alpha+2\delta = \frac{2\cdot 3\delta+(-3+3\alpha)}{3}, \quad -1+\alpha+\gamma+\delta = \frac{(-3+3\alpha)+3\gamma+3\delta}{3},$$
$$-1+\alpha+2\gamma = \frac{(-3+3\alpha)+2\cdot 3\gamma}{3}, \qquad \gamma+2\delta = \frac{2\cdot 3\delta+3\gamma}{3},$$

and so on, we deduce using the lemma that

$$\mathcal{UD}(G)|_{c=-1} = \max(-3+3\alpha, 3\beta, 3\gamma, 3\delta, -3+3\epsilon, -3+3\phi, -1+\alpha+\gamma+\epsilon,$$
$$\gamma+\delta+\epsilon, \gamma+2\epsilon, 2\gamma+\epsilon, -1+\gamma+\epsilon+\phi, \beta+\gamma+\epsilon).$$

Next, we describe the actions of $\tilde{f}_i$ ($i = 0, 1, 2$). Set $\Xi_j := \mathcal{UD}(\mathcal{C}_j)|_{c=-1}$, for $j = 1, \dots, 5$. Then we have

$$\Xi_1 = \max(-1+x_0-x_1,\ x_0+x_2-2x_1-x_3,\ x_0+x_2+x_4-2x_1-2x_3-x_5)$$
$$-\max(x_0-x_1,\ x_0+x_2-2x_1-x_3,\ x_0+x_2+x_4-2x_1-2x_3-x_5),$$
$$\Xi_3 = \max(-1+x_0-x_1,\ -1+x_0+x_2-2x_1-x_3,\ x_0+x_2+x_4-2x_1-2x_3-x_5)$$
$$-\max(-1+x_0-x_1,\ x_0+x_2-2x_1-x_3,\ x_0+x_2+x_4-2x_1-2x_3-x_5),$$
$$\Xi_5 = \max(-1+x_0-x_1,\ -1+x_0+x_2-2x_1-x_3,\ -1+x_0+x_2+x_4-2x_1-2x_3-x_5)$$
$$-\max(-1+x_0-x_1,\ -1+x_0+x_2-2x_1-x_3,\ x_0+x_2+x_4-2x_1-2x_3-x_5),$$
$$\Xi_2 = \max(-1+3x_1-x_2,\ 3x_1+3x_3-2x_2-x_4)$$
$$-\max(3x_1-x_2,\ 3x_1+3x_3-2x_2-x_4),$$
$$\Xi_4 = \max(-1+3x_1-x_2,\ -1+3x_1+3x_3-2x_2-x_4)$$
$$-\max(-1+3x_1-x_2,\ 3x_1+3x_3-2x_2-x_4).$$

Therefore, for $x \in \mathcal{X}$ we have

$$\tilde{f}_1(x) = (x_0,\, x_1+\Xi_1,\, x_2,\, x_3+\Xi_3,\, x_4,\, x_5+\Xi_5),$$
$$\tilde{f}_2(x) = (x_0,\, x_1,\, x_2+\Xi_2,\, x_3,\, x_4+\Xi_4,\, x_5).$$

We obtain the action $\tilde{e}_i$ $(i = 1, 2)$ by setting $c = 1$ in $\mathcal{UD}(\mathcal{C}_i)$.

Finally, we describe the action of $\tilde{f}_0$. Set

$$\begin{aligned}
\Psi_0 &:= \max(-2+\alpha,\, \beta,\, -1+\gamma,\, -1+\delta,\, -1+\epsilon,\, -1+\phi) - \max(\alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,\, \phi)+1,\\
\Psi_1 &:= \max(-1+\alpha,\, \beta,\, -1+\gamma,\, \delta,\, -1+\epsilon,\, -1+\phi) - \max(\alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,\, \phi)+1,\\
\Psi_2 &:= \max(-3+3\alpha,\, 3\beta,\, 3\gamma,\, 3\delta,\, -3+3\epsilon,\, -3+3\phi,\, -1+\alpha+\gamma+\epsilon,\, \gamma+\delta+\epsilon,\\
&\qquad \gamma+2\epsilon,\, 2\gamma+\epsilon,\, -1+\gamma+\epsilon+\phi,\, \beta+\gamma+\epsilon) - 3\max(\alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,\, \phi)+3,\\
\Psi_3 &:= \max(-2+\alpha,\, \beta,\, -1+\gamma,\, -1+\delta,\, -1+\epsilon,\, -1+\phi)\\
&\qquad + \max(-1+\alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,\, -1+\phi) - \max(\alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,\, \phi)\\
&\qquad - \max(-1+\alpha,\, \beta,\, -1+\gamma,\, \delta,\, -1+\epsilon,\, -1+\phi)+2,\\
\Psi_4 &:= 3\max(-2+\alpha,\, \beta,\, -1+\gamma,\, -1+\delta,\, -1+\epsilon,\, -1+\phi)\\
&\qquad - \max(-3+3\alpha,\, 3\beta,\, 3\gamma,\, 3\delta,\, -3+3\epsilon,\, -3+3\phi,\, -1+\alpha+\gamma+\epsilon,\, \gamma+\delta+\epsilon,\\
&\qquad \gamma+2\epsilon,\, 2\gamma+\epsilon,\, -1+\gamma+\epsilon+\phi,\, \beta+\gamma+\epsilon)+3,\\
\Psi_5 &:= \max(-2+\alpha,\, \beta,\, -1+\gamma,\, -1+\delta,\, -1+\epsilon,\, -1+\phi)\\
&\qquad - \max(1+\alpha,\, \beta,\, \gamma,\, \delta,\, \epsilon,\, -1+\phi)+1,
\end{aligned}$$

where $\alpha, \beta, \dots, \phi$ are as in (6-1). Therefore, by the explicit form of e_0^c as in the previous section, we have

$$\tilde{f}_0(x) = (x_0+\Psi_0,\, x_1+\Psi_1,\, x_2+\Psi_2,\, x_3+\Psi_3,\, x_4+\Psi_4,\, x_5+\Psi_5). \tag{6-3}$$

We have the explicit form of $\tilde{e}_0$ by setting $c = 1$ in $\mathcal{UD}(\mathcal{C}_i)$.

Proof of Theorem 6.1. Define the map

$$\begin{array}{rccc}
\Omega: & \mathcal{X} & \to & B_\infty\\
& (x_0, \dots, x_5) & \mapsto & (b_1, b_2, b_3, \bar{b}_3, \bar{b}_2, \bar{b}_1),
\end{array}$$

by

$$b_1 = x_5,\ \ b_2 = \tfrac{1}{3}x_4 - x_5,\ \ b_3 = x_3 - \tfrac{2}{3}x_4,\ \ \bar{b}_3 = \tfrac{2}{3}x_2 - x_3,\ \ \bar{b}_2 = x_1 - \tfrac{1}{3}x_2,\ \ \bar{b}_1 = x_0 - x_1,$$

and Ω^{-1} is given by

$$x_0 = b_1 + b_2 + \tfrac{1}{2}(b_3 + \bar{b}_3) + \bar{b}_2 + \bar{b}_1, \quad x_1 = b_1 + b_2 + \tfrac{1}{2}(b_3 + \bar{b}_3) + \bar{b}_2,$$
$$x_2 = 3b_1 + 3b_2 + \tfrac{3}{2}(b_3 + \bar{b}_3),\ \ x_3 = 2b_1 + 2b_2 + b_3,\ \ x_4 = 3b_1 + 3b_2,\ \ x_5 = b_1,$$

which means that Ω is bijective. Note that $\frac{3}{2}(b_3 + \bar{b}_3) \in \mathbb{Z}$ by the definition of B_∞ as on page 127. We shall show that Ω is commutative with actions of $\tilde{f}_i$ and preserves the functions wt_i and ε_i, that is,

$$\tilde{f}_i(\Omega(x)) = \Omega(\tilde{f}_i x), \quad \mathrm{wt}_i(\Omega(x)) = \mathrm{wt}_i(x), \quad \varepsilon_i(\Omega(x)) = \varepsilon_i(x) \quad (i = 0, 1, 2),$$

Indeed, the commutativity $\tilde{e}_i(\Omega(x)) = \Omega(\tilde{e}_i x)$ is shown by a similar way. First, let us check wt_i:

Set $b = \Omega(x)$ and let (z_1, z_2, z_3, z_4) be as in (4-2). By the explicit forms of wt_i on $\mathscr{X}$ and B_∞, we have

$$\begin{aligned}
\mathrm{wt}_0(\Omega(x)) &= \varphi_0(\Omega(x)) - \varepsilon_0(\Omega(x)) = 2z_1 + z_2 + z_3 + 3z_4 \\
&= 2(\bar{b}_1 - b_1) + (\bar{b}_2 - \bar{b}_3) + (b_3 - b_2) + \tfrac{3}{2}(\bar{b}_3 - b_3) \\
&= 2(\bar{b}_1 - b_1) + \bar{b}_2 - b_2 + \tfrac{1}{2}(\bar{b}_3 - b_3) = 2x_0 - x_1 - x_3 - x_5 = \mathrm{wt}_0(x), \\
\mathrm{wt}_1(\Omega(x)) &= \varphi_1(\Omega(x)) - \varepsilon_1(\Omega(x)) \\
&= b_1 + (b_3 - b_2 + (\bar{b}_2 - \bar{b}_3)_+)_+ - (\bar{b}_1 + (\bar{b}_3 - \bar{b}_2 + (b_2 - b_3)_+)_+) \\
&= b_1 - \bar{b}_1 - b_2 + \bar{b}_2 + b_3 - \bar{b}_3 = 2(x_1 + x_3 + x_5) - x_0 - x_2 - x_4 \\
&= \mathrm{wt}_1(x), \\
\mathrm{wt}_2(\Omega(x)) &= \varphi_2(\Omega(x)) - \varepsilon_2(\Omega(x)) \\
&= 3b_2 + \tfrac{3}{2}(\bar{b}_3 - b_3)_+ - 3\bar{b}_2 - \tfrac{3}{2}(b_3 - \bar{b}_3)_+ \\
&= 3b_2 - 3\bar{b}_2 + \tfrac{3}{2}(\bar{b}_3 - b_3) = 2(x_2 + x_4) - 3(x_1 + x_3 + x_5) = \mathrm{wt}_2(x).
\end{aligned}$$

Next, we check ε_i:

$$\begin{aligned}
\varepsilon_1(\Omega(x)) &= \bar{b}_1 + (\bar{b}_3 - \bar{b}_2 + (b_2 - b_3)_+)_+ \\
&= \max(\bar{b}_1, \bar{b}_1 + \bar{b}_3 - \bar{b}_2, \bar{b}_1 + \bar{b}_3 - \bar{b}_2 + b_2 - b_3) \\
&= \max(x_0 - x_1, x_0 - 2x_1 + x_2 - x_3, x_0 - 2x_1 + x_2 - 2x_3 + x_4 - x_5) = \varepsilon_1(x), \\
\varepsilon_2(\Omega(x)) &= 3\bar{b}_2 + \tfrac{3}{2}(b_3 - \bar{b}_3)_+ = \max\big(3\bar{b}_2, 3\bar{b}_2 + \tfrac{3}{2}(b_3 - \bar{b}_3)\big) \\
&= \max(3x_1 - x_2, 3x_1 - 2x_2 + 3x_3 - x_4) = \varepsilon_2(x).
\end{aligned}$$

Now let us see ε_0:

$$\begin{aligned}
&\varepsilon_0(\Omega(x)) \\
&\quad = -s(b) + \max A - (2z_1 + z_2 + z_3 + 3z_4) \\
&\quad = -x_0 + \max(0, z_1, z_1 + z_2, z_1 + z_2 + 3z_4, \\
&\qquad\qquad z_1 + z_2 + z_3 + 3z_4, 2z_1 + z_2 + z_3 + 3z_4) - (\alpha - \beta) \\
&\quad = -x_0 + \max(-2x_0 + x_1 + x_3 + x_5, -x_0 + x_3, -x_0 + x_1 - x_2 + 2x_3, \\
&\qquad\qquad -x_0 + x_1 - x_3 + x_4, -x_0 + x_1 + x_5, 0) \\
&\quad = -(3x_0 + x_2 + x_3) + \max(x_1 + x_2 + 2x_3 + x_5, x_0 + x_2 + 2x_3, x_0 + x_1 + 3x_3, \\
&\qquad\qquad x_0 + x_1 + x_2 + x_4, x_0 + x_1 + x_2 + x_3 + x_5, 2x_0 + x_2 + x_3) \\
&\quad = -(3x_0 + x_2 + x_3) + \max(\beta, \delta, \gamma, \epsilon, \phi, \alpha).
\end{aligned}$$

On the other hand, we have

$$\varepsilon_0(x) = -(3x_0 + x_2 + x_3) + \max(\alpha, \beta, \gamma, \delta, \epsilon, \phi).$$

which shows $\varepsilon_0(\Omega(x)) = \varepsilon_0(x)$.

Let us show $\tilde{f}_i(\Omega(x)) = \Omega(\tilde{f}_i(x))$ $(x \in \mathcal{X},\ i = 0, 1, 2)$. As for $\tilde{f}_1$, set

$$A = x_0 - x_1, \quad B = x_0 + x_2 - 2x_1 - x_3, \quad C = x_0 + x_2 + x_4 - 2x_1 - 2x_3 - x_5.$$

Then we obtain

$$\begin{aligned}
\Xi_1 &= \max(A-1, B, C) - \max(A, B, C),\\
\Xi_3 &= \max(A-1, B-1, C) - \max(A-1, B, C),\\
\Xi_5 &= \max(A-1, B-1, C-1) - \max(A-1, B-1, C).
\end{aligned}$$

Therefore, we have

$$\begin{array}{llll}
\Xi_1 = -1, & \Xi_3 = 0, & \Xi_5 = 0, & \text{if } A > B, C,\\
\Xi_1 = 0, & \Xi_3 = -1, & \Xi_5 = 0, & \text{if } A \le B > C,\\
\Xi_1 = 0, & \Xi_3 = 0, & \Xi_5 = -1, & \text{if } A, B \le C,
\end{array}$$

which implies

$$\tilde{f}_1(x) = \begin{cases} (x_0, x_1 - 1, x_2, \dots, x_5) & \text{if } A > B, C,\\ (x_0, \dots, x_3 - 1, x_4, x_5) & \text{if } A \le B > C,\\ (x_0, \dots, x_4, x_5 - 1) & \text{if } A, B \le C. \end{cases}$$

Since $A = \bar{b}_1$, $B = \bar{b}_1 + \bar{b}_3 - \bar{b}_2$ and $C = \bar{b}_1 + \bar{b}_3 - \bar{b}_2 + b_2 - b_3$, we get $(b = \Omega(x))$

$$\Omega(\tilde{f}_1(x)) = \begin{cases} (\dots, \bar{b}_2 - 1, \bar{b}_1 + 1) & \text{if } \bar{b}_2 - \bar{b}_3 > (b_2 - b_3)_+,\\ (\dots, b_3 - 1, \bar{b}_3 + 1, \dots) & \text{if } \bar{b}_2 - \bar{b}_3 \le 0 < b_3 - b_2,\\ (b_1 - 1, b_2 + 1, \dots) & \text{if } (\bar{b}_2 - \bar{b}_3)_+ \le b_2 - b_3, \end{cases}$$

which is the same as the action of $\tilde{f}_1$ on $b = \Omega(x)$ as on page 128. Hence, we have $\Omega(\tilde{f}_1(x)) = \tilde{f}_1(\Omega(x))$.

Let us see $\Omega(\tilde{f}_2(x)) = \tilde{f}_2(\Omega(x))$. Set

$$L = 3x_1 - x_2, \quad M = 3x_1 + 3x_3 - 2x_2 - x_4.$$

Then $\Xi_2 = \max(-1 + L, M) - \max(L, M)$ and $\Xi_4 = \max(-1 + L, -1 + M) - \max(-1 + L, M)$. Thus, one has

$$\begin{array}{lll}
\Xi_2 = -1, & \Xi_4 = 0 & \text{if } L > M,\\
\Xi_2 = 0, & \Xi_4 = -1 & \text{if } L \le M,
\end{array}$$

which means

$$\tilde{f}_2(x) = \begin{cases} (x_0, x_1, x_2 - 1, x_3, x_4, x_5) & \text{if } L > M, \\ (x_0, x_1, x_2, x_3, x_4 - 1, x_5) & \text{if } L \le M. \end{cases}$$

Since $L - M = x_2 - 3x_3 + x_4 = \frac{3}{2}(\bar{b}_3 - b_3)$, one gets

$$\Omega(\tilde{f}_2(x)) = \begin{cases} (\dots, \bar{b}_3 - \frac{2}{3}, \bar{b}_2 + \frac{1}{3}, \dots) & \text{if } \bar{b}_3 > b_3, \\ (\dots, b_2 - \frac{1}{3}, b_3 + \frac{2}{3}, \dots) & \text{if } \bar{b}_3 \le b_3, \end{cases}$$

where $b = \Omega(x)$. This action coincides with the one of $\tilde{f}_2$ on $b \in B_\infty$ on page 128. Therefore, we get $\Omega(\tilde{f}_2(x)) = \tilde{f}_2(\Omega(x))$.

Finally, we shall check $\tilde{f}_0(\Omega(x)) = \Omega(\tilde{f}_0(x))$. For the purpose, we shall estimate the values $\Psi_0, \dots, \Psi_5$ explicitly.

First, the following cases are investigated:

(f$_1$) $\beta \ge \gamma, \delta, \epsilon, \phi,\ \ \phi \ge \alpha,\ \ \delta \ge \alpha.$

(f$_2$) $\beta < \delta \ge \alpha, \gamma, \epsilon,\ \ \alpha > \phi,\ \ \beta \ge \phi.$

(f$_3$) $\beta, \delta < \gamma \ge \alpha, \epsilon, \phi.$

(f$_4$) $\beta, \delta < \epsilon \ge \alpha, \phi,\ \ \epsilon = \gamma + 1.$

(f$'_4$) $\beta, \delta < \epsilon \ge \alpha, \phi,\ \ \epsilon = \gamma + 2.$

(f$''_4$) $\beta, \delta < \epsilon \ge \alpha, \phi,\ \ \epsilon > \gamma + 2.$

(f$_5$) $\beta, \gamma, \epsilon < \phi \ge \alpha,\ \ \alpha > \delta,\ \ \beta \ge \delta.$

(f$_6$) $\alpha > \gamma, \delta, \epsilon, \phi,\ \ \delta, \phi > \beta.$

It is easy to see that each of these conditions are equivalent to the conditions (F_1)–(F_6) in (4-3); more precisely, we have (f$_i$) $\Longleftrightarrow$ (F_i) $(i = 1, 2, 3, 5, 6)$, (f$_4$) $\Longleftrightarrow$ (F_4) and $z_4 = \frac{1}{3}$, (f$'_4$) $\Longleftrightarrow$ (F_4) and $z_4 = \frac{2}{3}$ and (f$''_4$) $\Longleftrightarrow$ (F_4) and $z_4 \ne \frac{1}{3}, \frac{2}{3}$, and that (f$_1$)–(f$_6$) cover all cases and they have no intersection.

Let us show (f$_1$) $\Longleftrightarrow$ (F_1): the condition (f$_1$) means $\beta - \gamma = -(z_1 + z_2) \ge 0$, $\beta - \delta = -z_1 \ge 0$, $\beta - \epsilon = -(z_1 + z_2 + 3z_4) \ge 0$ and $\beta - \phi = -(z_1 + z_2 + z_3 + 3z_4) \ge 0$, which is equivalent to the condition $z_1 + z_2 \le 0$, $z_1 \le 0$, $z_1 + z_2 + 3z_4 \le 0$ and $z_1 + z_2 + z_3 + 3z_4 \le 0$. (Note that $\phi - \alpha = \beta - \delta$, $\delta - \alpha = \beta - \phi$.) This is just the condition (F_1). Other cases $i = 2, 3, 5, 6$ are shown similarly. Next, let us see the cases (f$_4$), (f$'_4$) and (f$''_4$). Indeed,

$$\epsilon - \gamma = x_2 - 3x_3 + x_4 = \tfrac{3}{2}(\bar{b}_3 - b_3) = 3z_4.$$

Thus, we can easily get that (f$_4$) $\Longleftrightarrow$ (F_4) and $z_4 = \frac{1}{3}$, (f$'_4$) $\Longleftrightarrow$ (F_4) and $z_4 = \frac{2}{3}$ and (f$''_4$) $\Longleftrightarrow$ (F_4) and $z_4 \ne \frac{1}{3}, \frac{2}{3}$.

Under the condition $(\mathrm{f}_1) \Longleftrightarrow (F_1)$, we have

$$\Psi_0 = \Psi_1 = \Psi_5 = 1,\ \Psi_2 = \Psi_4 = 3, \quad \Psi_3 = 2,$$

which means $\tilde{f}_0(x) = (x_0+1,\ x_1+1,\ x_2+3,\ x_3+2,\ x_4+3,\ x_5+1)$. Thus, we have

$$\Omega(\tilde{f}_0(x)) = (b_1+1, b_2, \ldots, \bar{b}_1),$$

which coincides with the action of $\tilde{f}_0$ under (F_1) given on page 128. Similarly, we have

$$\begin{aligned}
(\mathrm{f}_2) &\Rightarrow (\Psi_0,\ \Psi_1,\ \Psi_2,\ \Psi_3,\ \Psi_4,\ \Psi_5) = (0,\ 1,\ 3,\ 1,\ 0,\ 0)\\
&\Rightarrow \tilde{f}_0(x) = (x_0,\ x_1+1,\ x_2+3,\ x_3+1,\ x_4,\ x_5)\\
&\Rightarrow \Omega(\tilde{f}_0(x)) = (b_1, b_2, b_3+1, \bar{b}_3+1, \bar{b}_2, \bar{b}_1-1),
\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_2) on the same page;

$$\begin{aligned}
(\mathrm{f}_3) &\Rightarrow (\Psi_0,\ \Psi_1,\ \Psi_2,\ \Psi_3,\ \Psi_4,\ \Psi_5) = (0,\ 0,\ 3,\ 2,\ 0,\ 0)\\
&\Rightarrow \tilde{f}_0(x) = (x_0,\ x_1,\ x_2+3,\ x_3+2,\ x_4,\ x_5)\\
&\Rightarrow \Omega(\tilde{f}_0(x)) = (b_1, b_2, b_3+2, \bar{b}_3, \bar{b}_2-1, \bar{b}_1),
\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_3);

$$\begin{aligned}
(\mathrm{f}_4) &\Rightarrow (\Psi_0,\ \Psi_1,\ \Psi_2,\ \Psi_3,\ \Psi_4,\ \Psi_5) = (0,\ 0,\ 2,\ 2,\ 1,\ 0)\\
&\Rightarrow \tilde{f}_0(x) = (x_0,\ x_1,\ x_2+2,\ x_3+2,\ x_4+1,\ x_5)\\
&\Rightarrow \Omega(\tilde{f}_0(x)) = (b_1, b_2+\tfrac{1}{3}, b_3+\tfrac{4}{3}, \bar{b}_3-\tfrac{2}{3}, \bar{b}_2-\tfrac{2}{3}, \bar{b}_1),
\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_4) and $z_4 = \frac{1}{3}$;

$$\begin{aligned}
(\mathrm{f}'_4) &\Rightarrow (\Psi_0,\ \Psi_1,\ \Psi_2,\ \Psi_3,\ \Psi_4,\ \Psi_5) = (0,\ 0,\ 1,\ 2,\ 2,\ 0)\\
&\Rightarrow \tilde{f}_0(x) = (x_0,\ x_1,\ x_2+1,\ x_3+2,\ x_4+2,\ x_5)\\
&\Rightarrow \Omega(\tilde{f}_0(x)) = (b_1, b_2+\tfrac{2}{3}, b_3+\tfrac{2}{3}, \bar{b}_3-\tfrac{4}{3}, \bar{b}_2-\tfrac{1}{3}, \bar{b}_1),
\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_4) and $z_4 = \frac{2}{3}$;

$$\begin{aligned}
(\mathrm{f}''_4) &\Rightarrow (\Psi_0,\ \Psi_1,\ \Psi_2,\ \Psi_3,\ \Psi_4,\ \Psi_5) = (0,\ 0,\ 0,\ 2,\ 3,\ 0)\\
&\Rightarrow \tilde{f}_0(x) = (x_0,\ x_1,\ x_2,\ x_3+2,\ x_4+3,\ x_5)\\
&\Rightarrow \Omega(\tilde{f}_0(x)) = (b_1, b_2+1, b_3, \bar{b}_3-2, \bar{b}_2, \bar{b}_1),
\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_4) and $z_4 \neq \frac{1}{3}, \frac{2}{3}$;

$$\begin{aligned}
(\mathrm{f}_5) &\Rightarrow (\Psi_0,\ \Psi_1,\ \Psi_2,\ \Psi_3,\ \Psi_4,\ \Psi_5) = (0,\ 0,\ 0,\ 1,\ 3,\ 1)\\
&\Rightarrow \tilde{f}_0(x) = (x_0,\ x_1,\ x_2,\ x_3+1,\ x_4+3,\ x_5+1)\\
&\Rightarrow \Omega(\tilde{f}_0(x)) = (b_1+1, b_2, b_3-1, \bar{b}_3-1, \bar{b}_2, \bar{b}_1),
\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_5). Finally,

$$\begin{aligned}(\mathrm{f}_6) \;&\Rightarrow\; (\Psi_0, \Psi_1, \Psi_2, \Psi_3, \Psi_4, \Psi_5) = (-1, 0, 0, 0, 0, 0)\\ &\Rightarrow\; \tilde{f}_0(x) = (x_0 - 1, x_1, x_2, x_3, x_4, x_5)\\ &\Rightarrow\; \Omega(\tilde{f}_0(x)) = (b_1, b_2, b_3, \bar{b}_3, \bar{b}_2, \bar{b}_1 - 1),\end{aligned}$$

which coincides with the action of $\tilde{f}_0$ under (F_6), still on page 128. Now, we have $\Omega(\tilde{f}_0(x)) = \tilde{f}_0(\Omega(x))$. The proof of Theorem 6.1 has been completed. □

References

[Berenstein and Kazhdan 2000] A. Berenstein and D. Kazhdan, "Geometric and unipotent crystals", pp. 188–236 in *GAFA 2000* (Special volume of *Geom. Funct. Anal.*) (Tel Aviv, 1999), vol. I, edited by N. Alon et al., 2000. MR 2003b:17013 Zbl 1044.17006

[Fourier et al. 2009] G. Fourier, M. Okado, and A. Schilling, "Kirillov–Reshetikhin crystals for nonexceptional types", *Adv. Math.* **222**:3 (2009), 1080–1116. MR 2010j:17028 Zbl 05609507

[Fourier et al. 2010] G. Fourier, M. Okado, and A. Schilling, "Perfectness of Kirillov–Reshetikhin crystals for nonexceptional types", pp. 127–143 in *Quantum affine algebras, extended affine Lie algebras, and their applications*, edited by Y. Gao et al., Contemp. Math. **506**, Amer. Math. Soc., Providence, RI, 2010. MR 2011b:17031 Zbl 05901949

[Hatayama et al. 1999] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, and Y. Yamada, "Remarks on fermionic formula", pp. 243–291 in *Recent developments in quantum affine algebras and related topics* (Raleigh, NC, 1998), edited by N. Jing and K. C. Misra, Contemp. Math. **248**, Amer. Math. Soc., Providence, RI, 1999. MR 2001m:81129 Zbl 1032.81015

[Hatayama et al. 2002] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, and Z. Tsuboi, "Paths, crystals and fermionic formulae", pp. 205–272 in *MathPhys odyssey, 2001*, edited by M. Kashiwara and T. Miwa, Prog. Math. Phys. **23**, Birkhäuser, Boston, 2002. MR 2003e:17020 Zbl 1016.17011

[Igarashi and Nakashima 2010] M. Igarashi and T. Nakashima, "Affine geometric crystal of type $D_4^{(3)}$", pp. 215–226 in *Quantum affine algebras, extended affine Lie algebras, and their applications*, edited by Y. Gao et al., Contemp. Math. **506**, Amer. Math. Soc., Providence, RI, 2010. MR 2011h:17021 Zbl 05901953

[Kac 1990] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990. MR 92k:17038 Zbl 0716.17022

[Kac and Peterson 1983] V. G. Kac and D. H. Peterson, "Regular functions on certain infinite-dimensional groups", pp. 141–166 in *Arithmetic and geometry*, vol. II, edited by M. Artin and J. Tate, Progr. Math. **36**, Birkhäuser, Boston, 1983. MR 86b:17010 Zbl 0578.17014

[Kang et al. 1992a] S.-J. Kang, M. Kashiwara, K. C. Misra, T. Miwa, T. Nakashima, and A. Nakayashiki, "Affine crystals and vertex models", pp. 449–484 in *Infinite analysis* (Kyoto, 1991), edited by A. Tsuchiya et al., Adv. Ser. Math. Phys. **16**, World Sci. Publ., River Edge, NJ, 1992. MR 94a:17008 Zbl 0925.17005

[Kang et al. 1992b] S.-J. Kang, M. Kashiwara, K. C. Misra, T. Miwa, T. Nakashima, and A. Nakayashiki, "Perfect crystals of quantum affine Lie algebras", *Duke Math. J.* **68**:3 (1992), 499–607. MR 94j:17013 Zbl 0774.17017

[Kang et al. 1994] S.-J. Kang, M. Kashiwara, and K. C. Misra, "Crystal bases of Verma modules for quantum affine Lie algebras", *Compositio Math.* **92**:3 (1994), 299–325. MR 95h:17016 Zbl 0808.17007

[Kashiwara 1990] M. Kashiwara, "Crystalizing the q-analogue of universal enveloping algebras", *Comm. Math. Phys.* **133**:2 (1990), 249–260. MR 92b:17018 Zbl 0724.17009

[Kashiwara 1991] M. Kashiwara, "On crystal bases of the Q-analogue of universal enveloping algebras", *Duke Math. J.* **63**:2 (1991), 465–516. MR 93b:17045 Zbl 0739.17005

[Kashiwara 2002] M. Kashiwara, "On level-zero representations of quantized affine algebras", *Duke Math. J.* **112**:1 (2002), 117–175. MR 2002m:17013

[Kashiwara 2005] M. Kashiwara, "Level zero fundamental representations over quantized affine algebras and Demazure modules", *Publ. Res. Inst. Math. Sci.* **41**:1 (2005), 223–250. MR 2005i:17021 Zbl 1147.17306

[Kashiwara et al. 2007] M. Kashiwara, K. C. Misra, M. Okado, and D. Yamada, "Perfect crystals for $U_q(D_4^{(3)})$", *J. Algebra* **317**:1 (2007), 392–423. MR 2009b:17035 Zbl 1140.17012

[Kashiwara et al. 2008] M. Kashiwara, T. Nakashima, and M. Okado, "Affine geometric crystals and limit of perfect crystals", *Trans. Amer. Math. Soc.* **360**:7 (2008), 3645–3686. MR 2009e:17020 Zbl 1219.17010

[Kirillov and Reshetikhin 1987] A. N. Kirillov and N. Y. Reshetikhin, "Representations of Yangians and multiplicity of occurrence of the irreducible components of the tensor product of representations of simple Lie algebras", pp. 211–221 in *Anal. Teor. Chisel i Teor. Funktsii.*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **8**, 1987. In Russian; translated in *J. Sov. Math.* **52** (1990), 3156–3164. MR 89b:17012

[Kumar 2002] S. Kumar, *Kac–Moody groups, their flag varieties and representation theory*, Progress in Mathematics **204**, Birkhäuser, Boston, 2002. MR 2003k:22022 Zbl 1026.17030

[Lusztig 1990] G. Lusztig, "Canonical bases arising from quantized enveloping algebras", *J. Amer. Math. Soc.* **3**:2 (1990), 447–498. MR 90m:17023 Zbl 0703.17008

[Misra et al. 2010] K. C. Misra, M. Mohamad, and M. Okado, "Zero action on perfect crystals for $U_q(G_2^{(1)})$", *SIGMA* **6** (2010), Art. ID 022.

[Nakashima 2005a] T. Nakashima, "Geometric crystals on Schubert varieties", *J. Geom. Phys.* **53**:2 (2005), 197–225. MR 2006d:17016 Zbl 1156.17304

[Nakashima 2005b] T. Nakashima, "Geometric crystals on unipotent groups and generalized Young tableaux", *J. Algebra* **293**:1 (2005), 65–88. MR 2006j:20064 Zbl 1161.17319

[Nakashima 2007] T. Nakashima, "Affine geometric crystal of type $G_2^{(1)}$", pp. 179–192 in *Lie algebras, vertex operator algebras and their applications*, edited by Y.-Z. Huang and K. C. Misra, Contemp. Math. **442**, Amer. Math. Soc., Providence, RI, 2007. MR 2009e:17047 Zbl 1142.17010

[Nakashima 2010] T. Nakashima, "Ultra-discretization of the $G_2^{(1)}$-geometric crystals to the $D_4^{(3)}$-perfect crystals", pp. 273–296 in *Representation theory of algebraic groups and quantum groups*, edited by A. Gyoja et al., Progr. Math. **284**, Birkhäuser, New York, 2010. MR 2011m:17042 Zbl 05919687

[Okado and Schilling 2008] M. Okado and A. Schilling, "Existence of Kirillov–Reshetikhin crystals for nonexceptional types", *Represent. Theory* **12** (2008), 186–207. MR 2009c:17022 Zbl 05526467

[Peterson and Kac 1983] D. H. Peterson and V. G. Kac, "Infinite flag varieties and conjugacy theorems", *Proc. Nat. Acad. Sci. U.S.A.* **80**:6 i. (1983), 1778–1782. MR 84g:17017 Zbl 0512.17008

[Yamane 1998] S. Yamane, "Perfect crystals of $U_q(G_2^{(1)})$", *J. Algebra* **210**:2 (1998), 440–486. MR 2000f:17024 Zbl 0929.17013

Received October 5, 2010.

MANA IGARASHI
DEPARTMENT OF MATHEMATICS
SOPHIA UNIVERSITY
KIOICHO 7-1, CHIYODA-KU
TOKYO 102-8554
JAPAN
mana-i@hoffman.cc.sophia.ac.jp

KAILASH C. MISRA
DEPARTMENT OF MATHEMATICS
NORTH CAROLINA STATE UNIVERSITY
2311 STINSON DRIVE
RALEIGH, NC 27695-8205
UNITED STATES
misra@math.ncsu.edu

TOSHIKI NAKASHIMA
DEPARTMENT OF MATHEMATICS
SOPHIA UNIVERSITY
KIOICHO 7-1, CHIYODA-KU
TOKYO 102-8554
JAPAN
toshiki@sophia.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

CONNECTIVITY PROPERTIES FOR ACTIONS ON LOCALLY FINITE TREES

KEITH JONES

Given an action $G \overset{\rho}{\curvearrowright} T$ by a finitely generated group on a locally finite tree, we view points of the visual boundary ∂T as directions in T and use ρ to lift this sense of direction to G. For each point $E \in \partial T$, this allows us to ask whether G is $(n-1)$-connected "in the direction of E." Then the invariant $\Sigma^n(\rho) \subseteq \partial T$ records the set of directions in which G is $(n-1)$-connected. We introduce a family of actions for which $\Sigma^1(\rho)$ can be calculated through analysis of certain quotient maps between trees. We show that for actions of this sort, under reasonable hypotheses, $\Sigma^1(\rho)$ consists of no more than a single point. By strengthening the hypotheses, we characterize precisely when a given end point lies in $\Sigma^n(\rho)$ for any n.

1. Introduction

Let G be a group having type F_n,[1] and let M be a proper CAT(0) metric space.[2] Let $\rho : G \to \mathrm{Isom}(M)$ be an action by isometries. Bieri and Geoghegan [2003a] introduce a collection of geometric Σ-*invariants*, $\Sigma^n(\rho)$, $n \geq 0$. These arise naturally from the study of the Bieri–Neumann–Strebel–Renz (BNSR) invariants $\Sigma^n(G)$, which can then be viewed as a special case. Σ-invariants give topological insight into ρ and algebraic information about G. In particular, if ρ has discrete orbits and G is finitely generated, then $\Sigma^1(\rho) = \partial M$ if and only if the point stabilizers under ρ are finitely generated; more generally, if G has type F_n, then $\Sigma^n(\rho) = \partial M$ if and only if the point stabilizers under ρ have type F_n.[3]

MSC2010: 20E08, 20F65.
Keywords: BNSR invariant, controlled connectivity, Bieri–Geoghegan invariant, trees, finiteness properties, boundary at infinity.

[1] By definition, G has type F_n if and only if there exists a $K(G, 1)$-complex with finite n-skeleton. This is equivalent to saying that there is an n-dimensional $(n-1)$-connected CW-complex on which G acts freely and cocompactly by permuting cells. All groups have type F_0, while type F_1 is equivalent to finitely generated and type F_2 is equivalent to finitely presented [Geoghegan 2008, §7.2].

[2] A CAT(0) space is a geodesic metric space whose geodesic triangles are no fatter than the corresponding "comparison triangles" in the Euclidean plane, and a metric space is proper if every closed ball is compact [Bridson and Haefliger 1999, Chapter II.1].

[3] See [Bieri and Geoghegan 2003a, Theorem A and the Boundary Criterion]; the required condition "almost geodesically complete" is ensured by cocompactness [Ontaneda 2005, Theorem B].

The invariant $\Sigma^n(\rho)$ depends on a notion of *controlled connectivity*, which we describe briefly here.[4] The action ρ can be used to impose a sense of direction on G as follows. The space M has a CAT(0) boundary ∂M, which is in one-to-one correspondence with the collection of geodesic rays emanating from any particular point of M. Thus ∂M encompasses the set of directions in M in which one can go to infinity. For an end point $E \in \partial M$, there is a nested sequence of subsets of M (called horoballs about E). This nested sequence provides a filtration of M. Because G has type F_n, there is an n-dimensional $(n-1)$-connected CW-complex X on which G acts freely and cocompactly by permuting cells. One can then choose a G-equivariant "control" map $h : X \to M$. With $E \in \partial M$ fixed, h allows us to lift the sense of direction from M up to X (and therefore G by proxy) by taking the preimages of horoballs about E. If, roughly speaking, the preimages of the horoballs about E are $(n-1)$-connected, the action ρ is said to be *controlled* $(n-1)$*-connected* or CC^{n-1} over E.[5] The precise definition ensures that this is independent of choice of X or h, and is in fact a property of ρ [Bieri and Geoghegan 2003a, §3.2].

For $n \geq 0$, the invariant $\Sigma^n(\rho)$ consists of all those end points over which ρ is CC^{n-1}. These form a nested family

$$\Sigma^0(\rho) \supseteq \Sigma^1(\rho) \supseteq \Sigma^2(\rho) \dots .$$

The action ρ induces a topological action by G on ∂M, under which $\Sigma^n(\rho)$ is invariant. Those familiar with the BNSR invariant $\Sigma^n(G)$ may recall that the BNSR invariant is an open subset of the boundary, which in the original case is a sphere. The Bieri–Geoghegan invariant $\Sigma^n(\rho)$ is in general not open in ∂M.

Bieri and Geoghegan calculate Σ^n for the modular group acting on the hyperbolic plane [2003b], and provide information about Σ^n for actions on trees by metabelian groups of finite Prüfer rank [2003a, Chapter 10, Example C]. Rehn [2007] provides calculations for the natural action by $\mathrm{SL}_n(\mathbb{Z}[1/k])$ on the symmetric space for $\mathrm{SL}_n(\mathbb{R})$.

In the case where M is a locally finite simplicial tree, Bieri and Geoghegan [2003a] ask whether $\Sigma^1(\rho)$ is always either empty, a singleton, or the entire boundary of the tree. (The "entire boundary" case has been discussed above.) Lehnert [2009] gives an example for which this is not the case. However, here we illustrate that there does exist a class of actions for which Σ^n is either empty or a singleton.

Main result. All trees are assumed to be simplicial trees viewed as CAT(0) metric spaces, by giving each edge a length of 1. All actions under consideration are by simplicial automorphisms, and therefore are by isometries. Also, we assume

[4] The technical definition is provided in Section 2.

[5] For $n = 0$, we take (-1)-connected to mean nonempty.

that actions are without inversions — that is, an edge is stabilized if and only if it is fixed pointwise — since we can simply pass to the barycentric subdivision otherwise. Any tree exhibiting such an action by a group G is called a G-tree. We assume that all G-trees are infinite and that G is always finitely generated.

A group action on a tree is *minimal* if there exists no proper invariant subtree. A cocompact action on an infinite tree is minimal if and only if the tree has no leaves. We define a *morphism of trees* to mean a map between two trees that sends vertices to vertices and edges to edges and that preserves adjacency. All maps between G-trees are assumed to be G-equivariant morphisms of trees, and therefore continuous. The *star* of a vertex is the set of edges adjacent to that vertex, and a morphism is *locally surjective* or *locally injective* if, for each vertex of the domain tree, the corresponding map between stars is surjective or injective. See [Bass 1993] for further discussion. In the context of morphisms of trees (as opposed to graphs), local injectivity is equivalent to injectivity, and local surjectivity implies surjectivity. A tree is *locally finite* if the star of each vertex is finite; such trees are proper metric spaces.

Theorem 1.1 (Main Theorem). *Let G be a finitely generated group, T a locally finite tree, and $G \overset{\rho}{\curvearrowright} T$ a cocompact action by isometries. If there exists a minimal G-tree $\tilde{T}$ and a G-morphism $q : \tilde{T} \to T$ that is locally surjective, but not locally injective, then $\Sigma^1(\rho)$ consists of at most a single point of ∂T.*

We do not require $\tilde{T}$ to be locally finite, because it is irrelevant to us whether $\tilde{T}$ is proper as a metric space. Also, the map $q : \tilde{T} \to T$ does not generally extend to a map $\partial \tilde{T} \to \partial T$, because geodesic rays may be collapsed to finite paths by q.

As mentioned in the introduction, $\Sigma^1(\rho)$ is a G-invariant subset of ∂T. Hence, if the conditions of the Main Theorem apply and there does exist a point $E_0 \in \Sigma^1(\rho)$, then E_0 is necessarily fixed by ρ. In some cases, this allows us to easily determine that $\Sigma^1(\rho)$ is empty, as in the following examples.

Example 1.2. Let G be the group given by the presentation

$$G = \langle a, s, t \mid a^s = a^2, a^t = a^3 \rangle.$$

As is clear from this presentation, G can be realized as a fundamental group of a graph of groups, where the graph is a 2-rose (a single vertex with two loops). The Bass–Serre tree $\tilde{T}$ associated with this graph of groups decomposition is a regular 7-valent tree. Let N be the normal closure of a. Then N consists of all elements of G that stabilize a vertex in $\tilde{T}$. The quotient group G/N is free on two generators and acts freely on $T = N\backslash\tilde{T}$ with quotient a 2-rose of circles, so T is a regular 4-valent tree. Figure 1 demonstrates the collapsing on a neighborhood of a vertex in $\tilde{T}$. (One can take T to be the Cayley graph of G/N.) The natural quotient map

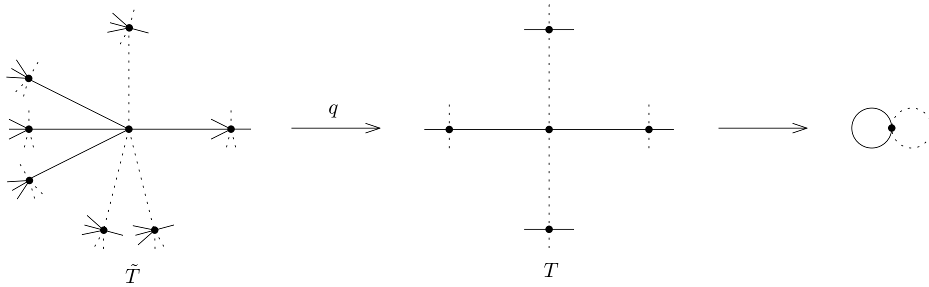


Figure 1. G admits a normal subgroup N, whose action on $\tilde{T}$ collapses $\tilde{T}$ to T.

$\tilde{T} \to T$ satisfies the conditions of the Main Theorem, and no end point $E \in \partial T$ is fixed by ρ. Hence $\Sigma^1(\rho) = \varnothing$.

This example can be generalized to any nonfree group with a graph of groups decomposition over a graph containing a single vertex. Such a group always has a free quotient obtained by collapsing the normal closure of the subgroup associated with the vertex, and as above, the Cayley graph of this free group can be viewed as the quotient of the original Bass–Serre tree.

Example 1.3. One of Lehnert's counterexamples to the question of whether Σ^1 must be either $\varnothing$, a singleton, or ∂T in the case of simplicial trees is closely related to the group G discussed in Example 1.2. Let $H = \mathbb{Z}[\frac{1}{6}] \rtimes F_2(x, y)$, where $F_2(x, y)$ is a free group generated by the letters x and y. One obtains H from G by adding relations corresponding to the commutator subgroup of N. The semidirect product structure is given by $t^x = t/2$ and $t^y = t/3$ for $t \in \mathbb{Z}[\frac{1}{6}]$. This group acts on the same tree T, by viewing it as the Cayley graph of its factor $F_2(x, y)$, and one can represent points in ∂T by infinite reduced words in $F_2(x, y)$. Any point represented by an infinite word eventually consisting of only x or only y does not lie in Σ^1 [Lehnert 2009]; this is a consequence of the interplay between the actions by $F_2(x, y)$ on $\mathbb{Z}[\frac{1}{6}]$ and on T. The author has a proof of this result in a paper currently in preparation, which is based on the topological construction of the Bass–Serre tree [Scott and Wall 1979; Geoghegan 2008, Chapter 6] and distinct in flavor from both the contents of this paper and the proof in [Lehnert 2009].

Evidently, for the action $H \curvearrowright T$, there exists no $\tilde{T}$ and $q : \tilde{T} \to T$ as described in the Main Theorem.

Example 1.4. Here is an example where $\tilde{T}$ is not locally finite. Let $K_4 = \mathbb{Z}_2 \oplus \mathbb{Z}_2$ be the Klein 4-group, and let $D_\infty = \mathbb{Z}_2 * \mathbb{Z}_2$ be the infinite dihedral group. Take the quotient map $\pi : D_\infty * D_\infty \twoheadrightarrow K_4 * K_4$, induced by performing the abelianization map $D_\infty \twoheadrightarrow K_4$ on each free factor of $D_\infty * D_\infty$. There is an action $\tilde{\rho} : D_\infty * D_\infty \to \mathrm{Aut}(\tilde{T})$, where $\tilde{T}$ (a regular ∞-valent tree) is the Bass–Serre tree corresponding to the given free product decomposition. There is also an action $\rho : D_\infty * D_\infty \to \mathrm{Aut}(T)$, where T, a regular 4-valent tree, is the Bass–Serre tree for $K_4 * K_4$; this action factors through π. We can realize T as a quotient of $\tilde{T}$ satisfying the conditions of the Main Theorem. Again, because no point of ∂T is fixed by ρ, it follows that $\Sigma^1(\rho)$ is empty. This example is of a kind initially pointed out to the author by Mike Mihalik.

This, too, can be generalized: if A_1 and A_2 are two finitely generated infinite groups that admit finite quotients Q_1 and Q_2, respectively, then $G = A_1 * A_2$ admits a quotient map $\pi : G \to Q_1 * Q_2$. While G acts on the Bass–Serre tree $\tilde{T}$ corresponding to the decomposition $A_1 * A_2$, it also acts on $\ker \pi \backslash \tilde{T}$, which is isomorphic to the Bass–Serre tree corresponding to $Q_1 * Q_2$.

Example 1.5. More generally, there is a notion of a *morphism of graphs of groups* (essentially, a morphism of graphs together with a collection of homomorphisms of vertex and edge groups that ensure that certain squares commute), which lifts to an equivariant morphism between the corresponding Bass–Serre trees [Bass 1993, Proposition 2.4], and one can determine whether the lift will be locally surjective and not locally injective [Bass 1993, Corollary 2.5]. This can be used to produce maps satisfying the conditions of the Main Theorem. For example, consider the Baumslag–Solitar groups $\mathrm{BS}(m, n) = \langle a, t \mid ta^m t^{-1} = a^n \rangle$. There is a projection map $\mathrm{BS}(2, 4) \twoheadrightarrow \mathrm{BS}(1, 2)$ obtained by adding the relation $tat^{-1}a^{-2}$. One can show that this corresponds to a morphism of graphs of groups that lifts to a map between the corresponding Bass–Serre trees and has the desired properties.

Applying [Bieri and Geoghegan 2003a, Theorems A and H], we have:

Corollary 1.6. *If $G \overset{\rho}{\curvearrowright} T$ satisfies the conditions of the Main Theorem, then for any point $z \in T$, the stabilizer G_z of z under the action ρ is not finitely generated.*

Collapsing pairs. Recall that, in the language of [Serre 1980, Chapter I.2], each geometric edge of T corresponds to two oriented edges, one pointing in either direction.

Remark 1.7. We use the lowercase e to refer to edges of T, oriented or not, and the uppercase E to refer to points of ∂T.

Definition 1.8. Under the hypotheses of the Main Theorem, let $(\tilde{e}_1, \tilde{e}_2)$ be a pair of adjacent distinct oriented edges in $\tilde{T}$ with common initial vertex $\tilde{v}$. If $q(\tilde{e}_1) = q(\tilde{e}_2)$, we call this a *collapsing pair (of edges)* under q. Let $e = q(\tilde{e}_1)$ be the resulting oriented edge in T. For a vertex $w \in T$ (or end point $E \in \partial T$), we say the pair $(\tilde{e}_1, \tilde{e}_2)$ *faces* w (resp. E) if e points toward w (resp. E). This is the same as saying that the geodesic from $q(\tilde{v})$ to w (resp. E) passes through e.

The proof of the Main Theorem follows from two facts: Proposition 3.8 states that because q is not locally injective, all end points of T (with the possible exception of a single end point) are faced by a collapsing pair, while Proposition 3.4 states that local surjectivity of q forces any end point of T faced by a collapsing pair to lie outside $\Sigma^1(\rho)$.

The case where stabilizers on $\tilde{T}$ have type F_n. If we add the condition that the stabilizers under $\tilde{\rho}$ have type F_n, then we can prove that a point $E \in \partial T$ that is *not* faced by a collapsing pair lies in $\Sigma^n(\rho)$.

Theorem 1.9. *Assume the conditions of the Main Theorem. Also, suppose that G has type F_n and that for each point $\tilde{z}$ of $\tilde{T}$, the stabilizer $G_{\tilde{z}}$ has type F_n, for $n > 0$. Then $E \in \partial T$ lies in $\Sigma^n(\rho)$ if and only if there is no collapsing pair facing E.*

Corollary 1.10. *Let the group H have type F_n, and let $\varphi : H \to H$ be injective, so that $G = \langle H, t \mid a^t = \varphi(a) \text{ for all } a \in H\rangle$ is an ascending HNN-extension. If $\chi : G \twoheadrightarrow \mathbb{Z}$ maps $t \mapsto 1$ and $\langle\langle H\rangle\rangle \mapsto 0$, then χ represents a point in $\Sigma^n(G)$.*

This corollary is not new [Meinert 1996; 1997], but the approach is. For further discussion on this result, see [Bieri et al. 2010].

2. Controlled connectivity

In a CAT(0) space M, there is a notion of a *(visual) boundary* ∂M, which is obtained by taking equivalence classes of geodesic rays [Bridson and Haefliger 1999, Chapter II.8]. This boundary carries a topology, called the cone topology, induced by the topology on M. We call points of ∂M *end points*. CAT(0) spaces are contractible, and the boundary of a proper CAT(0) space is a compact space. Let τ be a geodesic ray in M. Following [Bieri and Geoghegan 2003a], we define the *Busemann function* $\beta_\tau : M \to \mathbb{R}$ by

$$\beta_\tau(p) = \lim_{t\to\infty} (t - d(\tau(t), p)).$$

For $r \in \mathbb{R}$, the set $HB_r(\tau) = {\beta_\tau}^{-1}([r, \infty))$ is called a *horoball* around E. Horoballs in CAT(0) spaces are contractible. We can view $HB_r(\tau)$ as the nested union of closed balls $\bigcup_{k\geq\max\{0,r\}} \overline{B_{k-r}(\tau(k))}$.

Definition 2.1. Fix $n \in \mathbb{N}$. Let G be a group having type F_n, and let M be a proper CAT(0) space admitting an isometric action $G \overset{\rho}{\curvearrowright} M$. Choose an n-dimensional $(n-1)$-connected CW-complex X^n on which X acts freely and cocompactly, and choose a continuous G-map $h : X^n \to M$. We call h a *control map*; one can be found because the action by G on X^n is free and M is contractible. Fix a geodesic ray τ representing $E \in \partial M$. For a horoball $HB_r(\tau)$ about E, denote the largest subcomplex of X^n contained in $h^{-1}(HB_r(\tau))$ by $X_{(\tau,r)}$. Finally, we need a notion of *lag function*: any $\lambda(r) > 0$ satisfying $r - \lambda(r) \to \infty$ as $r \to \infty$ is called a lag.

We say ρ is *controlled* $(n-1)$*-connected*, or CC^{n-1}, *over* E if for all $r \in \mathbb{R}$ and all $-1 \leq p \leq (n-1)$, there exists a lag λ such that every map $f : S^p \to X_{(\tau,r)}$ extends to a map $\tilde{f} : B^{p+1} \to X_{(\tau,r-\lambda(r))}$.[6]

Definition 2.2. The Bieri–Geoghegan invariant $\Sigma^n(\rho)$ is the subset of ∂M consisting of all end points over which ρ is controlled $(n-1)$-connected.

Relationship to the BNSR invariant. If ρ fixes an end point E, then the pair (ρ, E) determines a homomorphism $\chi_{\rho,E} : G \to \mathbb{R}$, and E lies in $\Sigma^1(\rho)$ if and only if $\chi_{\rho,E}$ represents a point in $\Sigma^1(G)$ [Bieri and Geoghegan 2003a, §10.6]. In fact, we can obtain the classical BNSR invariant $\Sigma^n(G)$ as the special case where ρ is the action $G \curvearrowright G_{ab} \otimes \mathbb{R}$ [Bieri and Geoghegan 2003a, Chapter 10, Example A]. This is an action by translations on a finite-dimensional real vector space, so every end point is fixed, and $\partial(G_{ab} \otimes \mathbb{R}) \cong \mathrm{Hom}(G, \mathbb{R})$.

The question of finding a single technique for calculating Σ^1 for arbitrary group actions on trees seems out of reach currently. To see this, consider an action $G \overset{\rho}{\curvearrowright} T$ by translations, where T is a simplicial line. This corresponds to a homomorphism $\chi : G \twoheadrightarrow \mathbb{Z}$, and calculating $\Sigma^1(\rho)$ determines whether χ and $-\chi$ represent points of $\Sigma^1(G)$. However, it is known that $\ker \chi$ is finitely generated if and only if both do represent points of $\Sigma^1(G)$ [Bieri et al. 1987, Theorem B1]. Thus a method for calculating $\Sigma^1(\rho)$ even in the special case where the tree is a simplicial line would enable us to determine whether or not the kernel of an arbitrary homomorphism to $\mathbb{Z}$ is finitely generated.

3. Proof of the Main Theorem

An automorphism s of a tree T having no fixed point is said to be *hyperbolic*. For each such s, there is a unique line A_s, called the *axis* of s, stable under the action of the subgroup $\langle s \rangle$, that acts on A_s by translations. If e is an oriented edge of T, then s is said to *act coherently* on e if e and se are consistently oriented (that is, if they point in the same direction — neither toward each other nor away from each other). For an automorphism s, if $e \neq se$, then s acts coherently on e if and only if s is hyperbolic and both e and se lie on the axis of s [Serre 1980, Proposition 25].

Lemma 3.1. *Let T be a cocompact G-tree, and let $E \in \partial T$. Then for any geodesic ray τ representing E, any $r \in \mathbb{R}$, and any oriented edge e of T oriented toward E, there exists an element of the G-orbit of e that is oriented toward E and does not lie in $HB_r(\tau)$.*

Proof. The ray of oriented edges beginning at e and representing E, with all edges pointing toward E, contains infinitely many edges. Because the action is

[6] By convention, $S^{-1} = \varnothing$, and (-1)-connected means "nonempty".

cocompact, the pigeon-hole principle ensures that there must be edges e_1 and e_2 from this ray in the same G-orbit. Hence, there is an $h \in G$ with $he_1 = e_2$. Because e_1 and e_2 are consistently oriented, h is hyperbolic. Let v_1 be the terminus of e_1 (the vertex of e_1 where β_τ is maximized). By choosing $k \in \mathbb{Z}$ such that $|k| > \beta_\tau(v_1) - r$ and h^k moves e_1 away from E, we ensure that $h^k e_1$ is oriented toward E and does not lie in $HB_r(\tau)$. Thus $h^k e$ is the edge we seek. □

Observation 3.2. For trees $\tilde{T}$ and T, let $q : \tilde{T} \to T$ be locally surjective. If $\tau = (e_0, e_1, \dots)$ is a geodesic edge ray in T and $\tilde{e}_0$ is an edge of $\tilde{T}$ satisfying $q(\tilde{e}_0) = e_0$, then there exists a lift $\tilde{\tau}$ of τ to $\tilde{T}$ having initial edge $\tilde{e}_0$ and that is also a geodesic edge ray.

Observation 3.3. Given a nonempty connected G-graph Γ and minimal G-tree T, any G-morphism $h : \Gamma \to T$ is surjective.

Proposition 3.4. *Let T be a cocompact G-tree, and let $\tilde{T}$ be a minimal G-tree. Suppose $q : \tilde{T} \to T$ is a G-morphism that is locally surjective. If $E \in \partial T$ is such that there exists a collapsing pair facing E, then E does not lie in $\Sigma^1(\rho)$.*

Proof. Let Γ be a free cocompact G-graph, and choose any G-morphism $h : \Gamma \to \tilde{T}$. Then the composition $q \circ h$ is a suitable control map for determining controlled connectivity over E.

Let $\tau : [0, \infty) \to T$ be a geodesic edge ray representing E. We show that for any lag $\lambda > 0$, there exist points in the subgraph $\Gamma_{(\tau,0)}$ that cannot be connected via a path in $\Gamma_{(\tau,-\lambda)}$.

By Lemma 3.1, we can choose a collapsing pair $(\tilde{e}_1, \tilde{e}_2)$ facing E but whose image in T does not lie in $HB_{-\lambda}(\tau)$. Let $\tilde{v}$ be the vertex shared by $\tilde{e}_1$ and $\tilde{e}_2$, and let v be its image in T. Let γ be the geodesic ray representing E and emanating from v. By Observation 3.2, there exist two distinct lifts $\tilde{\gamma}_i$ $(i = 1, 2)$ of γ to $\tilde{T}$, with $\tilde{\gamma}_i$ having initial edge $\tilde{e}_i$. Because γ and τ both represent E, they eventually merge, so that γ intersects $HB_r(\tau)$ nontrivially for all $r \in \mathbb{R}$. Hence, both $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ intersect $q^{-1}(HB_r(\tau))$ for all r.

By design, $\tilde{\gamma}_1 \cap \tilde{\gamma}_2 = \tilde{v}$, and $\tilde{\gamma}_1 \cup \tilde{\gamma}_2$ is a line. By Observation 3.3, h is onto, so that $\tilde{\gamma}_1 \cup \tilde{\gamma}_2$ lies in the image of h. Choose a vertex $\tilde{y}_i \in \tilde{\gamma}_i \cap q^{-1}(HB_0(\tau))$, and choose $x_i \in h^{-1}(\tilde{y}_i)$. Then both x_i lie in $\Gamma_{(\tau,0)}$, but any path through $\Gamma_{(\tau,-\lambda)}$ joining x_1 to x_2 would be mapped to a path in $q^{-1}(HB_{-\lambda}(\tau))$ joining $\tilde{y}_1$ to $\tilde{y}_2$. Since $\tilde{T}$ is a tree, no such path exists. □

Lemma 3.5. *Let T be a minimal G-tree and let $\mathscr{E}$ be a nonempty G-invariant set of oriented edges. Then there is no vertex v in T such that all edges of $\mathscr{E}$ are oriented away from v.*

Proof. The full subtree of T on the vertex subset

$$\{v \mid \text{each edge of } \mathscr{E} \text{ is oriented away from } v\}$$

is a proper G-invariant subtree. By minimality, this set must be empty. □

Corollary 3.6. *Let T be a cocompact G-tree and let $\tilde{T}$ be a minimal G-tree. Let $q : \tilde{T} \to T$ be a G-morphism that is surjective but not locally injective. Then every vertex of T is faced by a collapsing pair.*

Proof. Let $\tilde{\mathscr{E}}$ be the set of oriented edges of $\tilde{T}$ that are part of a collapsing pair. This is a G-invariant set, and it is nonempty because q is not locally injective. By Lemma 3.5, each vertex $\tilde{v}$ of $\tilde{T}$ must therefore have an edge $\tilde{e}$ in $\tilde{\mathscr{E}}$ oriented toward $\tilde{v}$. Set $v = q(\tilde{v})$. Then if $q(\tilde{e})$ is not oriented toward v, the image of the path from $\tilde{e}$ to $\tilde{v}$ must contain points of backtracking. The point of backtracking closest to v gives rise to a collapsing pair facing v. Because q is surjective, all vertices of T are of this form. □

Observation 3.7. If a cocompact G-tree T has a nonempty G-invariant subtree T', then T is a Hausdorff neighborhood of T'. Hence, T and T' have the same set of end points.

Proposition 3.8. *Let T be a cocompact G-tree and let $\tilde{T}$ be a minimal G-tree. Suppose $q : \tilde{T} \to T$ is a G-morphism that is not locally injective. Then there exists at most one point $E_0 \in \partial T$ such that no collapsing pairs face E_0.*

Proof. By Observation 3.7, the ends of T and the ends of $q(\tilde{T})$ are the same, so we may assume q is surjective. By Corollary 3.6, each vertex of T is faced by a collapsing pair in $\tilde{T}$. If two points of ∂T were not faced by a collapsing pair, then no vertex on the line between them would be faced by a collapsing pair. Hence, there can be at most one point of ∂T not faced by a collapsing pair. □

This proposition has an interesting consequence. If such an end E_0 exists, it must clearly be fixed by ρ. Yet points of the boundary that are fixed by ρ correspond to homomorphisms $G \to \mathbb{R}$, and such an end point lies in $\Sigma^n(\rho)$ if and only if the corresponding homomorphism lies in the BNSR invariant $\Sigma^n(G)$, as discussed in Section 2. Since we only consider simplicial trees, such points in fact correspond to homomorphisms $G \twoheadrightarrow \mathbb{Z}$.

Corollary 3.9. *Under the conditions of Proposition 3.8, if an end point $E_0 \in \partial T$ is faced by no collapsing pair in $\tilde{T}$, then there exists a canonically associated discrete character $\chi : G \to \mathbb{Z}$ such that $E_0 \in \Sigma^n(\rho)$ if and only if $[\chi] \in \Sigma^n(G)$, the BNSR invariant.*

Proof of the Main Theorem. Because q is not locally injective, Proposition 3.8 ensures that there is at most one end point faced by a collapsing pair. Because q is locally surjective, Proposition 3.4 ensures that every end point faced by a collapsing pair lies outside $\Sigma^1(\rho)$. □

The case where stabilizers under $\tilde{\rho}$ have type F_n. Recall the topological construction of the Bass–Serre tree, discussed in [Geoghegan 2008, §6.2; Scott and Wall 1979]: the action $\tilde{\rho}$ corresponds to a graph of groups decomposition of G. From this we can build a $K(G, 1)$ X admitting the quotient $G\backslash\tilde{T}$ as a retract. Let $p : \tilde{X} \twoheadrightarrow X$ be the universal covering projection. There is a natural G-map $h : \tilde{X} \twoheadrightarrow \tilde{T}$, and it is clear from the construction of h that $h^{-1}(A) \subseteq \tilde{X}$ is contractible for any connected subset $A \subseteq \tilde{T}$. If for an integer $n \geq 1$ all point stabilizers under $\tilde{\rho}$ have type F_n, then we can take X to have compact n-skeleton. Hence, letting Γ be the n-skeleton of $\tilde{X}$, the composition $\bar{h} = q \circ h|_\Gamma : \Gamma \to T$ is an appropriate control map for ρ.

Definition 3.10. While the map q does not induce a map $\partial\tilde{T} \to \partial T$, each geodesic ray in T can be lifted to one or more geodesic rays in $\tilde{T}$ (see Observation 3.2) as long as q is locally surjective. Hence, given $E \in \partial T$, we can consider the set $q^{-1}(E) \subseteq \partial\tilde{T}$ of end points represented by lifts of rays representing E.

Lemma 3.11. *If q is locally surjective, then $q^{-1}(E)$ is a singleton if and only if there are no collapsing pairs facing E.*

Proof. Suppose that $q^{-1}(E)$ is not a singleton. Then for τ representing E, there exist two distinct lifts $\tilde{\tau}_1$ and $\tilde{\tau}_2$, representing distinct points $\tilde{E}_1$ and $\tilde{E}_2$ of $\partial\tilde{T}$. If these lifts are not disjoint, then where they split (as they must, eventually) there is a collapsing pair facing E. If they are disjoint, consider the geodesic path P through $\tilde{T}$ connecting them. The image of P in T is a finite subtree of T. Choose any vertex $v \neq \tau(0)$ that is a leaf of this subtree. This leaf and the corresponding edge lie under a collapsing pair of edges of P facing E.

Now suppose there is a collapsing pair $(\tilde{e}_1, \tilde{e}_2)$ of edges of $\tilde{T}$ facing E. Let e be their common image in T, and let ζ be the geodesic ray in T representing E and beginning with the edge e. Then there are distinct lifts $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$ of ζ, each representing a distinct end point of $\tilde{T}$. Hence $q^{-1}(E)$ is not a singleton. □

Proof of Theorem 1.9. If there is a collapsing pair facing E, then by Proposition 3.4, $E \notin \Sigma^1(\rho)$.

If there is no collapsing pair facing E, we take the control map $\bar{h}$ described above. By construction of $\bar{h}$, we need only show that $q^{-1}(HB_r(\tau))$ is connected for any horoball $HB_r(\tau)$ about E.

For $i = 1, 2$, let $\tilde{z}_i$ be a point in $q^{-1}(HB_r(\tau))$, and let z_i be its image in T. We find a path between $\tilde{z}_1$ and $\tilde{z}_2$ lying in $q^{-1}(HB_r(\tau))$.

There is a unique geodesic ray ζ_i in T that emanates from z_i and represents E. Let $\tilde{\zeta}_i$ be the lift of ζ_i to $\tilde{T}$ emanating from $\tilde{z}_i$. Since ζ_i lies in $HB_r(\tau)$, $\tilde{\zeta}_i$ lies in $q^{-1}(HB_r(\tau))$. Also, since $q^{-1}(E)$ is a singleton, $\tilde{\zeta}_1(\infty) = \tilde{\zeta}_2(\infty)$. Hence, $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$ must eventually merge. The closure of $(\operatorname{im}\tilde{\zeta}_1 \cup \operatorname{im}\tilde{\zeta}_2) - (\operatorname{im}\tilde{\zeta}_1 \cap \operatorname{im}\tilde{\zeta}_2)$ is the geodesic connecting $\tilde{z}_1$ to $\tilde{z}_2$. □

Acknowledgements

This paper is a development of part of the author's Ph.D. dissertation at SUNY Binghamton. The author is grateful to his Ph.D. advisor Ross Geoghegan for his constant support and encouragement. Also, the clarity and elegance of this paper have benefited significantly from the suggestions of anonymous referees, whose time and effort are greatly appreciated.

References

[Bass 1993] H. Bass, "Covering theory for graphs of groups", *J. Pure Appl. Algebra* **89**:1-2 (1993), 3–47. MR 94j:20028 Zbl 0805.57001

[Bieri and Geoghegan 2003a] R. Bieri and R. Geoghegan, "Connectivity properties of group actions on non-positively curved spaces", *Mem. Amer. Math. Soc.* **161**:765 (2003), xiv+83. MR 2004m:57001 Zbl 1109.20035

[Bieri and Geoghegan 2003b] R. Bieri and R. Geoghegan, "Topological properties of SL_2 actions on the hyperbolic plane", *Geom. Dedicata* **99** (2003), 137–166. MR 2004e:20068 Zbl 1039.20020

[Bieri et al. 1987] R. Bieri, W. D. Neumann, and R. Strebel, "A geometric invariant of discrete groups", *Invent. Math.* **90**:3 (1987), 451–477. MR 89b:20108 Zbl 0642.57002

[Bieri et al. 2010] R. Bieri, R. Geoghegan, and D. H. Kochloukova, "The sigma invariants of Thompson's group F", *Groups Geom. Dyn.* **4**:2 (2010), 263–273. MR 2595092 (2011h:20113 Zbl 1214.20048

[Bridson and Haefliger 1999] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften **319**, Springer, Berlin, 1999. MR 2000k:53038 Zbl 0988.53001

[Geoghegan 2008] R. Geoghegan, *Topological methods in group theory*, Graduate Texts in Mathematics **243**, Springer, New York, 2008. MR 2008j:57002 Zbl 1141.57001

[Lehnert 2009] R. Lehnert, *Kontrollierter zusammenhang von gruppenoperationen auf bäumen*, Diploma thesis, Goethe Universität Frankfurt am Main, 2009.

[Meinert 1996] H. Meinert, "The homological invariants for metabelian groups of finite Prüfer rank: a proof of the Σ^m-conjecture", *Proc. London Math. Soc.* (3) **72**:2 (1996), 385–424. MR 98b:20082 Zbl 0852.20042

[Meinert 1997] H. Meinert, "Actions on 2-complexes and the homotopical invariant Σ^2 of a group", *J. Pure Appl. Algebra* **119**:3 (1997), 297–317. MR 98g:20084 Zbl 0879.57010

[Ontaneda 2005] P. Ontaneda, "Cocompact CAT(0) spaces are almost geodesically complete", *Topology* **44**:1 (2005), 47–62. MR 2005m:57002 Zbl 1068.53026

[Rehn 2007] W. H. Rehn, *Kontrollierter Zusammenhang über symmetrischen Räumen*, Ph.D. thesis, Goethe Universität Frankfurt am Main, 2007, Available at http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/360.

[Scott and Wall 1979] P. Scott and C. T. C. Wall, "Topological methods in group theory", pp. 137–203 in *Homological group theory*, edited by C. T. C. Wall, London Math. Soc. Lecture Note Ser. **36**, Cambridge Univ. Press, 1979. MR 81m:57002 Zbl 0423.20023

[Serre 1980] J.-P. Serre, *Trees*, Springer, Berlin, 1980. MR 82c:20083 Zbl 0548.20018

Received April 29, 2011. Revised September 28, 2011.

Keith Jones
Department of Mathematics
Trinity College
300 Summit Street
Hartford, CT 06106
United States
keith.jones@trincoll.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

REMARKS ON THE CURVATURE BEHAVIOR AT THE FIRST SINGULAR TIME OF THE RICCI FLOW

NAM Q. LE AND NATASA SESUM

We study the curvature behavior at the first singular time of a solution to the Ricci flow

$$\frac{\partial}{\partial t} g_{ij} = -2R_{ij}, \quad t \in [0, T),$$

on a smooth, compact n-dimensional Riemannian manifold M. If the flow has uniformly bounded scalar curvature and develops Type I singularities at T, we show that suitable blow-ups of the evolving metrics converge in the pointed Cheeger–Gromov sense to a Gaussian shrinker by using Perelman's $\mathcal{W}$-functional. If the flow has uniformly bounded scalar curvature and develops Type II singularities at T, we show that suitable scalings of the potential functions in Perelman's entropy functional converge to a positive constant on a complete, Ricci flat manifold. We also show that if the scalar curvature is uniformly bounded along the flow in certain integral sense then the flow either develops a Type II singularity at T or it can be smoothly extended past time T.

1. Introduction

The Ricci flow and previous results. Let M be a smooth, compact n-dimensional Riemannian manifold without boundary equipped with a smooth Riemannian metric g_0, where $n \geq 3$. Let $g(t)$, $0 \leq t < T$, be a one-parameter family of metrics on M. The Ricci flow equation on M with initial metric g_0

$$\frac{\partial}{\partial t} g(t) = -2\,\mathrm{Ric}(g(t)), \quad g(0) = g_0. \tag{1-1}$$

was introduced in the seminal paper [Hamilton 1982]. It is a weakly parabolic system of equations whose short-time existence was proved by Hamilton using the Nash–Moser implicit function theorem in the same paper and after that simplified by DeTurck [1983]. The goal in the analysis of (1-1) is to understand the long-time behavior of the flow and possible singularity formation or convergence of the flow in the cases when we do have a long-time existence. In general, the behavior of the flow can give insight into the topology of the underlying manifold. One of the

Sesum was partially supported by NSF grant 0905749.
MSC2010: primary 53C44; secondary 35K10.
Keywords: Ricci flow, scalar curvature, evolution.

great successes is the resolution of the Poincaré conjecture by Perelman. In order to discuss the long-time behavior we have to understand what happens at the singular time and also what the optimal conditions for having a smooth solution are.

Hamilton [1995b] showed that if the norm of Riemannian curvature $|\mathrm{Rm}|(g(t))$ stays uniformly bounded in time for all $t \in [0, T)$ with $T < \infty$, then we can extend the flow (1-1) smoothly past time T. In other words, either the flow exists forever or the norm of Riemannian curvature blows up in finite time. Wang [2008] and Ye [2008] extended this result, assuming certain integral bounds on the Riemannian curvature. Namely, if

$$\int_0^T \int_M |\mathrm{Rm}|^\alpha \, d\mathrm{vol}_{g(t)} \, dt \le C \quad \text{for some } \alpha \ge \frac{n+2}{2},$$

then the flow can be extended smoothly past time T. Throughout the paper, $d\mathrm{vol}_g$ denotes the Riemannian volume density on (M, g). On the other hand, Sesum [2005] improved Hamilton's extension result and showed if the norm of Ricci curvature is uniformly bounded over a finite time interval $[0, T)$, then we can extend the flow smoothly past time T. Wang [2008] improved this even further, showing that if Ricci curvature is uniformly bounded from below and if the space-time integral of the scalar curvature is bounded, say

$$\int_0^T \int_M |R|^\alpha \, d\mathrm{vol}_{g(t)} \, dt \le C \quad \text{for } \alpha \ge \frac{n+2}{2},$$

where R is the scalar curvature, then we can extend the flow smoothly past time T. The requirement on Ricci curvature in [Wang 2008] is rather restrictive. Ricci flow does not in general preserve nonnegative Ricci curvature in dimensions $n \ge 4$. See [Knopf 2006] for noncompact examples starting in dimension $n = 4$ and [Böhm and Wilking 2007] for compact examples starting in dimension $n = 12$. Recently, Maximo [2011] brought the result of [Böhm and Wilking 2007] down to dimension 4 by showing that nonnegative Ricci curvature is not preserved under Ricci flow for closed compact manifolds of dimensions 4 and above. Without assuming the boundedness from below of Ricci curvature, Ma and Cheng [2010] proved that the norm of Riemannian curvature can be controlled given integral bounds on the scalar curvature R and the Weyl tensor W from the orthogonal decomposition of the Riemannian curvature tensor. Their bounds are of the form

$$\int_0^T \int_M (|R|^\alpha + |W|^\alpha) \, d\mathrm{vol}_{g(t)} \, dt \le C \quad \text{for } \alpha \ge \frac{n+2}{2}.$$

This is not surprising since Knopf [2009] has shown that the trace-free Ricci tensor is controlled pointwise by the scalar curvature and the Weyl tensor without any additional hypotheses. Zhang [2010] proved that the scalar curvature controls the Kähler Ricci flow $\frac{\partial}{\partial t} g_{i\bar{j}} = -R_{i\bar{j}} - g_{i\bar{j}}$ starting from any Kähler metric g_0.

Main results. The above results, in particular that of [Zhang 2010], support the belief that the scalar curvature should control the Ricci flow in the Riemannian setting as well. Enders, Müller and Topping [2010] justified this belief for Type I Ricci flow:

Theorem 1.1 [Enders et al. 2010]. *Let M be a smooth, compact n-dimensional Riemannian manifold equipped with a smooth Riemannian metric g_0 and $g(\cdot\,,t)$ be a solution to the Type I Ricci flow* (1-1) *on M. Assume there is a constant C so that* $\sup_M |R(\cdot\,,t)| \le C$ *for all $t \in [0,T)$ and $T < \infty$. Then we can extend the flow past time T.*

Their proof was based on a blow-up argument using Perelman's reduced distance and pseudolocality theorem.

Assume the flow (1-1) develops a singularity at $T < \infty$.

Definition 1.1. We say that (1-1) has a *Type I singularity at T* if there exists a constant $C > 0$ such that for all $t \in [0,T)$

$$\max_M |\mathrm{Rm}(\cdot\,,t)| \cdot (T-t) \le C. \tag{1-2}$$

Otherwise we say the flow develops *Type II singularity at T*. Moreover, the flow that satisfies (1-2) will be referred to as to the *Type I Ricci flow.*

In this paper, we also use a blow-up argument to study curvature behavior at the first singular time of the Ricci flow. We deal with both Type I and II singularities. Assume that the scalar curvature is uniformly bounded along the flow. If the flow develops Type I singularities at some finite time T then by using Perelman's entropy functional $\mathcal{W}$, we show that suitable blow-ups of the evolving metrics converge in the pointed Cheeger–Gromov sense to a Gaussian shrinker.

Theorem 1.2. *Let M be a smooth, compact n-dimensional Riemannian manifold ($n \ge 3$) and $g(\cdot\,,t)$ be a solution to the Ricci flow* (1-1) *on M. Assume there is a constant C so that* $\sup_M |R(\cdot\,,t)| \le C$ *for all $t \in [0,T)$ and $T < \infty$. Assume that at T we have a Type I singularity and the norm of the curvature operator blows up. Then by suitably rescaling the metrics, we get a Gaussian shrinker in the limit.*

A simple consequence of the proof of Theorem 1.2 is following result, which is also proved in [Naber 2010]. Instead of the reduced distance techniques used by Naber, we use Perelman's monotone functional $\mathcal{W}$.

Corollary 1.1. *Let M be a smooth, compact n-dimensional Riemannian manifold ($n \ge 3$) and $g(\cdot\,,t)$ be a solution to the Ricci flow* (1-1) *on M. If the flow has a Type I singularity at T, then a suitable rescaling of the solution converges to a gradient shrinking Ricci soliton.*

Naber [2010] proved that in the case of a Type I singularity, a suitable rescaling of the flow converges to gradient shrinking Ricci soliton. Enders, Müller and Topping [2010] recently showed that the limiting soliton represents a singularity model, that is, it is nonflat (see also [Cao and Zhang 2011]). The open question is whether using Perelman's $\mathcal{W}$-functional, one can produce in the limit a singularity model (*nonflat* gradient shrinking Ricci solitons). We prove some interesting estimates on the minimizers of Perelman's $\mathcal{W}$-functional which could be of independent interest.

On the other hand, if the flow develops Type II singularities at some finite time T, then we show that suitable scalings of the potential functions in Perelman's entropy functional converge to a positive constant on a complete, Ricci flat manifold which is the pointed Cheeger–Gromov limit of a suitably chosen sequence of blow-ups of the original evolving metrics.

Theorem 1.3. *Let M be a smooth, compact n-dimensional Riemannian manifold* ($n \geq 3$) *and $g(\cdot, t)$ be a solution to the Ricci flow* (1-1) *on M. Assume there is a constant C so that $\sup_M |R(\cdot, t)| \leq C$ for all $t \in [0, T)$ and $T < \infty$. Assume that at T we have a Type II singularity and the norm of the curvature operator blows up. Let ϕ_i be as in the proof of Theorem 1.2* (see, for example, (3-9)). *Then by suitably rescaling the metrics and ϕ_i, we get as a limit of ϕ_i a positive constant on a complete, Ricci flat manifold.*

We believe that Theorem 1.3 may play a role in proving the nonexistence of Type II singularities if the scalar curvature is uniformly bounded along the flow. We are still investigating this issue.

For a precise definition of ϕ_i, see Section 3.

There has been a striking analogy between the Ricci flow and the mean curvature flow for decades now. Around the same time Hamilton proved that the norm of the Riemannian curvature under the Ricci flow must blow up at a finite singular time, Huisken [1984] showed that the norm of the second fundamental form of an evolving hypersurface under the mean curvature flow must blow up at a finite singular time. The analogue of Wang's result holds for the mean curvature flow as well [Le and Sesum 2011], namely if the second fundamental form of an evolving hypersurface is uniformly bounded from below and if the mean curvature is bounded in a certain integral sense, then we can smoothly extend the flow. In the follow-up paper [Le and Sesum 2010] the authors show that given only the uniform bound on the mean curvature of the evolving hypersurface, the flow either develops a Type II singularity or can be smoothly extended. In the case the dimension of the evolving hypersurfaces is 2 they show that under some density assumptions one can smoothly extend the flow provided that the mean curvature is uniformly bounded. Finally, in contrast to the lower bound on the scalar curvature (2-3), at the first singular time of the mean curvature flow, the mean curvature can either tend to ∞ (as in the case of a round sphere) or $-\infty$ as in some examples of Type II singularities [Angenent and Velázquez 1997].

If we replace the pointwise scalar curvature bound in Theorem 1.1 with an integral bound, we can prove the following theorems.

Theorem 1.4. *If $g(\cdot, t)$ solves* (1-1) *and if*

$$\int_M |R|^\alpha(t)\, d\mathrm{vol}_{g(t)} \le C_\alpha \tag{1-3}$$

for all $t \in [0, T)$ where $\alpha > n/2$ and $T < \infty$, then either the flow develops a Type II singularity at T or the flow can be smoothly extended past time T.

Remark 1.1. The condition on α in Theorem 1.4 is optimal. Let (S^n, g_0) be the space form of constant sectional curvature 1. The Ricci flow on $M = S^n$ with initial metric g_0 has the solution $g(t) = (1 - 2(n-1)t)g_0$. Therefore $T = 1/(2(n-1))$ is the maximal existence time. Rewrite $g(t) = 2(n-1)(T-t)g_0$ to compute

$$\begin{aligned}\int_M |R|^\alpha(t)\, d\mathrm{vol}_{g(t)} &= \mathrm{vol}_{g(t)}(M)\Big(\frac{n}{2(T-t)}\Big)^\alpha \\ &= \mathrm{vol}_{g(0)}(M)(2(n-1)(T-t))^{n/2}\Big(\frac{n}{2(T-t)}\Big)^\alpha \\ &= \mathrm{vol}_{g(0)}(M)2^{n/2-\alpha}(n-1)^{n/2}n^\alpha \frac{1}{(T-t)^{\alpha-n/2}}.\end{aligned}$$

Hence $\int_M |R|^\alpha(t)\, d\mathrm{vol}_{g(t)}$ tends to ∞ as $t \to T$ if and only if $\alpha > n/2$.

Theorem 1.5. *If $g(\cdot, t)$ solves* (1-1) *and if we have the space-time integral bound*

$$\int_0^T \int_M |R|^\alpha(t)\, d\mathrm{vol}_{g(t)}\, dt \le C_\alpha \tag{1-4}$$

for $\alpha \ge (n+2)/2$, then the flow either develops a Type II singularity at T or can be smoothly extended past time T.

Remark 1.2. The condition on α in Theorem 1.5 is optimal. As in Remark 1.1 consider the Ricci flow on the round sphere. Following the computation in Remark 1.1 we get

$$\int_0^T \int_M |R|^\alpha\, d\mathrm{vol}_{g(t)}\, dt = \mathrm{vol}_{g(0)}(M)2^{n/2-\alpha}(n-1)^{n/2}n^\alpha \int_0^T \frac{1}{(T-t)^{\alpha-n/2}}\, dt,$$

and therefore the integral is ∞ if and only if $\alpha \ge (n+2)/2$.

For the mean curvature flow, a similar result to Theorem 1.5 has been obtained in [Le and Sesum 2010].

The rest of the paper is organized as follows. In Section 2 we give some necessary preliminaries. Section 3 is devoted to the statements and proofs of Theorems 1.2 and 1.3. In Section 4 we prove Theorems 1.4 and 1.5.

2. Preliminaries

In this section, we recall basic evolution equations during the Ricci flow and the definition of singularity formation. Then we recall Perelman's entropy functional $\mathcal{W}$ and in Lemma 2.1 prove one of its properties, nonpositivity of the μ-energy. The nonpositivity of the μ-energy turns out to be very crucial for the proof of Theorem 1.1.

Evolution equations and singularity formation. Consider the Ricci flow (1-1) on $[0, T)$. Then, the scalar curvature R and the volume form $\mathrm{vol}_{g(t)}$ evolve by

$$\frac{\partial}{\partial t} R = \Delta R + 2|\mathrm{Ric}|^2, \tag{2-1}$$

$$\frac{\partial}{\partial t} \mathrm{vol}_{g(t)} = -R\, \mathrm{vol}_{g(t)}\,. \tag{2-2}$$

Because $|\mathrm{Ric}|^2 \geq R^2/n$, the maximum principle applied to (2-1) yields

$$R(g(t)) \geq \frac{\min_M R(g(0))}{1 - (2 \min_M R(g(0))t)/n}. \tag{2-3}$$

If $T < +\infty$ and the norm of the Riemannian curvature $|\mathrm{Rm}|(g(t))$ becomes unbounded as t tends to T, we say the Ricci flow develops singularities as t tends to T and T is a singular time. It is well-known that the Ricci flow generally develops singularities.

If a solution $(M, g(t))$ to the Ricci flow develops singularities at $T < +\infty$, then according to [Hamilton 1995b], we say that it develops a *Type I singularity* if

$$\sup_{t\in[0,T)} (T - t) \max_M |\mathrm{Rm}(\,\cdot\,, t)| < +\infty,$$

and it develops a *Type II singularity* if

$$\sup_{t\in[0,T)} (T - t) \max_M |\mathrm{Rm}(\,\cdot\,, t)| = +\infty.$$

Clearly, the Ricci flow of a round sphere develops a Type I singularity in finite time. The existence of Type II singularities for the Ricci flow has been recently established in [Gu and Zhu 2008], proving the degenerate neckpinch conjecture of [Hamilton 1995b].

Finally, by the curvature gap estimate for Ricci flow solutions with a finite-time singularity (see, for example, [Chow et al. 2006, Lemma 8.7]), we have

$$\max_{x\in M} |\mathrm{Rm}(x, t)| \geq \frac{1}{8(T-t)}. \tag{2-4}$$

Perelman's entropy functional $\mathcal{W}$ and the μ-energy. Perelman [2002] introduced a very important functional, the entropy functional $\mathcal{W}$, for the study of the Ricci flow:

$$\mathcal{W}(g, f, \tau) = (4\pi\tau)^{-n/2}\int_M \big(\tau(R+|\nabla f|^2)+f-n\big)\, e^{-f}\, d\mathrm{vol}_g, \tag{2-5}$$

under the constraint $(4\pi\tau)^{-n/2}\int_M e^{-f}\, d\mathrm{vol}_g = 1$. The functional $\mathcal{W}$ is invariant under the parabolic scaling of the Ricci flow and invariant under diffeomorphism. Namely, for any positive number α and any diffeomorphism φ, we have $\mathcal{W}(\alpha\varphi^* g, \varphi^* f, \alpha\tau) = \mathcal{W}(g, f, \tau)$. Perelman showed that if $\dot\tau = -1$ and $f(\cdot\,, t)$ is a solution to the backwards heat equation

$$\frac{\partial f}{\partial t} = -\Delta f + |\nabla f|^2 - R + \frac{n}{2\tau}, \tag{2-6}$$

and if $g(\cdot\,, t)$ solves the Ricci flow (1-1) then

$$\frac{d}{dt}\mathcal{W}(g(t), f(t), \tau) = (2\tau)\cdot(4\pi\tau)^{-n/2}\int_M \left|R_{ij}+\nabla_i\nabla_j f - \frac{g_{ij}}{2\tau}\right|^2 e^{-f}\, d\mathrm{vol}_{g(t)} \geq 0.$$

The functional $\mathcal{W}$ is constant on metrics g with the property that

$$R_{ij}+\nabla_i\nabla_j f - \frac{g_{ij}}{2\tau} = 0$$

for a smooth function f. These metrics are called gradient shrinking Ricci solitons and appear often as singularity models, that is, limits of blown up solutions around finite-time singularities of the Ricci flow.

Let $g(t)$ be a solution to the Ricci flow (1-1) on $(-\infty, T)$. We call a triple $(M, g(t), f(t))$ on $(-\infty, T)$ with smooth functions $f: M \to \mathbb{R}$ a *gradient shrinking soliton in canonical form* if it satisfies

$$\mathrm{Ric}(g(t)) + \nabla^{g(t)}\nabla^{g(t)} f(t) - \frac{1}{2(T-t)}g(t) = 0 \quad \text{and} \quad \frac{\partial}{\partial t} f(t) = |f(t)|^2_{g(t)}. \tag{2-7}$$

Perelman also defines the μ-energy

$$\mu(g, \tau) = \inf \mathcal{W}(g, f, \tau) \quad \text{over } \left\{ f \mid (4\pi\tau)^{-n/2}\textstyle\int_M e^{-f}\, d\mathrm{vol}_g = 1\right\} \tag{2-8}$$

and shows that

$$\frac{d}{dt}\mu(g(\cdot\,, t), \tau) \geq (2\tau)\cdot(4\pi\tau)^{-n/2}\int_M \left|R_{ij}+\nabla_i\nabla_j - \frac{g_{ij}}{2\tau}\right|^2 e^{-f}\, d\mathrm{vol}_{g(t)} \geq 0, \tag{2-9}$$

where $f(\cdot\,, t)$ is the minimizer for $\mathcal{W}(g(\cdot\,, t), f, \tau)$ with the constraint on f as in (2-8). The μ-energy $\mu(g, \tau)$ corresponds to the best constant of a logarithmic Sobolev inequality. Adjusting some of Perelman's arguments to our situation we get the following lemma whose proof we include for the reader's convenience.

Lemma 2.1 (nonpositivity of the μ-energy). *If $g(t)$ is a solution to* (1-1) *for all $t \in [0, T)$, then $\mu(g(t), T - t) \le 0$ for all $t \in [0, T)$.*

Proof. We are assuming the Ricci flow exists for all $t \in [0, T)$. Fix $t \in [0, T)$. Define $\tilde g(s) = g(t+s)$ for $s \in [0, T-t)$. Pick any $\bar\tau < T - t$. Let $\tau_0 = \bar\tau - \varepsilon$ with $\varepsilon > 0$ small. Pick $p \in M$. We use normal coordinates about p on $(M, \tilde g(\tau_0))$ to define

$$f_1(x) = \begin{cases} |x|^2/4\varepsilon & \text{if } d_{\tilde g(\tau_0)}(x, x_0) < \rho_0, \\ \rho_0^2/4\varepsilon & \text{elsewhere,} \end{cases} \tag{2-10}$$

where $\rho_0 > 0$ is smaller than the injectivity radius. Note that $d\mathrm{vol}_{\tilde g(\tau_0)}(x) = 1 + O(|x|^2)$ near p. We compute

$$\begin{aligned}\int_M &(4\pi\varepsilon)^{-n/2} e^{-f_1}\, d\mathrm{vol}_{\tilde g(\tau_0)} \\ &= \int_{|x|\le\rho_0} (4\pi\varepsilon)^{-n/2} e^{-|x|^2/4\varepsilon}(1 + O(|x|^2))\,dx + O(\varepsilon^{-n/2}e^{-\rho_0^2/4\varepsilon}) \\ &= \int_{|y|\le\rho_0/\sqrt{\varepsilon}} (4\pi)^{-n/2} e^{-|y|^2/4}(1 + O(\varepsilon|y|^2))\,dy + O(\varepsilon^{-n/2}e^{-\rho_0^2/4\varepsilon}).\end{aligned}$$

The second term goes to zero as $\varepsilon \to 0$ while the first term converges to

$$\int_{\mathbb{R}^n} (4\pi)^{-n/2} e^{-|y|^2/4}\,dy = 1.$$

Writing the integral as e^C, then $C \to 0$ as $\varepsilon \to 0$. And $f = f_1 + C$ then satisfies the constraint $\int_M (4\pi\varepsilon)^{-n/2} e^{-f}\, d\mathrm{vol}_{\tilde g(\tau_0)} = 1$.

Solve Equation (2-6) backwards with initial value f at τ_0. Then

$$\begin{aligned}\mathcal{W}&(\tilde g(\tau_0), f(\tau_0), \bar\tau - \tau_0) \\ &= \int_{|x|\le\rho_0} \left(\varepsilon\left(\frac{|x|^2}{4\varepsilon^2} + R\right) + \frac{|x|^2}{4\varepsilon} + C - n\right)(4\pi\varepsilon)^{-n/2} e^{-|x|^2/4\varepsilon - C}(1 + O(|x|^2))\,dx \\ &\qquad + \int_{M - B(p,\rho_0)} \left(\frac{\rho_o^2}{4\varepsilon} + \varepsilon R + C - n\right)(4\pi\varepsilon)^{-n/2} e^{-r_0^2/4\varepsilon - C} \\ &= \mathrm{I} + \mathrm{II},\end{aligned}$$

where $\mathrm{I} = e^{-C}\int_{|x|\le\rho_0} (|x|^2/2\varepsilon - n)(4\pi\varepsilon)^{-n/2} e^{-|x|^2/4\varepsilon}(1 + O(|x|^2))\,dx$ and II contains all the remaining terms. It is obvious that $\mathrm{II} \to 0$ as $\varepsilon \to 0$ while

$$\begin{aligned}\mathrm{I} &= e^{-C}\int_{|y|\le\rho_0/\sqrt{\varepsilon}} \left(\frac{|y|^2}{2} - n\right)(4\pi)^{-n/2} e^{-|y|^2/4}(1 + O(\varepsilon|y|^2))\,dy \\ &\to \int_{\mathbb{R}^n} \left(\frac{|y|^2}{2} - n\right)(4\pi)^{-n/2} e^{-|y|^2/4}\,dy = 0 \quad \text{as } \varepsilon \to 0.\end{aligned}$$

Therefore $\mathcal{W}(\tilde{g}(\tau_0), f(\tau_0), \bar{\tau}-\tau_0) \to 0$ as $\tau_0 \to \bar{\tau}$. By the monotonicity of μ along the flow, $\mu(g(t), \bar{\tau}) = \mu(\tilde{g}(0), \bar{\tau}) \le \mathcal{W}(\tilde{g}(0), f(0), \bar{\tau}) \le \mathcal{W}(\tilde{g}(\tau_0), f(\tau_0), \bar{\tau}-\tau_0)$. Letting $\tau_0 \to \bar{\tau}$, we get $\mu(g(t), \bar{\tau}) \le 0$. Since $\bar{\tau} < T-t$ is arbitrary,

$$\mu(g(t), T-t) \le 0. \qquad \square$$

3. Uniform bound on scalar curvature

In this section, we prove Theorems 1.2 and 1.3.

Proof of Theorem 1.2. By our assumptions, there exists a sequence of times $t_i \to T$ so that $Q_i := \max_{M\times[0,t_i]} |\mathrm{Rm}|(x,t) \to \infty$ as $i \to \infty$. Assume that the maximum is achieved at $(p_i, t_i) \in M \times [0, t_i]$. Define a rescaled sequence of solutions

$$g_i(t) = Q_i \cdot g(t_i + t/Q_i). \tag{3-1}$$

We have that

$$|\mathrm{Rm}(g_i)| \le 1 \text{ on } M \times [-t_i Q_i, 0] \quad \text{and} \quad |\mathrm{Rm}(g_i)|(p_i, 0) = 1. \tag{3-2}$$

By Hamilton's compactness theorem [1995a] and Perelman's κ-noncollapsing theorem [2002] we can extract a pointed subsequence of solutions $(M, g_i(t), q_i)$, converging in the Cheeger–Gromov sense to a solution to (1-1), which we denote by $(M_\infty, g_\infty(t), q_\infty)$ for any sequence of points $q_i \in M$. In particular, if we take that sequence of points to be exactly $\{p_i\}$, we can guarantee the limiting metric is nonflat. The limiting metric has a sequence of nice properties: Since

$$|R(g_i(t))| = \frac{|R(g(t_i + t/Q_i))|}{Q_i} \le \frac{C}{Q_i} \to 0,$$

the limiting solution $(M_\infty, g_\infty(t))$ is scalar flat for each $t \in (-\infty, 0]$. Since it solves the Ricci flow (1-1) and $R_\infty := R(g_\infty)$ evolves by

$$\frac{\partial}{\partial t} R_\infty = \Delta R_\infty + 2|\operatorname{Ric}(g_\infty)|^2,$$

we have that $\operatorname{Ric}(g_\infty) \equiv 0$, that is, the limiting metric is Ricci flat. We will get a Gaussian shrinker by using Perelman's functional μ defined by (2-8). Recall that (see the computation in [Kleiner and Lott 2008])

$$\frac{d}{dt}\mu(g(t), \tau) \ge 2\tau \cdot (4\pi\tau)^{-n/2} \int_M \left| \operatorname{Ric} + \nabla\nabla f - \frac{g}{2\tau} \right|^2 e^{-f} \, d\mathrm{vol}_{g(t)},$$

where $f(\cdot, t)$ is the minimizer realizing $\mu(g(t), \tau)$, and $\tau = T - t$.

In this proof of Theorem 1.2, we take $s, v \in [-10, 0]$ *with* $s < v$. Then, by (3-2), $g_i(s)$ and $g_i(v)$ are defined for i sufficiently large. Then, by the invariant property of μ under the parabolic scaling of the Ricci flow, for $s < v \in [-10, 0]$ one has

$$\begin{aligned}
(3\text{-}3)\quad & \mu(g_i(v), Q_i(T-t_i)-v) - \mu(g_i(s), Q_i(T-t_i)-s) \\
&= \mu\Big(g\Big(t_i+\frac{v}{Q_i}\Big), T-t_i-\frac{v}{Q_i}\Big) - \mu\Big(g\Big(t_i+\frac{s}{Q_i}\Big), T-t_i-\frac{s}{Q_i}\Big) \\
&= \int_{t_i+s/Q_i}^{t_i+v/Q_i} \frac{d}{dt}\mu(g(t), T-t)\,dt \\
&\geq \int_{t_i+s/Q_i}^{t_i+v/Q_i} \int_M 2\tau(4\pi\tau)^{-n/2}\cdot \Big|\mathrm{Ric}+\nabla\nabla f - \frac{g}{2\tau}\Big|^2 e^{-f}\, d\mathrm{vol}_{g(t)}\, dt \\
&= 2\int_s^v \int_M \Big(m_i(r)(4\pi m_i(r))^{-n/2} \\
&\qquad\qquad \cdot \Big|\mathrm{Ric}(g_i(r)) + \nabla\nabla f - \frac{g_i}{2m_i(r)}\Big|^2 e^{-f}\Big)\, d\mathrm{vol}_{g_i(r)}\, dr,
\end{aligned}$$

where, for simplicity, $m_i(r) = Q_i(T-t_i) - r$.

Since we are assuming the flow develops a Type I singularity at T, we have

$$(3\text{-}4)\qquad \lim_{i\to\infty} Q_i(T-t_i) = a < \infty.$$

Thus, by (2-4), one has for $r \in [-10, 0]$,

$$(3\text{-}5)\qquad \lim_{i\to\infty} m_i(r) = a - r > 0.$$

By Lemma 2.1 and by the monotonicity of $\mu(g(t), T-t)$ (see (2-9)),

$$(3\text{-}6)\qquad \mu(g(0), T) \leq \mu(g(t), T-t) \leq 0.$$

Estimate (3-6) implies that there exists a finite $\lim_{t\to T}\mu(g(t), T-t)$ which implies that the left-hand side of (3-3) tends to zero as $i\to\infty$. Letting $i\to\infty$ in (3-3) and using (3-5), we get

$$\begin{aligned}
(3\text{-}7)\quad & \lim_{i\to\infty}\int_s^v \int_M \Big((a-r)(4\pi(a-r))^{-n/2} \\
&\qquad \times \Big|\mathrm{Ric}(g_i) + \nabla\nabla f - \frac{g_i}{2(a-r)}\Big|^2 e^{-f}\Big)\, d\mathrm{vol}_{g_i(r)}\, dr = 0.
\end{aligned}$$

We would like to say that we can extract a subsequence so that $f(\,\cdot\,, t_i + r/Q_i)$ converges smoothly to a smooth function $f_\infty(r)$ on $(M_\infty, g_\infty(r))$, which will then be a potential function for a limiting gradient shrinking Ricci soliton g_∞. In order to do that, we need some uniform estimates for $f(\,\cdot\,, t_i + r/Q_i)$. The equation satisfied by $f(t_i + r/Q_i)$ in (3-3) is

$$(3\text{-}8)\quad \Big(T-t_i-\frac{r}{Q_i}\Big)(2\Delta f - |\nabla f|^2 + R) + f - n = \mu\Big(g\Big(t_i+\frac{r}{Q_i}\Big), T-t_i-\frac{r}{Q_i}\Big).$$

Let $f_i(\cdot, r) = f(\cdot, t_i + r/Q_i)$. Then

$$(Q_i(T-t_i)-r)(2\Delta_{g_i(r)} f_i(r) - |\nabla_{g_i(r)} f_i(r)|^2 + R(g_i(r))) + f_i(r) - n \\ = \mu(g_i(r), Q_i(T-t_i)-r).$$

Define $\phi_i(\cdot, r) = e^{-f_i(\cdot, r)/2}$. This function $\phi_i(\cdot, r)$ satisfies a nice elliptic equation

$$(Q_i(T-t_i)-r)(-4\Delta_{g_i(r)} + R(g_i(r)))\phi_i \\ = 2\phi_i \log\phi_i + (\mu(g_i(r), Q_i(T-t_i)-r)+n)\phi_i. \tag{3-9}$$

Recall that, in this proof of Theorem 1.2, we consider $r \in [-10, 0]$. We take the liberty of suppressing certain dependencies on r whenever no confusion may arise.

Our first estimates are uniform global $W^{1,2}$ estimates for $\phi_i(r)$:

Lemma 3.1. *There exists a uniform constant C so that for all $r \in [-10, 0]$ and all i, one has*

$$\int_M \phi_i^2(\cdot, r)\, d\mathrm{vol}_{g_i(r)} + \int_M |\nabla_{g_i(r)}\phi_i(\cdot, r)|^2\, d\mathrm{vol}_{g_i(r)} \le C(Q_i(T-t_i)-r)^{n/2} \le \widetilde{C}.$$

Proof. The function $\phi_i(r)$ satisfies the L^2-constraint

$$\int_M (4\pi m_i(r))^{-n/2}(\phi_i(r))^2\, d\mathrm{vol}_{g_i(r)} = 1$$

and is in fact smooth [Rothaus 1981]. Here, we have used $m_i(r) = Q_i(T-t_i)-r$.

To simplify, let $F_i(r) = \phi_i(r)/c_i(r)$, where $c_i(r) = (4\pi m_i(r))^{n/4}$. Then

$$\int_M (F_i(r))^2\, d\mathrm{vol}_{g_i(r)} = 1,$$

and the equation for $F_i(r)$ becomes

$$m_i(r)(-4\Delta_{g_i(r)} + R(g_i(r)))F_i(r) \\ = 2F_i(r)\log F_i(r) + (\mu(g_i(r), m_i(r)) + n + 2\log c_i(r))F_i(r).$$

Introduce

$$\mu_i(r) = \mu(g_i(r), m_i(r)) + n + 2\log c_i(r).$$

Then

$$-\Delta_{g_i(r)} F_i = \frac{1}{2m_i(r)} F_i \log F_i + \Big(\frac{\mu_i(r)}{4m_i(r)} - \frac{1}{4}R(g_i(r))\Big)F_i.$$

Multiplying the above equation by $F_i(r)$ and integrating over M, we get

$$\int_M |\nabla_{g_i} F_i|^2\, d\mathrm{vol}_{g_i(r)} = \frac{1}{2m_i(r)}\int_M F_i^2 \log F_i\, d\mathrm{vol}_{g_i(r)} \\ + \int_M \Big(\frac{\mu_i(r)}{4m_i(r)} - \frac{1}{4}R(g_i)\Big)F_i^2\, d\mathrm{vol}_{g_i(r)}. \tag{3-10}$$

Because $\int_M (F_i(r))^2 \, d\mathrm{vol}_{g_i(r)} = 1$, by Jensen's inequality for the logarithm,

$$\begin{aligned} \int_M F_i^2 \log F_i \, d\mathrm{vol}_{g_i(r)} &= \frac{n-2}{4} \int_M F_i^2 \log F_i^{4/(n-2)} \, d\mathrm{vol}_{g_i(r)} \\ &\le \frac{n-2}{4} \log \int_M F_i^{2+4/(n-2)} \, d\mathrm{vol}_{g_i(r)} \\ &= \frac{n-2}{4} \log \int_M F_i^{(2n)/(n-2)} \, d\mathrm{vol}_{g_i(r)} \, . \end{aligned} \tag{3-11}$$

On the other hand, we recall the following Sobolev inequality (see also [Hebey 1999, Theorem 5.6]):

Theorem 3.1 [Hebey and Vaugon 1995]. *For any smooth, compact Riemannian n-manifold (M, g), where $n \ge 3$, such that*

$$|\mathrm{Rm}(g)| \le \Lambda_1, \quad |\nabla_g \mathrm{Rm}(g)| \le \Lambda_2, \quad \mathrm{inj}_{(M,g)} \ge \gamma,$$

there is a uniform constant $B(n, \Lambda_1, \Lambda_2, \gamma)$ so that for any $u \in W^{1,2}(M)$,

$$\begin{aligned} &\left(\int_M |u|^{(2n)/(n-2)} \, d\mathrm{vol}_g \right)^{(n-2)/n} \\ &\qquad \le C(n) \int_M |\nabla u|^2 \, d\mathrm{vol}_g + B(n, \Lambda_1, \Lambda_2, \gamma) \int_M u^2 \, d\mathrm{vol}_g \, . \end{aligned} \tag{3-12}$$

By Perelman's noncollapsing result, Theorem 3.1 applies to $(M, g_i(r))$ with uniform constants $\Lambda_1, \Lambda_2, \gamma$, independent of $r \in [-10, 0]$ and i. In particular, letting $u = F_i(r)$ in (3-12), we find that

$$\begin{aligned} &\int_M (F_i(r))^{(2n)/(n-2)} \, d\mathrm{vol}_{g_i(r)} \\ &\qquad \le C(n) \left(\int_M |\nabla_{g_i(r)} F_i(r)|^2 \, d\mathrm{vol}_{g_i(r)} \right)^{n/(n-2)} + B(n, \Lambda_1, \Lambda_2, \gamma). \end{aligned} \tag{3-13}$$

Combining (3-10), (3-11) and (3-13), we obtain

$$\begin{aligned} &\int_M |\nabla_{g_i} F_i|^2 \, d\mathrm{vol}_{g_i(r)} \\ &\quad \le \frac{n-2}{8m_i(r)} \log \int_M F_i^{(2n)/(n-2)} \, d\mathrm{vol}_{g_i(r)} \\ &\qquad\qquad + \int_M \left(\frac{\mu_i(r)}{4m_i(r)} - \frac{1}{4} R(g_i) \right) F_i^2 \, d\mathrm{vol}_{g_i(r)} \end{aligned} \tag{3-14}$$

$$\leq \frac{n-2}{8m_i(r)} \log\Bigg(C(n)\bigg(\int_M |\nabla F_i|^2\, d\mathrm{vol}_{g_i(r)}\bigg)^{n/(n-2)} + B(n, \Lambda_1, \Lambda_2, \gamma)\Bigg) + \int_M \bigg(\frac{\mu_i(r)}{4m_i(r)} - \frac{1}{4}R(g_i)\bigg) F_i^2\, d\mathrm{vol}_{g_i(r)}.$$

Recall that $R(g_i(r))$ is uniformly bounded by the scaling and furthermore

$$\lim_{i\to\infty} Q_i(T-t_i) = a \in [\tfrac{1}{8}, \infty).$$

Thus, if $r \in [-10, 0]$, then Equation (3-14) gives a global uniform bound for $\int_M |\nabla_{g_i(r)} F_i(r)|^2\, d\mathrm{vol}_{g_i(r)}$. Since $\phi_i(r) = c_i(r) F_i(r)$, we then have a global uniform bound for $\int_M |\nabla_{g_i(r)} \phi_i(r)|^2\, d\mathrm{vol}_{g_i(r)}$. □

Now, elliptic L^p theory gives uniform $C^{1,\alpha}$ estimates for $\phi_i(r)$ on compact sets [Gilbarg and Trudinger 2001]. We need higher order derivative estimates on $\phi_i(r)$ to conclude that for a suitably chosen sequence of points q_i around which we take the limit, we have $f_\infty(r) = -2\log\phi_\infty(r)$ for a smooth function $f_\infty(r)$ (where $f_\infty(r)$ is the limit of $f_i(r)$ and $\phi_\infty(r)$ is the limit of $\phi_i(r)$). For the higher order estimates, it is crucial to prove that $\{\phi_i(r)\}$ stay uniformly bounded from below on compact sets around q_i.

In (3-7), take $s = -10$ and $v = 0$. For each i, let $r_i \in [-10, 0]$ be such that

$$(a-r_i)(4\pi(a-r_i))^{-n/2} \bigg|\mathrm{Ric}(g_i(r_i)) + \nabla\nabla f\bigg(t_i + \frac{r_i}{Q_i}\bigg) - \frac{g_i}{2(a-r_i)}\bigg|^2 \times e^{-f(t_i + r_i/Q_i)}\, d\mathrm{vol}_{g_i(r_i)}$$
$$\leq (a-r)(4\pi(a-r))^{-n/2} \bigg|\mathrm{Ric}(g_i(r)) + \nabla\nabla f\bigg(t_i + \frac{r}{Q_i}\bigg) - \frac{g_i}{2(a-r)}\bigg|^2 \times e^{-f(t_i + r/Q_i)}\, d\mathrm{vol}_{g_i(r)}$$

for all $r \in [-10, 0]$. Take $q_i \in M$ at which the maximum of $\phi_i(r_i)$ over M has been achieved and denote also by $(M_\infty, g_\infty(t), q)$ the smooth pointed Cheeger–Gromov limit of the rescaled sequence of metrics $(M, g_i(t), q_i)$, defined as above. Take any compact set $K \subset M_\infty$ containing q. Let $\psi_i : K_i \to K$ be the diffeomorphisms from the definition of Cheeger–Gromov convergence of (M, g_i, q_i) to (M_∞, g_∞, q) and $K_i \subset M$. Following the previous notation, consider the functions $F_i(r_i)$, $\phi_i(r_i)$ and for simplicity denote them by F_i and ϕ_i, respectively. Also denote the metric $g_i(r_i)$ by g_i.

Lemma 3.2. *For any $\alpha \in (0, 1)$, there is a uniform constant $C(\alpha)$ so that*

$$\|F_i\|_{C^{1,\alpha}(M)} \leq C(\alpha). \tag{3-15}$$

Proof. The proof is via bootstrapping and rather standard for the equation satisfied by F_i:

$$-\Delta_{g_i} F_i = \frac{1}{2m_i(r)} F_i \log F_i + \left(\frac{\mu_i(r)}{4m_i(r)} - \frac{1}{4} R(g_i)\right) F_i. \tag{3-16}$$

The reason that bootstrapping works is simple. If F_i is uniformly bounded in $L^p(K_i)$, where $K_i \in M$ is a compact set, then $F_i \log F_i$ is uniformly bounded in $L^{p-\delta}(K_i)$ for any $\delta > 0$. Standard local parabolic estimates give (3-15), which is independent of a compact set since we have a uniform global $W^{1,2}$ bound on F_i. □

We now discuss how to get higher order derivatives estimates for F_i. Covariantly differentiating (3-16), commuting derivatives, and noting that

$$-\Delta_{g_i} \partial_l F_i = -\partial_l \Delta_{g_i} F_i - \mathrm{Ric}(g_i)_{lk} g_i^{kp} \partial_p F_i,$$

we get

$$\begin{aligned} -\Delta_{g_i} \partial_l F_i = \frac{1}{2m_i(r)} \partial_l F_i \log F_i + \left(\frac{2+\mu_i(r)}{4m_i(r)} - \frac{1}{4} R(g_i)\right) \partial_l F_i \\ - \frac{1}{4} \partial_l R(g_i) F_i - \mathrm{Ric}(g_i)_{lk} g_i^{kp} \partial_p F_i. \end{aligned} \tag{3-17}$$

The major obstacle in applying L^p theory to get uniform $C^{1,\alpha}$ estimates for $\partial_l F_i$ is the term $\partial_l F_i \log F_i$. This emanates from the potential smallness of $|F_i|$, though we have already found a nice uniform upper bound on it. Thus, to proceed further, we need to bound $|F_i|$ uniformly from below. Equivalently, we will prove in Lemma 3.3 that ϕ_i stays uniformly bounded from below on K_i.

As the first step, we bound $\phi_i(q_i)$ from below. This is simple. Applying the maximum principle to (3-8) gives $\min_M f_i \le C$, where $f_i = f_i(r_i)$ for a uniform constant C. This can be seen as follows. Define $\alpha_i = Q_i(T - t_i)$. At the minimum of f_i, we have

$$\frac{f_i - n}{\alpha_i - r_i} = \frac{\mu(g_i(r_i), \alpha_i - r_i)}{\alpha_i - r_i} - R(g_i(r_i)) - 2\Delta_{g_i(r_i)} f_i \le \frac{\mu(g_i(r_i), \alpha_i - r_i)}{\alpha_i - r_i} - R(g_i(r_i)).$$

Thus,

$$\begin{aligned} f_i &\le n + \mu(g_i(r_i), \alpha_i - r_i) - R(g_i(r_i))(\alpha_i - r_i) \\ &\le n + \mu(g_i(r_i), \alpha_i - r_i) + \frac{C}{Q_i}\big(Q_i(T - t_i) - r_i\big) \le C, \end{aligned} \tag{3-18}$$

where we have used the fact that $R(\cdot, t) \ge -C$ on M for all $t \in [0, T)$ (see (2-3)). This implies $\phi_i(q_i) \ge \delta > 0$ for all i, with a uniform constant δ.

Let $K \subset M_\infty$ and $K_i \subset M$ be compact sets as before. Also recall that $m_i(r_i) = Q_i(T - t_i) - r_i$.

Lemma 3.3. *For every compact set $K \subset M_\infty$ there exists a uniform constant $C(K)$ so that*

$$\phi_i \geq C(K) \text{ on } K_i \quad \text{for all } i.$$

Proof. Assume the lemma is not true and that there exist points $P_i \in K_i$ so that $\phi_i(P_i) \leq 1/i \to 0$ as $i \to \infty$. Assume $\psi_i(P_i)$ converge to a point $P \in K$. Then $\phi_\infty(P) = 0$. Take a smooth function $\eta \in C_0^\infty(M_\infty)$, compactly supported in $K \backslash \{P\}$. Then $\psi_i^* \eta \in C_0^\infty(M)$, compactly supported in $K_i \backslash \{P_i\}$. Multiplying (3-9) by $\psi_i^*\eta$, assuming $\lim_{i\to\infty} r_i = r_0$, and then integrating by parts, we get

$$\int_M \Big(m_i(r_i) \cdot (4\nabla\phi_i \nabla(\psi_i^*\eta) + R_i \phi_i \psi_i^* \eta) - 2\phi_i \psi_i^* \eta \ln \phi_i \\ - n\phi_i \psi_i^* \eta - \mu(g_i, m_i(r_i))\phi_i \psi_i^* \eta \Big)\, d\mathrm{vol}_{g_i(r_i)} = 0.$$

We now let $i \to \infty$ and observe that $\phi_i \to \phi_\infty$ $C^{1,\alpha}$ locally, that $\psi_i^*\eta \to \eta$ smoothly, that $\lim_{i\to\infty} R(g_i) = 0$, and that $a - r_0 := \lim_{i\to\infty} m_i(r_i) \equiv \lim_{i\to\infty}(Q_i(T - t_i) - r_i)$ is finite. Thus one finds that

$$\int_{M_\infty} \big(4(a-r_0)\nabla\phi_\infty \nabla\eta - 2\eta\phi_\infty \ln \phi_\infty - n\phi_\infty\eta - \mu(g_\infty, a - r_0)\eta\phi_\infty\big)\, d\mathrm{vol}_{g_\infty(r_0)} = 0.$$

Proceeding in the same manner as in [Rothaus 1981], we obtain $\phi_\infty \equiv 0$ in some small ball around P. Using the connectedness argument, $\phi_\infty \equiv 0$ everywhere in M_∞. That contradicts $\phi_\infty(q) \geq \delta > 0$. □

Having Lemma 3.3 and $C^{1,\alpha}$ uniform estimates on ϕ_i, we see that the right-hand side of (3-17) is uniformly bounded in $L^2(K_i)$. Because $\log F_i$ is uniformly bounded on K_i, we can bootstrap (3-17) to obtain $C^{1,\alpha}$ estimates for $|\nabla_{g_i} F_i|$. Hence, one has uniform $C^{2,\alpha}$ estimates for F_i on K_i. In terms of ϕ_i,

$$\|\phi_i\|_{C^{2,\alpha}(K_i)} \leq C(K, \alpha)(Q_i(T - t_i) - r_i)^{n/4}. \tag{3-19}$$

Differentiating (3-17) again gives all higher order derivative estimates on ϕ_i and therefore all higher order derivative estimates on $f_i = f_i(r_i) = -2\log\phi_i$. However, for our purpose, $C^{2,\alpha}$ estimates suffice.

Then, using (3-7), for $s = -10$ and $v = 0$,

$$\lim_{i\to\infty} \Bigg(10(a - r_i)(4\pi(a - r_i))^{-n/2} \\ \times \int_M \left| \mathrm{Ric}(g_i(r_i)) + \nabla\nabla f_i - \frac{g_i(r_i)}{2(a - r_i)} \right|^2 e^{-f_i}\, d\mathrm{vol}_{g_i(r_i)} \Bigg)$$

$$\leq \lim_{i\to\infty} \Bigg(\int_{-10}^{0} \int_M (a - r)(4\pi(a - r))^{-n/2} \\ \times \left| \mathrm{Ric}(g_i) + \nabla\nabla f_i - \frac{g_i}{2(a - r)} \right|^2 e^{-f_i}\, d\mathrm{vol}_{g_i(r)}\, dr \Bigg) = 0.$$

By Lemma 3.3 and (3-7), applying the Arzelà–Ascoli theorem on f_i results in

$$\mathrm{Ric}_\infty + \nabla\nabla f_\infty - \frac{g_\infty}{2(a-r_0)} = 0.$$

Since $\mathrm{Ric}_\infty \equiv 0$, we get

$$g_\infty = 2(a-r_0)\nabla\nabla f_\infty,$$

and therefore M_∞ is isometric to a standard Euclidean space $\mathbb{R}^n$; see, for example, [Ni 2005, Proposition 1.1]. It is now easy to see that

$$f_\infty = \frac{|x|^2}{4(a-r_0)}, \tag{3-20}$$

that is, the limiting manifold $(\mathbb{R}^n, g_\infty, q_\infty)$ is a Gaussian shrinker. □

Proof of Theorem 1.3. We will use many estimates and arguments developed in the proof of Theorem 1.2. Assume the flow does develop a Type II singularity at T. Then we can pick a sequence of times $t_i \to T$ and points $p_i \in M$ as in [Hamilton 1995b] so that the rescaled sequence of solutions $(M, g_i(t) := Q_i g(t_i + t/Q_i), p_i)$, converges in a pointed Cheeger–Gromov sense to a Ricci flat, nonflat, complete, eternal solution $(M_\infty, g_\infty(t), p_\infty)$. Here $Q_i := \max_{M\times[0,t_i]} |\mathrm{Rm}|(x,t) \to \infty$ as $i \to \infty$. The reasons for getting Ricci flat metric are the same as in the proof of Theorem 1.2. Define

$$\alpha_i := (T - t_i)Q_i.$$

Since we are assuming a Type II singularity occurs at T, we may assume that for a chosen sequence t_i we have $\lim_{i\to\infty} \alpha_i = \infty$.

By Lemma 2.1 and the monotonicity of μ, we have $|\mu(g(t), T-t)| \le C$ for all $t \in [0, T)$. Let $f_i(\,\cdot\,, s)$ be a smooth minimizer realizing

$$\mu\Big(g\Big(t_i+\frac{s}{Q_i}\Big), T-t_i-\frac{s}{Q_i}\Big) = \mu(g_i(s), \alpha_i - s) = \inf \mathcal{W}\Big(g\Big(t_i+\frac{s}{Q_i}\Big), f, T-t_i-\frac{s}{Q_i}\Big)$$

over the set of all smooth functions f satisfying

$$\Big(4\pi\Big(T-t_i-\frac{s}{Q_i}\Big)\Big)^{-n/2}\int_M e^{-f}\,d\mathrm{vol}_{g(t_i+s/Q_i)} = 1.$$

Then $f_i = f_i(\,\cdot\,, s)$ satisfies

$$2\Delta_{g_i(s)} f_i - |\nabla_{g_i(s)} f_i|^2 + R_i + \frac{f_i - n}{\alpha_i - s} = \frac{\mu(g_i(s), \alpha_i - s)}{\alpha_i - s}. \tag{3-21}$$

In terms of $\phi_i(x,s) = e^{-f_i(x,s)/2}$ this is equivalent to

$$\begin{aligned} -4\Delta_{g_i(s)}\phi_i(s) &+ R(g_i(s))\phi_i(s) \\ &= \frac{2\,\phi_i(s)\log\phi_i(s)}{\alpha_i - s} + \frac{(\mu(g_i(s), \alpha_i - s) + n)\phi_i(s)}{\alpha_i - s}, \end{aligned} \tag{3-22}$$

with

$$\int_M (\phi_i(s))^2 \, d\mathrm{vol}_{g_i(s)} = (4\pi(\alpha_i - s))^{n/2}. \tag{3-23}$$

In what follows, we fix $s = 0$. Define $\widetilde{\phi}_i(\,\cdot\,) := \phi_i(\,\cdot\,, 0)/\beta_i$, where

$$\beta_i := \max_M \left(\phi_i(x, 0) + |\nabla_{g_i(0)}\phi_i(x, 0)|\right). \tag{3-24}$$

This choice of β_i gives us uniform C^1 estimates for $\widetilde{\phi}_i$ on M. Thus, we can apply L^p theory to get uniform $C^{1,\alpha}$ estimates for $\widetilde{\phi}_i$ on compact sets around the points where the maxima in (3-24) are achieved. To be more precise, we proceed as follows.

Take $q_i \in M$ at which this maximum in (3-24) has been achieved and denote also by $(M_\infty, g_\infty(t), q)$ the smooth pointed Cheeger–Gromov limit of the rescaled sequence of metrics $(M, g_i(t), q_i)$, defined as above. Lemma 3.1, Theorem 3.1 and standard elliptic L^p estimates applied to (3-22) yield the estimates on β_i in terms of the $W^{1,2}$ norm of ϕ_i with respect to metric $g_i(0)$, that is, there exists a uniform constant C so that for all i, $\beta_i \le C\alpha_i^{n/4}$, which implies

$$\log\beta_i \le C_2 \log\alpha_i + C_2, \tag{3-25}$$

for some uniform constants C_1 and C_2. This can be proved the same way we obtained (3-19) in Theorem 1.2. After dividing (3-22) by β_i we get

$$-4\Delta_{g_i(0)}\widetilde{\phi}_i + R(g_i(0))\widetilde{\phi}_i = 2\widetilde{\phi}_i \cdot \frac{\log\widetilde{\phi}_i + \log\beta_i}{\alpha_i} + \frac{(\mu(g(t_i), T - t_i) + n)\widetilde{\phi}_i}{\alpha_i}. \tag{3-26}$$

Since $(M, g_i(t), q_i)$ converges to $(M_\infty, g_\infty(t), q)$ in the pointed Cheeger–Gromov sense, and $\|\widetilde{\phi}_i\|_{C^1(M, g_i(0))}$ is uniformly bounded, we can get uniform $C^{1,\alpha}$ estimates for $\widetilde{\phi}_i$ on compact sets around points q_i. By the Arzelà–Ascoli theorem, $\widetilde{\phi}_i$ converges uniformly in the C^1 norm on compact sets around points q_i to a smooth function $\widetilde{\phi}_\infty$. We will show in the next paragraph that $\widetilde{\phi}_\infty(\,\cdot\,)$ is a positive constant.

Indeed, if we apply the maximum principle to (3-21), similarly as in the proof of Theorem 1.2, we obtain $\min_M f_i(\,\cdot\,, 0) \le C$ for a uniform constant C. This implies $\log\beta_i \ge -C_1$ for a uniform constant C_1. In particular, there is a uniform constant $\delta > 0$ such that for all i, one has

$$\beta_i \ge \delta > 0. \tag{3-27}$$

This together with (3-25) and the $\lim_{i\to\infty}\alpha_i = \infty$ implies

$$\lim_{i\to\infty} \frac{\log\beta_i}{\alpha_i} = 0. \tag{3-28}$$

Multiplying (3-26) by any cut-off function $\eta_i = \psi_i^* \eta$ (where η is any cut-off function on M_∞ and ψ_i is a sequence of diffeomorphisms from the definition of Cheeger–Gromov convergence) and integrating by parts, we get

$$\begin{aligned} 4 \int_M \nabla \widetilde{\phi}_i \nabla \eta_i \, d\mathrm{vol}_{g_i(0)} &= -\int_M R(g_i(0)) \widetilde{\phi}_i \eta_i \, d\mathrm{vol}_{g_i(0)} \\ &\quad + 2 \int_M \eta_i \widetilde{\phi}_i \cdot \frac{\log \widetilde{\phi}_i + \log \beta_i}{\alpha_i} \, d\mathrm{vol}_{g_i(0)} - \frac{\mu(g(t_i), T - t_i) + n}{\alpha_i} \int_M \eta_i \widetilde{\phi}_i \, d\mathrm{vol}_{g_i(0)} . \end{aligned}$$

Let $i \to \infty$ in the previous identity. From (3-28) and the limits $\lim_{i\to\infty} \alpha_i = \infty$, $R(g_i(0)) \to 0$ uniformly on compact sets, and $\widetilde{\phi}_i \to \widetilde{\phi}_\infty$ in the C^1 sense, and using uniform bounds on $\mu(g(t), T - t)$, we obtain

$$\int_M \nabla \widetilde{\phi}_\infty \nabla \eta \, d\mathrm{vol}_{g_\infty(0)} = 0.$$

This means $\Delta \widetilde{\phi}_\infty = 0$ in the distributional sense. By Weyl's theorem, $\widetilde{\phi}_\infty$ is a harmonic function on M_∞. Since $(M_\infty, g_\infty(0))$ is a complete, Ricci flat manifold and $\phi_\infty \geq 0$, by the theorem of [Yau 1975], $\widetilde{\phi}_\infty = C_\infty$ is a constant function on M_∞. At the same time, from the definition of $\widetilde{\phi}_i$, we get for x in compact sets around points q_i,

$$1 = \lim_{i\to\infty} \left(\widetilde{\phi}_i(x) + |\nabla_{g_i(0)} \widetilde{\phi}_i(x)| \right) = \widetilde{\phi}_\infty(x) + |\nabla_{g_\infty(0)} \widetilde{\phi}_\infty(x)| \equiv C_\infty. \tag{3-29}$$

This implies, in particular $C_\infty \equiv 1 > 0$. □

4. Integral bounds on scalar curvature

In this section we will prove Theorem 1.4 and Theorem 1.5. Theorem 1.1 is a special case of Theorem 1.4 when $\alpha = \infty$ in the case with Type I singularities only. A crucial ingredient in our arguments is the following result.

Theorem 4.1 [Enders et al. 2010, Theorem 1.4]. *Let $g(t)$ be the solution to a Type I Ricci flow* (1-1) *on* $[0, T)$ *and suppose that the flow develops a Type I singularity at T. Then for every sequence $\lambda_j \to \infty$, the rescaled Ricci flows $(M, g_j(t))$ defined on $[-\lambda_j T, 0)$ by $g_j(t) := \lambda_j g(T + t/\lambda_j)$ subconverge in the Cheeger–Gromov sense to a normalized nontrivial gradient shrinking soliton in canonical form on* $(-\infty, 0)$.

Proof of Theorem 1.4. The proof is by contradiction. Assume the flow develops a Type I singularity at $p \in M$ at $T < \infty$. Consider any sequence $\lambda_j \to \infty$ and define $g_j(t) := \lambda_j g(T + t/\lambda_j)$ where $t \in [-\lambda_j T, 0)$. By Theorem 4.1, the rescaled Ricci flows $(M, g_j(t), p)$ defined on $[-\lambda_j T, 0)$ subconverge in the Cheeger–Gromov sense to a normalized nontrivial gradient shrinking soliton $(M_\infty, g_\infty(t), p_\infty)$ in canonical form on $(-\infty, 0)$. Under the condition (1-3), one has

$$\int_M |R(g_j(t))|^\alpha \, d\mathrm{vol}_{g_j(t)} = \frac{1}{\lambda_j^{\alpha-n/2}} \int_M \Big| R(g(T+\frac{t}{\lambda_j})) \Big|^\alpha d\mathrm{vol}_{g(T+t/\lambda_j)} \le \frac{C_\alpha}{\lambda_j^{\alpha-n/2}} \to 0.$$

Thus the limiting solution $(M_\infty, g_\infty(t), p_\infty)$ is scalar flat. Arguing as in the proof of Theorem 1.1, we see that M_∞ is isometric to a standard Euclidean space $\mathbb{R}^n$. However, this contradicts the nontriviality of M_∞. □

Proof of Theorem 1.5. By Hölder's inequality, it suffices to consider the case when $\alpha = (n+2)/2$. Then the integral bound is invariant under the usual parabolic scaling of the Ricci flow.

The proof is by contradiction. Assume the flow develops a Type I singularity at $p \in M$ at $T < \infty$. Consider any sequence $\lambda_j \to \infty$ and define $g_j(t) := \lambda_j g(T + t/\lambda_j)$ where $t \in [-\lambda_j T, 0)$. Then, by Theorem 4.1, the rescaled Ricci flows $(M, g_j(t), p)$ defined on $[-\lambda_j T, 0)$ subconverge in the Cheeger–Gromov sense to a normalized nontrivial gradient shrinking soliton $(M_\infty, g_\infty(t), p_\infty)$ in canonical form on $(-\infty, 0)$. Observe that

$$\int_{-1}^0 \int_M |R(g_j(t)|^\alpha \, d\mathrm{vol}_{g_j(t)} \, dt = \int_{T-1/\lambda_j}^T \int_M |R(g(s))|^\alpha \, d\mathrm{vol}_{g(s)} \, ds.$$

Since $\int_0^T \int_M |R(g(t))|^\alpha \, d\mathrm{vol}_{g(t)} \, dt < \infty$, letting $j \to \infty$, we obtain

$$\int_{-1}^0 \int_{M_\infty} |R(g_\infty(t)|^\alpha \, d\mathrm{vol}_{g_\infty(t)} \, dt \le \lim_{j\to\infty} \int_{T-1/\lambda_j}^T \int_M |R(g(s))|^\alpha \, d\mathrm{vol}_{g(s)} \, ds = 0,$$

which implies $R(g_\infty(t)) \equiv 0$ on M_∞ for $t \in [-1, 0]$. Thus the limiting solution $(M_\infty, g_\infty(t))$ is scalar flat. Arguing as in the proof of Theorem 1.1, we see that M_∞ is isometric to a standard Euclidean space $\mathbb{R}^n$. However, this contradicts the nontriviality of M_∞. □

Acknowledgement

The authors would like to thank John Lott for helpful conversations during the preparation of this paper.

References

[Angenent and Velázquez 1997] S. B. Angenent and J. J. L. Velázquez, "Degenerate neckpinches in mean curvature flow", *J. Reine Angew. Math.* **482** (1997), 15–66. MR 98k:58059 Zbl 0866.58055

[Böhm and Wilking 2007] C. Böhm and B. Wilking, "Nonnegatively curved manifolds with finite fundamental groups admit metrics with positive Ricci curvature", *Geom. Funct. Anal.* **17**:3 (2007), 665–681. MR 2008h:53050 Zbl 1132.53035

[Cao and Zhang 2011] X. Cao and Q. S. Zhang, "The conjugate heat equation and ancient solutions of the Ricci flow", *Adv. Math.* **228**:5 (2011), 2891–2919. Zbl 05969510

[Chow et al. 2006] B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci flow*, Graduate Studies in Mathematics **77**, American Mathematical Society, Providence, RI, 2006. MR 2008a:53068 Zbl 1118.53001

[DeTurck 1983] D. M. DeTurck, "Deforming metrics in the direction of their Ricci tensors", *J. Differ. Geom.* **18**:1 (1983), 157–162. MR 85j:53050 Zbl 0517.53044

[Enders et al. 2010] J. Enders, R. Müller, and P. M. Topping, "On type I singularities in Ricci flow", preprint, 2010. to appear in *Commun. Anal. Geom.* arXiv 1005.1624v1

[Gilbarg and Trudinger 2001] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Springer, Berlin, 2001. MR 2001k:35004 Zbl 1042.35002

[Gu and Zhu 2008] H.-L. Gu and X.-P. Zhu, "The existence of type II singularities for the Ricci flow on S^{n+1}", *Comm. Anal. Geom.* **16**:3 (2008), 467–494. MR 2009k:53169 Zbl 1152.53054

[Hamilton 1982] R. S. Hamilton, "Three-manifolds with positive Ricci curvature", *J. Differ. Geom.* **17**:2 (1982), 255–306. MR 84a:53050 Zbl 0504.53034

[Hamilton 1995a] R. S. Hamilton, "A compactness property for solutions of the Ricci flow", *Amer. J. Math.* **117**:3 (1995), 545–572. MR 96c:53056 Zbl 0840.53029

[Hamilton 1995b] R. S. Hamilton, "The formation of singularities in the Ricci flow", pp. 7–136 in *Surveys in differential geometry, Vol. II*, edited by C. C. Hsiung and S. T. Yau, International Press, Cambridge, MA, 1995. MR 97e:53075 Zbl 0867.53030

[Hebey 1999] E. Hebey, *Nonlinear analysis on manifolds: Sobolev spaces and inequalities*, Courant Lecture Notes in Mathematics **5**, Amer. Math. Soc., Providence, RI, 1999. MR 2000e:58011 Zbl 0981.58006

[Hebey and Vaugon 1995] E. Hebey and M. Vaugon, "The best constant problem in the Sobolev embedding theorem for complete Riemannian manifolds", *Duke Math. J.* **79**:1 (1995), 235–279. MR 96c:53057 Zbl 0839.53030

[Huisken 1984] G. Huisken, "Flow by mean curvature of convex surfaces into spheres", *J. Differ. Geom.* **20**:1 (1984), 237–266. MR 86j:53097 Zbl 0556.53001

[Kleiner and Lott 2008] B. Kleiner and J. Lott, "Notes on Perelman's papers", *Geom. Topol.* **12**:5 (2008), 2587–2855. MR 2010h:53098 Zbl 1204.53033

[Knopf 2006] D. Knopf, "Positivity of Ricci curvature under the Kähler–Ricci flow", *Commun. Contemp. Math.* **8**:1 (2006), 123–133. MR 2006k:53114 Zbl 1118.53045

[Knopf 2009] D. Knopf, "Estimating the trace-free Ricci tensor in Ricci flow", *Proc. Amer. Math. Soc.* **137**:9 (2009), 3099–3103. MR 2010e:53111 Zbl 1172.53043

[Le and Sesum 2010] N. Q. Le and N. Sesum, "The mean curvature at the first singular time of the mean curvature flow", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**:6 (2010), 1441–1459. MR 2012b:53142 Zbl 05838277

[Le and Sesum 2011] N. Q. Le and N. Sesum, "On the extension of the mean curvature flow", *Math. Z.* **267**:3-4 (2011), 583–604. MR 2776050 Zbl 1216.53060

[Ma and Cheng 2010] L. Ma and L. Cheng, "On the conditions to control curvature tensors of Ricci flow", *Ann. Global Anal. Geom.* **37**:4 (2010), 403–411. MR 2011d:53158 Zbl 1188.35033

[Máximo 2011] D. Máximo, "Non-negative Ricci curvature on closed manifolds under Ricci flow", *Proc. Amer. Math. Soc.* **139**:2 (2011), 675–685. MR 2012b:53143 Zbl 1215.53062

[Naber 2010] A. Naber, "Noncompact shrinking four solitons with nonnegative curvature", *J. Reine Angew. Math.* **645** (2010), 125–153. MR 2673425 Zbl 1196.53041

[Ni 2005] L. Ni, "Ancient solutions to Kähler–Ricci flow", *Math. Res. Lett.* **12**:5 (2005), 633–654. MR 2006i:53097 Zbl 1087.53061

[Perelman 2002] G. Perelman, "The entropy formula for the Ricci flow and its geometric applications", preprint, 2002. Zbl 1130.53001 arXiv math.DG/0211159v1

[Rothaus 1981] O. S. Rothaus, "Logarithmic Sobolev inequalities and the spectrum of Schrödinger operators", *J. Funct. Anal.* **42**:1 (1981), 110–120. MR 83f:58080b Zbl 0471.58025

[Sesum 2005] N. Sesum, "Curvature tensor under the Ricci flow", *Amer. J. Math.* **127**:6 (2005), 1315–1324. MR 2006f:53097 Zbl 1093.53070 arXiv math/0311397

[Wang 2008] B. Wang, "On the conditions to extend Ricci flow", *Int. Math. Res. Not.* **2008** (2008), Art. ID rnn012. MR 2009k:53176 Zbl 1148.53050

[Yau 1975] S. T. Yau, "Harmonic functions on complete Riemannian manifolds", *Comm. Pure Appl. Math.* **28**:2 (1975), 201–228. MR 55 #4042 Zbl 0291.31002

[Ye 2008] R. Ye, "Curvature estimates for the Ricci flow II", *Calc. Var. Partial Differ. Equ.* **31**:4 (2008), 439–455. MR 2009f:53107 Zbl 1142.53033

[Zhang 2010] Z. Zhang, "Scalar curvature behavior for finite-time singularity of Kähler–Ricci flow", *Mich. Math. J.* **59**:2 (2010), 419–433. MR 2011j:53128 Zbl 1198.53079

Received October 7, 2010. Revised February 23, 2011.

NAM Q. LE
DEPARTMENT OF MATHEMATICS
COLUMBIA UNIVERSITY
NEW YORK NY 10027
UNITED STATES
namle@math.columbia.edu

NATASA SESUM
DEPARTMENT OF MATHEMATICS
RUTGERS UNIVERSITY
PISCATAWAY NJ 10027
UNITED STATES
natasas@math.rutgers.edu
http://www.math.rutgers.edu/∼natasas

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

STABILITY OF CAPILLARY SURFACES WITH PLANAR BOUNDARY IN THE ABSENCE OF GRAVITY

PETKO I. MARINOV

We study immersed stable capillary surfaces with planar boundary in the absence of gravity. We assume that the surface approaches the boundary from one side. If the boundary of the capillary surface is embedded in a plane, we prove that the only immersed weakly stable capillary surface is the spherical cap.

Introduction

In this paper we study capillary surfaces with planar boundary in the absence of gravity. A comprehensive treatment of the theory of capillary surfaces can be found in [Finn 1986]. The problem we address arises from the related physical problem concerning a homogeneous liquid drop in contact with a smooth rigid boundary surface Σ. We call the free surface of the drop Ω and the angle of contact γ, and the wetted part of Σ we call Σ'. The liquid drop occupies a connected region in space, T, with a prescribed volume. The contact angle γ is measured relative to the interior of the liquid bounded by Ω and Σ. The problem is to describe the possible shapes of Ω if the liquid drop is in equilibrium.

There are three energies associated with this configuration. The first is the free surface energy, which is proportional to the area of Ω, with coefficient equal to the surface tension. The second is the wetting energy, which is a multiple of the area of Σ'. The third is the gravitational energy. Here we assume that there is no gravity acting, so the gravitational energy does not contribute. In order for the drop to be in equilibrium, it must be a critical point for the potential energy functional E. From this discussion we obtain a formula for E, that is, $E = \sigma \operatorname{Area}(\Omega) - \sigma\tau \operatorname{Area}(\Sigma')$, where σ is the surface tension and τ is the capillary constant. The constant τ is a physical quantity that is predetermined and, in equilibrium, equal to $\cos\gamma$. The wetting ability and the surface tension of the liquid are the two physical phenomena that cause the drop to become stationary. The above configuration is said to be in a stationary state if the first variation of E is zero for any volume-preserving perturbation. It is weakly stable (resp. stable) if it is stationary and the second

MSC2010: 49Q10, 53A10, 53C42.
Keywords: capillarity, mean curvature, Jacobi operator.

variation of E is nonnegative (resp. positive) for any nontrivial volume-preserving perturbation.

In this paper we study the stability problem when the fixed boundary surface Σ is plane. We denote the first variation of E by ∂E. If ∂E is zero subject to a volume constraint, one finds that the angle of contact γ must be constant along $\partial\Omega \subset \Sigma$, the mean curvature of Ω must be constant, and $\tau = \cos\gamma$. In the case when gravity is present, the mean curvature of Ω is proportional to the height. This discussion naturally leads to the study of constant mean curvature surfaces with boundary. If Ω forms a constant angle with Σ along $\partial\Omega$, one can ask what the possible shapes of Ω are. This question is hard to resolve [Earp et al. 1991]. There are a few known examples for planar or spherical Σ, including spherical caps, right cylinders and Delaunay surfaces. We generalize the problem and assume Ω to be immersed; that is, Ω could have self-intersections, which further complicates the discussion. For this reason we put an additional restriction on Ω and study the same problem. The physical discussion above leads us to consider the case when the second variation of the potential energy, $\partial^2 E$, is nonnegative, that is, Ω is weakly stable. Assuming stability, in the case of Σ being a plane we can say much more.

It is known that if Ω is bounded and embedded and sits on one side of the boundary plane, then it is a spherical cap [Wente 1980]. We assume only that $\partial\Omega$ is embedded and that the surface Ω comes close to the boundary from above, allowing the immersed Ω to be below Σ away from $\partial\Omega$. Our main theorem shows that the only possible stable configuration of this type is the spherical cap. The spherical cap is weakly stable, as shown in [Wente 1966]. For the proof of the main theorem we consider three cases. The first is when Ω is of disk type with genus zero. The proof of this case comes from a result of Nitsche [1985] (see also [Finn and McCuan 2000]) and does not assume stability. The second case is when Ω is of genus zero, but not of disk type. For this case we use an argument involving a Killing field, as suggested in [Ros and Souam 1997]. The third case is when the genus of Ω is positive. For this case we construct a perturbation that depends on the mean curvature of Ω and on the contact angle γ. A similar perturbation is used in [Barbosa and do Carmo 1984] to show that the round spheres are the only immersed stable constant mean curvature hypersurfaces in $\mathbb{R}^n$. The normal component of the constructed perturbation makes the second variation of the energy negative, while preserving the volume; therefore, Ω cannot be weakly stable.

1. Preliminaries

In this section we define capillarity and stability in terms of the energy and the volume for a given configuration. Throughout this paper, let Ω be an oriented compact surface immersed in $\mathbb{R}^3$ with nonempty planar boundary in the xy-plane.

Let it be given by a $C^{2,\alpha}$-immersion $x(u, v) : D \to \mathbb{R}^3$, with $x(D) = \Omega$ and $x(\partial D) = \partial\Omega$. Also, assume that $\partial\Omega$ is a finite collection of nonintersecting simple closed curves; that is, $\partial\Omega$ is embedded in $\mathbb{R}^3$. We denote the boundary $\partial\Omega$ by Γ, and the regions in $\mathbb{R}^2$ bounded by Γ we name Σ'. The boundary Γ is also oriented and it is assumed that Ω comes from above near the boundary. We denote the areas on the surface and the wetted area by $|\Omega|$ and $|\Sigma'|$, respectively. We denote the angle of contact between the surface Ω and the wetted region Σ' by γ. The surface area of Ω is given by

$$|\Omega| = \iint_D dS,$$

where dS is the surface element on $\Omega = x(D)$. We assume that Ω is extendable in a neighborhood of Γ, so we can compute tangent vectors, normal vectors, etc.

Definition. An immersed surface is called capillary if it has constant mean curvature and makes constant contact angle with the walls along its boundary.

Now we define our main object of interest: the energy.

Definition. The energy function of the above configuration, after dividing by the surface tension σ, is given by

$$E = |\Omega| - \tau|\Sigma'|,$$

with $-1 < \tau < 1$ being some predetermined constant.

Our main goal in the problems we consider is to minimize the energy subject to the natural constraints that arise. To do this we should look among all the nearby surfaces that are admissible. We get them if we apply a perturbation to the original surface. Thus we need to define what an admissible variation is; see for example [Ros and Souam 1997].

Definition. Admissible variation of x is a differentiable map $\Phi : (-\epsilon, \epsilon) \times D \to \mathbb{R}^3$ such that $\Phi_t(p) = \Phi(t, p)$ with $p \in D$ is an immersion and $\Phi_0 = x$.

As we assumed before, the surface Ω can be extended across its boundary. That will allow us to keep the boundary planar after applying an admissible variation.

Definition. The volume functional $V : (-\epsilon, \epsilon) \to \mathbb{R}$ is defined by

$$V(t) = \tfrac{1}{3} \iint_D (\Phi_t \cdot \xi_t)\, dS_t,$$

where ξ_t and dS_t are the unit outward normal and the surface element on $\Phi_t(D)$.

The corresponding variational field is $Y(p) = (\partial\Phi/\partial t)(p)\big|_{t=0}$, and we denote its normal part by ϕ. Now we need to write the first and the second variation formulae related to ϕ. We need to set the first variation equal to 0, subject to the volume constraint $V'(0) = 0$, and investigate the second variation. We also have a volume constraint, because the variation ϕ must preserve the volume. This means that we should introduce a Lagrange multiplier λ and compute the first and second variations for the expression $E+\lambda V$. For proofs of the first and second variation formulae subject to a volume constraint, one may check [Wente 1966] and [Ros and Souam 1997], respectively.

Theorem 1.1 (first variation formula). *Let $d\sigma$ be the line element on boundary Γ, and let dS be the surface element on Ω. The first variation formula for the energy of x in the direction of ϕ, subject to a volume constraint, implies that*

$$\partial(E)[\phi] \equiv \frac{d}{dt}E(t)\Big|_{t=0} = -2\iint_D H\phi\, dS + \oint_{\partial D}(-\tau\csc\gamma+\cot\gamma)\phi\, d\sigma, \tag{1}$$

$$\partial(V)[\phi] \equiv \frac{d}{dt}V(t)\Big|_{t=0} = \iint_D \phi\, dS \equiv 0. \tag{2}$$

Formula (2) represents the rate of change of the volume at time $t=0$, so if we want constant volume it must be zero. It follows from (1) and (2) that H and γ must be constants in order to have an extremal of E, subject to the volume being stationary. This follows from the observation that the constant τ must equal $\cos\gamma$ in order for the boundary integral to equal zero. This means that Ω must be a capillary surface.

Definition. A capillary surface is called *weakly stable* if the second variation is nonnegative for all admissible perturbations with normal components $\phi \neq 0$, and *stable* if the second variation is positive for all admissible perturbations.

The next theorem gives a formula for the second variation. This is our main object of interest. We assume that the configuration is weakly stable and choose a special ϕ, manipulating the formula to get a contradiction unless we have a spherical cap.

Theorem 1.2 (second variation formula). *With the notation above, the formula for the second variation of E is*

$$\partial^2(E)[\phi] \equiv \frac{d^2}{dt^2}E(t)\Big|_{t=0} = \iint_D \big[|\nabla\phi|^2-(k_1^2+k_2^2)\phi^2\big]\, dS + \oint_{\partial D} p\phi^2\, d\sigma, \tag{3}$$

where $\nabla\phi$ is the surface gradient of ϕ, k_1 and k_2 are the principal curvatures, and $p = K_\Omega\cot\gamma + K_\Sigma\csc\gamma$. Here K_Ω and K_Σ are the signed normal curvatures of Ω and Σ with respect to the boundary. Of course, condition (2) *should be fulfilled.*

In our case Σ is planar, so $K_\Sigma = 0$; and if we take a vertical slice and consider the profile curve, K_Ω will be its curvature. If the profile curve bends toward the boundary, the sign of that normal curvature is taken to be positive. In the proof of our main theorem, we use Green's identities to write this formula more concisely.

2. The main theorem

In this section we show that the only immersed weakly stable capillary surface with boundary embedded in a plane is the spherical cap. (Recall that there is no gravitational action involved.) To do this we consider three cases.

The first case is when Ω is an immersed disc type surface. It has been solved by Nitsche [1985] and Finn and McCuan [2000]. The first author proves a theorem that states that an immersed disk type surface in a ball that makes constant angle with the boundary sphere is a flat disk or a round spherical cap. He proves the theorem for a right angle but points out that the idea works for any angle. This is proved again in [Ros and Souam 1997]. Finn and McCuan have similar results for such surfaces with planar boundary. Notice that the stability condition is unnecessary for this case.

The second case is the general genus-zero case. Here the surface Ω has genus zero, but there could be possibly more than one boundary curve. We assume that the planar boundary $\Gamma = \partial\Omega$ is embedded, but the surface itself could be immersed. Let Γ belong to the plane $z = 0$. We adapt the method used in [Ros and Souam 1997] to our purposes. As before, Ω is given by mapping $x : D \to \Omega$. Let $p_0 \in \Omega$ be a point such that the Euclidean distance to the plane containing Γ is maximal. Obviously there is at least one point with that property. Let ξ be the unit normal to the surface. From our setup it follows that $\xi(p_0)$ is parallel to the z-axis. Later in this section, we see that the mean curvature of Ω must be negative, so the vector $\xi(p_0)$ points out in the positive z-direction. Denote by X the Killing field induced by rotations around the line directed by $\xi(p_0) = \boldsymbol{k}$; that is, $X = p \wedge \xi(p_0)$, where p is a point in $\mathbb{R}^3$ and $\wedge$ is the usual wedge product in $\mathbb{R}^3$. Consider the function $\phi(p) = \langle X(x(p)), \xi(p)\rangle$. Because of rotational invariance around $\xi(p)$, it follows that $\phi(p)$ is a Jacobi field on the surface. Using the notation from the previous sections, one has

$$\Delta\phi + (k_1^2 + k_2^2)\phi = 0,$$

with $\phi_\nu + p\phi = 0$ on Γ. Also $\phi(p_0) = 0$ and $\nabla\phi(p_0) = 0$. Therefore, the second variation of energy in the direction of ϕ is zero, and the volume constraint holds. As in [Ros and Souam 1997], with the Gauss–Bonnet theorem one can show that there are at least three nodal regions of ϕ; that is, $\Omega - \phi^{-1}(0)$ has at least three connected components. Let Ω_i, $i = 1, 2, 3$ be the nodal regions of ϕ, and let ϕ_i be equal to ϕ on Ω_i and zero elsewhere. Now construct $\widetilde{\phi} = \sum_{i=1}^{3} c_i\phi_i$ with c_1, c_2, c_3 constants. Then one can adjust the constants using the volume constraint to get a smaller number for the second variation, making it negative. Hence there are no surfaces of the assumed type with two or more connected boundary components. Thus we can conclude that the spherical caps are the only immersed weakly stable

CMC-surfaces with planar embedded boundary having genus $g = 0$ and constant contact angle along the boundary with the plane Σ.

The third and final case is when the genus is positive. Here we state the main theorem of this paper.

Main Theorem. *No weakly stable capillary surface with planar boundary exists that is immersed in* $\mathbb{R}^3$ *and has genus* $g > 0$.

We split the proof into several lemmas. Again we assume that the boundary of Ω is embedded; that is, it consists of a finite number of simple closed curves. Also we assume that the surface can be extended across its boundary. Thus we ensure that the boundary stays planar after a normal perturbation. Also we assume that Ω comes from above to Σ. We construct a special normal perturbation for which the second variation is negative and the volume is preserved. First we need to rewrite (3) using Green's first identity. We also assume that our mappings are $C^{2,\alpha}$ (in fact capillary surfaces are analytic by standard regularity theory), so we can compute derivatives at the boundary and extend the surface around the boundary Γ. The variation that we use does not necessarily keep the boundary planar. That is why we extend the surface across the boundary, so that after the perturbation, the new surface has planar boundary; that is, $\partial\Phi_t(D)$ belongs to the plane Σ. This is how it is done in [Wente 1966].

Applying Green's first identity to the second variation formula from Section 1, one gets

$$\iint_D |\nabla\phi|^2 - (k_1^2 + k_2^2)\phi^2\, dS = \iint_D \phi[-\Delta\phi - (k_1^2 + k_2^2)\phi]\, dS + \oint_{\partial D} \phi\phi_\nu\, d\sigma, \tag{4}$$

and for $\partial^2 E$ one obtains

$$\partial^2 E = \iint_D (-L\phi)\phi\, dS + \oint_{\partial D} (\phi_\nu + p\phi)\phi\, d\sigma, \tag{5}$$

where $L\phi = \Delta\phi + (k_1^2 + k_2^2)\phi$, $\partial V \equiv \iint_D \phi ds = 0$, and $p = K_\Omega \cot\gamma + K_\Sigma \csc\gamma$. The operator L is called the *Jacobi operator*.

Now we write the perturbation used to prove the main theorem. Let Φ be the perturbation that sends $x \to x + t\xi + Htx + ct\boldsymbol{k} + O(t^2)$. Here t lies in $[-\epsilon, \epsilon]$, $\boldsymbol{k} = (0, 0, 1)$ is the unit vertical vector, c is a constant, ξ is the outward unit normal on the surface, and H is the mean curvature of the surface. The normal part of Φ is $\phi = \xi \cdot (\partial\Phi/\partial t)(p)\big|_{t=0}$, where $p \in \Omega$. When we compute this quantity, we get $\phi = 1 + H(x \cdot \xi) + c(\boldsymbol{k} \cdot \xi)$, with c to be determined from (2).

Lemma 2.1. *Condition* (2) *implies that* $c = -\cos\gamma$; *that is*,

$$\phi = 1 + H(x \cdot \xi) - \cos\gamma(\boldsymbol{k} \cdot \xi),$$

in order to keep the volume fixed.

Proof. One needs to adjust c in $\phi = 1 + H(x\cdot\xi) + c(\boldsymbol{k}\cdot\xi)$ to get the integral of ϕ over the surface Ω to be zero.

$$0 = \iint_D \phi\, dS = \iint_D (1 + H(x\cdot\xi) + c(\boldsymbol{k}\cdot\xi))\, dS$$
$$= |\Omega| + H\iint_D (x\cdot\xi)\, dS + c\iint_D (\boldsymbol{k}\cdot\xi)\, dS,$$

where $|\Omega|$ is the area of the surface Ω. The quantity $\iint_D(\boldsymbol{k}\cdot\xi)\, dS$ is easily computed by the Divergence theorem. We know for the embedded case that

$$\iint_D (\boldsymbol{k}\cdot\xi)\, dS + \iint_{\Sigma'} (\boldsymbol{k}\cdot\xi)\, dS = \iiint_T \operatorname{div}\boldsymbol{k}\, dV = 0,$$

since $\boldsymbol{k}$ is a constant vector. Here Σ' is the wetted part bounded by Γ, T is the solid bounded by Ω and Σ, ξ is unit outward normal to $\partial T = \Omega\cup\Sigma'$, and dV is the volume element in $\mathbb{R}^3$. On Σ' the unit vector $\boldsymbol{k}$ is equal to $-\xi$, so

$$\iint_D (\boldsymbol{k}\cdot\xi)\, dS = -\iint_{\Sigma'} (\boldsymbol{k}\cdot\xi)\, dS = \iint_{\Sigma'} dS = |\Sigma'|.$$

For the immersed case there is not actually a solid T, but one can still apply the divergence theorem. In this case $\Omega\cup\Sigma'$ separates $\mathbb{R}^3$ into a finite number of connected regions, with one of them unbounded. On the bounded regions one can use the divergence theorem, and the calculation is the same as in the embedded case, since $\operatorname{div}\boldsymbol{k} = 0$ everywhere on $\mathbb{R}^3$. One can also apply Stokes's theorem to obtain the same result. Thus

$$\iint_D (\boldsymbol{k}\cdot\xi)\, dS = |\Sigma'|.$$

Next we compute $\iint_D H(x\cdot\xi)\, dS$. Assume conformal coordinates. It is well-known that in conformal coordinates one has $\Delta x = 2H\xi$ [Oprea 2007]; therefore

$$\iint_D H(x\cdot\xi)\, dS = \tfrac{1}{2}\iint_D (x\cdot\Delta x)\, dS = -\tfrac{1}{2}\iint_D |\nabla x|^2\, dS + \tfrac{1}{2}\oint_{\partial D} (x\cdot x_\nu)\, d\sigma.$$

Here Δx and ∇x are the vector surface Laplacian and the vector surface gradient of x. In conformal coordinates, the square of the surface gradient of x is

$$|\nabla x|^2 = \frac{1}{E}((x_u\cdot x_u) + (x_v\cdot x_v)) = \frac{1}{E}(E+E) = 2,$$

so

$$-\tfrac{1}{2}\iint_D |\nabla x|^2\, dS = -\tfrac{1}{2}\iint_D 2\, dS = -|\Omega|.$$

Also, if $\boldsymbol{n}$ is the unit normal of Γ in Σ, we have $x_\nu = (\cos\gamma)\boldsymbol{n} - (\sin\gamma)\boldsymbol{k}$. Therefore $(x\cdot x_\nu) = \cos\gamma(x\cdot\boldsymbol{n})$. It follows that

$$\tfrac{1}{2}\oint_{\partial D}(x\cdot x_\nu)\,d\sigma = \frac{\cos\gamma}{2}\oint_{\partial D}(x\cdot\boldsymbol{n})\,d\sigma = \cos\gamma|\Sigma'|. \tag{6}$$

The proof of the last equality in (6) can be seen in [Marinov 2010]. Combining the above results, we get

$$0 = |\Omega| - |\Omega| + \tfrac{1}{2}\oint_{\partial D}(x\cdot x_\nu)\,d\sigma + c|\Sigma'| = \cos\gamma|\Sigma'| + c|\Sigma'|.$$

This implies that $c = -\cos\gamma$, and therefore $\phi = 1 + H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)$ and $\iint_D \phi\,dS = 0$. □

For this particular ϕ, the boundary term in the second variation happens to be zero.

Lemma 2.2. *For $\phi = 1 + H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)$, we have $\phi_\nu + p\phi = 0$; that is,*

$$\partial^2 E = \iint_D (-L\phi)\phi\,dS.$$

Proof. One useful fact is that Γ is a line of curvature for both the plane Σ and the surface Ω, by the Terquem–Joachimsthal theorem [Spivak 1979]. On Γ, we have

$$\begin{aligned}\phi_\nu + p\phi &= (1 + H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi))_\nu + K_\Omega(\cot\gamma)\phi\\ &= H\frac{\partial}{\partial\nu}(x\cdot\xi) - \cos\gamma\frac{\partial}{\partial\nu}(\boldsymbol{k}\cdot\xi) + K_\Omega(\cot\gamma)\phi.\end{aligned}$$

Now we compute the normal derivative, taking into account that $(x_\nu\cdot\xi) = 0$ and $\boldsymbol{k}$ is a constant vector. Since Γ is a line of curvature, we get $(\partial/\partial\nu)(k\cdot\xi) = -K_\Omega(\boldsymbol{k}\cdot x_\nu)$ and $(\partial/\partial\nu)(x\cdot\xi) = -K_\Omega(x\cdot x_\nu)$. Substituting in the boundary expression from above, we have

$$\phi_\nu + p\phi = -K_\Omega H(x\cdot x_\nu) + (\cos\gamma)K_\Omega(\boldsymbol{k}\cdot x_\nu) + K_\Omega(\cot\gamma)\phi. \tag{7}$$

We use some more relations to rewrite this expression. Here $\boldsymbol{n}$ is the unit normal vector of Γ in the plane Σ.

$$(x\cdot x_\nu) = \cos\gamma(x\cdot\boldsymbol{n}),\qquad (x\cdot\xi) = \sin\gamma(x\cdot\boldsymbol{n}),\qquad (k\cdot x_\nu) = -\sin\gamma.$$

Using this and the fact that $\boldsymbol{k}\cdot\xi = \cos\gamma$ on Γ, we obtain from (7)

$$\begin{aligned}&K_\Omega\big[-H(\cos\gamma)(x\cdot\boldsymbol{n}) - \cos\gamma\sin\gamma + \cot\gamma\big(1 + H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)\big)\big]\\ &\quad= K_\Omega\big(-H(\cos\gamma)(x\cdot\boldsymbol{n}) - \cos\gamma\sin\gamma + (\cot\gamma)H\sin\gamma(x\cdot\boldsymbol{n}) + \cot\gamma - \cot\gamma\cos^2\gamma\big)\\ &\quad= K_\Omega(-\cos\gamma\sin\gamma + \cot\gamma - \cot\gamma\cos^2\gamma) = 0.\end{aligned}$$

This shows that $\phi_\nu + p\phi \equiv 0$ on Γ, and thus in the second variation formula (3), the boundary term is zero and the formula becomes

$$\partial^2 E = \iint_D (-L\phi)\phi \, dS. \tag{8}$$ □

Next is to rewrite (8).

Lemma 2.3.

$$\begin{aligned} \partial^2 E &= \iint_D (-L\phi)\phi \, dS \\ &= -\iint_D \frac{(k_1-k_2)^2}{2} \, dS - \oint_{\partial D} K_\Omega (\cos\gamma)[H(x\cdot \boldsymbol{n}) + \sin\gamma] \, d\sigma. \end{aligned} \tag{9}$$

Proof. For $\phi = 1 + H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)$, we have

$$(L\phi)\phi = (L\phi)(1 + H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)) = (L\phi) + (L\phi)(H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)),$$

so

$$\iint_D (-L\phi)\phi \, dS = -\iint_D (L\phi)\, dS - \iint_D (L\phi)(H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)) \, dS.$$

Now let's compute $L\phi$:

$$L\phi = L1 + L(H(x\cdot\xi)) - (\cos\gamma)L(\boldsymbol{k}\cdot\xi) = k_1^2 + k_2^2 + HL(x\cdot\xi).$$

Here $L1 = \Delta 1 + (k_1^2 + k_2^2)1 = k_1^2 + k_2^2$ and $L(\boldsymbol{k}\cdot\xi) = 0$, and it follows that $L(x\cdot\xi) = -2H$ [Barbosa and do Carmo 1984]. Taking this into account, we have

$$L\phi = k_1^2 + k_2^2 - 2H^2 = k_1^2 + k_2^2 - \frac{(k_1+k_2)^2}{2} = \frac{(k_1-k_2)^2}{2}.$$

Getting back to the integral of $(-L\phi)\phi$, we obtain

$$\begin{aligned} &\iint_D (-L\phi)\phi \, dS \\ &\qquad = -\iint_D \frac{(k_1-k_2)^2}{2} \, dS - \iint_D \frac{(k_1-k_2)^2}{2} (H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)) \, dS. \end{aligned}$$

We set $\psi = H(x\cdot\xi) - \cos\gamma(\boldsymbol{k}\cdot\xi)$; thus we need to compute $\iint_D (L\phi)\psi \, dS$. Green's second identity implies that

$$\iint_D (L\phi)\psi \, dS = \iint_D (L\psi)\phi \, dS + \oint_{\partial D} (\phi_\nu \psi - \psi_\nu \phi) \, d\sigma.$$

We know from the previous calculations that $L\psi = -2H^2$, so

$$\iint_D (L\phi)\psi \, dS = -2H^2 \iint_D \phi \, dS + \oint_{\partial D} (\phi_\nu \psi - \psi_\nu \phi) \, d\sigma; \tag{10}$$

but we know that $\iint_D \phi\, dS = \partial V = 0$. Using $\psi = \phi - 1$, this reduces (10) to

$$\iint_D (L\phi)\psi\, dS = \oint_D (\phi_\nu(\phi - 1) - (\phi - 1)_\nu \phi)\, d\sigma = -\oint_{\partial D} \phi_\nu\, d\sigma.$$

On Γ, we know from (7) that $\phi_\nu = -K_\Omega H(x \cdot x_\nu) + (\cos\gamma) K_\Omega(\boldsymbol{k} \cdot x_\nu)$. We also know that on the boundary, $(x \cdot x_\nu) = \cos\gamma(x \cdot \boldsymbol{n})$ and $(k \cdot x_\nu) = -\sin\gamma$; therefore

$$\phi_\nu = -K_\Omega H(\cos\gamma)(x \cdot \boldsymbol{n}) - (\cos\gamma) K_\Omega \sin\gamma$$

and

$$\iint_D (L\phi)\psi\, dS = \oint_{\partial D} K_\Omega [H(\cos\gamma)(x \cdot \boldsymbol{n}) + (\cos\gamma) K_\Omega \sin\gamma]\, d\sigma.$$

Now substituting $(L\phi)\psi$ back into the formula

$$-\iint_D (L\phi)\phi\, dS = -\iint_D \frac{(k_1 - k_2)^2}{2}\, dS - \iint_D (L\phi)\psi\, dS,$$

we get for the left-hand side the value

$$-\iint_D \frac{(k_1 - k_2)^2}{2}\, dS - \oint_{\partial D} [K_\Omega H(\cos\gamma)(x \cdot \boldsymbol{n}) + (\cos\gamma) K_\Omega \sin\gamma]\, d\sigma.$$

Thus we have (9), which was the statement of the lemma. □

Lemma 2.4. *Let* Σ' *be the region bounded by* Γ, *let* $|\Sigma'|$ *be its area, and let* $|\Gamma|$ *be the length of the boundary. The boundary may consist of several curves, so* Σ' *may not be connected. Let* d *be the number of boundary curves, that is, the number of components of* Γ. *Then*

$$\oint_{\partial D} (x \cdot \boldsymbol{n})\, d\sigma = 2|\Sigma'|, \tag{11}$$

$$\oint_{\partial D} k_\Gamma (x \cdot \boldsymbol{n})\, d\sigma = -|\Gamma|, \tag{12}$$

$$\left| \oint_{\partial D} k_\Gamma\, d\sigma \right| \le 2\pi d, \tag{13}$$

$$\sin\gamma |\Gamma| = -2H|\Sigma'|. \tag{14}$$

Proof. The proof of this lemma is given in [Marinov 2010]. Here we only prove formula (14), which is a version of the *balancing formula*, of which a general statement and proof can be found in [Earp et al. 1991]. Note that (14) implies that H is negative.

Choose conformal coordinates. In the proof of Lemma 2.1 we saw that

$$\iint_D (\boldsymbol{k} \cdot \xi)\, dS = |\Sigma'|$$

and that the surface Laplacian in conformal coordinates is $\Delta x = 2H\xi$ [Oprea 2007]. We use this and Green's first identity to get

$$\iint_D (\boldsymbol{k}\cdot\xi)\,dS = \frac{1}{2H}\iint_D (\boldsymbol{k}\cdot\Delta x)\,dS = \frac{1}{2H}\oint_{\partial D}(\boldsymbol{k}\cdot x_\nu)\,d\sigma.$$

This equality holds since $\boldsymbol{k}$ is a constant vector, and therefore its surface gradient is zero. From a previous discussion of the result that the boundary term in $\partial^2 E$ is zero, we know that $(\boldsymbol{k}\cdot x_\nu) = -\sin\gamma$ on Γ. Combining this fact with the above expressions for the integral of $(\boldsymbol{k}\cdot\xi)$ over Ω, we arrive at

$$|\Sigma'| = \iint_D (\boldsymbol{k}\cdot\xi)\,dS = \frac{1}{2H}\oint_{\partial D}(\boldsymbol{k}\cdot x_\nu)\,d\sigma = -\frac{1}{2H}\sin\gamma\,|\Gamma|.$$

Taking the first and last expressions above and multiplying by $-2H$, we get

$$-2H|\Sigma'| = \sin\gamma|\Gamma|,$$

which was the result to prove. This is an indication that the mean curvature H of Ω is negative for the immersed case, since all other quantities in (14) are positive. □

To continue, we rewrite (9). From Meusnier's theorem and Euler's theorem [Struik 1988], we know that $2H = K_\Omega + k_2 = K_\Omega + (\sin\gamma)k_\Gamma$ on Γ, where k_Γ is the curvature of the boundary. We have

$$K_\Omega = 2H - k_\Gamma \sin\gamma,$$

and therefore (9) becomes

$$\begin{aligned}\partial^2 E &= -\iint_D \frac{(k_1-k_2)^2}{2}\,dS - \oint_{\partial D}\cos\gamma\,(2H - k_\Gamma\sin\gamma)(H(x\cdot\boldsymbol{n}) + \sin\gamma)\,d\sigma\\ &= -\iint_D \frac{(k_1-k_2)^2}{2}\,dS\\ &\quad - \oint_{\partial D}\cos\gamma\,\bigl(2H^2(x\cdot\boldsymbol{n}) + 2H\sin\gamma - \sin\gamma\,k_\Gamma H(x\cdot\boldsymbol{n}) - \sin^2\gamma\,k_\Gamma\bigr)\,d\sigma.\end{aligned}$$

Lemma 2.5. *The following estimate holds for the second variation of energy*:

$$\partial^2 E < 4\pi(2-2g) - 2\pi d\bigl(2 - 2|\cos\gamma| - |\cos\gamma|\sin^2\gamma\bigr),$$

where g is the genus of Ω and d is the number of boundary components of Γ.

Proof. Using (11) and (12), we get

$$\begin{aligned}\partial^2 E = -\iint_D \frac{(k_1-k_2)^2}{2}\,dS&\\ + \cos\gamma\left(-4H^2|\Sigma'| - 2H|\Gamma|\sin\gamma - \sin\gamma\,H|\Gamma| + \sin^2\gamma\oint_{\partial D} k_\Gamma\,d\sigma\right)&,\end{aligned}$$

and if we use (14) we arrive at

$$\partial^2 E = -\iint_D \frac{(k_1-k_2)^2}{2}\,dS + \cos\gamma\left(2H\sin\gamma\,|\Gamma| - 2H\sin\gamma\,|\Gamma| - H\sin\gamma\,|\Gamma| + \sin^2\gamma \oint_{\partial D} k_\Gamma\,d\sigma\right).$$

After the obvious cancellation, we obtain

$$\partial^2 E = -\iint_D \frac{(k_1-k_2)^2}{2}\,dS + \cos\gamma\left(-\sin\gamma\,H|\Gamma| + \sin^2\gamma \oint_{\partial D} k_\Gamma\,d\sigma\right), \tag{15}$$

and by using (14) again, we get

$$\partial^2 E = -\iint_D \frac{(k_1-k_2)^2}{2}\,dS + \cos\gamma\left(2H^2|\Sigma'| + \sin^2\gamma \oint_{\partial D} k_\Gamma\,d\sigma\right). \tag{16}$$

We can easily see that this last expression is zero if Ω is the standard spherical cap. On a spherical cap all points are umbilical; that is, $k_1 = k_2$ everywhere, so the first integral is zero. Also we can see that $H^2|\Sigma'| = \pi \sin^2\gamma$ no matter what the scaling, and for a spherical cap we have $\oint_{\partial D} k_\Gamma\,d\sigma = -2\pi$, since we chose to work with the outward unit normal, and for us $k_\Gamma \le 0$. This means that the second expression is also zero, so the whole variation is zero. Thus, on a spherical cap this particular variation does not change the geometry.

One can express the first integral in this formula in another way:

$$\frac{(k_1-k_2)^2}{2} = \frac{k_1^2+k_2^2-2k_1k_2}{2} = \frac{(k_1+k_2)^2-4k_1k_2}{2} = 2H^2 - 2K.$$

This is the integrand in the Willmore energy. Continuing with the second variation and using the previous formula, we get

$$\partial^2 E = -2\iint_D H^2\,dS + 2\iint_D K\,dS + \cos\gamma\left(2H^2|\Sigma'| + \sin^2\gamma \oint_{\partial D} k_\Gamma\,d\sigma\right).$$

Using the Gauss–Bonnet formula we obtain for the right-hand side the value

$$-2\iint_D H^2\,dS + 4\pi\chi(\Omega) - 2\oint_{\partial D} k_g\,d\sigma + \cos\gamma\left(2H^2|\Sigma'| + \sin^2\gamma \oint_{\partial D} k_\Gamma\,d\sigma\right).$$

Again we use Meusnier's theorem and Euler's theorem to get $k_g = \pm(\cos\gamma)k_\Gamma$ on each Γ_i. From (13) it follows that

$$\left|\oint_{\partial D} k_g\,d\sigma\right| \le |\cos\gamma| 2\pi d.$$

Also, if Ω has d boundary curves and genus g, then $\chi(\Omega) = 2 - 2g - d$. This follows from the fact that one can attach flat discs to the surface at the boundary to make it closed, and reattaching the disks will decrease $\chi(\Omega)$ exactly with d. Taking all this into account, we have

$$\begin{aligned}
&\partial^2 E \\
&= -2H^2\left(\iint_D dS - \cos\gamma|\Sigma'|\right) + 4\pi\chi(\Omega) - 2\oint_{\partial D} k_g\,d\sigma + \cos\gamma\sin^2\gamma\oint_{\partial D} k_\Gamma\,d\sigma \\
&\le -2H^2\big(|\Omega| - \cos\gamma|\Sigma'|\big) + 4\pi(2-2g-d) + 4\pi d|\cos\gamma| + |\cos\gamma|\sin^2\gamma\cdot 2\pi d \\
&= -2H^2\big(|\Omega| - \cos\gamma\,|\Sigma'|\big) + 4\pi(2-2g) - 2\pi d\big(2 - 2|\cos\gamma| - |\cos\gamma|\sin^2\gamma\big).
\end{aligned}$$

The first term is always negative because $|\Omega| > |\Sigma'|$, since Σ' is a planar surface spanning Γ, so

$$\partial^2 E < 4\pi(2-2g) - 2\pi d\big(2 - 2|\cos\gamma| - |\cos\gamma|\sin^2\gamma\big). \qquad \square$$

The last lemma basically proves the main theorem of this paper. Through simple calculus, we can easily see that the expression

$$2 - 2|\cos\gamma| - |\cos\gamma|\sin^2\gamma$$

is nonnegative for any angle $\gamma \in (0,\pi)$. The computation is shown in [Marinov 2010]. This observation implies that

$$\partial^2 E < 4\pi(2-2g),$$

and if the genus is positive, the second variation of energy is negative. Therefore Ω cannot be weakly stable, which was the statement of the main theorem. This discussion fully resolves the case for an immersed stable capillary surface with planar embedded boundary.

References

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

[Earp et al. 1991] R. Earp, F. Brito, W. H. Meeks, III, and H. Rosenberg, "Structure theorems for constant mean curvature surfaces bounded by a planar curve", *Indiana Univ. Math. J.* **40**:1 (1991), 333–343. MR 93e:53009 Zbl 0759.53003

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

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

[Marinov 2010] P. I. Marinov, *Stability analysis of capillary surfaces with planar or spherical boundary in the absence of gravity*, Ph.D. thesis, The University of Toledo (Ohio), 2010, available at http://tinyurl.com/82epq4p. MR 2827297

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

[Oprea 2007] J. Oprea, *Differential geometry and its applications*, 2nd ed., Mathematical Association of America, Washington, DC, 2007. MR 2008k:53002 Zbl 1153.53001

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

[Spivak 1979] M. Spivak, *A comprehensive introduction to differential geometry*, 2nd ed., Publish or Perish, Berkeley, 1979. MR 82g:53003a Zbl 0439.53001

[Struik 1988] D. J. Struik, *Lectures on classical differential geometry*, 2nd ed., Dover, New York, 1988. MR 89b:53002 Zbl 0697.53002

[Wente 1966] H. C. Wente, *Existence theorems for surfaces of constant mean curvature and perturbations of a liquid globule in equilibrium*, Ph.D. thesis, Harvard University, 1966.

[Wente 1980] H. C. Wente, "The symmetry of sessile and pendent drops", *Pacific J. Math.* **88**:2 (1980), 387–397. MR 83j:49042a Zbl 0473.76086

Received February 23, 2011.

PETKO I. MARINOV
10A SVETI NAUM AP.9
SOFIA 1421
BULGARIA

Petko.Marinov@rockets.utoledo.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

SMALL HYPERBOLIC POLYHEDRA

SHAWN RAFALSKI

We classify the 3-dimensional hyperbolic polyhedral orbifolds that contain no embedded essential 2-suborbifolds, up to decomposition along embedded hyperbolic triangle orbifolds (turnovers). We give a necessary condition for a 3-dimensional hyperbolic polyhedral orbifold to contain an immersed (singular) hyperbolic turnover, we classify the triangle subgroups of the fundamental groups of orientable 3-dimensional hyperbolic tetrahedral orbifolds in the case when all of the vertices of the tetrahedra are nonfinite, and we provide a conjectural classification of all the triangle subgroups of the fundamental groups of orientable 3-dimensional hyperbolic polyhedral orbifolds. Finally, we show that any triangle subgroup of a (nonorientable) 3-dimensional hyperbolic reflection group arises from a triangle reflection subgroup.

1. Introduction

Let P be a finite volume 3-dimensional hyperbolic Coxeter polyhedron. That is, P is the finite volume intersection of a finite collection of half-spaces in hyperbolic 3-space $\mathbb{H}^3$ in which the bounding planes of each pair of intersecting half-spaces meet at an angle of the form π/n, where $n \geq 2$ is an integer (the geodesic of intersection is called an *edge* of P, and the angle of intersection is called the *dihedral angle* of P along this edge). Then the group of isometries of $\mathbb{H}^3$ generated by the reflections in the faces of P is a discrete group that acts on $\mathbb{H}^3$ with fundamental domain P. Let Γ be the subgroup of index two in this reflection group generated by all the rotations of the form rs, where r and s are the reflections through two intersecting planes that support P. We denote by $\mathbb{O}_P$ the quotient space $\mathbb{H}^3/\Gamma$. Then $\mathbb{O}_P$ is an orientable hyperbolic 3-orbifold called a *hyperbolic polyhedral orbifold*. The group Γ is sometimes denoted by $\pi_1(\mathbb{O}_P)$ and called the *fundamental group of* $\mathbb{O}_P$. We call P a *hyperbolic reflection polyhedron*.

A *small* hyperbolic reflection polyhedron corresponds to a hyperbolic 3-dimensional polyhedral orbifold that contains no embedded essential 2-suborbifolds, up

MSC2010: 52B10, 57M50, 57R18.
Keywords: hyperbolic polyhedra, 3-dimensional Coxeter polyhedra, triangle groups, hyperbolic orbifold, polyhedral orbifold, small orbifold, essential suborbifold, hyperbolic turnover.

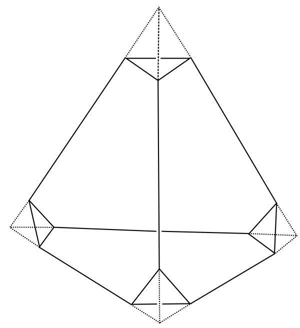


Figure 1. The small Coxeter polyhedra in $\mathbb{H}^3$.

to decomposition along embedded triangular 2-suborbifolds (Definition 2.1). We classify these polyhedra (see Figure 1):

Theorem 1.1. *A* 3*-dimensional hyperbolic reflection polyhedron is small if and only if it is a generalized tetrahedron.*

We also determine those hyperbolic polyhedral orbifolds that contain an immersed (singular) hyperbolic triangular 2-suborbifold. This result is a generalization of the partial classification of triangle groups inside of arithmetic hyperbolic tetrahedral reflection groups given in [Maclachlan 1996]. In Section 4, we will provide a conjectural list of all the possibilities for immersed turnovers in all polyhedral orbifolds:

Theorem 1.2. *If a hyperbolic polyhedral* 3*-orbifold contains a singular hyperbolic turnover that does not cover an embedded hyperbolic turnover, then at least one component of its Dunbar decomposition is a generalized tetrahedron, and the immersed turnover is contained in a unique such component. Furthermore, if T is a generalized tetrahedron with all nonfinite vertices and whose associated polyhedral* 3*-orbifold contains an immersed turnover, then, up to symmetry, T is of the form $T[2, m, q; 2, p, 3]$ (in the notation described in Section 4) with $m \geq 6$, $q \geq 3$ and $p \geq 6$, and the immersed turnover has singular points of orders m, q and p.*

We also determine the triangle subgroups of 3-dimensional hyperbolic reflection groups as arising from triangle reflection subgroups:

Theorem 1.3. *Any (orientable) hyperbolic triangle subgroup of a (nonorientable)* 3*-dimensional hyperbolic reflection group G arises as a subgroup of index two of a (nonorientable) hyperbolic triangle reflection subgroup of G.*

Essential surfaces play an integral role in low-dimensional topology and geometry. One of the most important instances of this fact is the proof of Thurston's hyperbolization theorem for Haken 3-manifolds [Thurston 1982; Morgan 1984]. In brief, Thurston's Theorem is proved by decomposing a given 3-manifold M

(which is called *Haken* if it contains an essential surface) along such surfaces as part of a finite-step process that ends in topological solid balls, from which the hyperbolic structure on M (whose existence is claimed by the theorem) is then, in a sense, reverse-engineered.

One difficulty that arises in attempting to extend the utility of essential surfaces to the orbifold setting is the possible presence of triangular hyperbolic 2-dimensional suborbifolds called *hyperbolic turnovers*. For example, whereas an irreducible 3-manifold with nonempty and nonspherical boundary always contains an essential surface, this is not always the case in the orbifold setting, with hyperbolic turnovers presenting the principal barrier. Thurston's original definition of a Haken 3-orbifold was given for nonorientable 3-orbifolds with underlying space the 3-ball and with singular locus equal to the boundary of the ball [Thurston 1979, Section 13.5, p. 324]. (The singular locus, in this instance, was meant to correspond to the boundary of a polyhedron.) Subsequent formulations of the definition of Haken ("sufficiently large" in [Dunbar 1988, Glossary] or "Haken" in [Boileau et al. 2003, Section 4.2, Remark]) were given for the orientable case and take into account the difficulties that arise from hyperbolic turnovers. Theorem 1.1, which is proved using the same observations that Thurston used to determine 3-orbifolds with the combinatorial type of a simplex as the original non-Haken polyhedral orbifolds, echoes Thurston's original result [1979, Proposition 13.5.2], with respect to this evolution of the language.

2. Definitions

There are several excellent references for orbifolds, such as [Boileau et al. 2003; Cooper et al. 2000]. All of the 3-orbifolds considered in this paper are either orientable hyperbolic polyhedral 3-orbifolds or the result of cutting an orientable hyperbolic polyhedral 3-orbifold along a finite set of totally geodesic hyperbolic turnovers or totally geodesic hyperbolic triangles with mirrored sides. A hyperbolic polyhedral 3-orbifold $\mathbb{O}_P$ is geometrically just two copies of its associated hyperbolic polyhedron P with the corresponding sides of the two copies identified. Therefore, $\mathbb{O}_P$ is a complete metric space of constant curvature -1 except along a 1-dimensional singular subset which is locally cone-like. If P is compact, $\mathbb{O}_P$ is topologically a 3-sphere together with a trivalent planar graph (corresponding to the 1-skeleton of P) with each edge marked by a positive integer to represent the submultiple of π of the dihedral angle at the corresponding edge of P. If P is noncompact with finite volume, then its ideal vertices correspond to trivalent or quadrivalent vertices in the planar graph (again, corresponding to the 1-skeleton of P) and the sum of the reciprocals of the incident edge marks at each such vertex is equal to one or two, according to whether the vertex is trivalent or quadrivalent.

In the noncompact case, $\mathbb{O}_P$ is topologically the result of taking a 3-sphere with this marked graph and removing a (closed) 3-ball neighborhood from each ideal vertex. The statements about the combinatorics of hyperbolic polyhedra in this paragraph are consequences of Andreev's Theorem [1970a; 1970b] (see also [Roeder et al. 2007; Thurston 1979, Section 13.6; 1992]).

A (closed) orientable 2-orbifold is topologically a closed orientable surface with some finite set of its points marked by positive integers (greater than one). Every such 2-orbifold can be realized as a complete metric space of constant curvature with cone-like singularities at the marked points, and where the sign of the curvature depends only on the topology of the underlying surface together with the markings. A 2-orbifold is called *spherical, Euclidean* or *hyperbolic* according to the sign of its constant curvature realization. A *turnover* is a 2-orbifold that is topologically a 2-sphere with three marked points, and a *hyperbolic turnover* is a turnover for which the reciprocal sum of the integer markings is less than one.

Although we will seldom deal with nonorientable objects, we define a *hyperbolic triangle with mirrored sides* as a topological closed disk whose boundary is marked with three distinct points, each point labeled by an integer greater than one and such that the sum of the reciprocals of these integers is less than one, and with the connecting intervals in the boundary between these points marked as "mirrors." Hyperbolic triangles with mirrored sides are nonorientable 2-orbifolds that are doubly covered by hyperbolic turnovers: they are the quotients of hyperbolic turnovers by an involution that fixes an embedded topological circle that passes through the marked points of the turnover. Every embedded hyperbolic turnover in a hyperbolic 3-orbifold can either be made totally geodesic by an isotopy in the 3-orbifold (in which case the preimage in $\mathbb{H}^3$ under the covering map of this totally geodesic 2-suborbifold is a collection of disjoint planes, each tiled by a hyperbolic triangle that is determined by the markings of the singular points — see [Maskit 1988, Chapter IX.C] or [Adams and Schoenfeld 2005, Theorem 2.1], for instance) or else can be moved by an isotopy to be the boundary of a regular neighborhood of a totally geodesic hyperbolic triangle with mirrored sides.

An embedded orientable 2-suborbifold of $\mathbb{O}_P$ is topologically a surface that meets the marked graph transversely. We note that any simple closed curve $C \subset \partial P$ that meets the 1-skeleton transversely determines such a 2-suborbifold by adjoining to C the two topological disks that it bounds, one to either side of $\partial P \subset \mathbb{O}_P$. A closed path on ∂P that is isotopic to a simple circuit in the dual graph to the 1-skeleton of P is called a *k-circuit*, where k is the number of edges the path crosses. An embedded hyperbolic triangle with mirrored sides occurs as a suborbifold of $\mathbb{O}_P$ whenever P has a triangular face all of whose edges are labeled two (in this case, the triangle with mirrored sides is topologically just the disc bounded by these three edges in the marked graph).

The terminology of this paragraph is introduced in terms of general orbifolds. A compact *n-orbifold* $\mathbb{O}$ with boundary is a metrizable topological space which is locally diffeomorphic either to the quotient of $\mathbb{R}^n$ by a finite group action or to the quotient of $\mathbb{R}^{n-1} \times [0, \infty)$ by a finite group action, with points of the latter type making up the *boundary* $\partial\mathbb{O}$ of $\mathbb{O}$ (itself an $(n-1)$-orbifold). We use the term *orbifold ball* (respectively, *orbifold disk*) to refer to the quotient of a compact 3-ball (respectively, 2-disk) by a finite group action. We say a compact 3-orbifold $\mathbb{O}$ is *irreducible* if every embedded spherical 2-suborbifold bounds an orbifold ball in $\mathbb{O}$. A 2-suborbifold $F \subset \mathbb{O}$ is called *compressible* if either F is spherical and bounds an orbifold ball or if there is a simple closed curve in F that does not bound an orbifold disk in F but that bounds an orbifold disk in $\mathbb{O}$, and *incompressible* otherwise. There is a relative notion of ∂-incompressibility (whose exact definition we do not require). We call F *essential* if it is incompressible, ∂-incompressible and not parallel to a boundary component of $\mathbb{O}$. We call a compact irreducible 3-orbifold *Haken* if it is either an orbifold ball, or a turnover crossed with an interval, or if it contains an essential 2-suborbifold but contains no essential turnover. A compact irreducible 3-orbifold is called *small* if it contains no essential 2-suborbifolds and has (possibly empty) boundary consisting only of turnovers. (We note that a compact, orientable and irreducible orbifold is both Haken and small if and only if it is either a cone on a spherical turnover or a product of a turnover with an interval.) These definitions extend to any arbitrary 3-orbifold that is diffeomorphic to the interior of a compact 3-orbifold with boundary.

We observe that Euclidean and hyperbolic turnovers are always incompressible because a simple closed curve on these objects always bounds an orbifold disk. As a consequence, in an irreducible 3-orbifold, any incompressible 2-orbifold (in fact, even any singular hyperbolic turnover) can be made disjoint from an embedded hyperbolic turnover.

Remark. It is a consequence of a theorem of Dunbar [1988] that a hyperbolic polyhedral 3-orbifold can be decomposed (uniquely, up to isotopy) along a system of essential, pairwise nonparallel hyperbolic turnovers into pieces that contain no essential (embedded) turnovers, and, moreover, that each component of the decomposition is either a Haken or a small 3-orbifold; see [Boileau et al. 2003, Theorem 4.8]. An embedded hyperbolic turnover in a hyperbolic polyhedral 3-orbifold $\mathbb{O}_P$ will correspond to a simple closed curve in ∂P that crosses exactly three edges whose dihedral angles sum to less than π. If such a curve is parallel in ∂P to a triangular face of P all of whose edges are labeled two, then the hyperbolic turnover corresponding to this curve is isotopic to the boundary of a regular neighborhood of a hyperbolic triangle with mirrored sides (the latter arising from the triangular face of P) in $\mathbb{O}_P$. In this case, one component of the Dunbar decomposition will consist of the regular neighborhood of this triangle with mirrored sides (in fact, this is a small 3-orbifold). The complement of this component in $\mathbb{O}_P$ is (orbifold) diffeomorphic to the complement of the triangle with mirrored sides in $\mathbb{O}_P$ (because the hyperbolic turnover collapses onto the mirrored triangle as the radius of the regular neighborhood goes to zero), and so, for convenience, we discard the component of the Dunbar decomposition corresponding to this regular neighborhood.

With the above convention in mind, we have the following:

Definition 2.1. A hyperbolic reflection polyhedron P is *small* if the Dunbar decomposition of $\mathbb{O}_P$ (with the convention of the preceding paragraph) consists of a single connected small component.

In the projective model of $\mathbb{H}^3$, consider a linearly independent set of four points, any or all of which may lie on the boundary of or outside of the projective ball. If the line segment between each pair of these points intersects the interior of the projective ball, then the points determine a *generalized tetrahedron*. This polyhedron is obtained by taking the (possibly infinite volume) polyhedron in $\mathbb{H}^3$ spanned by the points and truncating its infinite volume ends by the dual hyperplanes to the superideal vertices. The resulting polyhedron has finite volume and all of its vertices are either finite or ideal. The faces arising from truncated superideal vertices — which are called, along with the finite and ideal vertices, *generalized vertices* — are triangular, and the dihedral angle at each edge of these faces is $\pi/2$. In particular, if a generalized tetrahedron P is a Coxeter polyhedron, then any generalized vertex arising from a truncated face is a hyperbolic triangle that tiles (under the tiling associated to P) a geodesic plane in $\mathbb{H}^3$ (and thus gives rise to an embedded hyperbolic triangle with mirrored sides in $\mathbb{O}_P$).

3. Proof of Theorem 1.1

Let P be a 3-dimensional hyperbolic Coxeter polyhedron, and let $\mathbb{O}_P$ be its hyperbolic polyhedral 3-orbifold. First assume that P is a generalized tetrahedron. Then $\mathbb{O}_P$ is topologically the 3-sphere with a marked planar graph as in Figure 2.

Each dot in the figure represents a generalized vertex, and so is either a finite vertex, a triangle with mirrored sides or a Euclidean turnover cusp (the latter if the vertex is ideal). Any dot that represents a triangle corresponds to a nonseparating

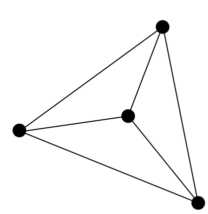


Figure 2. The graph associated to a generalized tetrahedron.

hyperbolic turnover of the Dunbar decomposition of $\mathbb{O}_P$. Moreover, since any two hyperbolic turnovers can be made disjoint by an isotopy, any other turnovers in the Dunbar decomposition occur as topological 2-spheres that intersect the graph from the figure in exactly three distinct edges. But the only possibility for such a 2-sphere is one that surrounds a dot, and that therefore is parallel to a generalized vertex of P. So the Dunbar decomposition of $\mathbb{O}_P$ (under the convention of Definition 2.1) has a single component.

To see that this component is small, we consider the graph of Figure 2 as the 1-skeleton of a tetrahedron in the 3-sphere. Using standard topology arguments, it can be shown that an incompressible 2-suborbifold intersects the interior of this tetrahedron in triangles and quadrilaterals. But a triangular intersection implies that the incompressible 2-suborbifold is isotopic to the hyperbolic turnover associated to a generalized vertex, and a quadrilateral intersection produces a compression. So P is small if it is a generalized tetrahedron.

Now assume that P is small. The rest of the proof of Theorem 1.1 depends on the following simple observation

Remark [Thurston 1979, Proposition 13.5.2]. Suppose that $C \subset \partial P$ is a simple closed curve that is transverse to, forms no bigons with, does not surround a single vertex of, and that crosses at least two distinct edges of the 1-skeleton of P. Then C determines an incompressible 2-suborbifold of $\mathbb{O}_P$ if and only if (1) it intersects any face in a connected set or not at all and (2) it intersects the common edge of two adjacent faces whenever its intersection with both faces is nonempty.

We begin with the following fact about triangular faces of P:

Lemma 3.1. *If T is a triangular face of P, then T corresponds to a hyperbolic turnover in $\mathbb{O}_P$ or P is a generalized tetrahedron.*

Proof. Suppose that T is as in Figure 3a (in this and all subsequent figures in this section, we depict P by a planar projection). If $1/p+1/q+1/r \geq 1$, then the three edges incident to the vertices of T must intersect (or meet at a Euclidean turnover)

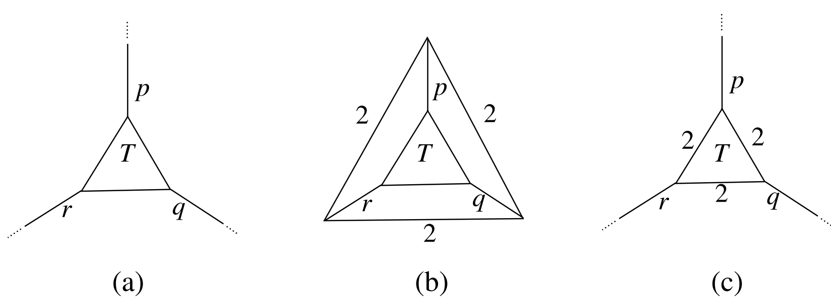


Figure 3. Triangular faces.

[Roeder et al. 2007, Lemmata 3.2 and 3.3], in which case P is a generalized tetrahedron (possibly with an ideal vertex). Otherwise, we have $1/p + 1/q + 1/r < 1$. Then the 3-circuit around this face determines a hyperbolic turnover in $\mathbb{O}_P$ whose associated triangle in P must be boundary-parallel (in P) because P is small. The two possibilities are shown in Figures 3b (in which the hyperbolic turnover collapses to the outermost face) and 3c (in which the hyperbolic turnover collapses to T). □

Throughout the rest of the proof, we will use the observation from the above lemma, i.e., that any 3-circuit in a small hyperbolic polyhedron surrounds a generalized vertex. In the case when the 3-circuit determines a hyperbolic turnover, this follows by the fact that a hyperbolic turnover in a hyperbolic 3-orbifold always corresponds to a totally geodesic 2-suborbifold (according to the second paragraph in Section 2; compare also with the incompressibility observation just before the remark on page 195): Because the polyhedron P is small, this totally geodesic 2-suborbifold of $\mathbb{O}_P$ cannot be an embedded hyperbolic turnover (because $\mathbb{O}_P$ has no boundary, and so such a turnover would have to be essential), and therefore must be a triangle with mirrored sides that corresponds to a triangular face of P.

Consider an n-sided face F of P, as in Figure 4. Assume that $n \geq 4$. The

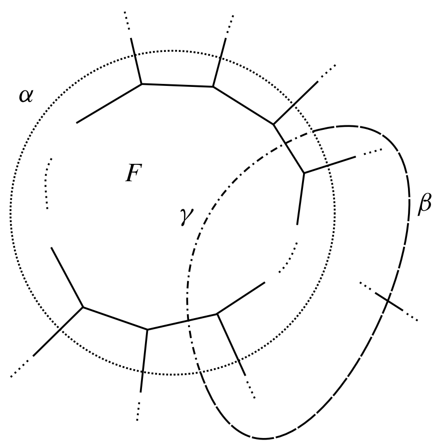


Figure 4. A face of P and a compression.

n-circuit α around F determines a 2-orbifold that must be compressible, with a compressing orbifold disk whose intersection with ∂P appears as the dashed arc β in the figure. Since $n \geq 4$, it must be that each side of the 3-circuit $\beta \cup \gamma$ contains at least two edges radiating outward from F (that is, edges meeting F only in vertices). Since $\mathbb{O}_P$ is small, $\beta \cup \gamma$ bounds a triangle $T \subset \partial P$. Figure 5 illustrates the two possibilities, depending on the side of $\beta \cup \gamma$ to which T lies. Of course, these differ only by the choice of projection of P into the plane.

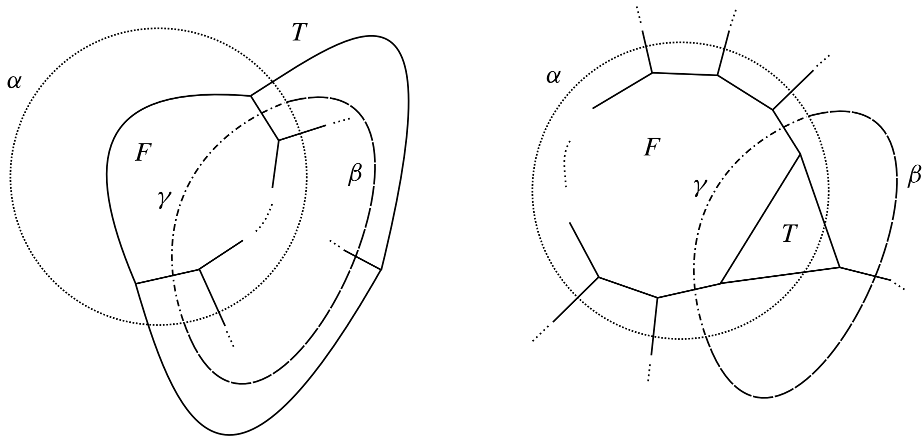


Figure 5. Two projections of a face of P with adjacent triangle.

We now consider all such compressions of this 2-orbifold, and all of the resulting adjacent triangles to F. Let α denote the k-circuit that encloses F and these triangles, as in Figure 6, left.

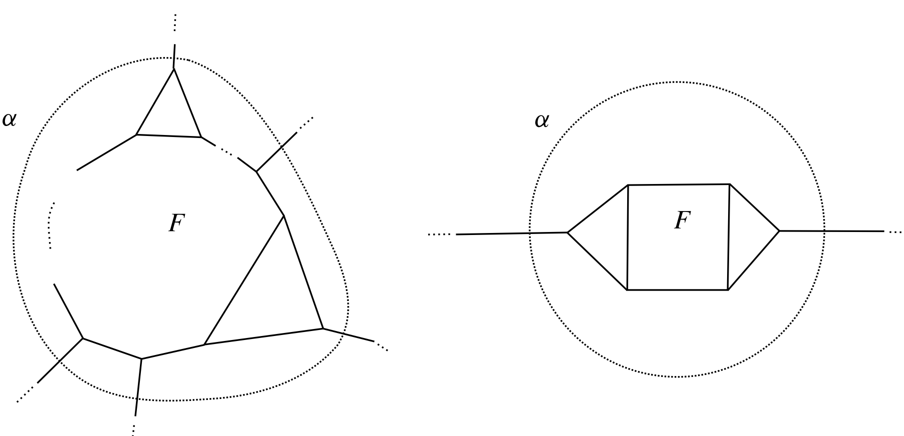


Figure 6. Left: a face of P with all of its adjacent triangles, and a k-circuit. Right: the same, with $k = 2$.

If $k = 2$, then F must be a quadrilateral with two triangles adjacent to it on opposite sides, in which case P is a triangular prism with one face that corresponds to a hyperbolic turnover in $\mathbb{O}_P$ as in Lemma 3.1, i.e., P is a generalized tetrahedron. See Figure 6, right.

If $k = 3$, then α surrounds a generalized vertex to the outside. In this case, the face F must be as in Figure 7, where each dot represents either a finite vertex, an ideal vertex or a hyperbolic triangle. Filling in the generalized vertex to the outside of α, we have that P is a generalized tetrahedron.

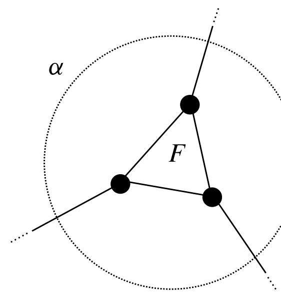


Figure 7. A face of P with all of its adjacent triangles, and a 3-circuit.

If $k > 3$, then the 2-orbifold determined by α has a compression. But any such compression would add an adjacent triangular face to F, and we have assumed that α encloses all such triangles. So $k \leq 3$. This completes the proof of Theorem 1.1.

4. Turnovers in hyperbolic polyhedra

In this final section, we prove Theorems 1.2 and 1.3, and provide a classification of the immersed hyperbolic turnovers in those tetrahedral orbifolds that arise from tetrahedra with no finite vertices. Although Theorem 1.3 does not follow from Theorem 1.2, we will provide the proof of the former in the midst of the proof of the latter, as it contains an observation that is necessary for both proofs.

It was shown in [Rafalski 2010] that if a hyperbolic 3-orbifold contains a singular hyperbolic turnover, then that turnover must be contained in a low-volume small 3-suborbifold. In particular:

Theorem 4.1 [Rafalski 2010, Theorem 1.1 and Corollary 1.3]. *Let Q be a compact*, *irreducible*, *orientable*, *atoroidal* 3*-orbifold. Then any immersion $f : \mathcal{T} \to Q$ of a hyperbolic turnover into Q is homotopic into a unique component of the Dunbar decomposition of Q*, *up to covers of parallel boundary components of the decomposition. Moreover*, *if f is a singular immersion that does not cover an embedded turnover or triangle with mirrored sides*, *then the component containing $f(\mathcal{T})$ is unique*, *and it is a small* 3*-orbifold.*

Proof of Theorem 1.2.. If $\mathbb{O}_P$ is a hyperbolic polyhedral 3-orbifold, then it is homeomorphic to the interior of an orbifold that satisfies the hypotheses of Theorem 4.1. If $\mathbb{O}_P$ contains a singular turnover, then this turnover is contained in a small component of the Dunbar decomposition of $\mathbb{O}_P$, and Theorem 1.1 classifies these small orbifolds as generalized tetrahedral orbifolds.

It remains to provide a classification of the generalized tetrahedra whose associated 3-orbifolds contain immersed turnovers. We will do so for generalized tetrahedra all of whose vertices are nonfinite. *See the summary at the end of the paper for the results of the classification.* The techniques we use to provide this

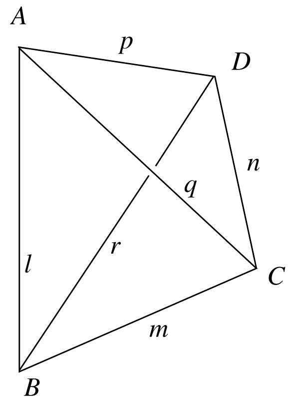


Figure 8. $T[l, m, q; n, p, r]$.

classification can be used to classify the immersed turnovers in all tetrahedral orbifolds, thereby extending and completing the classification begun by Maclachlan in the case of compact (nongeneralized) tetrahedral orbifolds [Maclachlan 1996], however, the case-by-case analysis required to complete this classification in general is somewhat excessive.

We let $T[l, m, q; n, p, r]$ denote the hyperbolic generalized tetrahedron $ABCD$ with dihedral angles π/l, π/m, π/q, π/n, π/p and π/r, as in Figure 8, with the convention that a vertex of T is truncated (respectively, ideal) if the dihedral angles of its three coincident edges sum to less than (respectively, equal to) π.

The conditions on l, m, q, n, p and r guaranteeing the existence of the tetrahedron $T[l, m, q; n, p, r]$ are known [Ushijima 2006]. In particular, there are nine compact (nontruncated) tetrahedra (see [Ratcliffe 1994, Chapter 7], for instance), all of whose associated orbifolds contain singular turnovers. We note that, of the nine compact (nontruncated) tetrahedra, eight yield arithmetic hyperbolic 3-orbifolds. As we noted above, Maclachlan classified almost all of the immersed turnovers in these arithmetic tetrahedral orbifolds using arithmetic methods. Our geometric technique can be considered as an alternative means to prove (and extend) those results, without appeal to arithmeticity.

Denote by $\mathbb{O}_T$ the 3-orbifold determined by $T[l, m, q; n, p, r]$. Recall from Section 2 that any hyperbolic turnover in a hyperbolic 3-orbifold that does not collapse onto a hyperbolic triangle with mirrored sides may be assumed to be totally geodesic. It also follows from the incompressibility of hyperbolic turnovers in irreducible orbifolds that an immersed turnover must be disjoint from any embedded turnover [Rafalski 2010, Lemma 5.3]. Consequently, if $\mathcal{T}$ is a hyperbolic turnover, then an immersion $f : \mathcal{T} \to \mathbb{O}_T$ lifts to the universal cover $\mathbb{H}^3$ as a collection of geodesic planes with some intersections — two or more of these planes will intersect whenever there is a covering transformation (i.e., an element of the fundamental group $\pi_1(\mathbb{O}_T)$ of $\mathbb{O}_T$, which is just the group of isometries of $\mathbb{H}^3$ that yields the quotient $\mathbb{O}_T$) that does not move one plane completely disjoint from some of the others, and this must occur if there is a singular immersion of a turnover in $\mathbb{O}_T$ — and, additionally, the collection of planes determined by an immersed turnover must be disjoint from the collection of planes determined by any turnover corresponding to a generalized vertex of T.

Proof of Theorem 1.3. Let $P \subset \mathbb{H}^3$ be a polyhedron that generates the nonorientable 3-dimensional hyperbolic polyhedral reflection group G, and let $S \subset G$ be an orientable triangle subgroup. Then S is generated by two elliptic elements in G and stabilizes a plane $\Pi_S \subset \mathbb{H}^3$. In particular, Π_S meets the axis of every element of S at a right angle, and the intersections of Π_S with these axes comprise the vertex set of a tiling of Π_S by hyperbolic triangles. Every such vertex will have k lines passing through it (where k is the order of the elliptic element stabilizing the vertex) that are the perpendicular intersections with Π_S of G-translates of a face of P. This set of lines and their intersections generates a tiling of Π_S by hyperbolic triangles that corresponds to a hyperbolic triangle with mirrored sides in the nonorientable hyperbolic orbifold $\mathbb{H}^3/G$, and this 2-orbifold is covered by the hyperbolic turnover corresponding to S. Therefore, S is contained in the triangle reflection subgroup of G that corresponds to this nonorientable triangle 2-orbifold. □

We take a moment to emphasize the observation from the above proof: Any maximal (orientable) triangle subgroup of 3-dimensional hyperbolic polyhedral reflection group has as a fundamental domain a triangle whose edges are contained in the faces of the corresponding polyhedral tiling of $\mathbb{H}^3$ (the edges may intersect multiple faces of the polyhedral tiling). This fact is used in the next paragraph.

Here is the strategy for classifying the immersed turnovers of $\mathbb{O}_T$. (The proof is long, but this paragraph contains the core idea.) Let $\mathcal{T}$ be a hyperbolic turnover. Up to conjugacy, there is a unique discrete orientation-preserving group of isometries of the hyperbolic plane $\mathbb{H}^2$ corresponding to the tiling of $\mathbb{H}^2$ by copies of the triangle that determines $\mathcal{T}$ (the fundamental group $\pi_1(\mathcal{T})$ of $\mathcal{T}$). If $f : \mathcal{T} \to \mathbb{O}_T$ is an immersion, then f may be assumed to have totally geodesic image. Consider a plane $\Pi_{\mathcal{T}}$ in the collection of planes in $\mathbb{H}^3$ corresponding to $f(\mathcal{T})$. This plane is stabilized by a copy of the fundamental group of some turnover (possibly a smaller turnover that is covered by $f(\mathcal{T})$, if the fundamental group of $f(\mathcal{T})$ is not maximal) — a subgroup Γ of the fundamental group of the orbifold $\mathbb{O}_T$ — for which there is a tiling of $\Pi_{\mathcal{T}}$ by hyperbolic triangles whose edges are a (possibly proper) subset of the intersections of $\Pi_{\mathcal{T}}$ with Γ-translates of the faces of T, and whose vertices are a (possibly proper) subset of the perpendicular intersections of $\Pi_{\mathcal{T}}$ with Γ-translates of the edges of T. We will locate all of the immersed turnovers in $\mathbb{O}_T$ by reversing this process, that is, by determining exactly the hyperbolic planes in the universal cover $\mathbb{H}^3$ that are stabilized by a triangle subgroup of $\pi_1(\mathbb{O}_T)$. Thus we choose an arbitrary edge e_1 of T and develop copies of T in $\mathbb{H}^3$ (by reflecting in faces) until we find another edge e_2 which is coplanar with but which shares no (generalized) vertex with e_1. Since we need only concern ourselves with maximal triangle subgroups, the observation following the proof of Theorem 1.3 allows to assume that the common plane, which we denote by Π_F (where F is a face of T incident to e_1), consist of developed faces of T. Let Π_1 be the plane containing another face of T incident with e_1, and let Π_2 be the plane containing another face of (a developed image of) T containing e_2. Suppose that Π_1 and Π_2 intersect Π_F at angles of π/a and π/b, respectively. If Π_1 and Π_2 intersect at an angle of π/c, and if $1/a+1/b+1/c<1$, then the rotations about edges e_1 and e_2 (of orders a and b, respectively), will generate a triangle subgroup of $\pi_1(\mathbb{O}_T)$, and the invariant plane for that subgroup will project to an immersed turnover in $\mathbb{O}_T$ (every developed edge of T that intersects the invariant plane for this triangle subgroup at an oblique angle will correspond to an immersion of the turnover). This determines a maximal triangle subgroup of $\pi_1(\mathbb{O}_T)$, and the type of the corresponding immersed turnover will be (a, b, c). In most cases, we will show that there can be no such edge e_2 that is both coplanar with e_1 and that has an incident face whose corresponding plane Π_2 intersects the plane Π_1, which rules out the possibility of an immersed turnover. In the other cases, we will find a turnover after a minimal development of T. Thus, our determination of the immersed turnovers in $\mathbb{O}_T$ will be complete.

We divide the remainder of the proof of Theorem 1.2 into subsections.

4.1. *The case when a single edge separates e_1 from e_2.*

4.1.1. *The single separating edge has order 2:* To begin, we determine the case in which the immersed turnover can be found after crossing only one edge between e_1 and e_2 (there must be at least one edge crossed, in this process, to ensure that the turnover is not parallel to a vertex). Consider Figure 9, which shows two

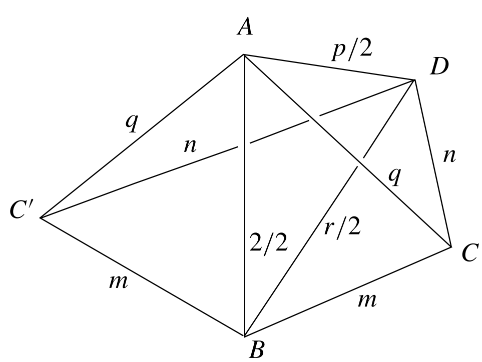


Figure 9. Two copies of the tetrahedron $T[2, m, q; n, p, r]$.

copies of the tetrahedron $T[2, m, q; n, p, r]$. Each edge is labeled according to the submultiple of π for the dihedral angle there (so, for example, the edge AD has a dihedral angle of $2\pi/p$). In particular, the points A, B, C and C' are coplanar. We use F to denote the face ABC of T and Π_F to denote the plane that contains F. We consider the edges $e_1 = AC'$ and $e_2 = BC$, and the planes $\Pi_1 = AC'D$, Π_F and $\Pi_2 = BCD$. Under the assumption that all of the vertices of T are nonfinite, we observe it is necessary for m, q, p and r to all be at least 3. From the figure, we see that Π_1 meets Π_F at an angle of π/q and that Π_F meets Π_2 at an angle of π/m, and so we are left to determine whether or not Π_1 and Π_2 intersect, and at what angle this possible intersection occurs.

The vertex D is either ideal or truncated. If it is ideal, then its link is the orbifold quotient of a horosphere by a Euclidean triangle group. If it is truncated, then it corresponds to a geodesic plane in the universal cover that is stabilized by a hyperbolic triangle group. In both cases, we illustrate the straightforward geometric determination of the conditions on n, p and r that ensure the intersection of Π_1 and Π_2 in the link of D, and determine the angle at which any intersection occurs [Rafalski 2010, Section 9.4].

Figure 10 illustrates part of the link of D *as viewed from* D (this is either a hyperbolic plane or a Euclidean plane corresponding to the horosphere centered at

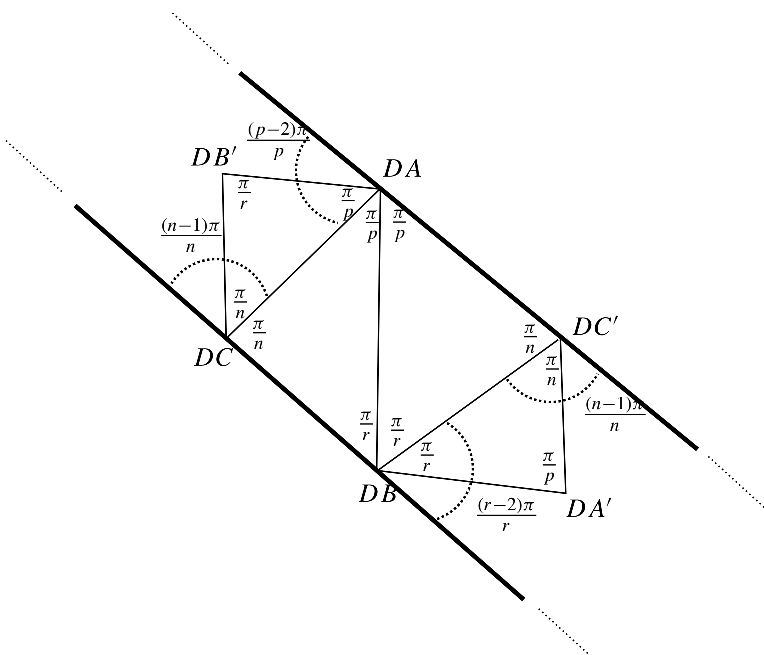


Figure 10. Part of the link of a nonfinite vertex of $T[2, m, q; n, p, r]$.

an ideal vertex). The vertices in the picture are labeled according to the edges of T that are incident at D (the labels DA' and DB' represent edges in the development of T that are the reflections of the edges DA and DB through the faces $BC'D$ and ACD, respectively, in Figure 9). Assume first that $n > 2$. Because p must be at least 3 (and similarly for r), we have the inequality $(p-2)\pi/p + (n-1)\pi/n \geq \pi$ (and similarly $(r-2)\pi/r + (n-1)\pi/n \geq \pi$). The angles with the measures from the previous sentence are indicated in the figure as the labels of the four *dotted* arcs (all other angles in the figure refer to the measure at the appropriate triangular vertex). Using this inequality, we conclude that the indicated bold rays directed northwest from DA and DC do not intersect, because the sum of the angles that these rays make with the segment from DC to DA is at least π (and similarly for the rays directed southeast from DC' and DB, because the sum of the angles that these rays make with the segment from DB to DC' is at least π). Consequently, the bold lines in the figure (and the corresponding planes Π_1 and Π_2) cannot intersect in this case. A similar argument implying that Π_1 and Π_2 do not intersect holds when $n = 2$ and both p and r are greater than 3: The rays directed northwest from DA and DC make angles with the segment between these two points of $(p-2)\pi/p \geq \pi/2$ and $\pi/2$, respectively, and so the sum of these angles will be at least π (when $n = 2$ and $r \geq 4$, the same argument proves that the southeast rays from DC' and DB do not intersect). Finally, if $n = 2$ and $p = 3$ (respectively, $r = 3$), then it is easily seen Π_1 and Π_2 intersect at an angle of π/r (respectively, π/p), and the line of intersection passes through the point DB' (respectively, DA').

We therefore have, when $l = 2$ and our search for a turnover crosses only one edge, that an immersed turnover only arises when $n = 2$ and either $r = 3$ or $p = 3$. If $r = 3$, then this yields a triple of planes intersecting pairwise in angles of π/q, π/m and π/p, with $q \geq 3$, $m \geq 6$ and $p \geq 6$. If $p = 3$, then the pairwise angles of intersection are π/q, π/m and π/r, with $q \geq 6$, $m \geq 3$ and $r \geq 6$. (The inequalities are induced by the assumption that all of the vertices of T are nonfinite.) By analyzing Table 1 (whose data is collected from [Singerman 1972]), we see that

supergroup	subgroup	index	normal	supergroup	subgroup	index	normal
$(3, 3, t)$	(t, t, t)	3	Yes	$(2, 3, 8)$	$(3, 8, 8)$	10	No
$(2, 3, 2t)$	(t, t, t)	6	Yes	$(2, 3, 9)$	$(9, 9, 9)$	12	No
$(2, s, 2t)$	(s, s, t)	2	Yes	$(2, 4, 5)$	$(4, 4, 5)$	6	No
$(2, 3, 7)$	$(7, 7, 7)$	24	No	$(2, 3, 4t)$	$(t, 4t, 4t)$	6	No
$(2, 3, 7)$	$(2, 7, 7)$	9	No	$(2, 4, 2t)$	$(t, 2t, 2t)$	4	No
$(2, 3, 7)$	$(3, 3, 7)$	8	No	$(2, 3, 3t)$	$(3, t, 3t)$	4	No
$(2, 3, 8)$	$(4, 8, 8)$	12	No	$(2, 3, 2t)$	$(2, t, 2t)$	3	No

Table 1. Triangle supergroups and subgroups.

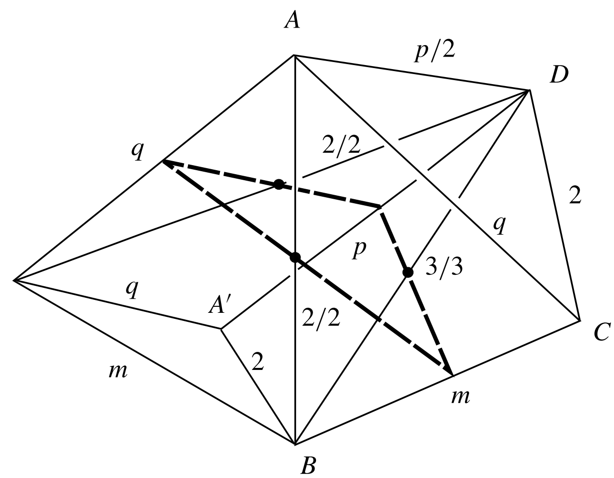


Figure 11. A (q, m, p) triangle in $T[2, m, q; 2, p, 3]$.

this triple of planes does not yield a triangle group that contains any other triangle group. By comparing the second column of the table with the first, we note that it is possible for this triple of planes to yield a triangle group that is contained in some larger triangle group. However, it is not possible for such a supergroup to be a subgroup of $\pi_1(\mathbb{O}_T)$. This follows from the observation in the paragraph following the proof of Theorem 1.3: Because such a supergroup would be a maximal triangle subgroup of $\pi_1(\mathbb{O}_T)$ stabilizing the plane that contains the (q, m, p) (or (q, m, r)) triangle, there would have to be edges in the development of T that intersect the interior of the (q, m, p) (or (q, m, r)) triangle perpendicularly (these intersections would be necessary for the corresponding orbifold covering of the smaller turnover by the larger (q, m, p) or (q, m, r) turnover). By construction, there are no such perpendicular intersections in the interior of the triangle. See Figure 11, which illustrates the case when $r = 3$. As can be seen in the figure, no developed edges of T intersect the interior of the (q, m, p) triangle (the intersections with this triangle that yield immersions of the corresponding turnover are indicated by the dots). Consequently, we can conclude that the (q, m, p) or (q, m, r) triangle determined by Π_1, Π_F and Π_2 is not parallel to any of the vertices of T, and therefore that it determines an immersed turnover in $\mathbb{O}_T$, because $\mathbb{O}_T$ is small. The observations of this paragraph are summarized in items (1) and (2) at the conclusion of the paper.

4.1.2. *The single separating edge has order 3:* We next turn to the case in which the immersed turnover can be found after crossing only one edge between e_1 and e_2, where the order of the crossed edge is $l = 3$. See Figure 12. Let $e_1 = AD'$, $e_2 = BC$, $\Pi_1 = AC'D'$ and $\Pi_2 = BCD$. We make several preliminary observations:

(1) Any two (distinct) planes that truncate developed vertices must be disjoint.

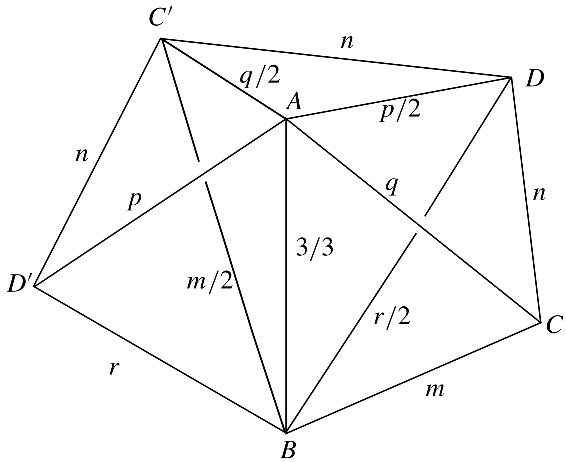


Figure 12. Three copies of the tetrahedron $T[3, m, q; n, p, r]$.

(2) By (1) and by the fact that T has no finite vertices, any two developed edges of the tetrahedron (whose corresponding geodesics in $\mathbb{H}^3$ are distinct) must be disjoint.

(3) It is always the case that the plane containing one face of a generalized hyperbolic tetrahedron will be disjoint from the plane that truncates the vertex opposite to that face.

(4) By (2), if two planes corresponding to two developed faces of T meet a third plane that corresponds to a developed face of T, then any intersection of the first two planes must occur on the side of the third plane where the two interior supplementary angles of intersection sum to less than π.

(5) Any two planes corresponding to two developed faces that both intersect a third plane that truncates a developed vertex intersect if and only if their intersections with that truncated plane (i.e., with the link of the generalized vertex) do so. A corresponding statement is also true in the case when the developed vertex is ideal, that is, that two planes corresponding to two developed faces that intersect at infinity in the case of an ideal vertex intersect in $\mathbb{H}^3$ if and only if their intersections with the link of the ideal vertex themselves intersect.

By (3), Π_2 is disjoint from the plane that truncates the vertex A. When $r = 2$, the planes Π_1 and Π_2 will intersect if and only if their intersections with the link of C' themselves intersect (by (4)). We will analyze the $r = 2$ case in a moment. When $r \geq 3$, we also have that Π_2 does not intersect the plane that truncates the vertex C', reasoned as follows. We will always choose the "inward" normal direction for a plane that contains a face of T by indicating the appropriate opposite vertex to that face in any of our diagrams. When $r = 3$, we observe that Π_2 contains the face of the tetrahedron (not pictured in the figure) that is the reflection of $ABDC'$ through the face BDC', and so Π_2 does not intersect the truncating plane of C' in this case (by (3)). When $r \geq 4$, then we consider the line containing the segment BD which divides Π_2. The half of Π_2 that meets C is prevented from intersecting the truncating plane for C' by the plane ABD, and the other half of Π_2 is prevented from intersecting the truncating plane at C' by the plane containing the reflection of ABD through the face BDC' (both of these follow from (3)).

Therefore, when $r \neq 2$, we have that Π_2 has no intersection with the planes that truncate the vertices A and C'. We observe now that these truncating planes at A and C' determine an open ball (i.e., the region between them in $\mathbb{H}^3$) which contains Π_2. We also note that the edge from A to C' is the only segment of the line of intersection of Π_1 with the planes ABC' and ADC' that lies in this ball. Using the convention for the inward normal direction given above, we conclude that, in order for Π_1 to intersect Π_2, it is necessary for that intersection to occur on the *outward* side of either ABC' (where inward is relative to D) or the *outward* side of ADC' (where inward is relative to B), and consequently that Π_2 must cross at least one of the planes ABC' or ADC'.

By considering the geometry of the generalized vertex B, we have that Π_2 meets ABC' if and only if $r = 2$, and so we analyze this case now. In this case, $\Pi_2 = BDC'$ (as planes) and Π_1 and Π_2 intersect if and only if their intersections with the link of C' intersect (by (5)). The conditions for this intersection in the link of C' are either $m = 2$ (not possible, since $r = 2$), or $n = 2$ and one of q or m equals 3 (not possible, since $r = 2$), or else $q = 2$. In the last case, the intersection of Π_1 and Π_2 occurs along the edge $C'D$ at an angle of π/n, and because $q = 2 = r$ we must have $m \geq 6$, $p \geq 6$ and $n \geq 3$. In this case, $T = T[3, m, 2; n, p, 2]$ contains an immersed (m, n, p) turnover, and this tetrahedron (and the set of conditions on m, n and p) is isometric to the tetrahedron $T[2, p, n; 2, m, 3]$, which appears in item (1) at the end of the paper (it is listed as item (3), additionally). The summary at the end of the paper gives exact conditions on the arrangements of l, m, q, n, p and r which yield isometric tetrahedra.

Otherwise, Π_2 must intersect ADC', and any possible intersection of Π_1 and Π_2 must occur on the outward side of ADC' (that is, the side opposite to vertex B). Using the geometry of the generalized vertex D, we conclude that either $r = 2$ (the case we just analyzed), or $p = 2$, or $n = 2$ and one of p or r equals 3. If $p = 2$, then $q \geq 6$ (using the vertex A), $n \geq 3$ (using the vertex D), and $ADC' = ACDC'$ (as planes). By item (2) above, the lines AC' and CD are disjoint lines in the plane $ACDC'$. These lines are also the intersections with $ACDC'$ of Π_1 and Π_2, respectively. We consider the side of $ACDC'$ that is outward from vertex B, and the interior angles of intersection $(q - 2)\pi/q$ (formed by Π_1 and $ACDC'$) and $(n - 1)\pi/n$ (formed by Π_2 and $ACDC'$) on this side of $ACDC'$ (that is,

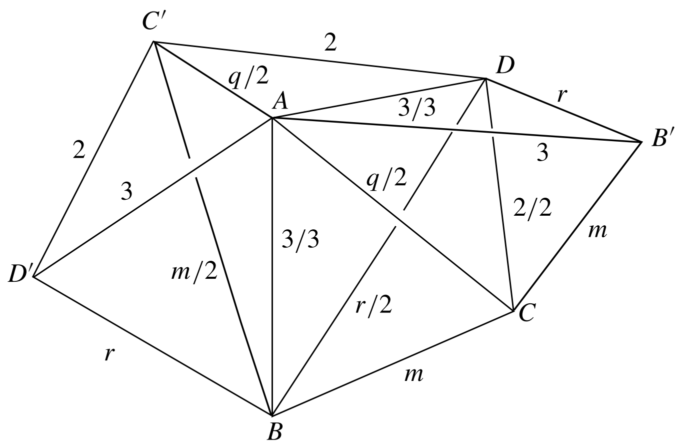


Figure 13. Four copies of the tetrahedron $T[3, m, q; 2, 3, r]$.

the two angles of intersection contained on this side of $ACDC'$ and in the same complementary component of these three planes). The conditions on n and p imply that $(n-1)\pi/n + (q-2)\pi/q \geq \pi$, and because it is only possible for Π_1 and Π_2 to intersect to the outward side of $ACDC'$ (relative to the inward B direction), we use item (4) above to conclude that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case.

In the remaining case, we have $n = 2$ and one of p or r equals 3. If $p = 3$, then $r \geq 6$ and q and m must both be bigger than 2 and also satisfy $1/q + 1/m \leq 1/2$. We modify Figure 12 by adjoining another copy of T to the face ACD. See Figure 13. In this case, $ADC' = AB'DC'$ as planes, and we consider, as in the previous case, the interior angles of intersection $(q-2)\pi/q \geq \pi/3$ and $(r-1)\pi/r \geq 5\pi/6$ formed by $AB'DC'$ with Π_1 and Π_2, respectively, on the outward side of this plane (again, relative to the inward B direction). Since $(r-1)\pi/r + (q-2)\pi/q > \pi$, and again because Π_1 and Π_2 can only intersect on the side of $AB'DC'$ opposite to B, we conclude that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case. The case when $n = 2$ and $r = 3$ is entirely similar, with the same conclusion.

4.1.3. *The single separating edge has order greater than 3:* We now handle the analogous cases to the previous two: when the search for an immersed turnover crosses a single edge between the planes Π_1 and Π_2, and when $l > 3$ (we will specify these planes in each example below, in an analogous way to the previous cases). We will show that no immersed turnovers can be found when $l > 3$.

We consider first the case when $l = 4$ and the vertex B has the Euclidean type $(2, 4, 4)$ with $m = 4$. See Figure 14 (we will, for the most part, drop references to the "link" of a vertex for the remainder of the paper, and assume that work done in, and figures referring to, the link of a vertex will be clear from the context). Referring to the lower half of this figure, we have $e_1 = AC''$, $\Pi_1 = AC''D'$, $e_2 = BC$

and $\Pi_2 = BCD$. The upper half of Figure 14 illustrates the view in the upper half-space model of $\mathbb{H}^3$ from the vertex B, which we have placed at the point at infinity. (This view, along with the similar figures in this section, was generated using the software *KaleidoTile* by Jeffrey Weeks [$\geq$ 2012].) Now Π_2 is represented in this diagram by the line CD, and the plane Π_1 must be represented by a circle (the circle is the boundary of a hemisphere in this model of $\mathbb{H}^3$). We claim that the circle representing Π_1 must be centered at some point in the triangle $AC''D'$, and that none of the three points A, C'' or D' can be contained in this circle's interior. To see this, suppose first that the vertex A of the tetrahedron is a truncated vertex. Then the plane truncating that vertex would appear as a circle in the figure. This circle would have to be centered at the point labeled A because the geodesic edge from B to this plane must meet the plane perpendicularly. Next, we observe that the circle representing Π_1 must intersect the circle centered at A at a right angle (because Π_1 intersects the plane that truncates the vertex A perpendicularly). This is only possible if the point labeled A lies outside of the circle representing Π_1. In the case when the vertex A of T is an ideal vertex, then the circle representing Π_1 would pass *through* the point labeled A. Since all of the vertices A, C'' and D' of

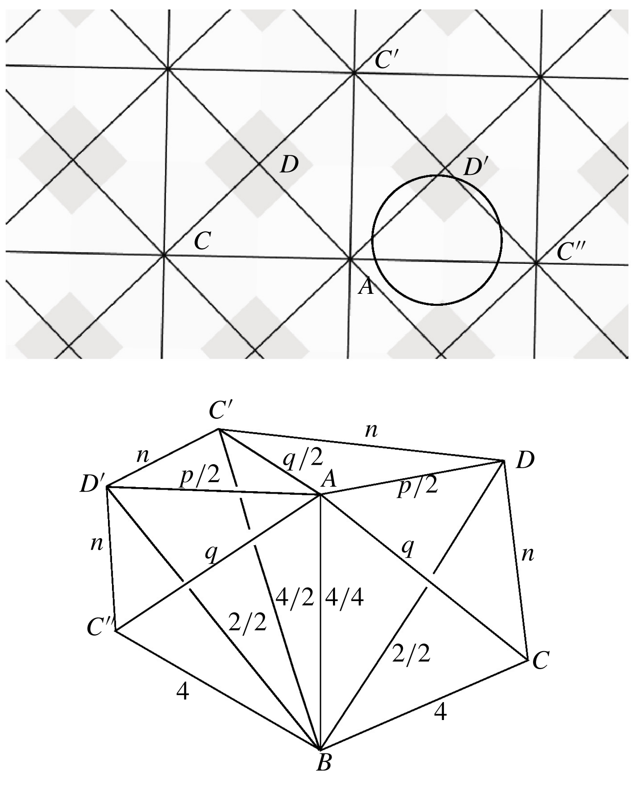
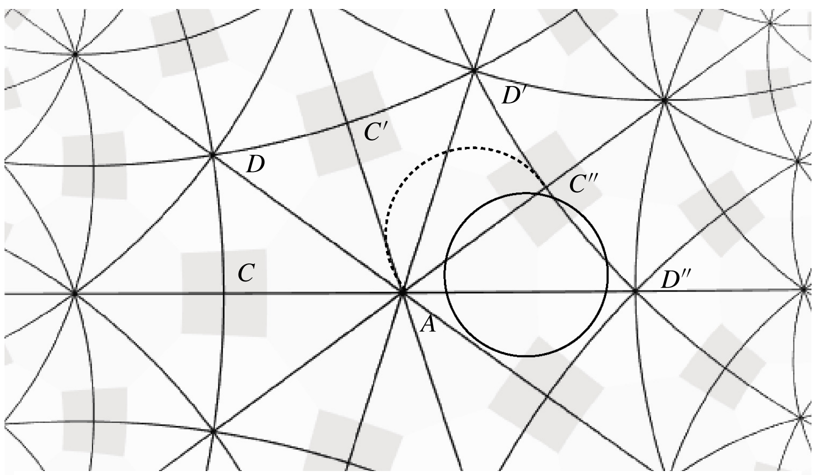


Figure 14. The view from the ideal vertex of type (2, 4, 4).

the tetrahedron are nonfinite, the circle representing Π_1 cannot contain the vertices of the triangle $AC''D'$ in its interior disk. Moreover, this circle must meet each line segment AD', AC'' and $C''D'$ (at angles of π/p, π/q and π/n, respectively) and so the center of this circle must be contained in the triangle $AC''D'$. Such a circle is depicted. Since any such circle cannot intersect the line CD, we conclude that $\Pi_1 \cap \Pi_2 = \varnothing$. An analogous argument can be used to show that we obtain no immersed turnover in this fashion, whenever the vertex B is Euclidean and l is not equal to 2 or 3; this occurs only when the triple (l, m, r) is one of $(4, 2, 4)$, $(6, 2, 3)$ or $(6, 3, 2)$.

We are left then to consider the case when $l \geq 4$ and the vertex B has a hyperbolic type. The argument is similar to the Euclidean vertex case, but we provide the details. Consider first the case of Figure 15. For the purposes of illustration, we have assumed that the vertex B has the type $(2, 4, 5)$, with $l = 5$, $m = 2$ and $r = 4$.



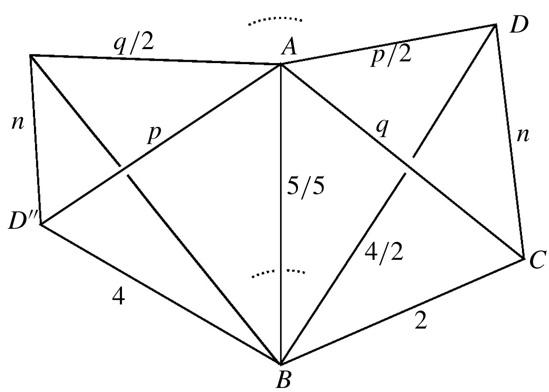


Figure 15. The view from the truncated vertex of type (2, 4, 5).

Here, we have

$$e_1 = AD'', \quad \Pi_1 = AD''C'', \quad e_2 = BC, \quad \Pi_2 = BCD.$$

We consider the hyperbolic plane Π_B that truncates vertex B as a hemisphere in the upper half-plane model, and wish to construct a "view from B" that is similar to the previous case when the B was an ideal vertex. The Poincaré disk $(2, 4, 5)$ tiling pattern of the figure results from projecting this *hemi*sphere to the bounding plane of $\mathbb{H}^3$ through the south pole of the *whole* sphere that contains it [Thurston 1997, Figure 2.12, p. 58]. An important observation about this projection is that it is equivalent to projecting every point $x \in \Pi_B$ to the bounding plane of half-space along the geodesic ray that is perpendicular to Π_B at x. In particular, as in the Euclidean vertex case, each line or circular arc in the figure is the ideal boundary of a plane (each plane corresponding to a face in the tiling of $\mathbb{H}^3$ by T) that meets Π_B perpendicularly, and this projection is conformal, so that the angle of intersection between two lines or circular arcs in the figure is equal to the angle of intersection of the corresponding planes in $\mathbb{H}^3$. We have indicated, in the projection of the figure, the images of the intersection of five copies of T with Π_B, labeled the endpoints of the lines emanating from B by the corresponding letters in the lower part of the figure, and applied an isometry so that A (or, in the case that the vertex A is truncated, the center of the circle that represents the truncating plane for the vertex A) is at the center of the Poincaré disk. The planes Π_1 and Π_2 are represented by a circle and the circular arc CD, respectively.

We observe that, if the vertex C'' is truncated, then the truncating plane $\Pi_{C''}$ for C'' will appear in the figure as a circle (not pictured) with center on the segment AC'', because the point labeled C'' is the endpoint of a semicircle in the half-space model that is perpendicular to both Π_B and $\Pi_{C''}$ (to see this, recall that we may consider the projection from Π_B to the bounding plane as a projection along arcs of such semicircles). As in the previous case, the point C'' cannot be contained in the interior of the circle that is the ideal boundary of Π_1, because then the arc of the semicircle from C'' to its inverse image in Π_B under the projection would meet Π_1, and this is impossible because this arc meets $\Pi_{C''}$ perpendicularly and $\Pi_{C''}$ and Π_1 are orthogonal (the contradiction arises because it would imply the existence of a triangle with two right angles). The same argument holds when either of A or D'' is a truncated vertex, and therefore, as in the previous case, the ideal boundary of Π_1 must bound a disk whose interior is disjoint from the points A, C'' and D'' (these points may lie on the ideal boundary of Π_1 if they are ideal vertices of T). The ideal boundary of Π_1 intersects the segments AC'' and AD'' and the circular arc $C''D''$ at angles of π/q, π/p and π/n, respectively, and the center of the circle representing this ideal boundary has its center contained in the hyperbolic triangle $AC''D''$ in the projection. This is the circle that is depicted in the figure. But such

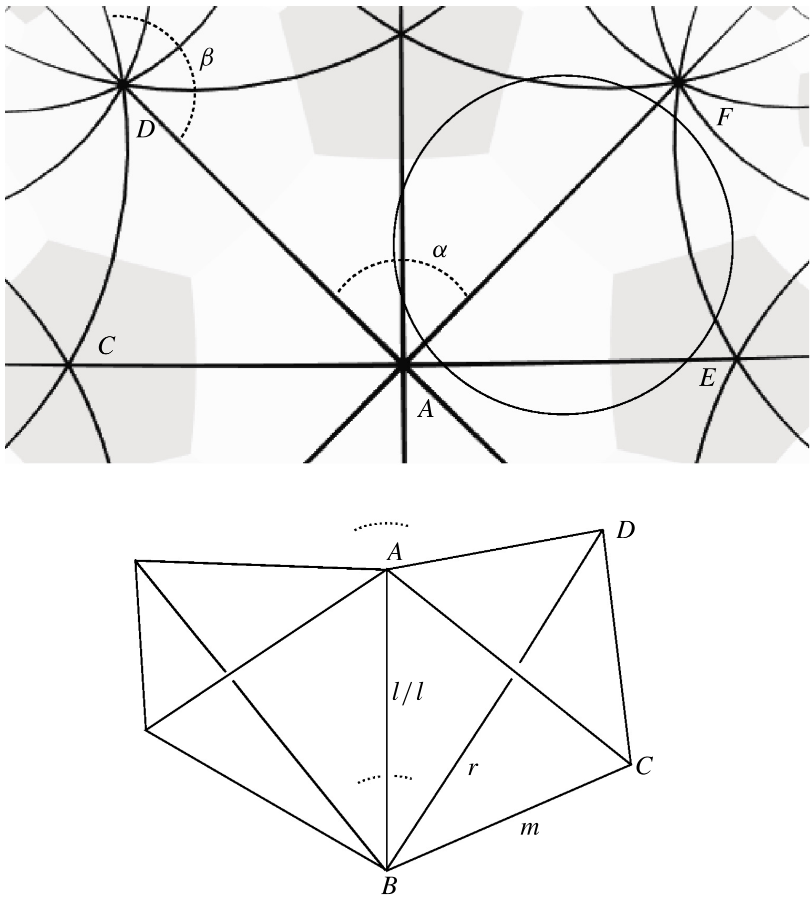


Figure 16. The view from the truncated vertex of generic hyperbolic type when none of l, m and r is 2.

a circle can have no points in the hyperbolic polygon $CAD''C''D'C'D$ that lie outside of the union of hyperbolic triangle $AC''D''$ and the circle with the segment AC'' as its diameter (pictured with a dashed arc in the figure). Consequently, this circle cannot meet any of the sides of this hyperbolic polygon other than AD'' and $D''C''$, and, in particular, we have $\Pi_1 \cap \Pi_2 = \varnothing$. An analogous argument works whenever B has hyperbolic type with one incident order 2 edge and $l \geq 4$.

The case when $l \geq 4$ and B has hyperbolic type with no incident order 2 edge is similar. See Figure 16, in which Π_2 is represented by the circular arc CD and Π_1 is represented as the circle pictured.

When $l \geq 4$, we observe that, in any similar picture (for example, Figure 17), the angles $\alpha = (l-2)\pi/l$ and $\beta = (r-1)\pi/r$ will always be at least $\pi/2$.

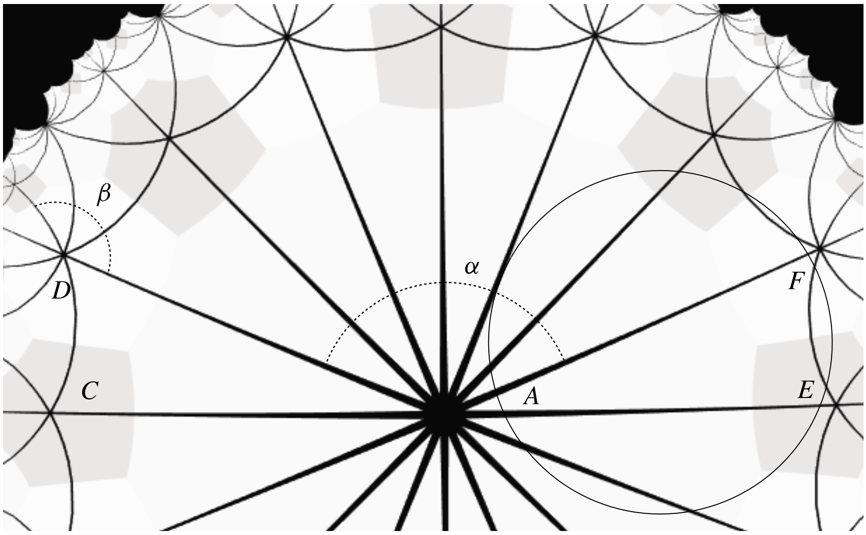


Figure 17. Another view from the truncated vertex of generic hyperbolic type when none of l, m and r is 2.

Hence, since $\alpha \geq \pi/2$ and because the center of the circle representing Π_1 is contained in the hyperbolic triangle AEF, this circle will be disjoint from the interior of the segment AD (it may pass through A, if the corresponding vertex is ideal). Also, noting that AD will always have Euclidean length equal to one of the lengths $|AF|$ or $|AE|$, the conditions on α and β imply that no point of the circle CD that lies above the line AD will be closer to the center of the circle representing Π_1 than any of the points A, E or F. Since A, E and F are not contained in the interior of this circle, we can conclude that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case.

4.2. ***The case when multiple edges separates e_1 from e_2.*** Recall that $\Pi_{\mathcal{T}}$ denotes the plane stabilized by a copy of a triangle subgroup in the fundamental group of $\mathbb{O}_T$, and that e_1 and e_2 denote two developed coplanar edges of T whose (perpendicular) intersections with $\Pi_{\mathcal{T}}$ correspond to two of the cone points of an immersed turnover (whose fundamental group is the triangle group stabilizing $\Pi_{\mathcal{T}}$) in $\mathbb{O}_T$.

Notation. For the remainder of the paper, Π_F refers to the plane containing e_1 and e_2. It is the development in $\mathbb{H}^3$ of one face F of T. The diagrams from Figures 18, 19 and 20 (along with several other figures later in this section) are all drawn with the convention that Π_F is the page containing the illustration. We use L_F to denote the intersection of $\Pi_{\mathcal{T}}$ with Π_F. Additionally, the phrase "the other side of Π_F" refers, in each of the relevant figures, to the side of Π_F that is behind the page (relative to the reader), and the use of the word "plane" at any edge in a diagram *always* refers to a plane that is the development of a face of T in $\mathbb{H}^3$ that passes through that edge.

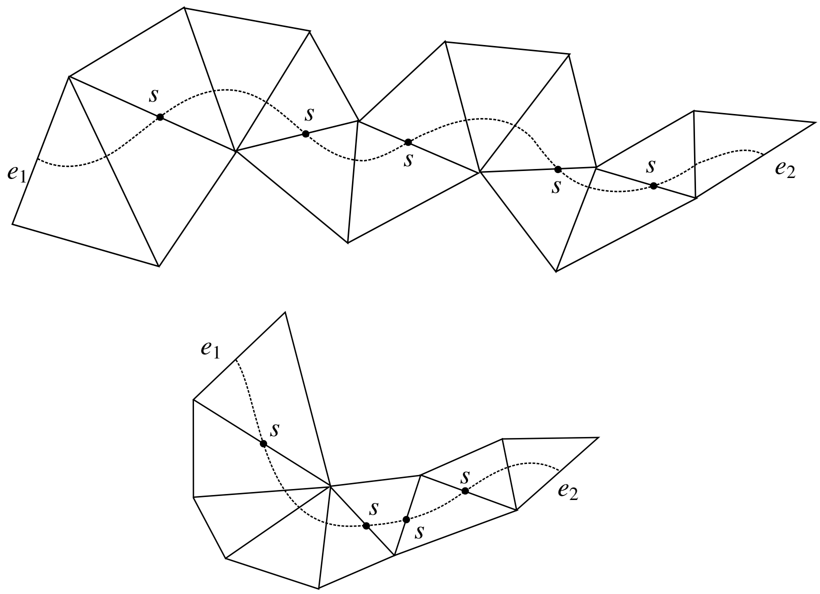


Figure 18. Schematic of some possible developments of a face of T, together with switches and the intersection of the plane $\Pi_{\mathcal{T}}$.

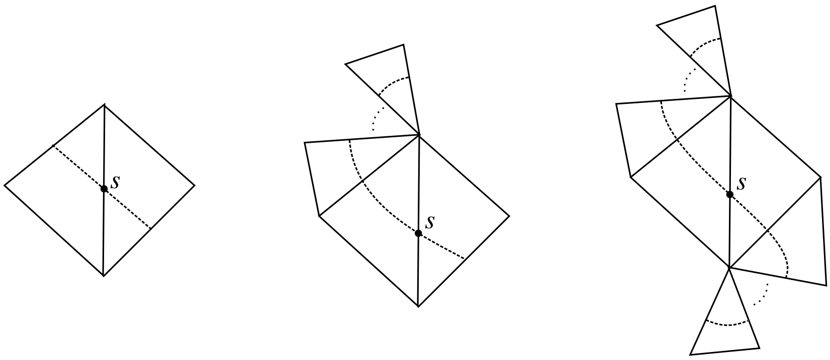
Now that we have determined the conditions on T which give rise to a turnover in $\mathbb{O}_T$ when $\Pi_{\mathcal{T}}$ intersects a single edge in the development of F between e_1 and e_2, we will show that it is impossible for there to be more than one such edge in the development of F between e_1 and e_2. This will complete the classification of immersed turnovers in tetrahedral orbifolds with no finite generalized vertices.

Figure 18 shows two possible schematic diagrams for this discussion. In each of the subfigures, the edges e_1 and e_2 are indicated, and the dotted line represents L_F. Notice that, in each triangle of the planar development of F, there is always a unique translate of a vertex of T that is separated from the other two by $\Pi_{\mathcal{T}}$. The edge translates of T labeled by s represent points at which this vertex switches.

We consider the following procedure for dividing any diagram of the type from Figure 18 into subdiagrams of the type (up to possible reflection or order two rotation) given in Figures 19 and 20:

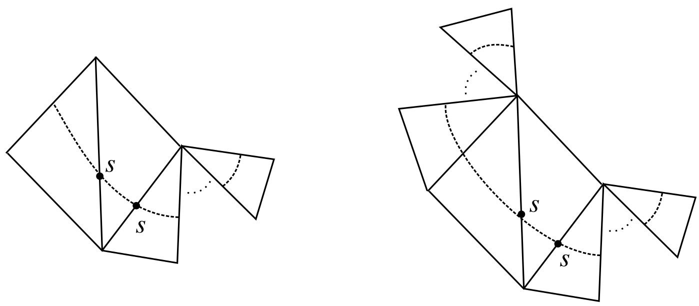
(1) Starting at the first edge of the diagram, we follow L_F until we arrive at the first switch. There must always be such a switch, for otherwise the supposed turnover would be parallel to a cover of an embedded turnover corresponding to one of the truncated vertices of T.

(2) If the switch is the only switch in the diagram, then our diagram looks like, up to reflection or rotation, one of the diagrams from Figure 19. In this case, we stop.



(a) (b) (c)

Figure 19. One type of possibility for the subdiagram components for a diagram of the type given in Figure 18.



(a) (b)

Figure 20. Another type of possibility for the subdiagram components for a diagram of the type given in Figure 18.

(3) If there is more than one switch and the diagram looks like, up to reflection or rotation, one of the diagrams from Figure 20, then we stop.

(4) If we have not halted in the previous two steps, then the diagram up to and including the first edge after the first switch looks like the diagram in either Figure 19(a) or 19(b). Call this portion a *subdiagram*.

(5) Starting at the last edge of the subdiagram from the previous step, we repeat this process with the remaining portion of the original diagram, starting from the first step, until we reach edge e_2.

This procedure divides our diagram into subdiagrams of the type illustrated in parts (a) and (b) of Figure 19(a), with the possible exception that the final subdiagram may be of the type in Figure 19(c) or one of the two types in Figure 20 (we note that this process can eliminate certain switches, in each of the resulting subdiagrams). Again, we denote by Π_1 and Π_2 the planes at e_1 and e_2, respectively, whose intersections with $\Pi_{\mathcal{T}}$ are supposed to form two of the sides of a triangle in the tiling of $\Pi_{\mathcal{T}}$. Our strategy is to use the subdiagrams of Figures 19 and 20 to find a sequence of planes in $\mathbb{H}^3$ — one or more planes at each of the two outermost edges of each subdiagram — that are pairwise disjoint on either side of Π_F and that therefore separate Π_1 from Π_2.

We first make two observations about the subdiagram from Figure 19(a). First, if either of the orders of the two edges separated by the switch is 2, then no plane at either edge can meet any of the planes at the other edge (excepting the plane Π_F). This fact follows from the extensive analysis done in Section 4.1. Second, if the two planes at the outer edges that are inclined closest toward the switch ("inclined closest" means closest, on the other side of Π_F, to the planes that pass through the switch edge) do meet (thus generating an immersed turnover in $\mathbb{O}_T$ with two singular points of order at least 6 and one singular point of order at least 3), then the next two planes (one at either outer edge) inclined away from the switch do not meet. This fact follows from an easy analysis of the patterns of line intersections in hyperbolic triangular tilings. See Figure 21 for the conditions on the vertex orders of an (a, b, c) hyperbolic triangular tiling under which such intersections can occur.

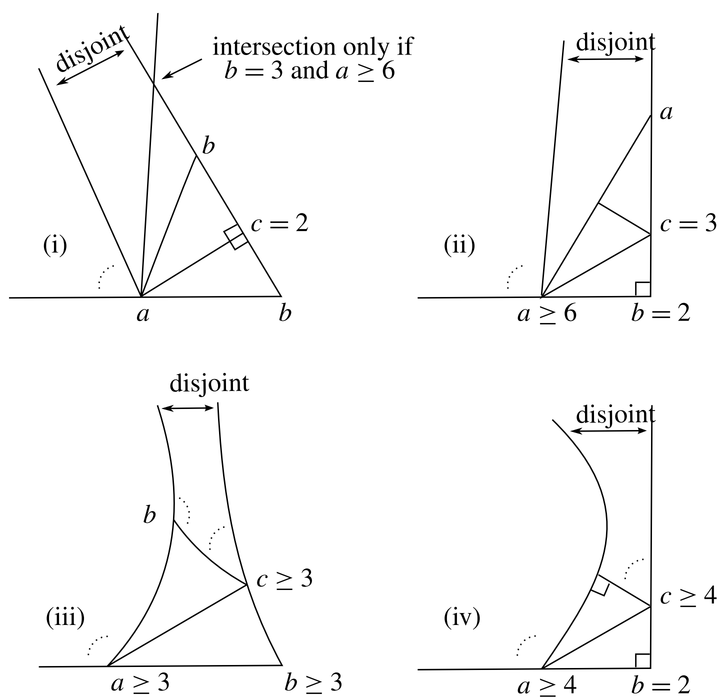


Figure 21. The possibilities for the intersection of lines in a triangular tiling of $\mathbb{E}^2$ or $\mathbb{H}^2$.

In this case (although this will not be the case for subsequent applications of this figure), Figure 21 should be thought of as depicting the plane which meets Π_F and the two southwest-to-northeast edges from Figure 19(a) perpendicularly, so that all the planes through these two edges appear as lines in Figure 21. In particular, in order for the next two planes inclined away from the switch in Figure 19(a) to meet, then one of the southwest-to-northeast edges must have order 2, which does not happen in this situation. Therefore, it is left to show that, for each of the remaining types of subdiagram, the two planes at the outer-most edges that are inclined closest to the single or double switch in the subdiagram do not intersect (again, "inclined closest" means closest, on the other side of Π_F, to the planes passing through the switch edge(s)). This will produce the sequence of planes that separates Π_1 and Π_2, and therefore complete the proof. We will show this by cases, which are indicated by their labels in the figures.

4.2.1. *Figure 19(b):* See Figure 22, in which we have supposed without loss of generality that F is the face ABC of the tetrahedron T, as in Figure 8. This picture only differs from Figure 19(b) by a 180° rotation. Observe that the edges incident at the vertices A and B have orders l, q, p and l, m, r (respectively).

We observe that the vertex B must have at least one order 2 edge incident to it. Otherwise, if B were of the type (x, y, z) with all orders at least 3, then it is readily seen, by using the information from Figure 21(iii) applied to vertex B, that Π_2 (the plane through e_2 that is inclined closest to the switch) cannot meet the plane at edge BC that is inclined closest to the switch. We indicate how this can be determined. Recall that we may construct the view from B as a triangular tiling of either the Euclidean or hyperbolic plane (in this case, a tiling by (x, y, z) triangles) such that Π_F appears as a horizontal line, and such that each edge incident to B appears as a point on that line and each plane through an edge incident to B appears as a line

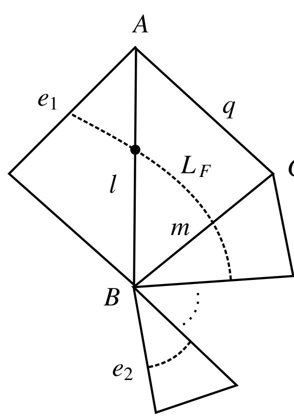


Figure 22. The case of Figure 19(b).

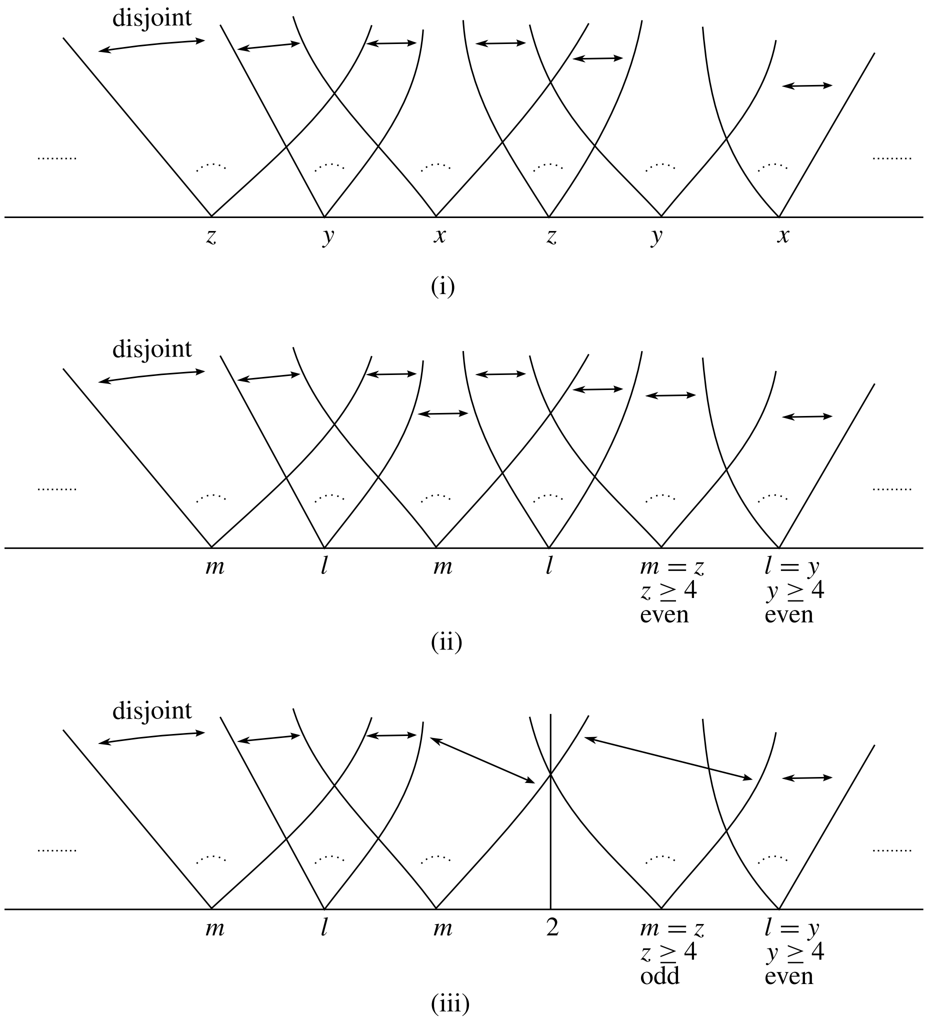


Figure 23. Patterns of intersections of certain lines corresponding to sides in a triangular tiling of $\mathbb{H}^2$ or $\mathbb{E}^2$. Double arrows indicate two lines that do not intersect above the horizontal line.

(or hyperbolic line, if B is superideal) passing through the corresponding point in the view from B. Using Figure 21(iii), we can conclude that the view from B, when B has no incident order 2 edge, looks schematically like Figure 23(i). This figure assumes that x, y and z are all odd; the other cases are similar. Suppose, for example, that the right-most point x in this figure represents the edge BC (x also indicates the order of that edge), and that the (schematic) line through this point inclined furthest to the right represents the plane through edge BC inclined closest to the switch edge AB. Then it is easily seen that no right-most inclined line through any subsequent point to the left along the horizontal can intersect with this line. Consequently, the planes to which these lines correspond cannot intersect on the other side of Π_F (i.e., the other side of the page in Figure 22). In particular, Π_2 cannot cross the plane through BC inclined closest to the switch, as we wished to show. Furthermore, by our analysis in the cases of Section 4.1, the only way that Π_1 can meet the plane through edge BC that is inclined closest to the switch is if B has an incident order 2 edge. Consequently, if there is no such order 2 edge at B, then we have $\Pi_1 \cap \Pi_2 = \varnothing$.

So B either has the type $(2, 3, x \geq 6)$ or $(2, y \geq 4, z \geq 4)$. In the latter case, if $l = y$ or $l = z$, then we have shown in Section 4.1.3 that Π_1 is disjoint from every plane through edge BC. If $l = y$ and $m = z$ and l and m are both even, it is a simple exercise, using Figure 21(i), to show that no plane that is inclined closest to the switch edge AB through any of the subsequent edges from BC toward e_2 along L_F can meet the plane through edge BC that is inclined closest to the switch, as in the argument of the previous paragraph (the schematic of the view from B in this case would be Figure 23(ii), with the edges AB and BC corresponding to the right-most points labeled l and m, respectively). So $\Pi_1 \cap \Pi_2 = \varnothing$ in this case. If $l = y$ and $m = z$ and m is odd, we can apply the same argument (but using the information from parts (i), (ii) and (iv) from Figure 21 to obtain the schematic view from B as depicted in Figure 23(iii)) to conclude that $\Pi_1 \cap \Pi_2 = \varnothing$. The analogous cases, where $l = y$ and $m = z$ and l and m are of mixed parity, are similar. The case when $l = y$ or $l = z$ and $m = 2$ requires more analysis. Here we use the geometry of the vertex A, the fact that $l \geq 4$ and the information from Figure 21 to conclude that Π_1 cannot intersect the plane through edge AB that is inclined closest to the edge BC. But Π_1 must intersect Π_F and it must intersect some of the planes through the switch edge AB. We refer to Figure 24, which depicts the schematic view from

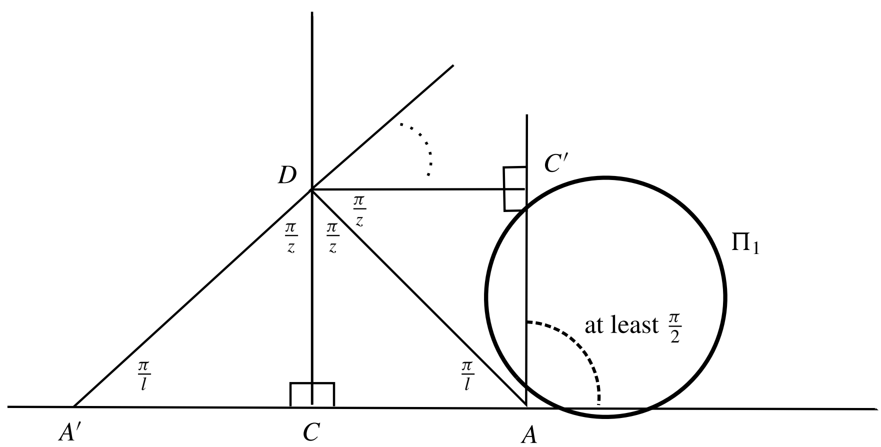


Figure 24. The schematic view from the vertex B in the case when $l = y \geq 4$, $m = 2$ and $r = z \geq 4$.

B in this case, with $l = y \geq 4$ and $m = 2$ (the third edge incident to B, which would have the label r in the tetrahedron T, is labeled by $z \geq 4$).

In this figure, the line segment AD corresponds to the plane through the switch edge AB of the tetrahedron that is inclined closest to the edge BC. As we have seen in previous cases, the ideal boundary of Π_1, in this view, is a circle that cannot contain any vertex of the triangulation in its interior disk. Since $z \geq 4$, we may conclude from the figure that the ideal boundary of Π_1 cannot intersect the line $A'D$. By noting that the line $A'D$ represents the plane inclined closest to the switch through the edge just after the edge BC along L_F toward e_2 in Figure 22, we may use the previous arguments from this paragraph to conclude that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case.

Referring to the first sentence of the previous paragraph, in the latter case and when $l = 2$ and $y = 4 = z$, we may show that $\Pi_1 \cap \Pi_2 = \varnothing$ by using the Euclidean vertex argument as in Figure 14. In the latter case and when $l = 2$ and one of y or z is greater than 4, it is again readily shown that the second closest plane to the switch through edge BC (recall that Π_1 must be disjoint from this plane, by the observation of the penultimate paragraph before the start of this subsection) misses the plane inclined closest to the switch at every subsequent edge that L_F crosses toward e_2. The argument uses the information of parts (i), (ii) and (iv) from Figure 21, and is similar to the arguments already presented in the previous two paragraphs. Thus, we have $\Pi_1 \cap \Pi_2 = \varnothing$ in the case that the type of vertex B is $(2, y \geq 4, z \geq 4)$.

This leaves us with the possibility that B has type $(2, 3, x \geq 6)$. When $l = x$, we are in a case that is similar to the first case in Section 4.1.3; that is, we have to consider a regular l-gon in either the Euclidean or hyperbolic plane and a circle

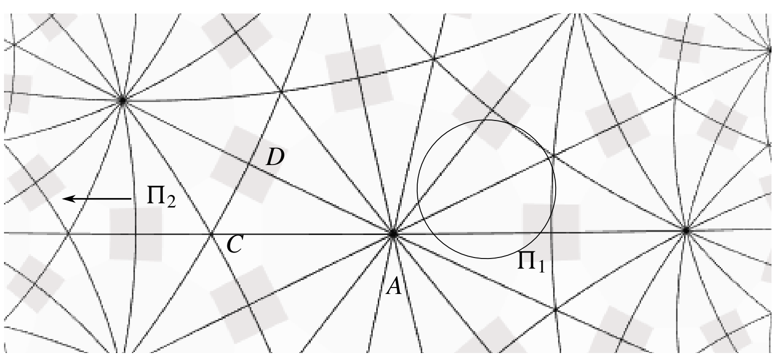


Figure 25. A view from the truncated vertex of hyperbolic type $(2, 3, 7)$. The arrow indicates that the plane Π_2 is represented by a circular arc that meets the horizontal somewhere to the left of the arc CD.

centered inside the polygon that does not contain in its interior the center of the polygon, any vertex of the polygon or any midpoint of a side. In this case, however, we observed that such a circle (representing Π_1) must be disjoint from all but two sides of the polygon. But the plane Π_2 will correspond in such a picture to a line or circular arc in the picture that does not meet the interior of this polygon, and so $\Pi_1 \cap \Pi_2 = \varnothing$ when $l = x$. See Figure 25 for an example illustration of this argument, in the case when $x = 7$.

The cases when $l = 2$ or $l = 3$ remain. In the case when $l = 3$, we refer to Figure 26. The upper half of this figure depicts the salient aspects of the view from vertex B, as in the previous cases we have considered. The lower half of the figure depicts part of the development of T in $\mathbb{H}^3$. In particular, in the lower half of the

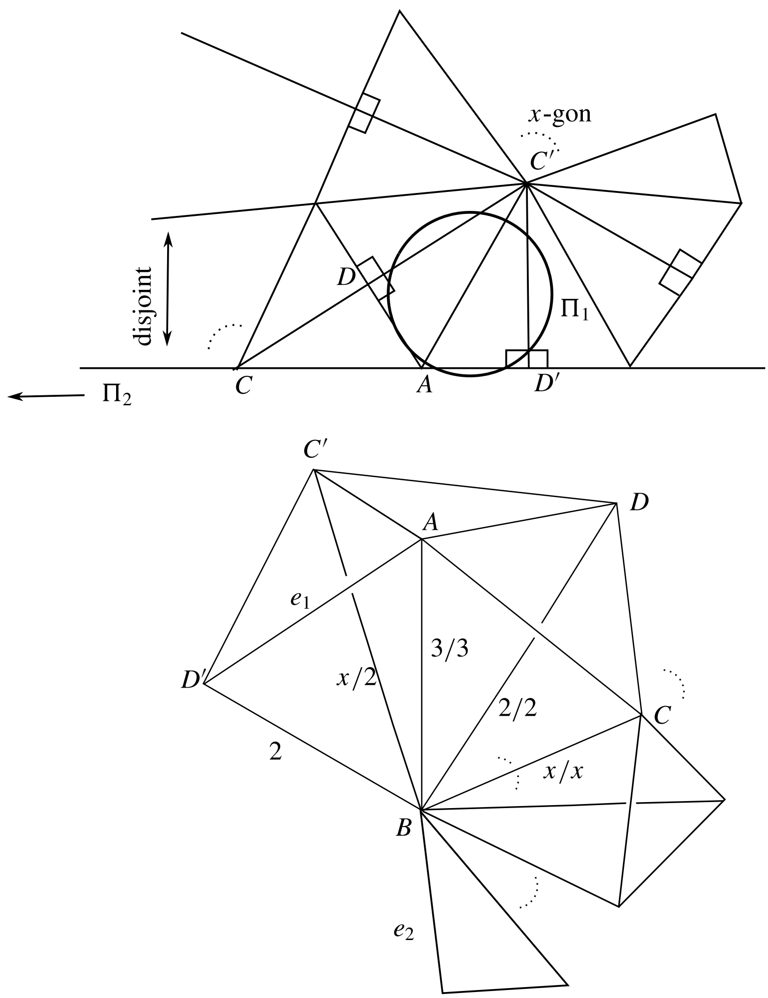


Figure 26. A view from the truncated vertex of type $(2, 3, x \geq 6)$.

figure, the triangle with edge e_2 and the lowest set of elliptical dots are both meant to lie in Π_F (which is the horizontal line CAD' in the upper half of the figure), and the plane Π_2 is not depicted, although $\Pi_1 = AC'D'$ is. In the upper half of the figure, Π_1 is represented by a circle centered at some point inside the triangle $AC'D'$ that cannot meet any vertex of the triangulation and that can only meet the sides AD and AD' of the x-gon centered at C' (the fact that this circle can meet no other sides of the x-gon centered at C' follows by an argument similar to that depicted in Figure 15 from Section 4.1.3). Since Π_2 must be represented by a line emanating from a vertex on the line CAD' which is further to the left than C (the direction, in the upper part of the figure, to which the line representing Π_2 must lie is indicated by the lower left arrow), and no such lines will enter the x-gon centered at C', we conclude that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case.

When $l = 2$, then the only way for which we are unable to apply the preceding argument is when $m = 3$. See Figure 27. This is because the angle $\angle A'DC'$ is less than $\pi/2$ when $x > 6$, and so it is, in principle, possible that the circle representing Π_1 (whose center must be contained in the triangle $AC'D$) may intersect the line

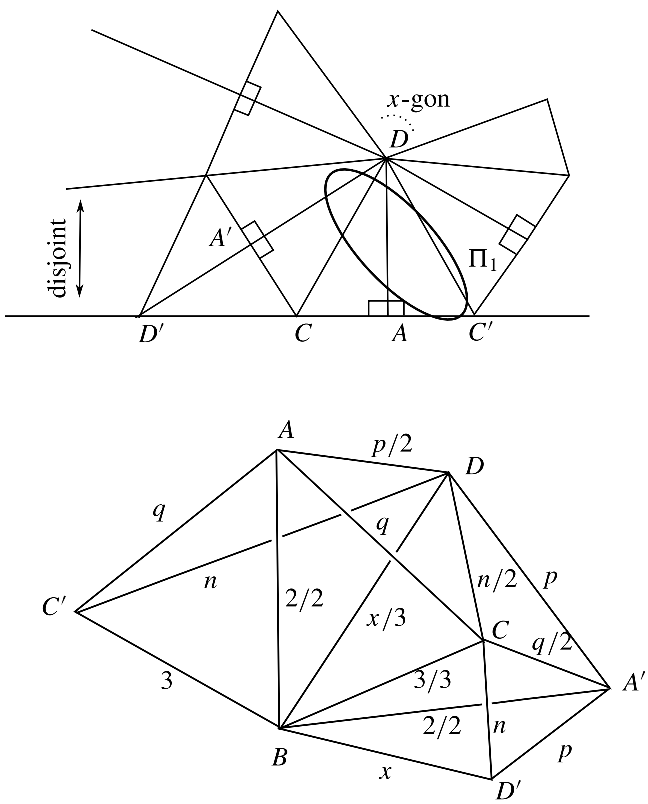


Figure 27. A view from the truncated vertex of type $(2, 3, x \geq 6)$.

representing Π_2 if $\Pi_1 = AC'D$ and $\Pi_2 = A'D'BD$ (we have drawn the circle as an ellipse in the view from B in order to indicate this possible intersection). However, using the accompanying tetrahedral illustration and the techniques of Section 4.1.1 (applied to vertex D), it is readily seen that we must have $p = 2$ and $n = 3$ in order for Π_1 and Π_2 to intersect. However, because we assume that T has no finite vertices and because $l = 2$, we do not allow $p = 2$. (*Note:* When $l = p = 2$ and $m = n = 3$ (so that the vertex A is finite), there is an immersed turnover of type (q, x, x) in T, provided that $q \geq 3$ and $x \geq 4$. See the conjectural classification at the end of this paper. In this case, $T = T[2, 3, q; 3, 2, x]$, which is isometric to the tetrahedron listed in item (6).)

4.2.2. *Figure 19(c):* See Figure 28, in which again we have supposed without loss of generality that F is the face ABC of the tetrahedron T, with the edges incident at the vertices A and B having orders l, q, p and l, m, r (respectively). We again denote by Π_1 and Π_2 the planes at the edges e_1 and e_2, respectively, that are inclined closest to the switch edge. The dotted curve in all of these figures, which we denote by L_F, represents the intersection of the planar development Π_F of F with the plane that (purportedly) contains the turnover determined by Π_F, Π_1 and Π_2.

Remark. The symbol * attached to a letter in this figure and in all subsequent figures is meant to indicate an ambiguity that may arise due to parity, and it is important for us to take note of it. For example, in Figure 28, if the order of the edge AB is even, then the vertex C^* is a developed copy of the vertex C, and the order of the edge AC^* is also q, i.e., the order of edge AC. However, if l is odd, then it would take an odd number l of tetrahedra developed around the edge AB to continue the development of the face ABC, making C^* a developed copy of the vertex D (recall that, behind the page, relative to the reader, lies the fourth vertex D of the tetrahedron), and making the order of the edge AC^* equal to p, i.e., the

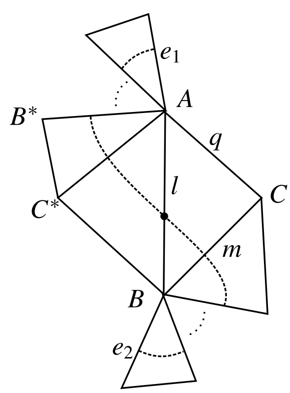
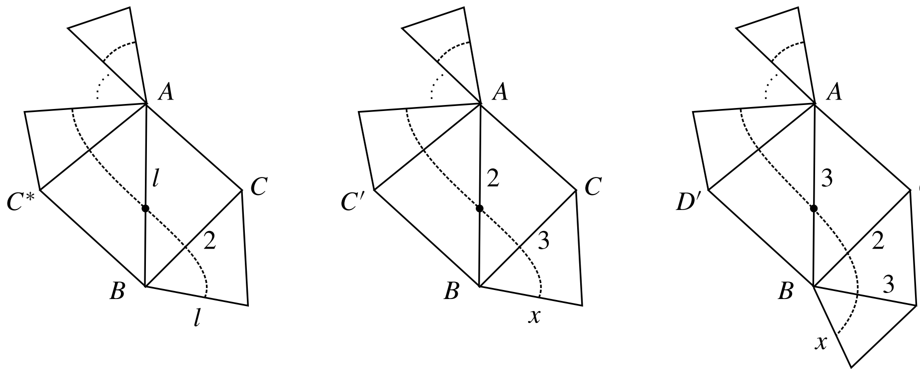


Figure 28. The case of Figure 19(c).



(i) $l \geq 3,\ m = 2$ (ii) $l = 2,\ m = 3,\ x \geq 6$ (iii) $l = 3,\ m = 2,\ x \geq 6$

Figure 29. After analysis, the remaining cases of Figure 19(c).

order of the edge AD (recall the notation $T[l, m, q; n, p, r]$ defined in Figure 8). We will avoid this notation whenever it is possible, although it will be necessary at times.

By the previous case, we know that Π_1 meets none of the planes through edge BC. It is therefore necessary, if Π_1 and Π_2 are to intersect, that Π_2 cross every plane through edge BC. As in the previous case, then, we can conclude that one of the edges incident at B must have order 2, for otherwise it is not possible for Π_2 to cross the plane through BC inclined closest to the switch.

Using Figure 21 and the fact that B must have an incident order 2 edge, we can reduce the cases that must be considered to those listed in Figure 29, as follows.

Referring to Figure 28, suppose first that $l = 2$ and $m = 3$. Recall that the dotted curve represents the line L_F. Then the next edge incident to B that L_F crosses after BC in the direction away from the switch should have order $x \geq 6$. A schematic of the view from B is pictured in Figure 30(i). The bold line in the figure represents *any* plane through a subsequent edge incident to B that L_F crosses after the edge with order x. Because the angle α, which is formed by the bold line and the line AC, will always be at least π/x, we conclude that the two lines indicated in the figure by the endpoints of the double arrow will not intersect above the line AC. Consequently, because the line AC represents the plane Π_F, we conclude that the planes represented by these lines will not intersect on the other side of Π_F (recall that the other side of Π_F refers to the side underneath the page in Figure 28). Therefore, we have reduced the case of showing that $\Pi_1 \cap \Pi_2 = \varnothing$ in Figure 28 to the case of Figure 29(ii), provided that $l = 2$ and $m = 3$. The case when $l = 2$ and m is even with $m \geq 4$ can be eliminated in an entirely similar fashion. See Figure 30(ii), which shows the pattern of intersections of lines that

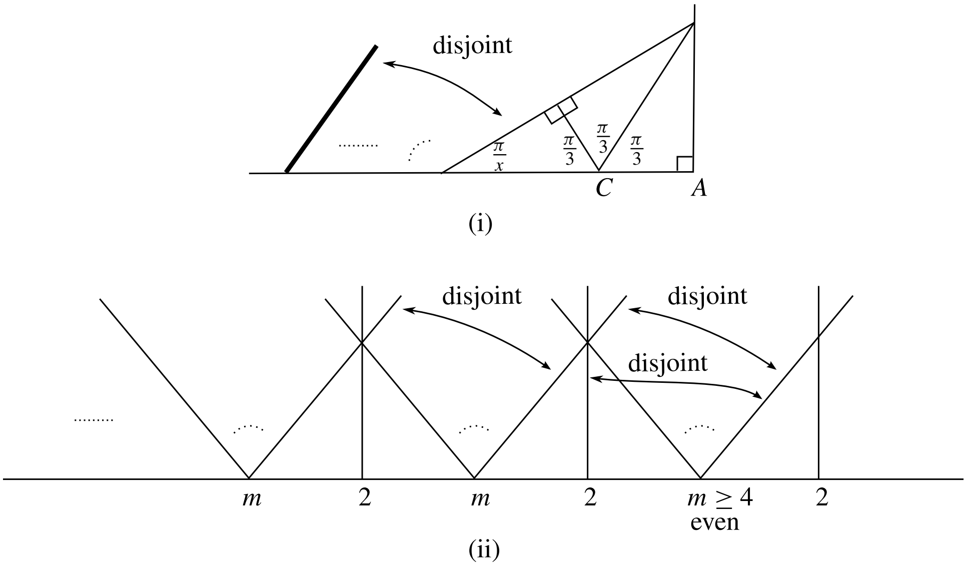


Figure 30. Patterns of intersections of certain lines corresponding to sides in a triangular tiling of $\mathbb{H}^2$ or $\mathbb{E}^2$. Double arrows indicate two lines that do not intersect above the horizontal line.

would result in the view from B. Here, we consider the right-most point on the horizontal (the horizontal represents Π_F in the view from B) with the label 2 as corresponding to the edge AB, and the right-most point on the horizontal with the label m as corresponding to the edge BC. It is readily seen from the figure that no lines passing through the labeled points on the horizontal to the left of the right-most point labeled m ever intersect the line through the latter point that is inclined closest to the switch point (i.e., the right-most point labeled 2). Therefore, no plane through an edge incident to B that is crossed by L_F after the edge BC can intersect the plane through BC inclined closest to the switch, when $l = 2$ and m is even and at least 4. Therefore, no plane through an edge incident to B that is crossed by L_F after the edge BC (such as Π_2) can intersect the plane Π_3 through BC inclined closest to the switch, when $l = 2$ and m is even and at least 4. Since Π_1 will also be disjoint from Π_3 (by Section 4.2.1), Π_1 will be separated from Π_2 by Π_3, which eliminates this case. In fact, all of the other reductions are arrived at in this way, that is, by using the information in Figure 21. The other cases that are *eliminated* by the methods of this paragraph are: (1) $l = 2$ and $m \geq 5$ with m odd, (2) $l = 3$ and $m \geq 6$ and (3) $l \geq 6$ and $m = 3$. The other cases that are *reduced* by the methods of this paragraph are: (4) $l \geq 3$ and $m = 2$ (which reduces to the case of Figure 29(i)) and (5) $l = 3$ and $m = 2$ (which reduces to the case of Figure 29(iii)). (We note that, when $l = 3$ and $m = 2$, case (i) of Figure 29 may seem to rule out case (iii).

However, the plane inclined closest to the switch through the edge labeled x in case (iii) intersects the plane inclined closest to the switch through the lower edge labeled 3 (this may be seen using the information of Figure 21). We therefore must show that $\Pi_1 \cap \Pi_2 = \varnothing$ in *both* the case that e_2 is the lower edge labeled $l = 3$ in (i) *and* in the case that e_2 is the lower edge labeled x in (iii).)

Now, we apply the arguments of the previous two paragraphs to the other direction along L_F from the switch. Specifically, referring to Figure 28, we know by the previous case that Π_2 meets none of the planes through the edge AC^*, and so we reduce the possibilities for the number of developed faces around the vertex A using the fact that Π_1 must intersect every plane through the edge AC^* in order for it to be possible for Π_1 and Π_2 to have nonempty intersection. The result of this further analysis leaves us to consider only the cases of Figure 31. We note the change from "$l \geq 3$" to "$l \geq 3$ odd" that occurs when reducing Figure 29(i) to

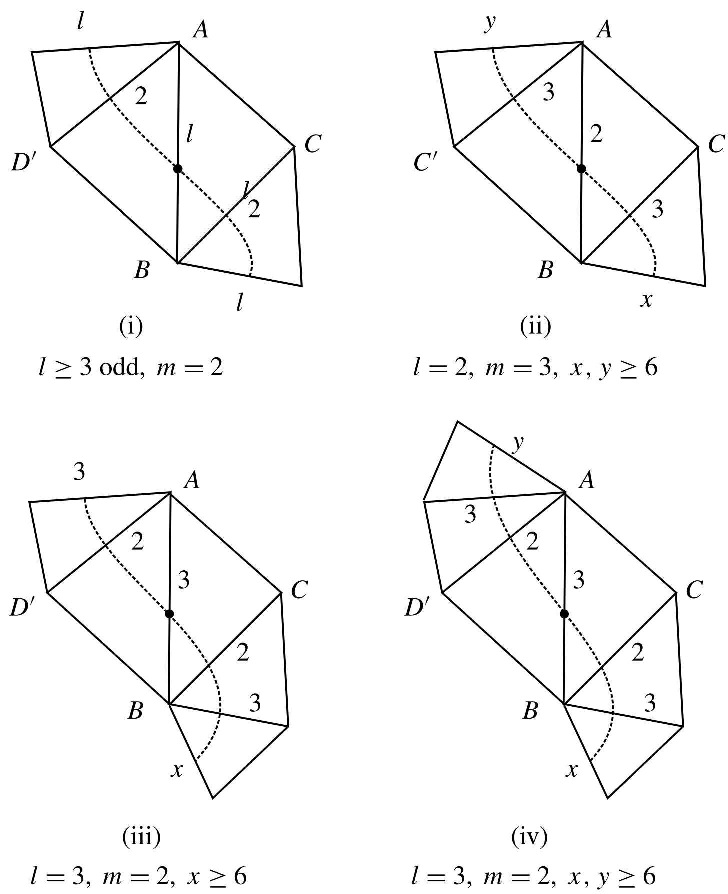


Figure 31. After further analysis, applied to the cases of Figure 29, these are the remaining cases of Figure 19(c) to consider.

Figure 31(i). This change is due to the fact that, when l is even, the edge label 2 for AD' in 31(i) must equal the edge label for AC. However, this would contradict our assumption that none of the vertices of T is finite, because C would have two incident edges, AC and BC, labeled 2.

So we are left to analyze the cases of Figure 31. We begin with case (iv). See Figure 32. The multiple parts of this figure are explained in the caption. Referring to the left side of the lower half of the figure, Π_1 is the plane through edge AC'' inclined closest to the switch edge AB and Π_2 is the plane through edge BD''

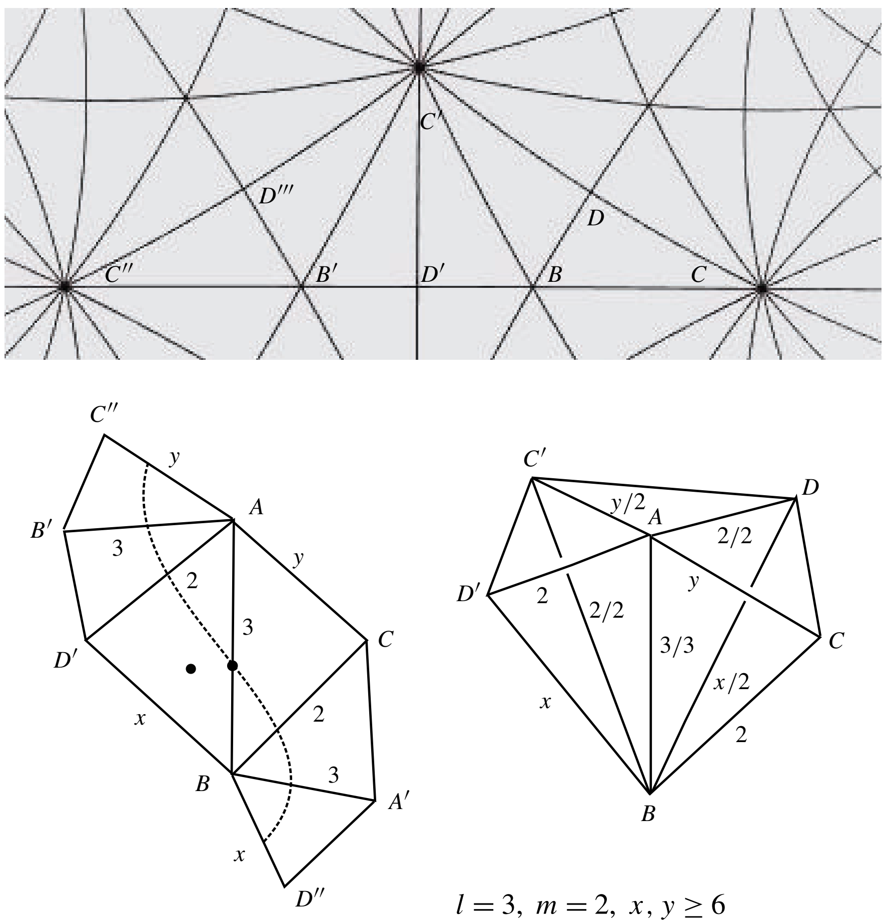


Figure 32. The case of Figure 31(iv). The upper half of the figure represents the view from the vertex A when $y = 7$. The lower half consists of a perspective image of the three copies of the tetrahedron $ABCD$ on the right, and several triangles in the development of the face ABC on the left.

inclined closest to the switch edge AB. We wish to show that $\Pi_1 \cap \Pi_2 = \varnothing$. We do so using the upper half of the figure, which shows the view from A under the assumption that $y = 7$ (the same argument we give here applies to any other value for $y \geq 6$). In the upper half of the figure, the plane Π_1 is represented by the line $C''D'''$, and the plane ACD — which is depicted in the right side of the lower half of the figure, and which is the plane through AC inclined closest to the switch edge AB in the left side of the lower half of the figure — is represented by the line CD. Recalling that Π_F is the plane containing the face ABC (and, therefore, the plane in which the left side of the lower half of the figure is drawn, as well as the horizontal line in the upper half of the figure), we observe that there are two planes, other than Π_F, that pass through AB. These planes are represented in the upper half of the figure by the lines BC' and BD. Using the upper half of the figure, we observe that any point of Π_1 that is on the same side of ACD as the vertex B is also on the same side of the plane ABC' (which is represented by the line BC') as the point D'. We now use the previous case (Section 4.2.1) to observe that $\Pi_2 \cap ACD = \varnothing$: namely, Π_2 and ACD are the planes through BD'' and AC, respectively, inclined closest to the *new* switch edge BC for the three triangles ABC, $A'BC$ and $A'BD''$ from the lower left half of Figure 32, to which Section 4.2.1 applies (to see this more clearly, turn these three triangles together so that the edge BC is vertical, and compare with Figure 22). In exactly the same way (i.e., using Section 4.2.1), we see that $\Pi_2 \cap AC'D' = \varnothing$, this time using AB as the switch edge. But now, since Π_2 is on the same side of ACD as the vertex B *and* on the same side of $AC'D'$ as the vertex B, we can use the upper half of Figure 32 to see that there is no part of Π_1 which is both on the B side of $AC'D'$ *and* on the B side of ACD. Therefore, $\Pi_1 \cap \Pi_2 = \varnothing$.

The argument of the previous paragraph can be used in case (iii) of Figure 31. See Figure 33. In the lower left half of this figure, Π_1 is the plane through the edge AB' inclined closest to the switch edge AB. In the lower right half, Π_1 is the plane $AC'B'A''$. In the upper half of the figure, which represents the view from A when $y = 7$ (the case when $y \geq 6$ is similar), Π_1 is represented as the line $B'C'$. Proceeding as in the previous paragraph, we have $\Pi_2 \cap ACD = \Pi_2 \cap AC'D' = \varnothing$ (by Section 4.2.1). Furthermore, Π_2 is on the B side of both ACD and $AC'D'$. But now, referring to the upper half of Figure 33, we see that there is no part of Π_1 that is on the B side of both ACD and $AC'D'$. So $\Pi_1 \cap \Pi_2 = \varnothing$.

We now address case (ii) of Figure 31. See Figure 34. In the upper part of this figure, Π_1 and Π_2 are the planes through the edges AD' and BD'', respectively, that are inclined closest to the switch edge AB. In the lower part of the figure, which depicts the development of multiple copies of the tetrahedron, Π_1 is the plane $ADB'D'$ and Π_2 is the plane $BDA'D''$. Because these two planes are both incident to the nonfinite vertex D, they intersect if and only if their intersections

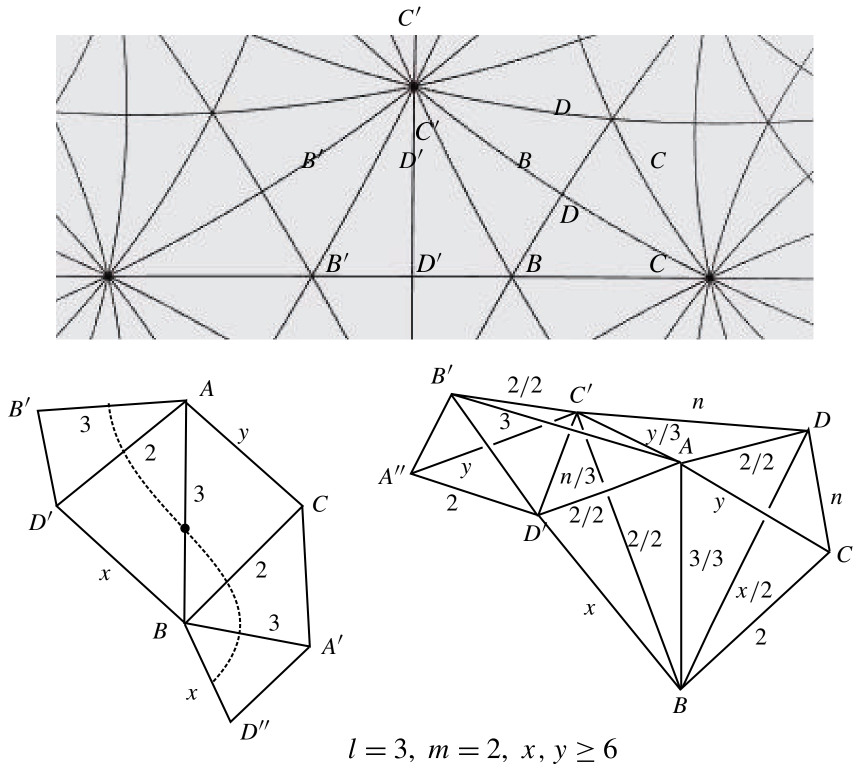


Figure 33. The case of Figure 31(iii). The upper half of the figure represents the view from vertex A when $y = 7$. The right side of the lower half of the figure depicts the development of several copies of the tetrahedron, and the left side of the lower half depicts the development of the face ABC.

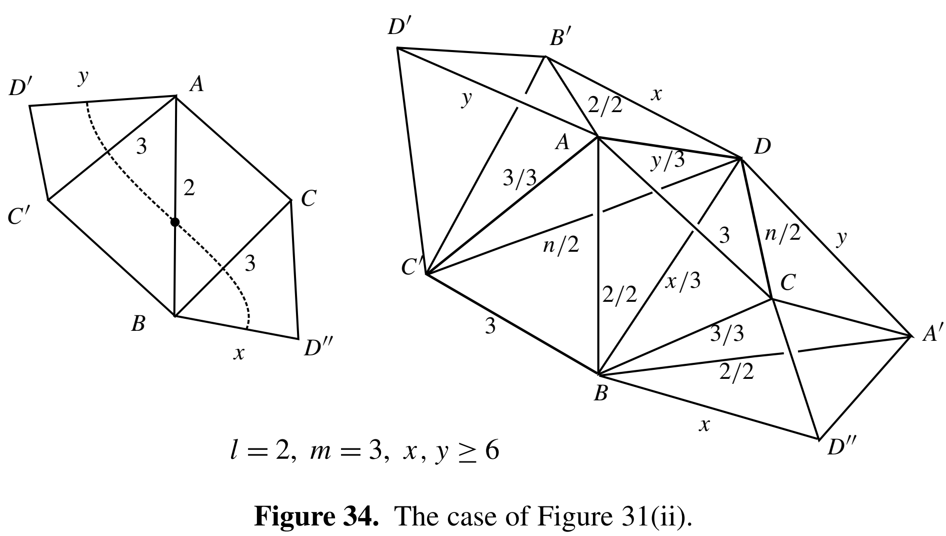
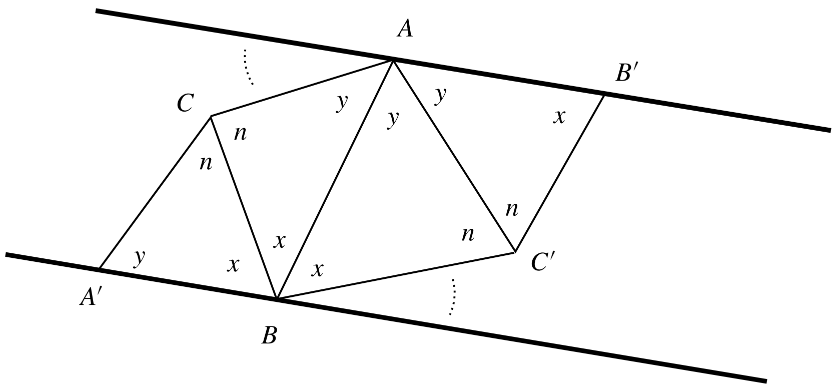


Figure 34. The case of Figure 31(ii).



$n \geq 3,\ x, y \geq 6$

Figure 35. The schematic view from the vertex D for Figure 34. The letters x, y and n represent the integral submultiples of π of the dihedral angles of the tetrahedra incident at D.

with the link of D intersect. See Figure 35, which schematically depicts the view of this link from the vertex D.

In the figure, Π_1 is represented by the bold line AB' and Π_2 by the bold line $A'B$. We have labeled the interior angles of the triangles in this view by their submultiples of π. Because vertex C has two edges of order 3 incident to it, we must have that $n \geq 3$. But since x and y must both be at least 6, we can use Figure 21(iii) (with base the segment AB) to conclude that the bold lines cannot intersect on either side of the line AB. So $\Pi_1 \cap \Pi_2 = \varnothing$ in the case of Figure 31(ii).

This leaves case (i) of Figure 31. We begin by assuming that $l \geq 5$ (recall that l must be odd). See Figure 36, which depicts the view from the vertex A with the projection centered at the vertex B, and Figure 37. For the purposes of illustration, we take the type of A to be (2, 4, 5), although the argument only depends on the

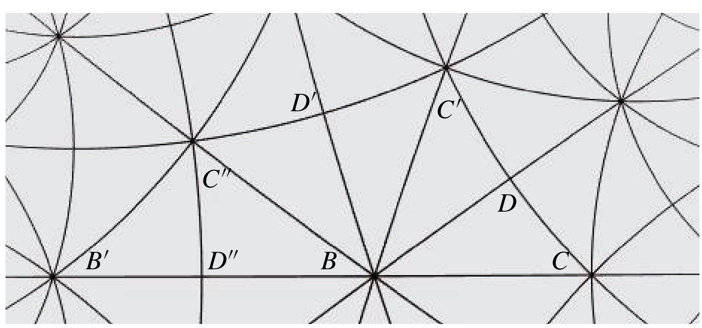


Figure 36. The case of Figure 31(i): view from the vertex A, in the case when A has type (2, 4, 5).

D

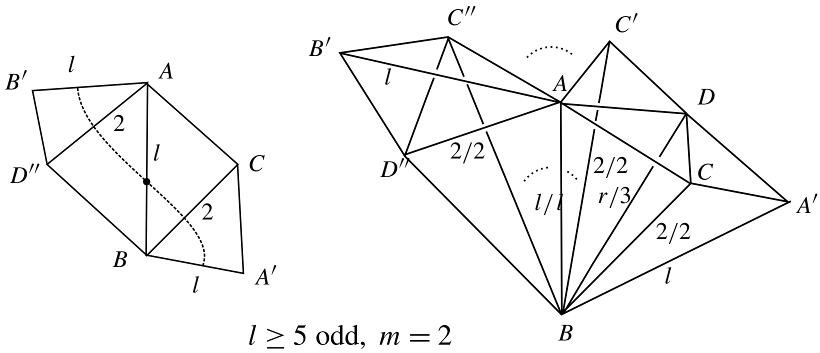
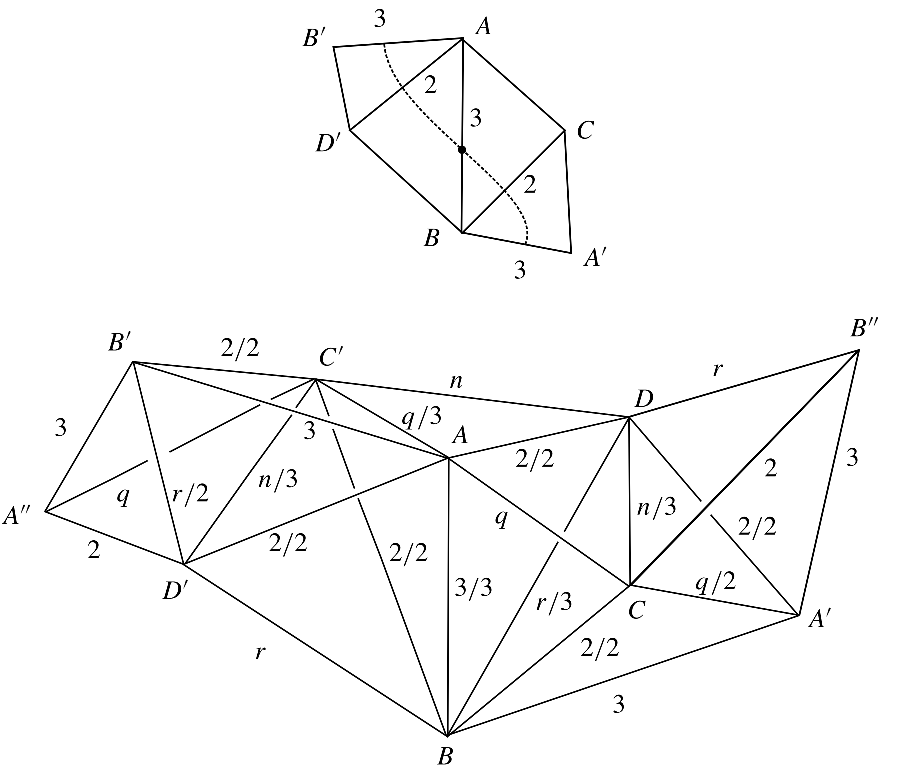


Figure 37. The case of Figure 31(i) (continued).

presence of the order 2 edge incident to A and the fact that the order of the edge AB is at least 5. The plane Π_1 is represented by the circular arc $B'C''$. We note that the plane $BC'D$ in the lower right part of the figure is represented in the upper half of the figure by a circle, centered on the line segment BC' because the planes $BC'D$ and ABC' are perpendicular, and whose interior disk does not contain any of the points B, D, C' or D'. As we have observed previously (see the argument depicted in Figure 15 from Section 4.1.3), the circle representing $BC'D$ can intersect at most two sides of the l-gon centered at B, and in this case those sides will always be $D'C'$ and DC'. It is clear that this circle is disjoint from the arc $B'C''$ representing Π_1, and hence that $\Pi_1 \cap BC'D = \varnothing$. Now referring to the lower right part of the figure, we observe that $\Pi_2 = A'BD$ (as planes) and that the part of Π_2 that is on the same side of BCD as A is also on the *opposite* side of $BC'D$ as A. Since Π_1 is disjoint both from $BC'D$ and BCD (the latter by the previous case of Section 4.2.1), and because Π_1 lies on the same side of these planes as A, we can conclude that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case when $l \geq 5$.

So we now assume that $l = 3$ in this case. We are not able to use the argument of the previous paragraph because some of the intersections ruled out in the previous paragraph can occur in this case. We refer to Figure 38. The possible values for q, n and r in the figure are based on the fact that the tetrahedron has no finite vertices. In this figure, $\Pi_1 = AC'A''B'$ and $\Pi_2 = A'B''DB$ (as planes). We determine that these planes are disjoint by applying the techniques of Sections 4.1.1 and 4.1.2. In particular, if $n \geq 4$, then we use the geometry of the link of vertex C' to conclude that Π_1 is disjoint from the plane $BDC'D'$ (it lies to the same side of $BDC'D'$ as the vertex A) and the geometry of the link of vertex D to conclude that Π_2 is disjoint from the plane $ACDC'$ (it lies to the same side of $ACDC'$ as vertex B). Now by considering the plane ABD and the geometry of the vertex B, we have that the part of Π_2 that is on the C' side of ABD is always on the opposite side of $BDC'D'$ to



$l = 3,\ m = 2,\ q, r \geq 6,\ n \geq 3$

Figure 38. The case of Figure 31(i) when $l = 3$.

Π_1. Similarly, we have that the part of Π_1 on the D side of ABC' is always on the opposite side of $ACDC'$ to Π_2. We conclude that $\Pi_1 \cap \Pi_2 = \varnothing$. When $n = 3$, the argument is similar, except that $ACDC' = ACB''DC'$ and $BDC'D' = BDC'A''D'$ (as planes), and Π_1 and Π_2 will form interior angles on the B side of $ACB''DC'$ of $3\pi/q \leq \pi/2$ and $\pi/r \leq \pi/6$, respectively (so that Π_1 and Π_2 cannot intersect on the side of this plane opposite to B), and interior angles on the A side of $BDC'A''D'$ of $\pi/q \leq \pi/6$ and $3\pi/r \leq \pi/2$, respectively (so that Π_1 and Π_2 cannot intersection on the side of this plane opposite to A). Again, we conclude that $\Pi_1 \cap \Pi_2 = \varnothing$. This completes case (i) of Figure 31, and concludes this subsection.

4.2.3. *Figure 20(a):* See Figure 39, and recall the significance of the symbol * from the remark on page 224. We must first address the case when $e_2 = A^*C$. There are two possibilities that we must consider in determining whether or not Π_1 and Π_2 can intersect: either (1) Π_1 meets the plane through BC that is closest in inclination to the switch edge AB (it cannot meet more planes through BC, by our previous observations) and Π_2 meets at least the second closest plane through

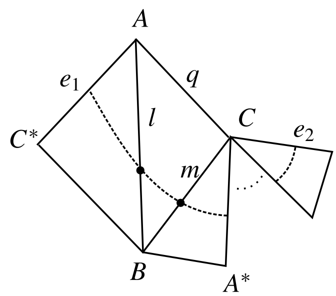


Figure 39. One case of Figure 20(a).

BC to the switch edge AB, or (2) Π_1 meets no planes passing through BC and Π_2 meets all of the planes passing through BC. We handle these two cases below:

(1) In order for Π_1 to meet a plane passing through BC, our tetrahedron must take one of the forms of items (1)–(3) in the summary at the conclusion of the paper. This follows from the extensive analysis of Section 4.1 (in fact, the pairwise intersections of Π_1, Π_F and the plane through BC inclined closest to the switch edge AB determine an immersed turnover in this case). We consider the case when $l = 3$, corresponding to item (3) in the summary. If $l = 3$, then $q = 2$, $m \geq 6$ and n (the order of the third edge associated to vertex C) is at least 3. It is then an easy analysis, using Figure 21 applied to the vertex C, to see that there is no choice of n and m for which Π_2 can intersect either of the two closest planes through BC toward the edge AB. So $\Pi_1 \cap \Pi_2 = \varnothing$. Exactly the same analysis holds if our tetrahedron takes the form of item (1) of the summary at the conclusion of the paper (in this case we have $l = 2$, $m \geq 6$, $n = 2$ and $q \geq 3$, and so the order of edge A^*C is either 2 or q, and there is no choice for m and q such that Π_2 meets either of the two planes through BC inclined closest to the switch). If our tetrahedron has the form of item (2) from the summary, then $l = 2$, $m \geq 3$ and $q \geq 6$. If m is odd and at least 5, then the order of edge A^*C is 2 and we can use Figure 21 applied to vertex C to conclude that Π_2 does not meet the two planes through BC inclined closest to the switch. If m is even and at least 6, then the order of A^*C is $q \geq 6$, and the conclusion of the previous sentence also holds. If $m = 4$, then we refer to Figure 40. Only the relevant edges are labeled in this figure, in which $\Pi_1 = AC^*D$ and $\Pi_2 = A^*D'A'C$. Because $q \geq 6$, we have that Π_1 and Π_2 form interior angles on the side of $ACA'D$ opposite to vertex B of $\pi/3$ and $(q-2)\pi/q \geq 2\pi/3$, respectively (these are interior angles with respect to the edge CD). Therefore, Π_1 and Π_2 do not intersect on the side of this plane opposite to B. But, as we have observed, $\Pi_1 \cap A'BC = \varnothing$. Since the part of Π_2 that is on the B side of $ACA'D$ is always on the opposite side

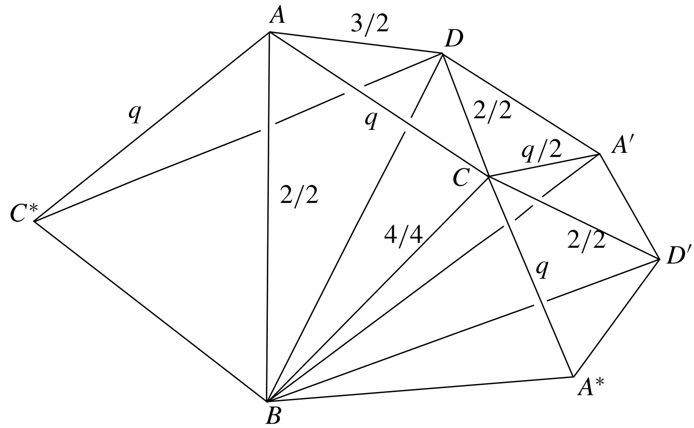


Figure 40. The case of Figure 39, when $l = 2$, $m = 4$ and $e_2 = A^*C$.

of $A'BC$ to Π_1, we have $\Pi_1 \cap \Pi_2 = \varnothing$. The case when $m = 3$ is exactly the same. These are all the possibilities for when the tetrahedron has one of the types (1)–(3) in the summary. So $\Pi_1 \cap \Pi_2 = \varnothing$ for this case.

(2) If the order of edge BC is greater than 4, then it is not possible to choose integers for the type of vertex C so that Π_2 crosses all the planes through BC. This follows by using the information of Figure 21 applied to the vertex C, as in the arguments that accompany Figure 23 in Section 4.2.1. The same statement is true (with the same argument) if the order of BC is 3 and the vertex C has no incident order 2 edge. So the order of edge BC is either 3 and C has the type $(2, 3, x \geq 6)$ or the order of edge BC is 2. Suppose that the edge BC has order 3. Then we can use the same argument as the one given at the end of the previous paragraph. Namely, it is readily shown that Π_1 and Π_2 meet the plane containing the face ACD at interior angles that sum to at least π on the opposite side of ACD of the vertex B, and since they do not meet on the B side of this plane, they must be disjoint. The same argument also works when the order of BC is 2. So $\Pi_1 \cap \Pi_2 = \varnothing$ in this case.

So we assume $e_2 \neq A^*C$. We observe that removing the sides AC^* and BC^* from the Figure 39 leaves a picture that is equivalent to the previous case of Section 4.2.1. We therefore know that Π_2 misses every plane through the switch edge AB. It follows, using Figure 21 applied to the vertex A, that l must be either 2 or 3, in order for Π_1 to cross every plane through this switch edge. Moreover, we must have, as in previous cases, that the type of vertices A and C must include an order 2 point. Suppose $l = 2$. This implies that neither m nor q is 2. If, in addition, neither m nor q is 3, then it is straightforward using the information in Figure 21 (applied to vertex C) to show that Π_2 cannot meet the plane through

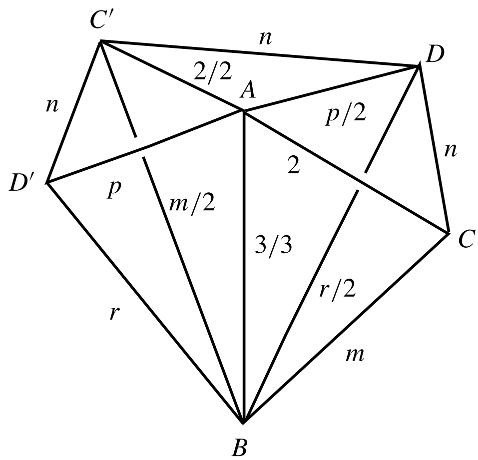


Figure 41. The case of Figure 39, when $l = 3$, $q = 2$ and $p \geq 6$.

A^*C inclined closest to Π_1, and so prove that $\Pi_1 \cap \Pi_2 = \varnothing$ in this case. So either $m = 3$ and $q \geq 6$ or $q = 3$ and $m \geq 6$, and in both cases $n = 2$. In either case, it is a straightforward application of the techniques already employed — specifically, the techniques involving developing tetrahedra from Sections 4.1.1 and 4.1.2 — to show that Π_1 and Π_2 do not intersect.

Now suppose $l = 3$. This implies that the order of edge e_1 is p. Because Π_1 must cross every plane through the switch edge AB, it is easily shown using Figure 21 (applied to vertex A) that the order p of edge e_1 is at least 6 and $q = 2$. Figure 41 shows three copies of T, with C^* relabeled as D'. Because $q = 2$, we must have $n \geq 3$ and $m \geq 3$. Since $n \neq 2$, analysis using the vertex D shows that Π_1, which is the plane $ADC'D'$, intersects the plane BCD if and only if $r = 2$. We analyze two cases:

Case $r \neq 2$: In this case, Π_1 does not intersect BCD, and so it is necessary for Π_2 to intersect BCD if Π_1 and Π_2 are to intersect. If $m \geq 4$ and even, then the edges emanating from the vertex C in Figure 39 — CA, CB, CA^*,..., e_2 — have labels that alternate $2, m, 2, \dots$. However, by using Figure 30(ii) applied to the vertex C, it is easily seen that no plane through any of the edges CA^*,..., e_2 that is inclined closest to the switch edge AB will intersect the plane BCD. Since Π_1 does not intersect BCD, the latter plane separates Π_1 from Π_2. So we are left to consider when $m \geq 3$ and odd. When $m \geq 5$ and odd, an application of the information from Figure 21 to the vertex C shows that no plane that is inclined closest to the switch edge AB through any of the edges from CA^* to e_2 can intersect with the plane BCD. So again, $\Pi_1 \cap \Pi_2 = \varnothing$. Finally, when $m = 3$, it is necessary for n (the label of the third edge of T that meets the vertex C, and the label of the edge

CA^*) to be at least 6. So the type of the vertex C is $(2, 3, n \geq 6)$, and no plane through any edge after CA^* and up to and including e_2 that is inclined closest to the switch edge AB will intersect the closest such inclined plane through the edge CA^* (as in Figure 30(i)). Since, by the observation of the first paragraph of this section, the closest inclined plane to the switch edge AB through CA^* is disjoint from Π_1, we again have $\Pi_1 \cap \Pi_2 = \varnothing$. This completes the analysis of the case when $r \neq 2$.

Case $r = 2$: In this case, Π_1 does intersect the plane BCD. Because $r = 2$ and $l = 3$, it is necessary that $m \geq 6$. We have previously observed that Π_1 cannot intersect with the second-closest inclined plane to the switch edge AB through BC (because the planes Π_1, ABC and BCD form pairwise angles of intersection π/p, π/m and π/n, with $p \geq 6$, $m \geq 6$ and $n \geq 3$). However, vertex C has type $(2, m \geq 6, n \geq 3)$, and it is easily seen using the information of Figure 21 applied to C that no plane that is inclined closest to the switch edge AB through any of the edges from CA^* to e_2 can intersect the second-closest inclined plane to AB through CB, *provided* that $m \geq 7$. So this second-closest inclined plane through CB separates Π_1 from Π_2, when $m \geq 7$. This leaves the case when $m = 6$. But this case is handled by an argument similar to the accompanying argument for Figure 26 in Section 4.2.1. This completes the case when $r = 2$, and concludes this subsection.

4.2.4. *Figure 20(b):* See Figure 42. By the result of Section 4.2.1, it is not possible for e_2 to equal CA^*. Because of this, it is not possible, also by the Section 4.2.1, for Π_2 to meet any of the planes through the edge AB. Nor is it possible, by Section 4.2.1, for Π_1 to meet any of the planes through the edge BC. Consequently, the intersection of Π_1 and Π_2 can only occur if Π_1 crosses every plane through AB and Π_2 crosses every plane through BC. The subsequent possibilities and arguments to rule them out are all straightforward to carry out, using the techniques we have employed to this point. This completes the proof of Theorem 1.2. □

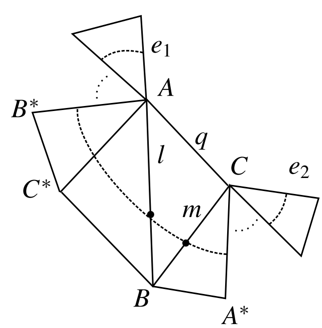


Figure 42. One case of Figure 20(b).

Summary

We provide a summary of the classification of immersed turnovers in the orbifold $\mathbb{O}_T$ associated to the generalized tetrahedron $T[l, m, q; n, p, r]$. These are listed in the order in which they appear in the proof, but isometric cases are indicated (the 24 isometric cases are determined by applying an element of the symmetric group S_4: any element of the symmetric group S_3 may be applied to both the first and second triples of $T[l, m, q; n, p, r]$, and any pair from one triple may be swapped with the corresponding pair of the other triple). We also include a conjectural list of all the immersed turnovers in hyperbolic tetrahedral orbifolds. All of these can be confirmed using the techniques of this paper, and while the author believes this list to be exhaustive, the necessary computations to determine the complete classification are somewhat extensive.

(1) $T[2, m, q; 2, p, 3]$. $\mathbb{O}_T$ contains an immersed (q, m, p) turnover, where $q \geq 3$, $m \geq 6$ and $p \geq 6$.

(2) $T[2, m, q; 2, 3, r]$ (isometric to item (1)). $\mathbb{O}_T$ contains an immersed (q, m, r) turnover, where $q \geq 6$, $m \geq 3$ and $r \geq 6$.

(3) $T[3, m, 2; n, p, 2]$ (isometric to item (1)). $\mathbb{O}_T$ contains an immersed (m, n, p) turnover, where $m \geq 6$, $n \geq 3$ and $p \geq 6$.

Conjectural list of all immersed turnovers in hyperbolic tetrahedral orbifolds:

(4) $T[2, m, q; 2, p, 3]$. $\mathbb{O}_T$ contains an immersed (q, m, p) turnover for any of the following values:
   - (a) $q = 2$, $m = 4$ and $p \geq 5$. In this case, $\mathbb{O}_T$ also contains
     - (i) a $(2, p, p)$ turnover,
     - (ii) a $(4, 4, 5)$ turnover if $p = 5$, and
     - (iii) a $(p/2, p, p)$ turnover if p is even.
   - (b) $q = 2$, $p = 4$ and $m \geq 5$ (isometric to item (4), with the same set of additional nonmaximal turnovers).
   - (c) $q = 2$, $m \geq 5$ and $p \geq 5$. In this case, $\mathbb{O}_T$ also contains
     - (i) a $(m, m, p/2)$ turnover if p is even, or
     - (ii) a $(m/2, p, p)$ turnover if m is even.
   - (d) q, m and p are all greater than 2, and at least one is greater than 3. In this case, if two of the values are 3, then $\mathbb{O}_T$ also contains a (x, x, x) turnover, where x is the integer that is greater than 3.

(5) $T[3, 2, 2; 2, p, 3]$. $\mathbb{O}_T$ contains an immersed $(2, p, p)$ turnover, where $p \geq 5$.

(6) $T[3, m, 2; 2, p, 3]$. $\mathbb{O}_T$ contains an immersed (m, p, p) turnover, where $m \geq 3$ and $p \geq 4$.

(7) $T[3, m, 3; 2, 3, 2]$. $\mathcal{O}_T$ contains an immersed $(3, m, m)$ turnover, where $m \geq 4$.

(8) $T[4, 3, q; 2, 2, 2]$. $\mathcal{O}_T$ contains an immersed $(q, q, 3)$ turnover, where $q \geq 4$.

(9) $T[2, 2, 4; n, 3, r]$. $\mathcal{O}_T$ contains an immersed turnover of type $(2, 4, r \geq 5)$ (as well as the additional nonmaximal turnovers listed in item (4)) if $n = 2$, an immersed turnover of type $(4, 4, r \geq 3)$ if $n = 3$, and immersed turnovers of types $(3, 3, 5)$, $(3, 5, 5)$ and $(5, 5, 5)$ if $n = 2$ and $r = 5$.

(10) $T[2, 3, q; 2, 3, r]$. $\mathcal{O}_T$ contains an immersed (q, r, r) turnover, where $q \geq 3$ and $r = 4$ or $r = 5$.

(11) $T[2, 2, q; 3, 5, 2]$. $\mathcal{O}_T$ contains an immersed $(q, q, 5)$ turnover, where $q \geq 3$.

(12) $T[2, 2, 5; 2, 3, 5]$. $\mathcal{O}_T$ contains an immersed $(3, 5, 5)$ turnover.

(13) $T[2, 2, 3; 3, p, 2]$. $\mathcal{O}_T$ contains immersed turnovers of type $(3, p, p)$ and (p, p, p), where $p = 5$ or $p = 6$ (also, $(2, p, p)$ by item (5) and $(3, 3, 5)$, when $p = 5$, by item (11)).

(14) $T[2, 2, 3; 2, p, 3]$. $\mathcal{O}_T$ contains immersed turnovers of type $(2, p, p)$, $(3, 3, p)$ and (p, p, p) if $p = 5$, and an immersed turnover of type $(3, p, p)$ if $p = 6$.

Acknowledgments

The author thanks Ian Agol for helpful conversations. Very special thanks to the referee for invaluable feedback and for recommending simplifications to some of the arguments.

References

[Adams and Schoenfeld 2005] C. Adams and E. Schoenfeld, "Totally geodesic Seifert surfaces in hyperbolic knot and link complements. I", *Geom. Dedicata* **116** (2005), 237–247. MR 2006j:57008 Zbl 1092.57003

[Andreev 1970a] E. M. Andreev, "On convex polyhedra in Lobachevskii spaces", *Mat. Sb.* (*N.S.*) **81** (1970), 445–478. In Russian; translated in *Math. USSR Sb.* **10** (1970), 413-440. MR 41 #4367 Zbl 0217.46801

[Andreev 1970b] E. M. Andreev, "On convex polyhedra of finite volume in Lobachevskii space", *Mat. Sb.* (*N.S.*) **83** (1970), 256–260. In Russian; translated in *Math. USSR Sb.* **12** (1971), 255-259. MR 42 #8388

[Boileau et al. 2003] M. Boileau, S. Maillot, and J. Porti, *Three-dimensional orbifolds and their geometric structures*, Panoramas et Synthèses **15**, Société Mathématique de France, Paris, 2003. MR 2005b:57030 Zbl 1058.57009

[Cooper et al. 2000] D. Cooper, C. D. Hodgson, and S. P. Kerckhoff, *Three-dimensional orbifolds and cone-manifolds*, MSJ Memoirs **5**, Math. Soc. of Japan, Tokyo, 2000. MR 2002c:57027 Zbl 0955.57014

[Dunbar 1988] W. D. Dunbar, "Hierarchies for 3-orbifolds", *Topology Appl.* **29**:3 (1988), 267–283. MR 89h:57008 Zbl 0665.57011

[Hodgson 1992] C. D. Hodgson, "Deduction of Andreev's theorem from Rivin's characterization of convex hyperbolic polyhedra", pp. 185–193 in *Topology '90*, edited by B. Apanasov et al., Ohio State Univ. Math. Res. Inst. Publ. **1**, de Gruyter, Berlin, 1992. MR 93h:57022 Zbl 0765.52013

[Maclachlan 1996] C. Maclachlan, "Triangle subgroups of hyperbolic tetrahedral groups", *Pacific J. Math.* **176**:1 (1996), 195–203. MR 98d:20056 Zbl 0865.20031

[Maskit 1988] B. Maskit, *Kleinian groups*, Grundlehren der Mathematischen Wissenschaften **287**, Springer, Berlin, 1988. MR 90a:30132 Zbl 0627.30039

[Morgan 1984] J. W. Morgan, "On Thurston's uniformization theorem for three-dimensional manifolds", pp. 37–125 in *The Smith conjecture* (New York, 1979), edited by J. W. Morgan and H. Bass, Pure Appl. Math. **112**, Academic Press, Orlando, FL, 1984. MR 758464 Zbl 0599.57002

[Rafalski 2010] S. Rafalski, "Immersed turnovers in hyperbolic 3-orbifolds", *Groups Geom. Dyn.* **4**:2 (2010), 333–376. MR 2011a:57036 Zbl 1194.57024

[Ratcliffe 1994] J. G. Ratcliffe, *Foundations of hyperbolic manifolds*, Graduate Texts in Mathematics **149**, Springer, New York, 1994. MR 95j:57011 Zbl 0809.51001

[Roeder et al. 2007] R. K. W. Roeder, J. H. Hubbard, and W. D. Dunbar, "Andreev's theorem on hyperbolic polyhedra", *Ann. Inst. Fourier* (*Grenoble*) **57**:3 (2007), 825–882. MR 2008e:51011 Zbl 1127.51012

[Singerman 1972] D. Singerman, "Finitely maximal Fuchsian groups", *J. London Math. Soc.* (2) **6** (1972), 29–38. MR 48 #529 Zbl 0251.20052

[Thurston 1979] W. P. Thurston, "The geometry and topology of three-manifolds", lecture notes, Princeton University, 1979, available at http://msri.org/publications/books/gt3m.

[Thurston 1982] W. P. Thurston, "Three-dimensional manifolds, Kleinian groups and hyperbolic geometry", *Bull. Amer. Math. Soc.* (*N.S.*) **6**:3 (1982), 357–381. MR 83h:57019 Zbl 0496.57005

[Thurston 1997] W. P. Thurston, *Three-dimensional geometry and topology*, vol. 1, Princeton Mathematical Series **35**, Princeton Univ. Press, Princeton, NJ, 1997. MR 97m:57016 Zbl 0873.57001

[Ushijima 2006] A. Ushijima, "A volume formula for generalised hyperbolic tetrahedra", pp. 249–265 in *Non-Euclidean geometries: János Bolyai memorial volume* (Budapest, 2002), edited by A. Prékopa and E. Molnár, Math. Appl. (N. Y.) **581**, Springer, New York, 2006. MR 2007h:52008 Zbl 1096.52006

[Weeks ≥ 2012] J. Weeks, "Kaleidotile", available at http://www.geometrygames.org/KaleidoTile.

Received February 1, 2011. Revised October 28, 2011.

SHAWN RAFALSKI
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
FAIRFIELD UNIVERSITY
1073 NORTH BENSON ROAD
15 BANNOW SCIENCE CENTER
FAIRFIELD, CT 06825-5195
UNITED STATES

srafalski@fairfield.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 255, No. 1, 2012

HURWITZ SPACES OF COVERINGS WITH TWO SPECIAL FIBERS AND MONODROMY GROUP A WEYL GROUP OF TYPE B_d

FRANCESCA VETRO

Let $d \geq 3$, $n_1 > 0$ and $n_2 > 0$ be integers. Let $\underline{e} = (e_1, \ldots, e_r)$ and $\underline{q} = (q_1, \ldots, q_s)$ be two partitions of d. Let X, X' and Y be smooth, connected, projective complex curves. In this paper we study coverings that decompose into a sequence

$$X \xrightarrow{\pi} X' \xrightarrow{f} Y,$$

where π is a degree-two coverings with n_1 branch points and branch locus D_π and f is a degree-d coverings with n_2 points of simple branching and two special points whose local monodromy is given by $\underline{e}$ and $\underline{q}$, respectively. Furthermore the covering f has monodromy group S_d and $f(D_\pi) \cap D_f = \varnothing$ where D_f denotes the branch locus of f. We prove that the corresponding Hurwitz spaces are irreducible under the hypothesis $n_2 - s - r \geq d + 1$.

Introduction

In this paper we study Hurwitz spaces that parametrize branched coverings with two special fibers whose monodromy group is a Weyl group of type B_d.

We notice that the irreducibility of Hurwitz spaces, parametrizing branched coverings of a smooth, connected, projective complex curve Y with monodromy group S_d and with at most two special fibers, has been well studied both when $Y \simeq \mathbb{P}^1$ and when Y has positive genus. The case of simple coverings was studied in [Berstein and Edmonds 1984; Hurwitz 1891], the case of coverings with one special fiber in addition to points of simple branching was studied in [Kanev 2004; Kluitmann 1988; Natanzon 1991; Vetro 2006] and the case of two special fibers in addition to points of simple branching was studied in [Vetro 2010; Wajnryb 1996].

S_d is the Weyl group of a root system of type A_{d-1} and so it is interesting to study coverings with monodromy group a Weyl group different by S_d. Furthermore coverings of this type are interesting, for example, because they appear in the study of spectral curves and of Prym–Tyurin varieties.

MSC2010: primary 14H30; secondary 14H10.
Keywords: Hurwitz spaces, special fibers, branched coverings, Weyl group of type B_d, monodromy, braid moves.

Hurwitz spaces parametrizing coverings of this type were studied in [Biggers and Fried 1986; Kanev 2006; Vetro 2007; 2008a; 2008b; 2009]. Biggers and Fried proved the irreducibility of Hurwitz spaces parametrizing coverings of $\mathbb{P}^1$ whose monodromy group is a Weyl group of type D_d and whose local monodromies are all reflections. Kanev extended the result to Hurwitz spaces of Galois coverings of $\mathbb{P}^1$ whose Galois group is an arbitrary Weyl group.

Let X and X' be smooth, connected, projective complex curves. We studied Hurwitz spaces of coverings that decompose into a sequence of coverings of type $X \xrightarrow{\pi} X' \xrightarrow{f} Y$, where π is a degree-two covering and f is a degree $d \geq 3$ covering with one special fiber and with monodromy group S_d. We analyzed in [Vetro 2007; 2008a] the case that π is branched, and in [Vetro 2008b; 2009] the unramified case.

In this paper we continue the study of coverings of type $X \xrightarrow{\pi} X' \xrightarrow{f} Y$, with π a degree-two covering and f a degree-d covering. Let $\underline{e} = (e_1, \ldots, e_r)$ and $\underline{q} = (q_1, \ldots, q_s)$ be two partitions of d and let b_0 be a point of Y. In particular we study equivalence classes of pairs $[X \xrightarrow{\pi} X' \xrightarrow{f} Y, \phi]$ satisfying the following conditions:

- π is branched in n_1 points and has branch locus D_π, f is simply branched in n_2 points and has two special points with local monodromy given by $\underline{e}$ and $\underline{q}$, respectively;
- f has monodromy group S_d and $f(D_\pi) \cap D_f = \varnothing$, where D_f denotes the branch locus of f;
- $f \circ \pi$ is unramified in b_0 and $\phi : (f \circ \pi)^{-1}(b_0) \to \{-d, \ldots, -1, 1, \ldots, d\}$ is a bijection.

We study the irreducibility of the corresponding Hurwitz spaces both when $Y \simeq \mathbb{P}^1$ and when Y has genus > 0. We prove that, in both the cases, these spaces are irreducible under the hypothesis $n_2 - s - r \geq d + 1$. This condition is necessary in [Vetro 2010] in order to prove the irreducibility of the Hurwitz spaces $H^o_{d,n_2,\underline{e},\underline{q}}(Y, b_0)$ that parametrize equivalence classes of pairs $[f, \varphi]$ where f is a coverings as above and $\varphi : f^{-1}(b_0) \to \{1, \ldots, d\}$ is a bijection. Here, we also use the results of [Vetro 2010].

Notation. Two degree-d branched coverings of Y, $f_1 : X_1 \to Y$ and $f_2 : X_2 \to Y$, are *equivalent* if there exists a biholomorphic map $p : X_1 \to X_2$ such that $f_2 \circ p = f_1$. Two sequences of coverings,

$$X_1 \xrightarrow{\pi_1} X'_1 \xrightarrow{f_1} Y \quad \text{and} \quad X_2 \xrightarrow{\pi_2} X'_2 \xrightarrow{f_2} Y,$$

are *equivalent* if there exist two biholomorphic maps $p : X_1 \to X_2$ and $p' : X'_1 \to X'_2$ such that $p' \circ \pi_1 = \pi_2 \circ p$ and $f_2 \circ p' = f_1$. The equivalence class containing $f \circ \pi$ is denoted by $[f \circ \pi]$. The natural action of S_d on $\{1, \ldots, d\}$ is on the right.

1. Preliminaries

Throughout this section, d and n denote positive integers.

1.1. *Weyl groups of type B_d.* (Refer to [Bourbaki 1968; Carter 1972] for details.) Let $\{\varepsilon_1, \dots, \varepsilon_d\}$ be the standard base of $\mathbb{R}^d$ and let R be the root system

$$\{\pm\varepsilon_i,\ \pm\varepsilon_i \pm \varepsilon_j : 1 \le i, j \le d\}.$$

Let us denote by $W(B_d)$ the group generated by the reflections s_{ε_i}, with $1 \le i \le d$, and by the reflections $s_{\varepsilon_i - \varepsilon_j}$, with $1 \le i < j \le d$. We call $W(B_d)$ a Weyl group of type B_d.

We notice that the reflection $s_{\varepsilon_i - \varepsilon_j}$ exchanges ε_i with ε_j and $-\varepsilon_i$ with $-\varepsilon_j$, leaving fixed each ε_h with $h \ne i, j$. The reflection s_{ε_i} exchanges ε_i with $-\varepsilon_i$ and fixes all the ε_h with $h \ne i$. Thus if we identify $\{\pm\varepsilon_i : 1 \le i \le d\}$ with $\{\pm 1, \dots, \pm d\}$ by the map $\pm\varepsilon_i \to \pm i$, we can easily define an injective homomorphism from $W(B_d)$ into S_{2d} such that

$$s_{\varepsilon_i - \varepsilon_j} \to (i\ j)(-i\ -j), \quad s_{\varepsilon_i} \to (i\ -i), \quad s_{\varepsilon_i + \varepsilon_j} = s_{\varepsilon_i} s_{\varepsilon_j} s_{\varepsilon_i - \varepsilon_j} \to (i\ -j)(-i\ j).$$

Let $\mathbb{Z}_2^d$ be the set of the functions from $\{1, \dots, d\}$ into $\mathbb{Z}_2$ equipped with the sum operation. We will use $\bar{1}_j$ to denote the function in $\mathbb{Z}_2^d$ defined by

$$\bar{1}_j(j) = \bar{1} \quad \text{and} \quad \bar{1}_j(h) = \bar{0} \quad \text{for each } h \ne j$$

and we will write z_{ij} to denote the function in $\mathbb{Z}_2^d$ defined by

$$z_{ij}(i) = z_{ij}(j) = z \quad \text{and} \quad z_{ij}(h) = \bar{0} \quad \text{for each } h \ne i, j \text{ and } z \in \mathbb{Z}_2.$$

Let Ψ be the homomorphism from S_d into $\operatorname{Aut}(\mathbb{Z}_2^d)$ that assigns to $t \in S_d$ the element $\Psi(t) \in \operatorname{Aut}(\mathbb{Z}_2^d)$, where $[\Psi(t)\,a]\,(j) := a(j^t)$ for each $a \in \mathbb{Z}_2^d$.

Let $\mathbb{Z}_2^d \times^s S_d$ be the semidirect product of $\mathbb{Z}_2^d$ and S_d through the homomorphism Ψ. Given $(a'; t_1), (a''; t_2) \in \mathbb{Z}_2^d \times^s S_d$, we put

$$(a'; t_1) \cdot (a''; t_2) := (a' + \Psi(t_1)a''; t_1 t_2).$$

It is easy to check that the homomorphism from $W(B_d) \to \mathbb{Z}_2^d \times^s S_d$ defined by

$$s_{\varepsilon_i - \varepsilon_j} \to (0; (i\ j)), \quad s_{\varepsilon_i} \to (\bar{1}_i; \mathrm{id}), \quad s_{\varepsilon_i + \varepsilon_j} \to (\bar{1}_{ij}; (i\ j))$$

is an isomorphism. We will identify $W(B_d)$ with $\mathbb{Z}_2^d \times^s S_d$ via this isomorphism.

Definition 1. Let k be a positive integer. Let $(c; \xi)$ be an element of $W(B_d)$ such that ξ is a k-cycle of S_d and c is a function that sends to $\bar{0}$ all the indexes fixed by ξ. We call an such element a *positive k-cycle* if c is either zero or a function which sends to $\bar{1}$ an even number of indexes. We call it *negative k-cycle* if it is not positive.

We notice that two cycles $(c;\xi)$ and $(c';\xi')$ in $W(B_d)$ are disjoint if ξ and ξ' are disjoint. Furthermore, all the elements in $W(B_d)$ can be expressed as a product of disjoint positive and negative cycles. The lengths of such disjoint cycles together with their signs determine the signed cycle type of the elements of $W(B_d)$. Two elements of $W(B_d)$ are conjugate if and only if they have the same signed cycle type [Carter 1972].

Braid group actions on Hurwitz systems. (Refer to [Birman 1969; Fadell and Neuwirth 1962; Graber et al. 2002; Hurwitz 1891; Kanev 2004; Scott 1970].) Let Y be a smooth, connected, projective complex curve of genus g and let $b_0 \in Y$. Let $(Y-b_0)^{(n)}$ be the n-fold symmetric product of $(Y-b_0)$ and let Δ be the codimension 1 locus of $(Y-b_0)^{(n)}$ consisting of non simple divisors. The generators of the braid group $\pi_1((Y-b_0)^{(n)}-\Delta, D)$ were studied in [Birman 1969; Fadell and Neuwirth 1962; Scott 1970]. They are the elementary braids σ_i, with $1 \le i \le n-1$, and the braids $\rho_{jk},\ \tau_{jk}$, with $1 \le j \le n$ and $1 \le k \le g$.

Definition 2. Let G be a subgroup of S_h. An ordered sequence of elements of G

$$(\underline{t};\underline{\lambda},\underline{\mu}) := (t_1,\dots,t_n;\lambda_1,\mu_1,\dots,\lambda_g,\mu_g)$$

such that $t_i \neq \mathrm{id}$ for each i and $t_1\cdots t_n = [\lambda_1,\mu_1]\cdots[\lambda_g,\mu_g]$ is called a *Hurwitz system with values in G*. The subgroup of G generated by $t_1,\dots,t_n,\lambda_1,\mu_1,\dots,\lambda_g,\mu_g$ is called the *monodromy group* of the Hurwitz system.

Remark 3. An ordered sequence $\underline{t} := (t_1,\dots,t_n)$ of elements of G, with $t_i \neq \mathrm{id}$ for each i, is a Hurwitz system if $t_1\cdots t_n = \mathrm{id}$.

To each generator of $\pi_1((Y-b_0)^{(n)}-\Delta, D)$ one associates a pair of braid moves. We denote by σ_i' and $\sigma_i'' = (\sigma_i')^{-1}$ the moves associated with σ_i, and we call them elementary moves. Similarly, ρ_{jk}' and $\rho_{jk}'' = (\rho_{jk}')^{-1}$ denote the moves associated to ρ_{jk}, and likewise for τ_{jk}.

The moves σ_i' and σ_i'' fix all the λ_k, all the μ_k and all the t_h with $h \neq i, i+1$. The elementary move σ_i' transforms (t_i, t_{i+1}) into $(t_i t_{i+1} t_i^{-1}, t_i)$, while the move σ_i'' transforms (t_i, t_{i+1}) into $(t_{i+1}, t_{i+1}^{-1} t_i t_{i+1})$; see [Hurwitz 1891].

The braid moves ρ_{jk}' and ρ_{jk}'' fix all the λ_l, all the t_h with $h \neq j$ and all the μ_l with $l \neq k$. They modify t_j and μ_k. Analogously the braid moves τ_{jk}' and τ_{jk}'' modify t_j and λ_k, leaving unchanged μ_l for all l, λ_l with $l \neq k$ and t_h with $h \neq j$.

The braid moves $\rho_{jk}',\ \rho_{jk}'',\ \tau_{jk}'$ and τ_{jk}'' transform t_j to an element belonging to the same conjugate class (see Theorem 1.8, [Kanev 2004]).

By [Kanev 2004, Corollary 1.9], when $\lambda_1 = \cdots = \lambda_k = \mu_1 = \cdots = \mu_{k-1} = \mathrm{id}$, the braid move ρ_{1k}' transforms μ_k into $t_1^{-1}\mu_k$.

Analogously when $\lambda_1 = \cdots = \lambda_{k-1} = \mu_1 = \cdots = \mu_{k-1} = \mathrm{id}$, the braid move τ_{1k}'' transforms λ_k into $t_1^{-1}\lambda_k$.

Definition 4. Two Hurwitz systems with values in G are *braid-equivalent* if one is obtained from the other by a finite sequence of braid moves σ_i', ρ_{jk}', τ_{jk}', σ_i'', ρ_{jk}'', τ_{jk}'', where $1 \leq i \leq n-1$, $1 \leq j \leq n$ and $1 \leq k \leq g$. Two ordered sequences of elements of G, $(t_1, \ldots, t_l)$ and $(t_1', \ldots, t_l')$, are *braid-equivalent* if $(t_1', \ldots, t_l')$ is obtained from $(t_1, \ldots, t_l)$ by a finite sequence of braid moves of type σ_i', σ_i''. We denote braid equivalence by $\sim$.

2. The Hurwitz spaces $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ and $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$

Let X, X' and Y be smooth, connected, projective complex curves. Let $d \geq 3$, $n_1 > 0$ and $n_2 > 0$ be integers. Let $\underline{e} = (e_1, \ldots, e_r)$ and $\underline{q} = (q_1, \ldots, q_s)$ be two partitions of d with $e_1 \geq e_2 \geq \cdots \geq e_r \geq 1$ and $q_1 \geq q_2 \geq \cdots \geq q_s \geq 1$. Let b_0 be a point of Y and let g be the genus of Y. In this paper we study equivalence classes of pairs $[X \xrightarrow{\pi} X' \xrightarrow{f} Y, \phi]$ satisfying the following conditions:

(a) π is a degree-two coverings with n_1 branch points and branch locus D_π;

(b) f is a degree-d coverings with n_2 points of simple branching and two special points whose local monodromy has cycle type given by $\underline{e}$ and $\underline{q}$, respectively;

(c) the covering f has monodromy group S_d and $f(D_\pi) \cap D_f = \varnothing$ where D_f denotes the branch locus of f;

(d) $f \circ \pi$ is unramified in b_0 and $\phi : (f \circ \pi)^{-1}(b_0) \to \{-d, \ldots, -1, 1, \ldots, d\}$ is a bijection such that if $f^{-1}(b_0) = \{y_1, \ldots, y_d\}$ then $\pi^{-1}(y_i) = \{\phi^{-1}(i), \phi^{-1}(-i)\}$ for each $i = 1, \ldots, d$.

$H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ will denote the Hurwitz space that parametrizes equivalence classes of pairs $[f \circ \pi, \phi]$ satisfying conditions (a)–(d).

$H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$ will denote the Hurwitz space that parametrizes equivalence classes of coverings $f \circ \pi$ satisfying conditions (a)–(c).

Definition 5. A $(n_1, n_2, \underline{e}, \underline{q})$*-Hurwitz system* is a Hurwitz system with values in $\mathbb{Z}_2^d \times^s S_d$, $(t_1, \ldots, t_{n_1+n_2+2}; \underline{\lambda}, \underline{\mu})$, such that n_1 of $t_1, \ldots, t_{n_1+n_2+2}$ are of the form $(\underline{1}_*; \text{id})$, n_2 are of the form $(z_{hk}; (hk))$, one is a product of r disjoint positive cycles whose lengths are given by the elements of the partition $\underline{e}$, and one is a product of s disjoint positive cycles whose lengths are given by the elements of the partition $\underline{q}$.

Let $D = f(D_\pi) \cup D_f$ and let $m : \pi_1(Y - D, b_0) \to S_{2d}$ be the monodromy homomorphism associated to $[f \circ \pi, \phi]$. Let $(\gamma_1, \ldots, \gamma_{n_1+n_2+2}, \alpha_1, \beta_1, \ldots, \alpha_g, \beta_g)$ be a standard generating system for $\pi_1(Y - D, b_0)$. The images under m of γ_1, $\ldots$, $\gamma_{n_1+n_2+2}$, α_1, β_1, $\ldots$, α_g, β_g determine an $(n_1, n_2, \underline{e}, \underline{q})$-Hurwitz system with monodromy group $W(B_d)$.

In the sequel we will denote by $A^o_{n_1,n_2,\underline{e},\underline{q},g}$ the set of all $(n_1, n_2, \underline{e}, \underline{q})$-Hurwitz systems with monodromy group $W(B_d)$. When $g = 0$ we will write $A^o_{n_1,n_2,\underline{e},\underline{q}}$ instead of $A^o_{n_1,n_2,\underline{e},\underline{q},g}$.

Let $\delta : H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0) \to (Y - b_0)^{(n_1+n_2+2)} - \Delta$ be the map that assigns to each pair $[f \circ \pi, \phi]$ the branch locus of $f \circ \pi$. By Riemann's existence theorem we can identify the fiber of δ over D with $A^o_{n_1,n_2,\underline{e},\underline{q},g}$. There is a unique topology on $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ such that δ is a topological covering map; see [Fulton 1969]. Therefore the braid group $\pi_1((Y-b_0)^{(n_1+n_2+2)}-\Delta, D)$ acts on $A^o_{n_1,n_2,\underline{e},\underline{q},g}$. If this action is transitive, $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ is connected and hence, since $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ is smooth, it is also irreducible.

Remark 6. The forgetful map $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0) \to H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$ defined by $[f \circ \pi, \phi] \to [f \circ \pi]$ is a morphism, whose image is a dense subset of $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$. This ensures that if $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ is irreducible also $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$ is irreducible.

3. The results

We denote by ϵ the following element in S_d having cycle type $\underline{e}$:

(1) $$(1\ 2\ \dots\ e_1)(e_1{+}1\ \dots\ e_1{+}e_2)\cdots((e_1{+}\cdots{+}e_{r-1}){+}1\ \dots\ d)\,.$$

We denote by ν the following element in S_d having cycle type $\underline{q}$:

(2) $$(1\ d\ d{-}1\ \dots\ d{-}q_1{+}2)(d{-}q_1{+}1\ \dots\ d{-}(q_1{+}q_2){+}2)\\ \cdots\ (d{-}(q_1{+}\cdots{+}q_{s-1}){+}1\ \dots 2).$$

Lemma 7. *Let $(t_1, \dots, t_i, t_{i+1}, \dots, t_l)$ be a sequence of permutations in S_d where t_i and t_{i+1} are two equal transpositions of S_d. Then we can move to the right and to the left the pair (t_i, t_{i+1}) leaving unchanged the other permutations of the sequence.*

Proof. Applying the elementary moves $\sigma''_{i-1}, \sigma''_i$ we obtain

$$(t_{i-1}, t_i, t_{i+1}) \sim (t_i, t_i^{-1}t_{i-1}t_i, t_{i+1}) \sim (t_i, t_{i+1}, t_{i-1});$$

applying the moves σ'_{i+1}, σ'_i we have

$$(t_i, t_{i+1}, t_{i+2}) \sim (t_i, t_{i+1}t_{i+2}t_{i+1}^{-1}, t_{i+1}) \sim (t_{i+2}, t_i, t_{i+1}).$$

Hence using sequences of elementary moves of type either $\sigma''_{j-1}, \sigma''_j$ or σ'_{j+1}, σ'_j we can move respectively on the left and on the right the pair (t_i, t_{i+1}), leaving unchanged the other permutations of the sequence. □

Lemma 8. *Let $(t_1, \dots, t_l, \tau, \tau)$ be a sequence of permutations of S_d, with τ a transposition. Let H be the subgroup of S_d generated by $t_1, \dots, t_l$. Then, for each $h \in H$, one has*

$$(t_1, \dots, t_l, \tau, \tau) \sim (t_1, \dots, t_l, h^{-1}\tau h, h^{-1}\tau h).$$

Proof. Let $h \in H$, then $h = h_1 h_2 \cdots h_k$ where h_i or h_i^{-1}, with $i = 1, \dots, k$, belonging to $\{t_1, \dots, t_l\}$. If h_1 is equal to t_j for some $j \in \{1, \dots, l\}$, we use Lemma 7 to bring the pair (τ, τ) to the left of t_j and then we act by the moves $\sigma''_{j+1}, \sigma''_j$ in order to replace (τ, τ, t_j) with $(t_j, t_j^{-1}\tau t_j, t_j^{-1}\tau t_j)$.

On the contrary, if h_1 is equal to t_j^{-1} for some $j \in \{1, \dots, l\}$, we use Lemma 7 to shift the pair (τ, τ) on the right of t_j and then we apply σ'_j, σ'_{j+1}. In this way we replace (t_j, τ, τ) with $(t_j \tau t_j^{-1}, t_j \tau t_j^{-1}, t_j)$.

For h_2 we reason as above but we bring the pair $(h_1^{-1}\tau h_1, h_1^{-1}\tau h_1)$ to the left or to the right of t_n depending on whether h_2 is equal to t_n or to t_n^{-1}.

Following this line for each h_i, with $i = 3, \dots, k$, we obtain the claim. □

Proposition 9 [Vetro 2010, Proposition 2]. *Let $\underline{t} = (t_1, \dots, t_{n_2+2})$ be a Hurwitz system of permutations of S_d with monodromy group S_d such that one of $t_1, \dots, t_{n_2+2}$ has cycle type $\underline{e}$, one has cycle type $\underline{q}$ and the other n_2 permutations in $t_1, \dots, t_{n_2+2}$ are transpositions. If $n_2 - s - r \geq d + 1$, $\underline{t}$ is braid-equivalent to the Hurwitz system*

$(\epsilon, \tilde{t}_2, \dots, \tilde{t}_{n_2+2-s}, \nu)$ *if* $s = 1$,

$(\epsilon, \tilde{t}_2, \dots, \tilde{t}_{n_2+2-s}, \nu, (1\ d - q_1 + 1), \dots, (1\ d - (q_1 + \cdots + q_{s-1}) + 1))$ *if* $s > 1$,

where ϵ and ν are the permutations defined in (1) *and* (2), *and where the sequence $(\tilde{t}_2, \dots, \tilde{t}_{n_2+2-s})$ is equal to*

$((1\,2), \dots, (1\,2))$ *if* $r = 1$,

$((1\ e_1 + 1), \dots, (1\ (e_1 + \cdots + e_{r-1}) + 1), (1\,2), \dots, (1\,2))$ *if* $r > 1$

with the transposition $(1\,2)$ *appearing an even number of times.*

Remark 10. Seeing that $d \geq 3$, the hypothesis $n_2 - s - r \geq d + 1$ ensures that in the sequence $(\tilde{t}_2, \dots, \tilde{t}_{n_2+2-s})$ there are more than 3 transpositions (12).

3.1. *Irreducibility of $H_{W(B_d), n_1, n_2, \underline{e}, \underline{q}}(\mathbb{P}^1, b_0)$ and $H_{W(B_d), n_1, n_2, \underline{e}, \underline{q}}(\mathbb{P}^1)$.*

We next show that, if $n_2 - s - r \geq d + 1$, the braid group $\pi_1((\mathbb{P}^1 - b_0)^{(n_1+n_2+2)} - \Delta, D)$ acts transitively on $A^o_{n_1, n_2, \underline{e}, \underline{q}}$. To prove this we show that each $(n_1, n_2, \underline{e}, \underline{q})$-Hurwitz system in $A^o_{n_1, n_2, \underline{e}, \underline{q}}$ is braid-equivalent to a given normal form.

Proposition 11. *If $n_2 - s - r \geq d + 1$, each Hurwitz system in $A^o_{n_1, n_2, \underline{e}, \underline{q}}$ is braid-equivalent to a Hurwitz system of the form*

$\big(\tilde{t}_1, \dots, \tilde{t}_{n_2+2-s}, (0; \nu), (\bar{1}_1; \mathrm{id}), \dots, (\bar{1}_1; \mathrm{id})\big)$ *if* $s = 1$,

$$\Big(\tilde{t}_1, \dots, \tilde{t}_{n_2+2-s}, (0; \nu), (0; (1\ d - q_1 + 1)), \dots, \Big(0; \Big(1\ d - \sum_{h=1}^{s-1} q_h + 1\Big)\Big), (\bar{1}_1; \mathrm{id}), \dots, (\bar{1}_1; \mathrm{id})\Big) \text{ if } s > 1,$$

where $(\bar{1}_1; \mathrm{id})$ appears n_1 times and where $(\tilde{t}_1, \dots, \tilde{t}_{n_2+2-s})$ is the sequence

$\big((0; \epsilon), (0; (1\,2)), \dots, (0; (1\,2))\big)$ *if* $r = 1$,

$$\big((0;\epsilon),(0;(1e_1+1)),\dots,\big(0;\big(1\ \sum_{i=1}^{r-1}e_i+1\big)\big),(0;(12)),\dots,(0;(12))\big)\ \textit{if}\ r>1,$$

with $(0;(1\,2))$ *appearing an even number of times.*

Proof. Step 1. Let $\underline{t}\in A^o_{n_1,n_2,\underline{e},\underline{q}}$. We prove first that $\underline{t}$ is braid-equivalent to a Hurwitz system of either the form

$$\big(\dots,(0;\nu),(\bar{1}_1;\mathrm{id}),\dots,(\bar{1}_1;\mathrm{id})\big)$$

or the form

$$\big(\dots,(0;\nu),(0;(1\ d-q_1+1)),\dots,\big(0;\big(1\ d-\sum_{h=1}^{s-1}q_h+1\big)\big),(\bar{1}_1;\mathrm{id}),\dots,(\bar{1}_1;\mathrm{id})\big),$$

depending on whether $s=1$ or $s>1$, where $(\bar{1}_1;\mathrm{id})$ appears n_1 times.

Acting by elementary moves σ'_j we shift on the right the elements of the form $(\bar{1}_*;\mathrm{id})$ obtaining that $\underline{t}$ is braid-equivalent to

$$\big(\hat{t}_1,\dots,\hat{t}_{n_2+2},(\bar{1}_h;\mathrm{id}),\dots,(\bar{1}_k;\mathrm{id})\big),$$

where $\hat{t}_i=(*;t'_i)$. We notice that $(t'_1,\dots,t'_{n_2+2})$ is a Hurwitz system of permutations of S_d with monodromy group S_d such that one of $t'_1,\dots,t'_{n_2+2}$ has cycle type given by $\underline{e}$, one has cycle type given by $\underline{q}$ and the other n_2 permutations are transpositions. Since $n_2-s-r\geq d+1$, by Proposition 9, the system $(t'_1,\dots,t'_{n_2+2})$ is braid-equivalent to either

$$\big(\epsilon,\dots,(1\,2),\dots,(1\,2),(1\,2),(1\,2),\nu\big)$$

or

$$\big(\epsilon,\dots,(1\,2),\dots,(1\,2),(1\,2),(1\,2),\nu,(1\ d-q_1+1),\dots,\big(1\ d-\sum_{h=1}^{s-1}q_h+1\big)\big)$$

depending on whether $s=1$ or $s>1$.

We notice that from

$$\epsilon\cdots(1\,2)\cdots(1\,2)(1\,2)(1\,2)=(12\dots d)$$

it follows that the group generated by the permutations $\epsilon,\dots,(1\,2)$ is all of S_d. Hence, by Lemma 8, the sequence $(\epsilon,\dots,(1\,2),\dots,(1\,2),(1\,2),(1\,2))$ is braid-equivalent to a sequence of the form $(\epsilon,\dots,(1\,2),\dots,(1\,2),\tau,\tau)$, where τ is an arbitrary transposition of S_d.

This ensures that $\underline{t}$ is braid-equivalent to a system of type either

$$\big(\bar{t}_1,\dots,\bar{t}_{n_2+2-s},(b;\nu),(\bar{1}_h;\mathrm{id}),\dots\big)$$

or

$$\big(\bar{t}_1, \dots, \bar{t}_{n_2+2-s}, (b;\nu), (z^1_{1\,d-q_1+1}; (1\; d-q_1+1)), \dots, \\ \big(z^{s-1}_{1\,d-\sum_{h=1}^{s-1} q_h+1}; \big(1\; d-\sum_{h=1}^{s-1} q_h+1\big)\big), (\bar{1}_h; \mathrm{id}), \dots\big),$$

depending on whether $s = 1$ or $s > 1$, where $\bar{t}_i = (*; t''_i)$ and

$$(t''_1, \dots, t''_{n_2+2-s}) = (\epsilon, \dots, (12), \dots, (1\,2), \tau, \tau).$$

Furthermore we can affirm that our system is braid-equivalent to either

$$\big(\bar{t}_1, \dots, \check{t}_{n_2+2-s}, (\bar{1}_u; \mathrm{id}), (b;\nu), (\bar{1}_*; \mathrm{id}), \dots\big)$$

or

$$\big(\bar{t}_1, \dots, \check{t}_{n_2+2-s}, (\bar{1}_u; \mathrm{id}), (b;\nu), \dots, \\ \big(z^{s-1}_{1\,d-\sum_{h=1}^{s-1} q_h+1}; \big(1\; d-\sum_{h=1}^{s-1} q_h+1\big)\big), (\bar{1}_*; \mathrm{id}), \dots\big),$$

depending on whether $s = 1$ or $s > 1$, where u is an arbitrary index in $\{1, \dots, d\}$ and $\check{t}_{n_2+2-s} = (\star'; \tau)$.

In fact, acting by elementary moves of the form σ'_j we can bring to the left of $(b;\nu)$ one element of type $(\bar{1}_*; \mathrm{id})$. We choose $\tau = (u\;*)$ and then we act by σ'_{n_2+2-s} two times to replace $((\star;\tau), (\bar{1}_*; \mathrm{id}))$ by $((\star';\tau), (\bar{1}_u; \mathrm{id}))$.

Now we analyze separately the cases $s = 1$ and $s > 1$.

<u>*Case $s = 1$*</u>. Let $i_1, i_2, \dots, i_l$ be the indexes that b sends to $\bar{1}$. We suppose that $i_1 > i_2 > \cdots > i_{l-1} > i_l$. Since our system is braid-equivalent to

$$\big(\bar{t}_1, \dots, \bar{t}_{n_2}, \check{t}_{n_2+1}, (\bar{1}_{i_l}; \mathrm{id}), (b;\nu), (\bar{1}_*; \mathrm{id}), \dots\big),$$

acting two times by the move σ'_{n_2+2} we can replace the pair $((\bar{1}_{i_l}; \mathrm{id}), (b;\nu))$, with $((\bar{1}_{i_{l+1}}; \mathrm{id}), (\hat{b};\nu))$ where $\hat{b}$ is a function that sends to $\bar{1}$ the indexes $i_1, i_2, \dots, i_{l-1}$, $i_l + 1$, where $i_l + 1$ is the index that precedes i_l in ν. Observe that if there are h indexes among i_{l-1} and i_l, it is sufficient to use the move σ'_{n_2+2} another $2h$ times, to replace the pair $((\bar{1}_{i_{l+1}}; \mathrm{id}), (\hat{b};\nu))$ with $((\bar{1}_{i_{l-1}}; \mathrm{id}), (\check{b};\nu))$ where $\check{b}$ is a function that sends to $\bar{1}$ the indexes $i_1, i_2, \dots, i_{l-2}$.

Since b is a function that sends to $\bar{1}$ an even number of indexes (see Definition 1), following this line we can replace the pair $((\bar{1}_*; \mathrm{id}), (\check{b};\nu))$ with $((\bar{1}_\star; \mathrm{id}), (0;\nu))$. Now, we use σ''_{n_2+2} to shift $(0;\nu)$ to the place $n_2 + 2$.

We notice that if all the elements of the form $(\bar{1}_*; \mathrm{id})$ in our system are equal to $(\bar{1}_1; \mathrm{id})$ we have the claim. Otherwise we place the elements $(\bar{1}_1; \mathrm{id})$ to the last places and then we act by σ'_{n_2+2} to bring one element of type $(\bar{1}_*; \mathrm{id})$ to the left of

$(0;\nu)$. By Lemma 8 and by using σ'_{n_2+1} two times, we can replace our system by a system of type

$$\big((*;\epsilon),\dots,(*;(1\,2)),(*;\tau'),(*;\tau'),(\bar{1}_2;\mathrm{id}),(0;\nu),(\bar{1}_*;\mathrm{id}),\dots\big).$$

Thus, acting by the elementary move σ''_{n_2+2}, we can replace the pair $((\bar{1}_2;\mathrm{id}),(0;\nu))$ with $((0;\nu),(\bar{1}_1;\mathrm{id}))$. Now, acting with elementary moves of type σ'_j, we bring $(\bar{1}_1;\mathrm{id})$ next to the other elements $(\bar{1}_1;\mathrm{id})$.

Reasoning in this way for each $(\bar{1}_*;\mathrm{id})$ such that $*\neq 1$ we obtain the claim.

Case $s>1$. Our system is braid-equivalent to a system of the form

$$\Big(\dots,\bar{t}_{n_2+1-s},\check{t}_{n_2+2-s},(\bar{1}_1;\mathrm{id}),(b;\nu),(z^1_{1\,d-q_1+1};(1\,d-q_1+1)),\dots,\\ \big(z^{s-1}_{1\,d-\sum_{h=1}^{s-1}q_h+1};\big(1\,d-\sum_{h=1}^{s-1}q_h+1\big)\big),(\bar{1}_*;\mathrm{id}),\dots\Big),$$

so if $z^{s-1}=\bar{1}$ we can use the moves $\sigma'_{n_2+3-s},\sigma'_{n_2+4-s},\dots,\sigma'_{n_2+1},\sigma'_{n_2+2}$ in order to replace it by

$$\Big(\dots,\check{t}_{n_2+2-s},(b';\nu),(\hat{z}^1_{1\,d-q_1+1};(1\,d-q_1+1)),\dots,\\ \big(0;\big(1\,d-\sum_{h=1}^{s-1}q_h+1\big)\big),(\bar{1}_1;\mathrm{id}),\dots\Big).$$

Since this system is braid-equivalent to a system of type

$$\Big((*;\epsilon),\dots,(*;(1\,2)),(*;\tau'),(*;\tau'),(\bar{1}_1;\mathrm{id}),(b';\nu),\\ (\hat{z}^1_{1\,d-q_1+1};(1\,d-q_1+1)),\dots,\big(0;\big(1\,d-\sum_{h=1}^{s-1}q_h+1\big)\big),\dots\Big),$$

we can reason as above for all the elements

$$(*;(1\,d-q_1+1)),\quad\dots,\quad\big(*;(1\,d-\sum_{h=1}^{s-2}q_h+1)\big)$$

such that $*$ is a function different from 0. In this way, after at most $s-2$ steps, we transform our system into

$$\Big(\dots,(\bar{1}_1;\mathrm{id}),(\hat{b};\nu),(0;(1\,d-q_1+1)),\dots,\big(0;\big(1\,d-\sum_{h=1}^{s-1}q_h+1\big)\big),\dots\Big).$$

Now if $\hat{b}\neq 0$, it is sufficient to proceed as in the case $s=1$ in order to obtain the system

$$\Big((*;\epsilon),\dots,(*;(1\,2)),(*;\tau),(*;\tau),(\bar{1}_*;\mathrm{id}),(0;\nu),\\ (0;(1\,d-q_1+1)),\dots,\big(0;\big(1\,d-\sum_{h=1}^{s-1}q_h+1\big)\big),\dots\Big).$$

Using elementary moves σ'_j, we move to the left of $(0; \nu)$ all the elements of type $(\bar{1}_*; \mathrm{id})$, so we replace our system with

$$\Big(\dots, (*; \tau), (*; \tau), (\bar{1}_{h_1}; \mathrm{id}), \dots, (\bar{1}_{h_{n_1}}; \mathrm{id}), (0; \nu), (0; (1\ d - q_1 + 1)), \dots, \Big(0; \Big(1\ d - \sum_{h=1}^{s-1} q_h + 1\Big)\Big)\Big).$$

By Lemma 8 we can choose $\tau = (1\ h_1)$. We apply σ'_{n_2+2-s} two times in order to replace $(\bar{1}_{h_1}; \mathrm{id})$ with $(\bar{1}_1; \mathrm{id})$. Now we use elementary moves σ'_j to bring $(\bar{1}_1; \mathrm{id})$ next to $(0; \nu)$. We repeat this reasoning for all $(\bar{1}_{h_i}; \mathrm{id})$ such that $h_i \neq 1$. Since by the Hurwitz formula n_1 is even, we obtain the claim using the sequence of moves $\sigma'_{n_2+n_1+2-s}, \sigma'_{n_2+n_1+1-s}, \dots, \sigma'_{n_2+3-s}, \sigma'_{n_2+n_1+3-s}, \sigma'_{n_2+n_1+2-s}, \dots, \sigma'_{n_2+4-s}, \dots, \sigma'_{n_2+n_1+1}, \dots, \sigma'_{n_2+2}$.

Step 2. By Step 1 and by Lemma 8, $\underline{t}$ is braid-equivalent to either

$$\big((a; \epsilon), (z^1_{12}; (1\,2)), \dots, (z^l_{12}; (1\,2)), (0; \nu), \dots, (\bar{1}_1; \mathrm{id})\big)$$

or

$$\Big((a; \epsilon), (v^1_{1e_1+1}; (1e_1 + 1)), \dots, \Big(v^{r-1}_{1\sum_{i=1}^{r-1} e_i+1}; \Big(1 \sum_{i=1}^{r-1} e_i + 1\Big)\Big), (z^1_{12}; (1\,2)), \dots, (z^l_{12}; (1\,2)), (0; \nu), \dots, (\bar{1}_1; \mathrm{id})\Big),$$

depending on whether $r = 1$ or $r > 1$. We analyze separately the two cases.

<u>Case $r = 1$</u>. From

$$(a; \epsilon)(z^1_{12}; (1\,2)) \cdots (z^l_{12}; (1\,2))(0; \nu) \cdots (\bar{1}_1; \mathrm{id}) = (0; \mathrm{id})$$

it follows that

$$a + z^1_{1\,d} + \cdots + z^l_{1\,d} + \bar{1}_1 + \cdots + \bar{1}_1 = 0.$$

Since in our system there are n_1 elements of type $(\bar{1}_1; \mathrm{id})$ and n_1 is even, by the Hurwitz formula we can affirm that a is either 0 or $\bar{1}_{1\,d}$ depending on whether the number of z^i equal to $\bar{1}$ is even or odd. Acting by moves of type σ'_j we move the elements of the form $(0; (1\,2))$ to the left of $(0; \nu)$. Successively, acting by sequences of moves of type $\sigma''_j, \sigma''_{j+1}$, we shift a pair of type $((\bar{1}_1; \mathrm{id}), (\bar{1}_1; \mathrm{id}))$ to the right of the elements $(\bar{1}_{12}; (1\,2))$.

If the function a is equal to 0 and the elements of type $(\bar{1}_{12}; (1\,2))$ are in the places $r + 1, \dots, h$, it is sufficient to use the sequence of moves $\sigma''_h, \sigma''_{h-1}, \dots, \sigma''_{r+1}, \sigma''_{r+1}, \dots, \sigma''_h$ to obtain the system

$$\big((0; \epsilon), (0; (1\,2)), \dots, (0; (1\,2)), (\bar{1}_1; \mathrm{id}), (\bar{1}_1; \mathrm{id}), (0; (1\,2)), \dots, (0; (1\,2)), (0; \nu), \dots\big).$$

The claim follows by using the sequence of moves $\sigma'_{h+2}, \sigma'_{h+1}, \dots, \sigma'_{n_2+3}, \sigma'_{n_2+2}$.

On the contrary, if $a = \bar{1}_{1\,d}$ and the elements of type $(\bar{1}_{12};\,(1\,2))$ are in the places $r+1, \dots, h$, we use the sequence of moves $\sigma''_h, \sigma''_{h-1}, \dots, \sigma''_{r+2}, \sigma'_{r+1}$ to bring our system to the form

$$\big((\bar{1}_{1\,d};\epsilon), (\bar{1}_2;\mathrm{id}), (\bar{1}_{12};(1\,2)), (0;(1\,2)), \dots, (0;(1\,2)),\\ (\bar{1}_1;\mathrm{id}), (0;(1\,2)), \dots, (0;\nu), \dots\big).$$

We use σ'_1 to replace the pair $((\bar{1}_{1\,d};\epsilon), (\bar{1}_2;\mathrm{id}))$ with $((\bar{1}_1;\mathrm{id}), (\bar{1}_{1\,d};\epsilon))$ and then we apply the moves σ'_1, σ'_2 to replace $((\bar{1}_1;\mathrm{id}),\ (\bar{1}_{1\,d};\epsilon), (\bar{1}_{12};(1\,2)))$ by

$$\big((0;\epsilon), (0;(1\,2)), (\bar{1}_1;\mathrm{id})\big).$$

Now we obtain the claim acting by the sequence of elementary moves $\sigma''_{r+2}, \sigma''_{r+3}, \dots, \sigma''_h, \sigma'_{h+2}, \sigma'_{h+1}, \dots, \sigma'_{n_2+3}, \sigma'_{n_2+2}$.

Case $r > 1$. Seeing that

$$(a;\epsilon)(v^1_{1e_1+1};(1e_1+1))\cdots(z^1_{12};(1\,2))\cdots(0;\nu)\cdots(\bar{1}_1;\mathrm{id}) = (0;\mathrm{id}),$$

one has

$$a + v^1_{e_1(e_1+e_2)} + v^2_{(e_1+e_2)(e_1+e_2+e_3)} + \cdots + v^{r-1}_{(e_1+\cdots+e_{r-1})d} + z^1_{1\,d} + \cdots + \bar{1}_1 + \cdots + \bar{1}_1 = 0.$$

Since a is a function that sends to $\bar{1}$ at most an even number of indexes moved by every disjoint cycle of which is product ϵ, the equality above ensures that a is either 0 or $\bar{1}_{1e_1}$.

If $a = 0$, we have $v^1 = v^2 = \cdots = v^{r-1} = 0$. Furthermore there is an even number of z^i equal to $\bar{1}$. So in order to obtain the claim, it is sufficient to act as in the case $r = 1$ and $a = 0$.

On the contrary, if $a = \bar{1}_{1e_1}$ we have $v^1 = v^2 = \cdots = v^{r-1} = \bar{1}$; furthermore, there is an odd number of z^i equal to $\bar{1}$. Then we act as in the case $r = 1$ and $a = \bar{1}_{1\,d}$ to replace our system with the braid-equivalent system

$$\Big((\bar{1}_{1e_1};\epsilon), \dots, \big(\bar{1}_{1\sum_{i=1}^{r-1} e_i+1};\big(1\sum_{i=1}^{r-1} e_i+1\big)\big), (\bar{1}_2;\mathrm{id}), (\bar{1}_{12};(1\,2)),\\ (0;(1\,2)), \dots, (0;(1\,2)), (\bar{1}_1;\mathrm{id}), (0;(1\,2)), \dots, (0;\nu), \dots\Big).$$

Using the moves $\sigma'_r, \sigma'_{r-1}, \dots, \sigma'_2, \sigma'_1$ we transform the sequence

$$\Big((\bar{1}_{1e_1};\epsilon), \dots, \big(\bar{1}_{1\sum_{i=1}^{r-1} e_i+1};(1\ \sum_{i=1}^{r-1} e_i+1)\big), (\bar{1}_2;\mathrm{id}), (\bar{1}_{12};(1\,2))\Big)$$

into

$$\Big((\bar{1}_1;\mathrm{id}), (\bar{1}_{1e_1};\epsilon), \dots, \big(\bar{1}_{1\sum_{i=1}^{r-1} e_i+1};\big(1\ \sum_{i=1}^{r-1} e_i+1\big)\big), (\bar{1}_{12};(1\,2))\Big).$$

Now in order to obtain the claim it is sufficient to act by the sequence of moves $\sigma'_1, \dots, \sigma'_r, \sigma'_{r+1}, \sigma''_{r+2}, \dots, \sigma''_h, \sigma'_{h+2}, \sigma'_{h+1}, \dots, \sigma'_{n_2+3}, \sigma'_{n_2+2}$. □

The following result is a direct consequence of Proposition 11.

Theorem 12. *If* $n_2 - s - r \geq d + 1$, *the Hurwitz space* $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(\mathbb{P}^1, b_0)$ *is irreducible.*

Combining Theorem 12 and Remark 6, we derive the following result.

Corollary 13. *If* $n_2 - s - r \geq d + 1$, *the Hurwitz space* $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(\mathbb{P}^1)$ *is irreducible.*

3.2. ***Irreducibility of*** $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ ***and*** $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$**.** Let Y be a smooth, connected, projective complex curve of genus $g \geq 1$.

Theorem 14. *If* $n_2 - s - r \geq d + 1$, *the Hurwitz space* $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ *is irreducible.*

Proof. To prove the irreducibility of $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y, b_0)$ it is sufficient to show that each $(n_1, n_2, \underline{e}, \underline{q})$-Hurwitz system in $A^o_{n_1,n_2,\underline{e},\underline{q},g}$ is braid-equivalent to a system of the form

$$(\hat{\underline{t}}; (0; \mathrm{id}), \dots, (0; \mathrm{id})).$$

In fact, $\hat{\underline{t}} \in A^o_{n_1,n_2,\underline{e},\underline{q}}$ and so the theorem follows by Proposition 11.

Let $(\underline{t}; \underline{\lambda}, \underline{\mu}) \in A^o_{n_1,n_2,\underline{e},\underline{q},g}$. Acting by elementary moves of type σ'_j we shift to the right the elements of the form $(\bar{1}_*; \mathrm{id})$ transforming our system into

$$\left(\tilde{t}_1, \dots, \tilde{t}_{n_2+2}, (\bar{1}_*; \mathrm{id}), \dots, (\bar{1}_*; \mathrm{id}); \lambda_1, \mu_1, \dots, \lambda_g, \mu_g\right),$$

where $\tilde{t}_i = (*; t'_i)$, $\lambda_k = (*; \lambda'_k)$ and $\mu_k = (*; \mu'_k)$.

We notice that $(t'_1, \dots, t'_{n_2+2}; \lambda'_1, \mu'_1, \dots, \lambda'_g, \mu'_g)$ is the Hurwitz system of a covering of Y of degree $d \geq 3$, with monodromy group S_d and with $n_2 + 2$ branch points, n_2 of which are points of simple branching, one is a special point whose local monodromy is given by $\underline{e}$ and one is a special point whose local monodromy is given by $\underline{q}$.

Since $n_2 - s - r \geq d + 1$, the Hurwitz space $H^o_{d,n_2,\underline{e},\underline{q}}(Y, b_0)$ is irreducible (see [Vetro 2010], Theorem 2) and then the Hurwitz system

$$(t'_1, \dots, t'_{n_2+2}; \lambda'_1, \mu'_1, \dots, \lambda'_g, \mu'_g)$$

is braid-equivalent to one of the form

$$(t''_1, \dots, t''_{n_2+2}; \mathrm{id}, \mathrm{id}, \dots, \mathrm{id}, \mathrm{id}).$$

Hence it follows that $(\underline{t}; \underline{\lambda}, \underline{\mu})$ is braid-equivalent to a system of type

$$(\bar{t}_1, \dots, \bar{t}_{n_2+2}, (\bar{1}_*; \mathrm{id}), \dots; (a_1; \mathrm{id}), (b_1; \mathrm{id}), \dots, (a_g; \mathrm{id}), (b_g; \mathrm{id})).$$

We notice that if $a_h = 0$ and $b_k = 0$ for each $1 \leq h, k \leq g$ the theorem follows by Proposition 11. So let $a_1 \neq 0$ and i be one of the indexes that a_1 sends to $\bar{1}$.

Since it is not restrictive to suppose that among the element of type $(\bar{1}_*; \mathrm{id})$ in our system there is $(\bar{1}_i; \mathrm{id})$ (see Step 1, Proposition 11), acting by elementary moves of type σ''_j we can transform our system into

$$\big((\bar{1}_i; \mathrm{id}), \dots; (a_1; \mathrm{id}), (b_1; \mathrm{id}), \dots, (a_g; \mathrm{id}), (b_g; \mathrm{id})\big).$$

Now we use the move τ''_{11} to replace $(a_1; \mathrm{id})$ with $(\bar{1}_i; \mathrm{id})(a_1; \mathrm{id})$, where $\bar{1}_i + a_1$ is a function that sends i to $\bar{0}$.

So reasoning for all the indexes that a_1 sends to $\bar{1}$, after a finite number of steps, we obtain a new Hurwitz system with (0; id) at the place $(n_2 + n_1 + 3)$.

On the contrary, if $a_1 = 0$, $b_1 \neq 0$ and b_1 sends i to $\bar{1}$, we at first use elementary moves of type σ''_j to bring to the first place $(\bar{1}_i; \mathrm{id})$ and then we act by the braid move ρ'_{11} in order to transform $(b_1; \mathrm{id})$ into $(\bar{1}_i; \mathrm{id})(b_1; \mathrm{id})$ where the function $\bar{1}_i + b_1$ sends i to $\bar{0}$. Following this line for all the indexes that b_1 sent to $\bar{1}$, we can replace $(\bar{1}_i + b_1; \mathrm{id})$ by (0; id).

We notice that if $a_k \neq 0$ and $a_l = b_l = 0$, for each $l \leq k - 1$, in order to obtain the claim one can reason in the same way but this time applying the braid move τ'_{1k}. Analogously if $b_k \neq 0$, $a_l = b_l = 0$, for each $l \leq k - 1$, and $a_k = 0$ one can apply the braid move ρ'_{1k} to transform $(b_k; \mathrm{id})$ into (0; id). □

From Theorem 14 and Remark 6 we deduce the following result.

Corollary 15. *If $n_2 - s - r \geq d + 1$, the Hurwitz space $H_{W(B_d),n_1,n_2,\underline{e},\underline{q}}(Y)$ is irreducible.*

References

[Berstein and Edmonds 1984] I. Berstein and A. L. Edmonds, "On the classification of generic branched coverings of surfaces", *Illinois J. Math.* **28**:1 (1984), 64–82. MR 85k:57004 Zbl 0551.57001

[Biggers and Fried 1986] R. Biggers and M. Fried, "Irreducibility of moduli spaces of cyclic unramified covers of genus g curves", *Trans. Amer. Math. Soc.* **295**:1 (1986), 59–70. MR 87f:14011 Zbl 0601.14022

[Birman 1969] J. S. Birman, "On braid groups", *Comm. Pure Appl. Math.* **22**:1 (1969), 41–72. MR 38 #2764 Zbl 0157.30904

[Bourbaki 1968] N. Bourbaki, *Groupes et algèbres de Lie, IV–VI*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. MR 39 #1590 Zbl 0186.33001

[Carter 1972] R. W. Carter, "Conjugacy classes in the Weyl group", *Compositio Math.* **25**:1 (1972), 1–59. MR 47 #6884 Zbl 0254.17005

[Fadell and Neuwirth 1962] E. Fadell and L. Neuwirth, "Configuration spaces", *Math. Scand.* **10** (1962), 111–118. MR 25 #4537 Zbl 0136.44104

[Fulton 1969] W. Fulton, "Hurwitz schemes and irreducibility of moduli of algebraic curves", *Ann. of Math.* (2) **90**:3 (1969), 542–575. MR 41 #5375 Zbl 0194.21901

[Graber et al. 2002] T. Graber, J. Harris, and J. Starr, "A note on Hurwitz schemes of covers of a positive genus curve", preprint, 2002. arXiv math.AG/0205056

[Hurwitz 1891] A. Hurwitz, “Ueber Riemann’sche Flächen mit gegebenen Verzweigungspunkten”, *Math. Ann.* **39**:1 (1891), 1–60. MR 1510692 Zbl 23.0429.01

[Kanev 2004] V. Kanev, “Irreducibility of Hurwitz spaces”, preprint 241, Dipartimento di Matematica ed Applicazioni, Università degli Studi di Palermo, 2004. arXiv AG/0509154

[Kanev 2006] V. Kanev, “Hurwitz spaces of Galois coverings of $\mathbb{P}^1$, whose Galois groups are Weyl groups”, *J. Algebra* **305**:1 (2006), 442–456. MR 2007g:14032 Zbl 1118.14034

[Kluitmann 1988] P. Kluitmann, “Hurwitz action and finite quotients of braid groups”, pp. 299–325 in *Braids* (Santa Cruz, CA, 1986), edited by J. S. Birman and A. Libgober, Contemp. Math. **78**, Amer. Math. Soc., Providence, RI, 1988. MR 90d:20071 Zbl 0701.20019

[Natanzon 1991] S. M. Natanzon, “Topology of 2-dimensional coverings and meromorphic functions on real and complex algebraic curves”, *Tr. Semin. Vektorn. Tenzorn. Anal.* **23–24** (1991), 79–132. In Russian; translated in *Selecta Math. Soviet.* **12**:3 (1993), 251–291. MR 95f:57005 Zbl 0801.30034

[Scott 1970] G. P. Scott, “Braid groups and the group of homeomorphisms of a surface”, *Proc. Cambridge Philos. Soc.* **68**:3 (1970), 605–617. MR 42 #3786 Zbl 0203.56302

[Vetro 2006] F. Vetro, “Irreducibility of Hurwitz spaces of coverings with one special fiber”, *Indag. Math.* (*N.S.*) **17**:1 (2006), 115–127. MR 2008j:14054 Zbl 1101.14040

[Vetro 2007] F. Vetro, “Irreducibility of Hurwitz spaces of coverings with monodromy groups Weyl groups of type $W(B_d)$”, *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat.* (8) **10**:2 (2007), 405–431. MR 2008f:14043 Zbl 1178.14029

[Vetro 2008a] F. Vetro, “Connected components of Hurwitz spaces of coverings with one special fiber and monodromy groups contained in a Weyl group of type B_d”, *Boll. Unione Mat. Ital.* (9) **1**:1 (2008), 87–103. MR 2009b:57004 Zbl 1200.14053

[Vetro 2008b] F. Vetro, “Irreducibility of Hurwitz spaces of coverings with one special fiber and monodromy group a Weyl group of type D_d”, *Manuscripta Math.* **125**:3 (2008), 353–368. MR 2008j:14055 Zbl 1139.14023

[Vetro 2009] F. Vetro, “On Hurwitz spaces of coverings with one special fiber”, *Pacific J. Math.* **240**:2 (2009), 383–398. MR 2010k:14045 Zbl 1198.14026

[Vetro 2010] F. Vetro, “On irreducibility of Hurwitz spaces of coverings with two special fibers”, 2010. To appear in *Georgian Math. J.*

[Wajnryb 1996] B. Wajnryb, “Orbits of Hurwitz action for coverings of a sphere with two special fibers”, *Indag. Math.* (*N.S.*) **7**:4 (1996), 549–558. MR 99c:14040 Zbl 0881.57001

Received February 1, 2011.

FRANCESCA VETRO
DIPARTIMENTO DI MATEMATICA E INFORMATICA
UNIVERSITÀ DEGLI STUDI DI PALERMO
VIA ARCHIRAFI, 34
90123 PALERMO
ITALY

fvetro@math.unipa.it

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. 1 January 2012

0030-8730(201201)255:1;1-D