Counting and equidistribution in quaternionic Heisenberg groups

By Jouni Parkkonen

Department of Mathematics and Statistics, P.O. Box 35
40014 University of Jyväskylä, FINLAND
e-mail: `jouni.t.parkkonen@jyu.fi`

and Frédéric Paulin

Laboratoire de mathématique d'Orsay, UMR 8628 CNRS
Université Paris-Saclay, 91405 ORSAY Cedex, FRANCE
e-mail: `frederic.paulin@universite-paris-saclay.fr`

(*Received*)

Abstract

We develop the relationship between quaternionic hyperbolic geometry and arithmetic counting or equidistribution applications, that arises from the action of arithmetic groups on quaternionic hyperbolic spaces, especially in dimension 2. We prove a Mertens counting formula for the rational points over a definite quaternion algebra A over $\mathbb{Q}$ in the light cone of quaternionic Hermitian forms, as well as a Neville equidistribution theorem of the set of rational points over A in quaternionic Heisenberg groups. [1]

1. *Introduction*

The two main arithmetic results of this paper are a counting theorem and an equidistribution theorem of rational points with error estimates in quaternionic Heisenberg groups, see Theorems 1·1 and 1·2 below. The proofs use methods and results from quaternionic hyperbolic geometry, arithmetic groups and ergodic theory of the geodesic flow in negatively curved spaces. We refer for instance to [**Bre**, **BeQ**, **Kim**] for related results, and especially to the introductions of [**PaP3**, **PaP6**] for motivations, going back to the Mertens and Neville counting and equidistribution results of Farey fractions. The case of the standard Heisenberg group has been treated in [**PaP6**], but new tools have to be developped in this paper besides dealing with noncommutativity issues.

Let $\mathbb{H}$ be Hamilton's quaternion algebra over $\mathbb{R}$, with $x \mapsto \overline{x}$ its conjugation, $\mathtt{n} : x \mapsto x\overline{x}$ its reduced norm, $\mathtt{tr} : x \mapsto x + \overline{x}$ its reduced trace. Let A be a definite ($A \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{H}$) quaternion algebra over $\mathbb{Q}$, with discriminant D_A and class number h_A. Let $m_A = 24$ if D_A is even, and $m_A = 1$ otherwise. Let $\mathscr{O}$ be a maximal order in A. We denote by ${}_{\mathscr{O}}\langle a, \alpha, c\rangle$ the left ideal of $\mathscr{O}$ generated by $a, \alpha, c \in \mathscr{O}$. See [**Vig**] and Section 2 for

[1] **Keywords:** counting, equidistribution, Mertens formula, quaternionic Heisenberg group, Cygan distance, sub-Riemannian geometry, common perpendicular, quaternionic hyperbolic geometry. **AMS codes:** 11E39, 11F06, 11N45, 20G20, 53C17, 53C22, 53C55

definitions. The 2-step nilpotent group

$$\mathscr{N}(\mathscr{O}) = \{(w_0, w) \in \mathscr{O} \times \mathscr{O} \;:\; \mathtt{tr}(w_0) = \mathtt{n}(w)\} \tag{1·1}$$

with law

$$(w_0, w)(w_0', w') = (w_0 + w_0' + \overline{w}\, w', w + w') \tag{1·2}$$

acts on $\mathscr{O} \times \mathscr{O} \times \mathscr{O}$ by the shears

$$(w_0, w)(a, \alpha, c) = (a + \overline{w}\, \alpha + w_0\, c, \alpha + w\, c, c)\ . \tag{1·3}$$

THEOREM 1·1. *There exists $\kappa > 0$ such that as $s \to +\infty$,*

$$\begin{aligned}\mathrm{Card}\ \ & \mathscr{N}(\mathscr{O})\backslash\big\{(a, \alpha, c) \in \mathscr{O} \times \mathscr{O} \times \mathscr{O} \;:\; {}_{\mathscr{O}}\langle a, \alpha, c\rangle = \mathscr{O},\ \mathtt{tr}(\overline{a}\, c) = \mathtt{n}(\alpha),\ \mathtt{n}(c) \leqslant s\big\} \\ &= \frac{2^3 \cdot 3^6 \cdot 5 \cdot 7\ D_A^4}{\pi^8\ m_A\ |\mathscr{O}^\times|\ \prod_{p|D_A}(p-1)(p^2+1)(p^3-1)}\ s^5\,(1 + \mathrm{O}(s^{-\kappa}))\,.\end{aligned}$$

The quaternionic Heisenberg group

$$\mathbb{H}\mathrm{eis}_7 = \{(w_0, w) \in \mathbb{H} \times \mathbb{H} \;:\; \mathtt{tr}\ w_0 = \mathtt{n}(w)\}\,,$$

with the group law defined by Equation (1·2) is the Lie group of $\mathbb{R}$-points of a $\mathbb{Q}$-group whose group of $\mathbb{Q}$-points is $\mathbb{H}\mathrm{eis}_7 \cap (A \times A)$, and in which $\mathscr{N}(\mathscr{O}) = \mathbb{H}\mathrm{eis}_7 \cap (\mathscr{O} \times \mathscr{O})$ is a (uniform) lattice. We endow it with its Haar measure $\mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}$ normalised in such a way that the total mass of the induced measure on $\mathscr{N}(\mathscr{O})\backslash\mathbb{H}\mathrm{eis}_7$ is $\frac{D_A^2}{4}$. We will explain later on this normalisation. Theorem 1·1 is a counting result of rational points $(ac^{-1}, \alpha c^{-1})$ (analogous to Farey fractions) in $\mathbb{H}\mathrm{eis}_7$, and the following result is a related equidistribution theorem. In this paper, we denote by Δ_x the unit Dirac mass at a point x.

THEOREM 1·2. *As $s \to +\infty$, we have*

$$\begin{aligned}&\frac{\pi^8\ m_A\ |\mathscr{O}^\times|\ \prod_{p|D_A}(p-1)(p^2+1)(p^3-1)}{2^5 \cdot 3^6 \cdot 5 \cdot 7\ D_A^2}\ s^{-5}\ \times \\ &\qquad\qquad \sum_{\substack{(a,\,\alpha,\,c)\in\mathscr{O}\times\mathscr{O}\times\mathscr{O},\ 0<\mathtt{n}(c)\leqslant s \\ \mathtt{tr}(a\,\overline{c})=\mathtt{n}(\alpha),\ {}_{\mathscr{O}}\langle a,\,\alpha,\,c\rangle=\mathscr{O}}} \Delta_{(ac^{-1},\, \alpha c^{-1})} \ \overset{*}{\rightharpoonup}\ \mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}\ .\end{aligned}$$

We refer to Theorems 8·2 and 8·3 in Section 8 for more general results with added congruence properties, and to Remark 8·5 for counting and equidistribution results in higher dimensional quaternionic Heisenberg groups.

The proof of the above arithmetic results strongly rely on quaternionic hyperbolic geometry that we recall and develop in Sections 3 and 6 (see also for instance [**All**, **KiP**, **CaP1**, **Kim**, **Phi**, **CaP2**, **EmK**]). Let q be the quaternionic Hermitian form on $\mathbb{H}^3$ defined by

$$q(z_0, z_1, z_2) = -\,\mathtt{tr}(\,\overline{z_0}\, z_2) + \mathtt{n}(z_1)\ .$$

Let PU_q be its projective unitary group, which is the isometry group of the quaternionic hyperbolic plane $\mathbf{H}^2_{\mathbb{H}}$, realized as the negative cone of q in the right projective plane $\mathbb{P}^2_{\mathrm{r}}(\mathbb{H})$, and normalised in order to have maximal sectional curvature -1. The proofs of Theorems 1·1 and 1·2 use the following two results of independent interests.

The subgroup $\mathrm{PU}_q(\mathscr{O}) = P(\mathrm{GL}_3(\mathscr{O}) \cap U_q)$ of PU_q is an arithmetic lattice, and hence by a standard result of Borel-Harish-Chandra, the orbifold $\mathrm{PU}_q(\mathscr{O})\backslash\mathbf{H}^2_{\mathbb{H}}$ has finitely many ends (also called *cusps*). By adapting Zink's method [**Zin**], we compute their number in Section 4.

THEOREM 1·3. *The number of ends of the quaternionic hyperbolic orbifold* $\mathrm{PU}_q(\mathscr{O})\backslash\mathbf{H}^2_{\mathbb{H}}$ *is equal to the class number* h_A *of* A.

Using Prasad's formula [**Pra**] and adapting Emery's Appendix in [**PaP2**], we compute in Section 5 the volume of the quaternionic hyperbolic manifolds $\mathrm{PU}_q(\mathscr{O})\backslash\mathbf{H}^2_{\mathbb{H}}$. This computation (equivalent by Hirzebruch's proportionality theorem to the computation of its Euler-Poincaré characteristic), is close to, and uses argument from, the paper [**EmK**] of Emery-Kim.

THEOREM 1·4. *The volume of the quaternionic hyperbolic orbifold* $\mathrm{PU}_q(\mathscr{O})\backslash\mathbf{H}^2_{\mathbb{H}}$ *is equal to*

$$\mathrm{Vol}(\mathrm{PU}_q(\mathscr{O})\backslash\mathbf{H}^2_{\mathbb{H}}) = \frac{\pi^4\ m_A}{2^{13}\cdot 3^5\cdot 5^2\cdot 7}\ \prod_{p|D_A}(p-1)(p^2+1)(p^3-1)\ .$$

The analytical tools for the proof of Theorems 1·1 and 1·2 are developped in Section 7, where we give various computations of the measures that appear in the application of the ergodic tools of [**PaP5**], see also [**BrPP**, Chap. 12] where these results are announced.

The Cygan distance on the quaternionic Heisenberg group, the Poisson kernel, the Patterson measures introduced and computed in Section 7, and related quantities, should be useful in potential theory on the quaternionic Heisenberg group and for the study of the hypoelliptic Laplacian in sub-Riemannian geometry, see for instance [**FS**, **Kra**] for the (complex) Heisenberg group.

In the subsequent paper [**PaP7**], we will give geometrical applications of this paper to counting and equidistribution of quaternionic chains in the boundary at infinity of the quaternionic hyperbolic plane. A quaternionic chain is the boundary at infinity of a quaternionic geodesic line (as defined in Section 3). We will also prove a Cartan-type theorem of rigidity for the bijections of $\partial_\infty\mathbf{H}^n_{\mathbb{H}}$ preserving the set of quaternionic chains.

2. *A reminder on quaternion algebras*

In this section, we recall the basic definitions and a few facts on quaternion algebras (4-dimensional central simple algebras), quaternionic linear algebra, and the ideal theory in maximal orders in quaternion algebras. Our main reference for this material is [**Vig**].

Given a skew field D and $n\in\mathbb{N}$, we denote by $\mathbb{P}^n_{\mathrm{r}}(D)$ the right projective space of D of dimension n, that is, the space of lines of the right vector space D^{n+1} over D.[2]

Let $\mathbb{H}$ be Hamilton's quaternion algebra over $\mathbb{R}$, with $x\mapsto\overline{x}$ its conjugation, $\mathtt{n}: x\mapsto x\overline{x}$ its reduced norm, $\mathtt{tr}: x\mapsto x+\overline{x}$ its reduced trace. Note that $\mathtt{n}(xy)=\mathtt{n}(x)\,\mathtt{n}(y)$, $\mathtt{tr}(\overline{x})=\mathtt{tr}(x)$ and $\mathtt{tr}(xy)=\mathtt{tr}(yx)$ for all $x,y\in\mathbb{H}$. Let

$$\mathrm{Im}\ \mathbb{H}=\{x\in\mathbb{H}\ :\ \mathtt{tr}\,x=0\}$$

be the $\mathbb{R}$-subspace of purely imaginary quaternions of $\mathbb{H}$, so that every $x\in\mathrm{Im}\ \mathbb{H}$ satisfies $\overline{x}=-x$. For every $x\in\mathbb{H}$, let $\mathrm{Im}\ x=x-\frac{1}{2}\ \mathtt{tr}(x)=\frac{1}{2}(x-\overline{x})$, which is a purely imaginary quaternion.

[2] For all $x_0,x_1,\dots,x_n\in D$, we have $(x_1,\dots,x_n)x_0=(x_1x_0,\dots,x_nx_0)$.

For every $N \in \mathbb{N} - \{0\}$, we consider the right vector space $\mathbb{H}^N$ over $\mathbb{H}$, on which the group $\mathrm{GL}_N(\mathbb{H})$ acts linearly on the left. For all $w = (w_1, \ldots, w_N)$ and $w' = (w'_1, \ldots, w'_N)$ in $\mathbb{H}^N$, we denote by $\overline{w} \cdot w' = \sum_{p=1}^N \overline{w_p}\, w'_p$ their standard quaternionic Hermitian product,[3] and we define $\mathtt{n}(w) = \overline{w} \cdot w = \sum_{p=1}^N \mathtt{n}(w_p)$. We endow $\mathbb{H}^N$ with the standard Euclidean structure $(w, w') \mapsto \frac{1}{2}\,\mathtt{tr}(\,\overline{w} \cdot w')$. In particular, $\mathbb{H}$ and $\mathrm{Im}\ \mathbb{H}$ are endowed with the Euclidean structure making their standard basis $(1, i, j, k)$ and (i, j, k) orthonormal.

On the right vector space $\mathbb{H} \times \mathbb{H}^{n-1} \times \mathbb{H}$ over $\mathbb{H}$ with coordinates (z_0, z, z_n), let q be the nondegenerate quaternionic Hermitian form[4]

$$q(z_0, z, z_n) = -\,\mathtt{tr}(\,\overline{z_0}\, z_n) + \mathtt{n}(z) \tag{2·1}$$

of Witt signature $(1, n)$, and let $\Phi : \mathbb{H}^{n+1} \times \mathbb{H}^{n+1} \to \mathbb{H}$, defined by

$$\Phi : ((z_0, z, z_n), (z'_0, z', z'_n)) \mapsto -\overline{z_0}\, z'_n - \overline{z_n}\, z'_0 + \overline{z} \cdot z'\,, \tag{2·2}$$

be the associated quaternionic sesquilinear form. An element $x \in \mathbb{H}^{n+1}$ is *isotropic* if $q(x) = 0$.

Throughout this paper, A is a quaternion algebra over $\mathbb{Q}$, which is *definite*, that is, $A \otimes_{\mathbb{Q}} \mathbb{R}$ is isomorphic with $\mathbb{H}$. We fix an identification of $A \otimes_{\mathbb{Q}} \mathbb{R}$ and $\mathbb{H}$ and, accordingly, consider A as a $\mathbb{Q}$-subalgebra of $\mathbb{H}$.

The *reduced discriminant* D_A of A is the product of the primes $p \in \mathbb{N}$ such that $A \otimes_{\mathbb{Q}} \mathbb{Q}_p$ is a division algebra. Two definite quaternion algebras over $\mathbb{Q}$ are isomorphic if and only if they have the same reduced discriminant, which can be any product of an odd number of primes (see [**Vig**, page 74]).

A $\mathbb{Z}$*-lattice* I in A is a finitely generated $\mathbb{Z}$-submodule of A generating A as a $\mathbb{Q}$-vector space. An *order* in A is a unitary subring $\mathscr{O}$ of A which is a $\mathbb{Z}$-lattice. It is invariant by conjugation, since for every $x \in \mathscr{O}$, we have $\overline{x} = -x + \mathtt{tr}\, x$ and by [**Vig**, pages 19-20], we have $\mathtt{tr}\,\mathscr{O} \subset \mathbb{Z} \subset \mathscr{O}$ since every element of $\mathscr{O}$ is integral over $\mathbb{Z}$ and $\mathscr{O}$ is a unitary subring. It is contained in a *maximal order* (for the inclusion). The *left order* of a $\mathbb{Z}$-lattice I is

$$\mathscr{O}_\ell(I) = \{x \in A \ :\ xI \subset I\}\,.$$

From now on, let $\mathscr{O}$ be a maximal order in A. It is well known that the trace map $\mathtt{tr} : \mathscr{O} \to \mathbb{Z}$ is surjective (see, for instance, the proof of Prop. 16 in [**ChP**]). A *left fractional ideal* of $\mathscr{O}$ is a $\mathbb{Z}$-lattice of A whose left order is $\mathscr{O}$. A *left (integral) ideal* of $\mathscr{O}$ is a left fractional ideal of $\mathscr{O}$ contained in $\mathscr{O}$. For any subset B of A, we denote by ${}_{\mathscr{O}}\langle B\rangle$ the left fractional ideal of $\mathscr{O}$ generated by the elements of B. Right fractional ideals are defined analogously. The *inverse* of a left fractional ideal $\mathfrak{m}$ of $\mathscr{O}$ is the right fractional ideal

$$\mathfrak{m}^{-1} = \{x \in A \ :\ \mathfrak{m}\, x\, \mathfrak{m} \subset \mathfrak{m}\} = \{x \in A \ :\ \mathfrak{m}\, x \subset \mathscr{O}\}\,.$$

For all $u, v \in \mathscr{O} - \{0\}$, we have

$$(\mathscr{O}u + \mathscr{O}v)^{-1} = u^{-1}\mathscr{O} \cap v^{-1}\mathscr{O}\,. \tag{2·3}$$

If M is a right $\mathscr{O}$-module, then endowed with the pointwise multiplication by $\mathscr{O}$ on

[3] We have $\overline{w\lambda} \cdot (w'\mu) = \overline{\lambda}\,(\overline{w} \cdot w')\mu$ for all $w, w' \in \mathbb{H}^N$ and $\lambda, \mu \in \mathbb{H}$.

[4] It satisfies $q(x\lambda) = \mathtt{n}(\lambda)\, q(x)$ for all $\lambda \in \mathbb{H}$ and $x \in \mathbb{H}^{n+1}$, since $\mathtt{tr}(uv) = \mathtt{tr}(vu)$ for all $u, v \in \mathbb{H}$.

the left, the $\mathbb{Z}$-module $\operatorname{Hom}_{\mathscr{O}}(M,\mathscr{O})$ (of morphisms of right $\mathscr{O}$-modules from M to $\mathscr{O}$) is a left $\mathscr{O}$-module. We denote by $\widetilde{M}$ the left $\mathscr{O}$-module equal to the $\mathbb{Z}$-module M endowed with the left multiplication by $\mathscr{O}$ defined by $(\lambda, v)\mapsto v\,\overline{\lambda}$. If $\mathfrak{m}$ is a right fractional ideal of $\mathscr{O}$, then the map from $\mathfrak{m}^{-1}$ to $\operatorname{Hom}_{\mathscr{O}}(\mathfrak{m},\mathscr{O})$ defined by $x\mapsto\{y\mapsto xy\}$ is an isomorphism of left $\mathscr{O}$-modules, see for instance [**Rei**, page 192].

Two left fractional ideals $\mathfrak{m}$ and $\mathfrak{m}'$ of $\mathscr{O}$ are isomorphic as left $\mathscr{O}$-modules if and only if $\mathfrak{m}'=\mathfrak{m}c$ for some $c\in A^\times$. A *(left) ideal class* of $\mathscr{O}$ is an equivalence class of left fractional ideals of $\mathscr{O}$ for this equivalence relation. We will denote by ${}_{\mathscr{O}}\mathscr{I}$ the set of ideal classes of $\mathscr{O}$. The *class number* h_A of A is the number of ideal classes of $\mathscr{O}$. It is finite and independent of the maximal order $\mathscr{O}$.

We denote by $\mathscr{O}^\times$ the group of invertible elements (or equivalently of norm 1 elements) of $\mathscr{O}$. Its order is 2, 4 or 6 except that $|\mathscr{O}^\times|=24$ when $D_A=2$ and $|\mathscr{O}^\times|=12$ when $D_A=3$ (see [**Eic**, page 103] for a formula when $h_A=1$).

By for instance [**KO**, Lem. 5.5], the covolume of the $\mathbb{Z}$-lattice $\mathscr{O}$ in the Euclidean vector space $\mathbb{H}$ is

$$\operatorname{Vol}(\mathscr{O}\backslash\mathbb{H})=\frac{D_A}{4}\,. \tag{2·4}$$

3. *Quaternionic hyperbolic space*

In this section, we recall some background on the quaternionic hyperbolic spaces, as mostly contained in [**KiP**], see also [**Phi**]. Note, however, that our conventions differ from those of these references in the sesquilinearity properties of Hermitian products, in the choice of the Hermitian form of Witt signature $(1,n)$, and in the normalisation of the curvature.

We fix $n\in\mathbb{N}-\{0,1\}$. The *Siegel domain* model of the quaternionic hyperbolic n-space $\mathbf{H}^n_{\mathbb{H}}$ is

$$\big\{(w_0,w)\in\mathbb{H}\times\mathbb{H}^{n-1}\;:\;\mathtt{tr}\;w_0-\mathtt{n}(w)>0\big\}\,,$$

endowed with the Riemannian metric

$$ds^2_{\mathbf{H}^n_{\mathbb{H}}}=\frac{1}{(\mathtt{tr}\;w_0-\mathtt{n}(w))^2}\big(\,\mathtt{n}(dw_0-\overline{dw}\cdot w)+(\mathtt{tr}\;w_0-\mathtt{n}(w))\;\mathtt{n}(dw)\,\big)\,. \tag{3·1}$$

Note that this metric is normalised so that its sectional curvatures are in $[-4,-1]$, instead of in $[-1,-\frac14]$ as in [**KiP**] and [**Phi**]. This will facilitate in Section 8 the references to works using that normalisation. Its boundary at infinity is

$$\partial_\infty\mathbf{H}^n_{\mathbb{H}}=\big\{(w_0,w)\in\mathbb{H}\times\mathbb{H}^{n-1}\;:\;\mathtt{tr}\;w_0-\mathtt{n}(w)=0\big\}\cup\{\infty\}\,.$$

A *quaternionic geodesic line* in $\mathbf{H}^n_{\mathbb{H}}$ is the image by an isometry of $\mathbf{H}^n_{\mathbb{H}}$ of the intersection of $\mathbf{H}^n_{\mathbb{H}}$ with the quaternionic line $\mathbb{H}\times\{0\}$. With our normalisation of the metric, a quaternionic geodesic line is a totally geodesic submanifold of real dimension 4 and constant sectional curvature -4.

The Siegel domain $\mathbf{H}^n_{\mathbb{H}}$ embeds in the right quaternionic projective n-space $\mathbb{P}^n_{\rm r}(\mathbb{H})$ by the map (using homogeneous coordinates)

$$(w_0,w)\mapsto[w_0:w:1]\,.$$

By this map, we identify $\mathbf{H}^n_{\mathbb{H}}$ with its image, which when endowed with the isometric Riemannian metric, is called the *projective model* of $\mathbf{H}^n_{\mathbb{H}}$. Note that this image is the *negative cone* of the quaternionic Hermitian form q defined in Equation (2·1), that is

$\{[z_0 : z : z_n] \in \mathbb{P}^n_{\mathrm{r}}(\mathbb{H}) \ : \ q(z_0, z, z_n) < 0\}$. This embedding extends continuously to the boundary at infinity, by mapping the point $(w_0, w) \in \partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\}$ to $[w_0 : w : 1]$ and ∞ to $[1 : 0 : 0]$, so that the image of $\partial_\infty \mathbf{H}^n_{\mathbb{H}}$ in $\mathbb{P}^n_{\mathrm{r}}(\mathbb{H})$ is the *isotropic cone* of q, that is $\{[z_0 : z : z_n] \in \mathbb{P}^n_{\mathrm{r}}(\mathbb{H}) \ : \ q(z_0, z, z_n) = 0\}$.

The distance between two points in the Siegel domain of $\mathbf{H}^n_{\mathbb{H}}$ has an explicit expression using the projective model: If $(w_0, w), (w'_0, w') \in \mathbf{H}^n_{\mathbb{H}}$, with Φ defined in Equation (2·2), then

$$\cosh^2 d((w_0, w), (w'_0, w')) = \frac{\Phi((w_0, w, 1), (w'_0, w', 1)) \ \Phi((w'_0, w', 1), (w_0, w, 1))}{q(w_0, w, 1) \ q(w'_0, w', 1)} , \qquad (3·2)$$

see for example [**Mos**] with the same normalisation of the metric as ours, [**KiP**, page 292] and [**Phi**, Sect. 1.2] with a discussion of the different normalizations of the curvature.

For every $N \in \mathbb{N}$, let I_N be the identity $N \times N$ matrix. Let

$$J = \begin{pmatrix} 0 & 0 & -1 \\ 0 & I_{n-1} & 0 \\ -1 & 0 & 0 \end{pmatrix} ,$$

which differs only up to signs with the matrix J in [**KiP**]. Given a quaternionic matrix $X = (x_{p,p'})_{1 \leqslant p \leqslant r, \, 1 \leqslant p' \leqslant s} \in \mathscr{M}_{r,s}(\mathbb{H})$, let $X^* = (x^*_{p,p'} = \overline{x_{p',p}}\,)_{1 \leqslant p \leqslant s, \, 1 \leqslant p' \leqslant r} \in \mathscr{M}_{s,r}(\mathbb{H})$ be its *conjugate-transpose* matrix. Let

$$\mathrm{U}_q = \{g \in \mathrm{GL}_{n+1}(\mathbb{H}) \ : \ q \circ g = q\} = \{g \in \mathrm{GL}_{n+1}(\mathbb{H}) \ : \ g^* J \, g = J\}$$

be the *unitary group* of q. Its linear action on $\mathbb{H}^{n+1}$ induces a projective action on $\mathbb{P}^n_{\mathrm{r}}(\mathbb{H})$ with kernel its center, which is reduced to $\{\pm I_{n+1}\}$. The *projective unitary group*

$$\mathrm{PU}_q = \mathrm{U}_q \, / \{\pm I_{n+1}\}$$

of q acts faithfully on $\mathbb{P}^n_{\mathrm{r}}(\mathbb{H})$, preserving $\mathbf{H}^n_{\mathbb{H}}$, and its restriction to $\mathbf{H}^n_{\mathbb{H}}$ is the full isometry group of $\mathbf{H}^n_{\mathbb{H}}$. The connected (almost-)simple real Lie groups U_q and PU_q are also denoted by $\mathrm{Sp}(1, n)$ and $\mathrm{PSp}(1, n)$, when the dependence on the choice of q is not important.

We identify any element of $\mathbb{H}^{n-1}$ with its column matrix. If

$$X = \begin{pmatrix} a & \gamma^* & b \\ \alpha & A & \beta \\ c & \delta^* & d \end{pmatrix} \in \mathrm{GL}_{n+1}(\mathbb{H})$$

is a matrix with $a, b, c, d \in \mathbb{H}$, $\alpha, \beta, \gamma, \delta \in \mathbb{H}^{n-1}$ and $A \in \mathscr{M}_{n-1,n-1}(\mathbb{H})$, then

$$J X^* J = \begin{pmatrix} \overline{d} & -\beta^* & \overline{b} \\ -\delta & A^* & -\gamma \\ \overline{c} & -\alpha^* & \overline{a} \end{pmatrix} .$$

The matrix X belongs to U_q if and only if X is invertible with inverse $J X^* J$. In particular,

X belongs to U_q if and only if

$$\begin{cases} c\,\overline{d} - \delta^*\delta + d\,\overline{c} & = 0 \\ a\,\overline{b} - \gamma^*\gamma + b\,\overline{a} & = 0 \\ -\alpha\beta^* + AA^* - \beta\alpha^* & = I_{n-1} \\ c\,\overline{b} - \delta^*\gamma + d\,\overline{a} & = 1 \\ \alpha\,\overline{d} - A\delta + \beta\,\overline{c} & = 0 \\ \alpha\,\overline{b} - A\gamma + \beta\,\overline{a} & = 0\,. \end{cases} \tag{3·3}$$

These equations are the same ones as in the complex hyperbolic case in [**PaP1**, §6.1], up to being careful with the orders of the products; see also [**CaP1**, **CaP2**] with different sign conventions.

By for instance [**KiP**] or the set of equations (3·3), an element $g \in \mathrm{U}_q$ fixes $\infty \in \partial_\infty \mathbf{H}^n_{\mathbb{H}}$ if and only if its $(3,1)$ entry vanishes, or, equivalently, if g is block upper triangular (this is the reason, besides rationality problems, why we chose the quaternionic Hermitian form q rather than a diagonal one). We denote by

$$\mathrm{Sp}(1) = \{x \in \mathbb{H} \ : \ \mathtt{n}(x) = 1\}$$

the subgroup of units (elements of norm one) of $\mathbb{H}^\times$, and

$$\mathrm{Sp}(n-1) = \{g \in \mathrm{GL}_{n-1}(\mathbb{H}) \ : \ g^* g = I_{n-1}\}$$

the compact symplectic group in dimension $n-1$. An easy computation shows that the block upper triangular subgroup of U_q is

$$\mathrm{B}_q = \left\{ \begin{pmatrix} \mu\, r & \zeta^* & \frac{1}{2r}(\mathtt{n}(\zeta) + u)\mu \\ 0 & U & \frac{1}{r}\, U\,\zeta\,\mu \\ 0 & 0 & \frac{\mu}{r} \end{pmatrix} \ : \ \begin{array}{c} \zeta \in \mathbb{H}^{n-1},\ u \in \mathrm{Im}\,\mathbb{H}, \\ U \in \mathrm{Sp}(n-1), \mu \in \mathrm{Sp}(1), r > 0 \end{array} \right\}.$$

Its image $\mathrm{PB}_q = \mathrm{B}_q/\{\pm I_{n+1}\}$ in PU_q is equal to the stabiliser of ∞ in PU_q.

4. *The number of cusps of* $\mathrm{PU}_q(\mathscr{O})$

Let A be a definite quaternion algebra over $\mathbb{Q}$, and let $\mathscr{O}$ be a maximal order in A. Let q be the quaternionic Hermitian form defined in Equation (2·1), whose associated quaternionic sesquilinear form Φ is defined by Equation (2·2). Let $\mathrm{U}_q(\mathscr{O}) = \mathrm{U}_q \cap \mathrm{GL}_{n+1}(\mathscr{O})$, which is (see below) an arithmetic lattice in U_q, and let $\mathrm{PU}_q(\mathscr{O})$ be its image in PU_q. The aim of this section is to describe precisely the structure of the set of ends of the finite volume quaternionic hyperbolic orbifold $\mathrm{PU}_q(\mathscr{O})\backslash \mathbf{H}^n_{\mathbb{H}}$ when $n = 2$.

In this section and the following one, we will need to make explicit the arithmetic structure of $\mathrm{U}_q(\mathscr{O})$. Since J has rational coefficients, we consider the linear algebraic group $\underline{G}$ defined over $\mathbb{Q}$, such that $\underline{G}(\mathbb{Q}) = \{g \in \mathrm{GL}_{n+1}(A) \ : \ g^* J\, g = J\}$, and $\underline{G}(K) = \{g \in \mathrm{GL}_{n+1}(A \otimes_{\mathbb{Q}} K) \ : \ g^* J\, g = J\}$ for every commutative field K with characteristic 0. In particular, $\underline{G}(\mathbb{R}) = \mathrm{U}_q = \mathrm{Sp}(1,n)$ and $\underline{G}(\mathbb{C}) \simeq \mathrm{Sp}_{n+1}(\mathbb{C})$.[5]

Note that the elements of $\underline{G}(\mathbb{Q})$ automatically have reduced norm one in the right A-algebra $\mathscr{M}_{n+1}(A)$. The involution $\tau : x \mapsto \overline{x}$ of the central skew field A over $\mathbb{Q}$ is of the first kind (that is $\tau\,|_{\mathbb{Q}} = \mathrm{id}_{\mathbb{Q}}$) and second type (that is, $A^\tau = \mathbb{Q}$: the fixed point set of τ in A is its center $\mathbb{Q}$). By [**PlR**, §2.3.3] and in particuliar Prop. 2.15 (1) of loc. cit.,

[5] This group is sometimes denoted by $\mathrm{Sp}_{2(n+1)}(\mathbb{C})$, for instance in [**PlR**].

the algebraic group $\underline{G}$ is absolutely connected, (quasi-)simple, simply connected of type C_{n+1}, and denoted by $\mathbf{SU}_{n+1}(A,\Phi)$ in loc. cit. Note that here SU means having reduced norm one. See also [**EmK**, Rem. 2.2].

Considering $\mathscr{O}^{n+1}$ as a $\mathbb{Z}$-lattice of $\mathbb{H}^{n+1}$, we endow $\underline{G}$ with the natural $\mathbb{Z}$-form such that $\underline{G}(\mathbb{Z}) = \mathrm{U}_q(\mathscr{O})$ and $\underline{G}(\mathbb{Z}_p) = \{g \in \mathrm{GL}_{n+1}(\mathscr{O}_p) \;:\; g^* J g = J\}$ for every prime p, where $\mathscr{O}_p = \mathscr{O} \otimes_{\mathbb{Z}} \mathbb{Z}_p$.

Let us recall a few facts that follow from the work of Borel and Harish-Chandra (see for instance [**Bor**, Th. 1.10]). The discrete group $\underline{G}(\mathbb{Z})$ is a lattice in $\underline{G}(\mathbb{R})$. If $\underline{P}$ is a minimal parabolic subgroup of $\underline{G}$ defined over $\mathbb{Q}$ (for instance the stabiliser of ∞), then the set $\mathrm{Par}_{q,\mathscr{O}}$ of parabolic fixed points of $\underline{G}(\mathbb{Z})$ in $\underline{G}(\mathbb{R})/\underline{P}(\mathbb{R}) = \partial_\infty \mathbf{H}^n_{\mathbb{H}}$ is exactly

$$\mathrm{Par}_{q,\mathscr{O}} = \underline{G}(\mathbb{Q})\underline{P}(\mathbb{R}) = \partial_\infty \mathbf{H}^n_{\mathbb{H}} \cap \mathbb{P}^n_{\mathrm{r}}(A)\ .$$

This is the set of isotropic rational projective points in $\mathbb{P}^n_{\mathrm{r}}(\mathbb{H})$, on which $\underline{G}(\mathbb{Z})$ acts with finitely many orbits. In particular, the *set of cusps* $\mathrm{PU}_q(\mathscr{O})\backslash\mathrm{Par}_{q,\mathscr{O}}$ is in bijection with $\underline{G}(\mathbb{Z})\backslash\underline{G}(\mathbb{Q})/\underline{P}(\mathbb{Q})$.

For every right $\mathscr{O}$-submodule M of $\mathscr{O}^{n+1}$, with $\Phi_A : A^{n+1} \times A^{n+1} \to A$ the restriction over A of the (integral over $\mathscr{O}$) sesquilinear form Φ over $\mathbb{H}$, we denote by

$$M^\perp = \{y \in \mathscr{O}^{n+1} \;:\; \forall\, x \in M,\ \Phi_A(x,y) = 0\}$$

the right $\mathscr{O}$-submodule of $\mathscr{O}^{n+1}$ orthogonal to M. Note that $\Phi_A(\mathscr{O}^{n+1} \times \mathscr{O}^{n+1}) = \mathscr{O}$. The Hermitian $\mathscr{O}$-module $(\mathscr{O}^{n+1}, \Phi_A)$ is *unimodular*, that is, the map

$$\begin{aligned} \Theta :\ \widetilde{\mathscr{O}^{n+1}} &\to \mathrm{Hom}_{\mathscr{O}}(\mathscr{O}^{n+1}, \mathscr{O}) \\ z &\mapsto \{z' \mapsto \Phi_A(z,z')\} \end{aligned} \tag{4·1}$$

is an isomorphism of left $\mathscr{O}$-modules. It is indeed clearly an injective morphism of left $\mathscr{O}$-modules. Its surjectivity comes from the fact that the coordinate forms $z' \mapsto z'_n$, $z' \mapsto z'_0$, $z' \mapsto z'_i$ for $1 \leqslant i \leqslant n-1$ are up to signs the images by Θ of the canonical basis elements e_0, e_n and e_i for $1 \leqslant i \leqslant n-1$ respectively.

For every $x = (x_0, x_1, \dots, x_n) \in A^{n+1}$, let ${}_{\mathscr{O}}\langle x\rangle = \mathscr{O}x_0 + \mathscr{O}x_1 + \dots + \mathscr{O}x_n$ be the left fractional ideal of $\mathscr{O}$ generated by $x_0, x_1, \dots, x_n$. The proof of the following result adapts arguments of [**Zin**], where $\mathscr{O}$ is replaced by the ring of integers of an imaginary quadratic field.

PROPOSITION 4·1. *Assume that $n = 2$. For all nonzero isotropic elements x, x' in A^{n+1}, there exists an element in $\mathrm{PU}_q(\mathscr{O})$ sending the image of x in $\mathbb{P}^n_{\mathrm{r}}(A)$ to the one of x' if and only if the left fractional ideals ${}_{\mathscr{O}}\langle x\rangle$ and ${}_{\mathscr{O}}\langle x'\rangle$ have the same class.*

We do not know whether the result remains valid when $n \geqslant 3$.

Proof. The direct implication is immediate. Conversely, let $x = (x_0, \dots, x_n)$ and x' be nonzero isotropic elements of $\mathscr{O}^{n+1}$ such that ${}_{\mathscr{O}}\langle x\rangle$ and ${}_{\mathscr{O}}\langle x'\rangle$ are in the same left ideal class. This means that there is some $c \in A^\times$ such that ${}_{\mathscr{O}}\langle x\rangle = {}_{\mathscr{O}}\langle x'\rangle c = {}_{\mathscr{O}}\langle x'c\rangle$. In particular, this implies that $x'c \in \mathscr{O}^{n+1}$. As we are interested in the images of x and x' in $\mathbb{P}^n_{\mathrm{r}}(A)$, it is therefore sufficient to prove that there exists an element of $\mathrm{U}_q(\mathscr{O})$ sending x to x' if ${}_{\mathscr{O}}\langle x\rangle = {}_{\mathscr{O}}\langle x'\rangle$.

Let $\mathfrak{a} = \{a \in A \;:\; xa \in \mathscr{O}^{n+1}\}$ and $\mathfrak{a}' = \{a \in A \;:\; x'a \in \mathscr{O}^{n+1}\}$, which are (nonzero) right fractional ideals of $\mathscr{O}$, containing 1 since $x, x' \in \mathscr{O}^{n+1}$. By Equation (2·3), we have (omitting $x_i^{-1}\mathscr{O}$ and $\mathscr{O}x_i$ if $x_i = 0$)

$$\mathfrak{a} = x_0^{-1}\mathscr{O} \cap \cdots \cap x_n^{-1}\mathscr{O} = (\mathscr{O}x_0 + \cdots + \mathscr{O}x_n)^{-1} = ({}_{\mathscr{O}}\langle x\rangle)^{-1} = ({}_{\mathscr{O}}\langle x'\rangle)^{-1} = \mathfrak{a}' \ .$$

Composing the map Θ defined in Equation (4·1) with the restriction map to $x\mathfrak{a}$, we have a surjective morphism of left $\mathscr{O}$-modules from $\widetilde{\mathscr{O}^{n+1}}$ to $\mathrm{Hom}_{\mathscr{O}}(x\mathfrak{a}, \mathscr{O})$. Its kernel is the orthogonal subspace $(x\mathfrak{a})^{\perp} = \{z \in \mathscr{O}^{n+1} \ : \ \Phi_A(x, z) = 0\}$, which contains $x\mathfrak{a}$ since x is isotropic.

Let $y \in A$ be such that $\Phi_A(x, y) \neq 0$, which exists since Φ_A is nondegenerate. Up to replacing y by $y\,\Phi_A(x, y)^{-1}$, we may assume that $\Phi_A(x, y) = 1$. Let $\mathfrak{m}$ be the right fractional ideal of $\mathscr{O}$ such that $\mathscr{O}^{n+1} = (x\mathfrak{a})^{\perp} \oplus y\mathfrak{m}$. Composing the explicit isomorphisms of right $\mathscr{O}$-modules

$$\mathfrak{m} \simeq y\mathfrak{m} \simeq \mathscr{O}^{n+1}/(x\mathfrak{a})^{\perp} \simeq \widetilde{\mathrm{Hom}}_{\mathscr{O}}(x\mathfrak{a}, \mathscr{O}) \simeq \widetilde{\mathrm{Hom}}_{\mathscr{O}}(\mathfrak{a}, \mathscr{O}) \simeq \breve{\mathfrak{a}}^{\,-1} \ ,$$

we have $\mathfrak{m} = \breve{\mathfrak{a}}^{\,-1}$. Since $xA \cap \mathscr{O}^{n+1} = x\mathfrak{a} \subset (x\mathfrak{a})^{\perp}$, there exists a right $\mathscr{O}$-submodule M of $\mathscr{O}^{n+1}$ such that

$$\mathscr{O}^{n+1} = x\mathfrak{a} \oplus y\,\breve{\mathfrak{a}}^{\,-1} \oplus M \ .$$

Note that the map $\mathtt{tr} : A \to \mathbb{Q}$ is onto. Since $\Phi_A(y + x\lambda, y + x\lambda) = \Phi_A(y, y) + \mathtt{tr}\,\lambda$, up to replacing y by $y + x\lambda$ for some $\lambda \in A$ such that $\mathtt{tr}\,\lambda = -\Phi_A(y, y)$, which is possible since $\Phi_A(y, y) \in A \cap \mathbb{R} = \mathbb{Q}$, we may assume that $q(x) = q(y) = 0$ and $\Phi_A(x, y) = 1$.

Since $x\mathfrak{a} \oplus y\,\breve{\mathfrak{a}}^{\,-1}$ is unimodular, we may take $M = (x\mathfrak{a} \oplus y\,\breve{\mathfrak{a}}^{\,-1})^{\perp}$. Since $n = 2$, we may write $M = z\mathfrak{b}$ for some $z \in A^{n+1}$ such that $\Phi_A(x, z) = \Phi_A(y, z) = 0$ and some (nonzero) right fractional ideal $\mathfrak{b}$ of $\mathscr{O}$. Since $(\mathscr{O}^{n+1}, \Phi_A)$ is unimodular and $z\mathfrak{b}$ is orthogonal to $x\mathfrak{a} \oplus y\,\breve{\mathfrak{a}}^{\,-1}$, we have $\Phi_A(z\mathfrak{b}, z\mathfrak{b}) = \overline{\mathfrak{b}}\,\mathfrak{b}\,q(z)$ contains 1 and is contained in $\mathscr{O}$, hence is equal to $\mathscr{O}$. Therefore $q(z) = \frac{1}{\mathtt{n}(\mathfrak{b})}$.

Similarly, we have $\mathscr{O}^{n+1} = x'\mathfrak{a} \oplus y'\,\breve{\mathfrak{a}}^{\,-1} \oplus z'\mathfrak{b}'$ with

$$q(x') = q(y') = \Phi_A(x', z') = \Phi_A(y', z') = 0, \quad \Phi_A(x', y') = 1 \quad \text{and} \quad q(z') = \frac{1}{\mathtt{n}(\mathfrak{b}')} \ .$$

In order to prove that $\mathfrak{b} = \mathfrak{b}'$ by applying [**Frö**, Theo. 1], we need an idèlic interpretation of the set $\mathscr{I}_{\mathscr{O}}$ of right fractional ideal classes, see for instance [**Vig**, §III.5.B]. For every prime $p \geqslant 2$, let $A_p = A \otimes_{\mathbb{Q}} \mathbb{Q}_p$ and $\mathscr{O}_p = \mathscr{O} \otimes_{\mathbb{Z}} \mathbb{Z}_p$. Let $J(A)$ be the idèle group of elements $(x_p)_p \in \prod_p A_p^{\times}$ such that $x_p \in \mathscr{O}_p^{\times}$ for all but finitely many primes p. Its subgroup $J(\mathscr{O}) = \prod_p \mathscr{O}_p^{\times}$ acts by translation on the right, and the multiplicative group $A^{\times}$ acts by pointwise multiplication on the left. To every nonzero right fractional ideal I corresponds an idèle $\widehat{I} = (x_p)_p \in J(A)$, where $I \otimes_{\mathbb{Z}} \mathbb{Z}_p = x_p\mathscr{O}_p$ for every prime p, well defined modulo multiplication on the right by an element of $J(\mathscr{O})$. The map $I \mapsto \widehat{I}$ induces a bijection between the set $\mathscr{I}_{\mathscr{O}}$ of right fractional ideal classes and the double coset $A^{\times}\backslash J(A)/J(\mathscr{O})$. Note that $\mathscr{O}^{n+1}$ is a free (hence locally free) $\mathscr{O}$-module.

Now [**Frö**, Theo. 1 (ii)] implies that for all $m \geqslant 2$ and all right fractional ideals $\mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}_1, \ldots, \mathfrak{b}_m$ of $\mathscr{O}$, the right $\mathscr{O}$-modules $\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_m$ and $\mathfrak{b}_1 \oplus \cdots \oplus \mathfrak{b}_m$ are isomorphic if and only if the products of idèles $\widehat{\mathfrak{a}_1} \ldots \widehat{\mathfrak{a}_m}$ and $\widehat{\mathfrak{b}_1} \ldots \widehat{\mathfrak{b}_m}$ represent the same ideal class. Now since $x\mathfrak{a} \oplus y\,\breve{\mathfrak{a}}^{\,-1} \oplus z\mathfrak{b} = \mathscr{O}^{n+1} = x'\mathfrak{a} \oplus y'\,\breve{\mathfrak{a}}^{\,-1} \oplus z'\mathfrak{b}'$, the right $\mathscr{O}$-modules $\mathfrak{a} \oplus \breve{\mathfrak{a}}^{\,-1} \oplus \mathfrak{b}$ and $\mathfrak{a} \oplus \breve{\mathfrak{a}}^{\,-1} \oplus \mathfrak{b}'$ are isomorphic, hence by cancellation, the right fractional ideals $\mathfrak{b}$ and $\mathfrak{b}'$ have the same ideal class. Up to changing $\mathfrak{b}$ and $\mathfrak{b}'$ in their equivalence class, we may assume that $\mathfrak{b} = \mathfrak{b}'$, hence in particular $q(z) = q(z')$.

The map $xa + yc + zb \mapsto x'a + y'c + z'b$ for all $a \in \mathfrak{a}$, $c \in \breve{\mathfrak{a}}^{\,-1}$ and $b \in \mathfrak{b}$ is an isomorphism of right $\mathscr{O}$-modules from $\mathscr{O}^{n+1}$ to itself, preserving the quaternionic Hermitian form q and sending x to x', as wanted. □

From now on, we fix an integral ideal $\mathfrak{m}$ of $\mathscr{O}$, which is bilateral and stable by the conjugation $x \mapsto \overline{x}$, as for instance $\mathfrak{m} = (1+i)\mathscr{O}$ if $\mathscr{O} = \mathbb{Z}[\frac{1+i+j+k}{2}, i, j, k]$ is the Hurwitz maximal order in the Hamilton quaternion algebra $A = \big(\frac{-1,-1}{\mathbb{Q}}\big)$ over $\mathbb{Q}$. The quotient $\mathbb{Z}$-module $\mathscr{O}/\mathfrak{m}$ is then a ring endowed with an anti-involution again denoted by $x \mapsto \overline{x}$. We denote by $\mathrm{U}_q(\mathscr{O}/\mathfrak{m})$ the finite group of $(n+1)\times(n+1)$ invertible matrices in $\mathscr{O}/\mathfrak{m}$, preserving the Hermitian form $-\overline{z_0}z_n - \overline{z_n}z_0 + \sum_{i=1}^{n-1}\overline{z_i}z_i$ on $(\mathscr{O}/\mathfrak{m})^{n+1}$. Let $\mathrm{B}_q(\mathscr{O}/\mathfrak{m})$ be its upper triangular subgroup by blocks $1 \times (n-1) \times 1$. We denote by $\Gamma_{\mathfrak{m}}$ the *Hecke congruence subgroup* of $\mathrm{U}_q(\mathscr{O})$ modulo $\mathfrak{m}$, that is, the preimage of $\mathrm{B}_q(\mathscr{O}/\mathfrak{m})$ by the group morphism $\mathrm{U}_q(\mathscr{O}) \to \mathrm{U}_q(\mathscr{O}/\mathfrak{m})$ of reduction modulo $\mathfrak{m}$. For every subgroup H of U_q, we denote by $\mathrm{P}H$ its image in PU_q.

PROPOSITION 4·2. *If $n = 2$, then*

(1) *the set of parabolic fixed points of $\mathrm{P}\Gamma_{\mathfrak{m}}$ is the set of points in $\partial_\infty \mathbf{H}^n_{\mathbb{H}}$, which is the isotropic cone of q in $\mathbb{P}^n_{\mathrm{r}}(\mathbb{H})$, having homogeneous coordinates that are elements in $\mathscr{O}$;*

(2) *the orbit $\mathrm{P}\Gamma_{\mathfrak{m}} \cdot \infty$ is the set of points in $\partial_\infty \mathbf{H}^n_{\mathbb{H}}$ having homogeneous coordinates in $\mathbb{P}^n_{\mathrm{r}}(\mathbb{H})$ of the form $[a : \alpha : c]$ with $(a, \alpha, c) \in \mathscr{O} \times \mathfrak{m}^{n-1} \times \mathfrak{m}$, $\mathtt{tr}(\overline{a}c) = \mathtt{n}(\alpha)$ and ${}_{\mathscr{O}}\langle a, \alpha, c\rangle = \mathscr{O}$;*

(3) *the map which associates to $[a : \alpha : c] \in \mathbb{P}^n_{\mathrm{r}}(A)$ the class of the left fractional ideal ${}_{\mathscr{O}}\langle a, \alpha, c\rangle$ generated by its homogeneous coordinates induces a bijection from the set of cusps $\mathrm{PU}_q(\mathscr{O})\backslash \mathrm{Par}_{q,\mathscr{O}}$ of $\mathrm{PU}_q(\mathscr{O})$ to the set of left ideal classes ${}_{\mathscr{O}}\mathscr{I}$ of $\mathscr{O}$.*

The number of cusps of $\mathrm{PU}_q(\mathscr{O})$ is hence exactly the class number h_A of A, and in particular is equal to 1 if and only if $D_A = 2, 3, 5, 7, 13$ (see [**Vig**, page 155]). Since the simple real Lie group PU_q has rank one, the set of ends of the quaternionic hyperbolic orbifold $\mathrm{PU}_q(\mathscr{O})\backslash\mathbf{H}^n_{\mathbb{H}}$ is in bijection with the set of cusps of $\mathrm{PU}_q(\mathscr{O})$, and Theorem 1·3 in the Introduction follows.

Proof. (1) By the previously mentioned results of Borel and Harish-Chandra, the result is true if $\mathfrak{m} = \mathscr{O}$, since any element in $\mathbb{P}^n_{\mathrm{r}}(A)$ may be represented by an element of $\mathscr{O}^{n+1}$. As $\mathrm{P}\Gamma_{\mathfrak{m}}$ has finite index in $\mathrm{PU}_q(\mathscr{O})$, the general case follows since a discrete group and a finite index subgroup have the same set of parabolic fixed points.

Since ${}_{\mathscr{O}}\langle 1, 0, 0\rangle = \mathscr{O}$, the assertions (2) when $\mathfrak{m} = \mathscr{O}$ and (3) follow from Assertion (1) and Proposition 4·1. Assertion (2) for any $\mathfrak{m}$ follows by the definition of $\Gamma_{\mathfrak{m}}$, since the image of $(1,0,0)$ by a matrix in $\mathrm{GL}_3(\mathbb{H})$ is its first column, and since a matrix $\begin{pmatrix} a & \gamma^* & b \\ \alpha & A & \beta \\ c & \delta^* & d \end{pmatrix} \in \mathrm{U}_q(\mathscr{O})$ belongs to $\Gamma_{\mathfrak{m}}$ if and only if $\alpha, c \in \mathfrak{m}$, by reducing modulo $\mathfrak{m}$ the equations (3·3). □

5. *The covolume of* $\mathrm{PU}_q(\mathscr{O})$

In this section, we prove Theorem 1·4 in the Introduction, using Prasad's volume formula in [**Pra**] and arguments from [**EmK**].

Let $\mathscr{P}$ be the set of positive primes in $\mathbb{Z}$. For every $p \in \mathscr{P}$, the order $\mathscr{O}_p = \mathscr{O} \otimes_{\mathbb{Z}} \mathbb{Z}_p$ is a maximal order in the quaternion algebra $A_p = A \otimes_{\mathbb{Q}} \mathbb{Q}_p$ over $\mathbb{Q}_p$ (see for instance [**Vig**, page 84]). Let us denote by $\mathtt{v}_p$ the p-adic valuation of $\mathbb{Q}_p$ and by $\mathtt{n}_p$ the reduced norm on A_p. For every $p \in \mathscr{P}$, recall that by the definition of the discriminant D_A of A, if p does not divide D_A, then A_p is isomorphic to $\mathscr{M}_2(\mathbb{Q}_p)$ and otherwise A_p is the (unique up to isomorphism) quaternion algebra over $\mathbb{Q}_p$ that is a division algebra. Furthermore, let us consider the discrete valuation $\nu_p = \frac{1}{2}\,\mathtt{v}_p \circ \mathtt{n}_p$ on A_p (with value group $\frac{1}{2}\,\mathbb{Z}$). It coincides with $\mathtt{v}_p$ on $\mathbb{Q}_p$, which is the reason of the factor $\frac{1}{2}$. The unique maximal order $\mathscr{O}_p$ is equal to the valuation ring of ν_p (see for instance [**Vig**, page 34], which does not have the factor $\frac{1}{2}$, but this does not change the valuation ring $\{x \in A_p : \nu_p(x) \geqslant 0\}$). We fix a uniformiser $\pi_p \in \mathscr{O}_p$ for ν_p: we have $\mathtt{n}_p(\pi_p) = p$ and $\nu_p(\pi_p) = \frac{1}{2}$.

As in the beginning of Section 4, let $\underline{G}$ be the absolutely connected, (quasi-)simple, simply connected algebraic group over $\mathbb{Q}$, endowed with a $\mathbb{Z}$-form, such that $\underline{G}(\mathbb{Z}) = \mathrm{U}_q(\mathscr{O})$ and $\underline{G}(\mathbb{R}) = \mathrm{U}_q$. We assume that $n = 2$ throughout Section 5.

Note that $\underline{G}$ is a $\mathbb{Q}$-form of the split (hence quasi-split) algebraic group $\mathscr{G} = \mathrm{Sp}_3$ over $\mathbb{Q}$, whose type is C_3, by [**PlR**, page 89], since the involution $\tau : x \mapsto \overline{x}$ is of the first kind and second type, Φ is Hermitian and $\underline{G} = \mathbf{SU}_{n+1}(A, \Phi)$ with the notation of loc. cit.. The $\mathbb{Q}$-group $\underline{G}$ is an inner form of $\mathscr{G}$ since the type C_3 has no symmetries in its diagram, by [**PlR**, page 67].

Recall that (see for instance [**Bou**, §V.6.2]) the *exponents* $m_1, \dots, m_r$ of an (irreducible, finite) Coxeter system (W, S) are the positive integers m such that $e^{2i\pi \frac{m}{h}}$ is an eigenvalue of a *Coxeter element* of (W, S) (the product of the elements of S, which has order h) under the standard reflexion representation. The *absolute rank* r of $\mathscr{G}$ and the exponents $m_1, \dots, m_r$ of the (irreducible, finite) Coxeter system (W, S) having the same type as $\mathscr{G}$ are given by

$$r = 3 \quad \text{and} \quad m_1 = 1,\ m_2 = 3,\ m_3 = 5 \tag{5·1}$$

(see for instance [**Pra**, page 96] or the C_n tables of [**Bou**]). Note that (see for instance [**Pra**, §1.5]) the dimension of $\mathscr{G}$ is

$$\dim(\mathscr{G}) = r + \sum_{i=1}^{r} m_i = 21\,. \tag{5·2}$$

Let $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ be the Bruhat-Tits building of $\underline{G}$ over $\mathbb{Q}_p$ (see for instance [**Tit2**, **BrT2**] for the necessary background on Bruhat-Tits theory).

Recall that a subgroup of $\underline{G}(\mathbb{Q}_p)$ is *parahoric* if it is the (full) stabiliser of a simplex of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$, and *maximal parahoric* if it is the stabiliser of a vertex of the building $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$. A vertex $x \in \mathscr{I}_{\underline{G},\mathbb{Q}_p}$ is *special* (see for instance [**Tit2**, §1.9]) if the affine Weyl group of an apartment of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ through x is the semidirect product of its translation subgroup and the stabiliser of x in it. A vertex of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ is *hyperspecial* (see for instance [**Tit2**, §1.10]) if is special and stays special in the building $\mathscr{I}_{\underline{G},\overline{\mathbb{Q}_p}}$ for an unramified closure $\overline{\mathbb{Q}_p}$ of $\mathbb{Q}_p$ for which $\underline{G}$ splits. A maximal parahoric subgroup is *special* (resp. *hyperspecial*) if it is the stabiliser of a special (resp. hyperspecial) vertex of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$

A *coherent family of parahoric subgroups* of $\underline{G}$ is a family $(Y_p)_{p\in\mathscr{P}}$, where Y_p is a parahoric subgroup of $\underline{G}(\mathbb{Q}_p)$ for every p and $Y_p = \underline{G}(\mathbb{Z}_p)$ for p big enough. The *principal lattice* associated with this family is (see [**Pra**, § 3.4]) the subgroup of $\underline{G}(\mathbb{Q})$ consisting of its elements which, when considered as elements of $\underline{G}(\mathbb{Q}_p)$, belong to Y_p, for every $p \in \mathscr{P}$.

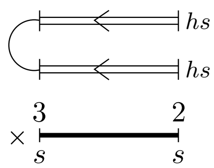
Let $p \in \mathscr{P}$. First assume that p does not divide D_A. Then $\underline{G}$ is isomorphic to the algebraic group $\mathscr{G} = \mathrm{Sp}_3$ over $\mathbb{Q}_p$. The vertices of the building $\mathscr{I}_{\mathscr{G},\mathbb{Q}_p}$ are (see for instance [**BrT2**] or [**She**]) the homothety classes of $\mathbb{Z}_p$-lattices in ${\mathbb{Q}_p}^6$ generated as $\mathbb{Z}_p$-module by the union $\mathscr{B}$ of the standard basis of three orthogonal hyperbolic planes, as for instance with $\mathscr{B}$ the canonical basis for the standard symplectic form on ${\mathbb{Q}_p}^6$. The special vertices of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ are hyperspecial, and correspond to the endpoints of the Dynkin diagram of type C_3, see for instance [**Tit2**, page 60].

LEMMA 5·1. *If p does not divide D_A, then $\underline{G}(\mathbb{Z}_p)$ is hyperspecial parahoric.*

See also [**EmK**, Lem. 5.5] at least when $p \neq 2$: their hypothesis on $(L = \mathscr{O}^3, h = \Phi)$ is not satisfied here, since our form Φ is not diagonal, and making it diagonal requires to invert 2 (which is possible in $\mathbb{Q}_p$ if $p \neq 2$), but their proof only uses the properties of the pair $(L \otimes_{\mathbb{Q}} \mathbb{Q}_p, h)$ at the place p, that stay valid for the diagonalisation of our form Φ (with matrix $\begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$).

Proof. In what follows, we denote by $X \mapsto {}^{t_n}X$ the transposition map of $n \times n$ matrices. Note that $A_p = \mathscr{M}_2(\mathbb{Q}_p)$ and $\mathscr{O}_p = \mathscr{M}_2(\mathbb{Z}_p)$ since p does not divide D_A, and that the conjugation map of the quaternion algebra $\mathscr{M}_2(\mathbb{Q}_p)$ is $x = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto x^\sigma = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ (see for instance [**Vig**, page 3]). Let $J_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so that for every $x \in \mathscr{M}_2(\mathbb{Q}_p)$, we have $x^\sigma = J_0^{-1}\ {}^{t_2}x\, J_0$. Let $J_1 = \begin{pmatrix} 0 & 0 & -I_2 \\ 0 & I_2 & 0 \\ -I_2 & 0 & 0 \end{pmatrix}$ and $J_2 = \begin{pmatrix} J_0 & 0 & 0 \\ 0 & J_0 & 0 \\ 0 & 0 & J_0 \end{pmatrix}$. Considering 3×3 matrices with coefficients in $\mathscr{M}_2(\mathbb{Q}_p)$ as 6×6 matrices, an easy computation shows that for every X in $\mathscr{M}_3(\mathscr{M}_2(\mathbb{Q}_p))$, with X^σ the matrix whose coefficients are the conjugates of the coefficients of X, we have ${}^{t_3}X^\sigma = J_2^{-1}\ {}^{t_6}X\ J_2$. Thus X belongs to $U_q(A_p)$, that is, ${}^{t_3}X^\sigma\ J_1\ X = J_1$, if and only if ${}^{t_6}X\ J_3\ X = J_3$ where $J_3 = J_2 J_1 = \begin{pmatrix} 0 & 0 & -J_0 \\ 0 & J_0 & 0 \\ -J_0 & 0 & 0 \end{pmatrix}$. Note that J_3 is (up to a harmless signed permutation of the canonical basis) the matrix of the standard symplectic product defining $\mathscr{G} = \mathrm{Sp}_3$. We hence have $\underline{G}(\mathbb{Z}_p) = \mathrm{U}_q(\mathscr{O}_p) = \mathrm{Sp}_3(\mathbb{Z}_p)$, which is the stabiliser of the class of the standard $\mathbb{Z}_p$-lattice ${\mathbb{Z}_p}^6$. This class is a special vertex by [**Tit2**, §3.4.2], or since its stabilizer in the affine Weyl group of the apartment defined by the standard symplectic basis of $(\mathbb{Q}_p^6, J_3)$ is the spherical Weyl group of signed permutations of this standard symplectic basis. Since special vertices are hyperspecial for the type C_3, the result follows. $\square$

Now assume that p divides D_A. Then $\underline{G}(\mathbb{Q}_p) = U_q(A_p)$ has local type ${}^2C_3$ in Tits' classification, see [**Tit2**, page 67]. Note that SU = U in our case, as mentioned in [**EmK**, Rem. 2.2]. Its local index is shown below (see [**Tit2**, page 63]):



In particular, $\underline{G}(\mathbb{Q}_p)$ has relative rank 1. By [**Tit2**, Ex. 2.7], the building $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ is a biregular tree of degrees p^3+1 and p^2+1, and all its vertices are special. Hence all maximal parahoric subgroups are special ones.

LEMMA 5·2. *If p divides D_A and $p \neq 2$, then $\underline{G}(\mathbb{Z}_p)$ is special parahoric.*

See also [**EmK**, Lem. 5.6] with a different proof (and the same comment as previously concerning the verification of the hypotheses), our proof being useful in order to deal with the case $p = 2 \mid D_A$.

Proof. We will use the interpretation of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ as the set of minimaximizing norms on A_p^3 (see [**BrT1**, **BrT2**], which uses a $-\log_p$ version of them, in order to allow for infinite residual fields).

With the notation of [**BrT2**, §1.2], we take $K = L := \mathbb{Q}_p$, $D := A_p$ (so that the center of the quaternion algebra D is L), $\sigma : x \mapsto x^\sigma := \overline{x}$ the quaternion conjugation in A_p (so that $L^\sigma = K$), $\varepsilon := +1$, $D^0 = \{x \in A_p \ : \ \mathtt{tr}\ x = 0\}$ (so that D^0 is the set of elements $\xi \in D$ such that $\xi^\sigma + \varepsilon\xi = 0$), $X := A_p^3$ considered as a right vector space over A_p, and $b : X \times X \to D$ the form induced by extension of scalars to $\mathbb{Q}_p$ of the restriction $\Phi_A : A^3 \times A^3 \to A$ to A of the quaternionic sesquilinear form Φ defined in Equation (2·2) with $n = 2$. Hence the form $q : X \to D/D^0 = K$ of loc. cit. (defined by $q(x) = b(x,x) + D^0$) coincides with the extension of scalars to $\mathbb{Q}_p$ of our form $q : A^3 \to \mathbb{Q}$ over $\mathbb{Q}$ defined in Equation (2·1).

With the notation of [**BrT2**, §1.15], we take $\omega_L = \omega := \mathtt{v}_p$, which is a discrete normalised (that is, $\omega(K^\times) = \mathbb{Z}$) valuation on $K = L = \mathbb{Q}_p$ (such that $\omega_L(x^\sigma) = w_L(x)$ for every $x \in L$) and $\omega_D = \nu_p$, which is a discrete valuation (with value group $\frac{1}{2}\mathbb{Z}$) on $D = A_p$ extending $\mathtt{v}_p$ (as explained at the beginning of Section 5).

Let $\lambda \mapsto |\lambda|_p = p^{-\nu_p(\lambda)}$, which is a map from A_p to $[0,+\infty[$, be the unique extension to A_p of the absolute value of $\mathbb{Q}_p$. A *norm* on the right A_p-vector space $X = A_p^3$ is a map $\alpha : X \to [0,+\infty[$ such that

- $\alpha(x) = 0$ if and only if $x = 0$
- $\alpha(x\lambda) = |\lambda|_p\ \alpha(x)$ for all $x \in X$ and $\lambda \in A_p$,
- $\alpha(x+y) \leqslant \max\{\ \alpha(x),\ \alpha(y)\ \}$ for all $x, y \in X$.[6]

Let f (keeping the notation of [**BrT2**, §1.2, Eq. (6)]) be the map from $X \times X$ to $\mathbb{Q}_p$ defined by $f(x,y) = q(x+y) - q(x) - q(y)$. As defined in [**BrT2**, §2.1], a norm α on X *maximizes* f if for all $x, y \in X$, we have

$$|f(x,y)|_p \leqslant \alpha(x)\alpha(y)\ .$$

A norm α on X *maximizes* the pair (f,q) if it maximizes f and if furthermore, for every $x \in X$, we have $|q(x)|_p \leqslant \alpha(x)^2$. Note that if $p \neq 2$ so that $|2|_p = 1$, since $f(x,x) = 2\,q(x)$, then a norm α maximizes f if and only if it maximizes (f,q). A norm α on X *minimaximizes* the pair (f,q) if it is minimal among the norms that maximizes (f,q). The linear action of $\underline{G}(\mathbb{Q}_p) = \mathrm{U}_q(A_p)$ on X induces a left action on the set of norms on X that are minimaximizing for (f,q), by $(g,\alpha) \mapsto \alpha \circ g^{-1}$.

A *Witt basis* of X is a basis (e_{-1}, e_0, e_1) of the right A_p-vector space X such that

$$q(e_{\pm 1}) = f(e_0, e_{\pm 1}) = 0, \quad \text{and} \quad q(e_0) = f(e_1, e_{-1}) = 1\ .$$

[6] Compare with [**BrT1**, §1.1]: a map $\alpha : X \to [0,+\infty[$ is a norm as defined here if and only if $-\log_p \alpha$ is a norm in the sense of [**BrT1**].

By [**BrT2**, Theo. 2.12], there exists a $\underline{G}(\mathbb{Q}_p)$-equivariant bijection from the building $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ to the set of norms on X that are minimaximizing for (f, q). Moreover, by [**BrT2**, §2.9, Prop.], using the fact that the value group of ν_p is $\frac{1}{2}\mathbb{Z}$, for every Witt basis (e_{-1}, e_0, e_1) of X, the sequence $(\alpha_n)_{n\in\mathbb{Z}}$ of norms

$$\alpha_n : \sum_{i=-1}^{1} e_i\lambda_i \mapsto \max\{\, p^{-\frac{n}{2}}|\lambda_{-1}|_p,\ |\lambda_0|_p,\ p^{\frac{n}{2}}|\lambda_1|_p \,\}\,,$$

for $n \in \mathbb{Z}$, is the sequence of norms that are minimaximizing for (f, q) associated with the sequence of vertices $(x_n)_{n\in\mathbb{Z}}$ along an apartment of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$, such that α_0 is associated with x_0. Furthermore, let

$$\mathscr{X}_n = e_{-1}\pi_p^{-n}\mathscr{O}_p + e_0\mathscr{O}_p + e_1\pi_p^{n}\mathscr{O}_p$$

be the right $\mathscr{O}_p$-lattice generated by the Witt basis $(e_{-1}\pi_p^{-n}, e_0\mathscr{O}_p, e_1\pi_p^{n})$, which is the unit ball of the norm α_n, since $|\pi_p|_p = p^{-\frac{1}{2}}$. Then by §3.9 page 180 of [**BrT2**], the smooth affine group scheme $\mathscr{G}_{x_0}$ over $\mathbb{Z}_p$ associated with the vertex x_0 is the schematic closure of $\underline{G}$ in the $\mathbb{Z}_p$-form of the general linear group $\mathrm{GL}_3(A_p)$ over $\mathbb{Q}_p$ defined by the $\mathbb{Z}_p$-lattice $\mathscr{X}_0$, and $\mathscr{G}_{x_0}(\mathbb{Z}_p)$ is the stabiliser of the vertex x_0 in $\underline{G}(\mathbb{Q}_p)$.

Since $p \neq 2$, the element -2 is an invertible element of $\mathbb{Z}_p$ hence of $\mathscr{O}_p$. If (e'_{-1}, e'_0, e'_1) is the canonical basis of $X = A_p^3$, which satisfies $q(e'_{\pm1}) = f(e'_0, e'_{\pm1}) = 0$, $q(e'_0) = 1$ and $f(e'_1, e'_{-1}) = -2$, then $(e_{-1}, e_0, e_1) = (e'_{-1}, e'_0, e'_1\,\frac{1}{-2})$ is a Witt basis of X, and generates the same right $\mathscr{O}_p$-lattice $\mathscr{X}_0$ as the canonical basis. Therefore, $U_q(\mathscr{O}_p) = \underline{G}(\mathbb{Q}_p) \cap \mathrm{GL}_3(\mathscr{O}_p) = \mathscr{G}_{x_0}(\mathbb{Z}_p)$ is maximal parahoric, hence special parahoric. □

REMARK 5·3. *When p divides D_A and $p = 2$, the group $U_q(\mathscr{O}_p)$ is not parahoric. Indeed, again with (e'_{-1}, e'_0, e'_1) the canonical basis of X, since $\mathtt{n}_2(\pi_2) = 2$, the basis $(e_{-1}, e_0, e_1) = (e'_{-1}\pi_2^{-1}, e'_0, -e'_1\pi_2^{-1})$ is a Witt basis of X. With the above notation, $U_q(\mathscr{O}_2)$ is the subgroup of the stabiliser of the special vertex x_0 in $\underline{G}(\mathbb{Q}_2)$ fixing the two edges with origin x_0 and endpoints $x_{\pm1}$, since $\mathscr{X}_{-1} \cap \mathscr{X}_1 = e'_{-1}\mathscr{O}_p + e'_0\,\mathscr{O}_p + e'_1\,\mathscr{O}_p$.*

Thus, by definition, if D_A is odd, the family $(Y_p)_{p\in\mathscr{P}}$ with $Y_p = \underline{G}(\mathbb{Z}_p)$ for every $p \in \mathscr{P}$, is a coherent family of (maximal) parahoric subgroups of $\underline{G}$, and

$$\mathrm{U}_q(\mathscr{O}) = \underline{G}(\mathbb{Z}) = \{g \in \underline{G}(\mathbb{Q}) \;:\; \forall p \in \mathscr{P}, g \in \underline{G}(\mathbb{Z}_p)\}$$

is its associated principal lattice. If D_A is even, the family $(Y_p)_{p\in\mathscr{P}}$ with $Y_p = \underline{G}(\mathbb{Z}_p)$ if $p \neq 2$ and Y_2 the stabiliser in $\underline{G}(\mathbb{Q}_2)$ of the point x_0 defined in Remark 5·3, is a coherent family of special parahoric subgroups of $\underline{G}$. We will compute below the index of $\mathrm{U}_q(\mathscr{O})$ in the associated principal lattice when D_A is even.

For every $p \in \mathscr{P}$,

- let y_p (respectively $\mathfrak{y}_p$) be the vertex of $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ (respectively $\mathscr{I}_{\mathscr{G},\mathbb{Q}_p}$) stabilised by the subgroup $\mathrm{U}_q(\mathscr{O}_p)$ (respectively $\mathrm{Sp}_3(\mathbb{Z}_p)$), such that if D_A is even, then y_2 is the point x_0 defined in Remark 5·3,
- let $\overline{M}_p$ (respectively $\overline{\mathscr{M}}_p$) be the maximal reductive quotient, defined over the residual field $\mathbb{F}_p = \mathbb{Z}_p/p\mathbb{Z}_p$, of the identity component of the reduction modulo p of the smooth affine group scheme over $\mathbb{Z}_p$ associated with y_p (respectively $\mathfrak{y}_p$); see for instance [**Tit2**, §3.5] with $\Omega = \{v\}$.

Note that $\overline{M}_p = \overline{\mathscr{M}}_p$ if p does not divide D_A, and that for every $p \in \mathscr{P}$ the algebraic group $\overline{\mathscr{M}}_p$ is isomorphic to Sp_3 (of type C_3) over $\mathbb{F}_p$. In particular $\overline{\mathscr{M}}_p(\mathbb{F}_p) = \mathrm{Sp}_3(\mathbb{F}_p)$

and thus, for every $p \in \mathscr{P}$, by Equation (5·2) and the orders of the finite groups of Lie type being listed for example in [**Ono**, Table 1], we have

$$\dim \overline{\mathscr{M}}_p = 21 \quad\text{and}\quad |\,\overline{\mathscr{M}}_p(\mathbb{F}_p)\,| = p^9(p^2-1)(p^4-1)(p^6-1)\ . \tag{5·3}$$

Assume now that p divides D_A and $p \neq 2$. Let us consider the pair $(L = \mathscr{O}^3, h = \Phi_{\mathscr{O}})$, where $\Phi_{\mathscr{O}} : L \times L \to \mathscr{O}$ is the restriction to $L \times L$ of the map Φ defined by Equation (2·2) with $n = 2$. Recall that $\mathscr{O}$ is a maximal order in $A = \mathscr{O} \otimes_{\mathbb{Z}} \mathbb{Q}$. The pair (L, h) is a Hermitian right $\mathscr{O}$-module with Witt signature $(1, 2)$, which is unimodular over p (called *regular* over p in [**EmK**]). Recall that this means that the map from $\breve{\mathscr{O}}_p^3$ to $\mathrm{Hom}_{\mathscr{O}_p}(\mathscr{O}_p^3, \mathscr{O}_p)$ is an isomorphism of left $\mathscr{O}_p$-modules. This property is indeed satisfied by [**EmK**, Lem. 5.1] since $p \neq 2$. This restriction $p \neq 2$ is needed since putting h in diagonal form as in [**EmK**, Eq. (5.1)] requires to invert 2 in $\mathscr{O}_p$.

With the notation of [**EmK**, Lem. 4.1], let us consider $k = \mathbb{Q}$, $v = p$, $k_v = \mathbb{Q}_p$, $\mathbf{G} = \underline{G}$ (which does not split over p since p divides D_A), and $P_v^0 = \mathrm{U}_q(\mathscr{O}_p)$, which is a special parahoric group by Lemma 5·2, with type the vertex of the Tits index ${}^2C_3$ distinct from one α_0 coming from the hyperspecial vertices of the Tits index C_3. See also [**EmK**, Lem. 5.6], with the same caveat about the hypotheses as before. Then [**EmK**, Lem. 4.1, Eq. (4.8)] says that

$$p^{(\dim \overline{M}_p - \dim \overline{\mathscr{M}}_p)/2}\ \frac{|\,\overline{\mathscr{M}}_p(\mathbb{F}_p)|}{|\,\overline{M}_p(\mathbb{F}_p)|} = (p-1)(p^2+1)(p^3-1)\ . \tag{5·4}$$

Assume finally that $p = 2$ divides D_A. The Tits index of $\overline{M}_p$, as computed by the rule of [**Tit2**, §3.5.2] is ${}^2A_2$, and by [**Tit2**, §3.5.4], the link of the vertex y_p in the tree $\mathscr{I}_{\underline{G},\mathbb{Q}_p}$ canonically identifies with the spherical building of $\overline{M}_p$ over $\mathbb{F}_p$. By [**Tit1**, page 55], $\overline{M}_p$ is hence the group $\mathrm{U}_q(\mathbb{F}_{p^2})$ where the involution on $\mathbb{F}_{p^2}$ is its Frobenius automorphism $x \mapsto \overline{x} = x^p$. The spherical building of $\overline{M}_p$ is the finite set of isotropic points in the projective plane over $\mathbb{F}_{p^2}$. The vertices of the link of $y_2 = x_0$ corresponding to x_{-1} and x_1 with the notation of Remark 5·3 are the projective points defined by the (isotropic) first and last vectors of the canonical basis of $(\mathbb{F}_{p^2})^3$. Let H be the intersection of the stabilisers in $\mathrm{U}_q(\mathbb{F}_{p^2})$ of the two isotropic points $[1:0:0]$ and $[0:0:1]$. An easy computation shows that H consists of the diagonal matrices $\begin{pmatrix} a & 0 & 0\\ 0 & U & 0\\ 0 & 0 & d\end{pmatrix}$, with $a, U, d \in \mathbb{F}_{p^2}^\times$, $d = \frac{1}{\overline{a}} = a^{-p}$ and $U^{p+1} = U\,\overline{U} = 1$. Since the multiplicative group $\mathbb{F}_{p^2}^\times$ is isomorphic to the additive cyclic group $\mathbb{Z}/((p^2-1)\mathbb{Z})$, which contains exactly $p+1$ elements x such that $(p+1)x = 0$, the order of H is equal to $(p^2-1)(p+1)$. The center of $\mathrm{U}_q(\mathbb{F}_{p^2})$ has order $p+1$, and the quotient by its center is the finite group called the *Steinberg group* ${}^2A_2(p^2)$ (which is simple if $p \neq 2$, and solvable if $p = 2$), whose order is[7]

$$\frac{1}{(3, p+1)}\ p^3(p^2-1)(p^3+1)$$

Hence the index of H in $\mathrm{U}_q(\mathbb{F}_{p^2})$ is $\frac{1}{(3,p+1)}\ p^3(p^3+1)$. By Remark 5·3 and since $p = 2$, we hence have that the index of $U_q(\mathscr{O}_2)$ in the parahoric subgroup Y_2 is

$$[Y_2 : U_q(\mathscr{O}_2)] = [U_q(\mathbb{F}_{2^2}) : H] = 24\ . \tag{5·5}$$

[7] See Table I on page 8 of [**GLS**] where the group ${}^2A_2(p^2)$ in our notation is denoted by ${}^2A_2(q)$ for $q = p$.

Now, let μ be the Haar measure on $\underline{G}(\mathbb{R}) = \mathrm{U}_q = \mathrm{Sp}(1,2)$ normalized as in [**Pra**, §3.6]. The next lemma relates it to the Riemannian measure coming from the choice made in Section 3 of the sectional curvature on $\mathbf{H}^2_{\mathbb{H}}$.

LEMMA 5·4. *We have* $\mathrm{Vol}(\mathrm{PU}_q(\mathscr{O}_q)\backslash\mathbf{H}^2_{\mathbb{H}}) = \frac{\pi^4}{120}\,\mu(\mathrm{U}_q(\mathscr{O}_q)\backslash\mathrm{U}_q)$.

Proof. The proof is similar to the one in [**Eme**] or Emery's appendix of [**PaP2**].

By the definition [**Pra**, §3.6 and §1.4] of μ, if w is the top degree exterior form on the real Lie algebra of $\underline{G}(\mathbb{R})$ whose associated invariant differential form on $\underline{G}(\mathbb{R})$ defines the measure μ and if $G_u = \underline{G}_u(\mathbb{R})$ is the compact real form of $\underline{G}(\mathbb{C})$, then the complexification $w_{\mathbb{C}}$ of w on the complex Lie algebra of $\underline{G}(\mathbb{C}) = \underline{G}_u(\mathbb{C})$ defines a top degree exterior form w_u on the real Lie algebra of G_u, whose associated invariant differential form on G_u defines a measure μ_u, and we require that $\mu_u(G_u) = 1$.

Let μ' be the Haar measure on the noncompact real Lie group $\mathrm{Sp}(1,2)$ that disintegrates by the fibration $\mathrm{Sp}(1,2) \to \mathrm{Sp}(1,2)/(\mathrm{Sp}(1)\times\mathrm{Sp}(2)) = \mathbf{H}^2_{\mathbb{H}}$ with measures on the fibers of total mass one and measure on the base the Riemannian measure $d\mathrm{vol}_{\mathbf{H}^2_{\mathbb{H}}}$ of the Riemannian metric with sectional curvatures contained in $[-4,-1]$, as in Section 3. In particular,

$$\mathrm{Vol}(\mathrm{PU}_q(\mathscr{O}_q)\backslash\mathbf{H}^2_{\mathbb{H}}) = \mu'(\mathrm{U}_q(\mathscr{O}_q)\backslash\mathrm{U}_q)\ .$$

Let μ'_u be the Haar measure on the compact real Lie group $\mathrm{Sp}(3)$ that disintegrates by the fibration $\mathrm{Sp}(3) \to \mathrm{Sp}(3)/(\mathrm{Sp}(1)\times\mathrm{Sp}(2)) = \mathbb{P}^2_{\mathrm{r}}(\mathbb{H})$ with measures on the fibers of total mass one and measure on the base the Riemannian measure $d\mathrm{vol}_{\mathbb{P}^2_{\mathrm{r}}(\mathbb{H})}$ of the Riemannian metric with sectional curvatures contained in $[1,4]$. By [**BGM**, page 112 and Ex.A.III.8], this Riemannian metric is the standard Fubini-Study metric, and

$$\mu'_u(\mathrm{Sp}(3)) = \mathrm{Vol}(\mathbb{P}^2_{\mathrm{r}}(\mathbb{H})) = \frac{\pi^4}{120}\ .$$

The duality between irreducible symmetric spaces of noncompact type endowed with a left invariant Riemannian metric and the ones of compact type sends $\mathbf{H}^2_{\mathbb{H}}$ to $\mathbb{P}^2_{\mathrm{r}}(\mathbb{H})$, with opposite signs on the range of the sectional curvatures (see for instance [**Hel**, Ch. 5]), and hence $\mu' = \frac{\pi^4}{120}\,\mu$. The result follows. □

We now want to apply Prasad's volume formula [**Pra**, Theo. 3.7]. For the notation of this theorem, we take

- $k = \mathbb{Q}$

so that its set of infinite places is $V_\infty = \{\infty\}$ with associated completion $k_\infty = \mathbb{R}$, its set of finite places is $V_f = \mathscr{P}$ with associated (nonarchimedean) completions $k_v = \mathbb{Q}_p$ for every $v = p \in V_f$, with valuation ring $\mathfrak{o}_v = \mathbb{Z}_p$, maximal ideal $\mathfrak{m}_v = p\mathbb{Z}_p$, and the order q_v of the residual field $\mathfrak{f}_v = \mathbb{Z}_p/p\mathbb{Z}_p = \mathbb{F}_p$ is p, and its set of places is $V = V_\infty \cup V_f$,

- $\mathbf{G} = \underline{G}$,

which is an absolutely quasi-simple, simply connected algebraic k-group, which is an inner form of the absolutely quasi-simple, simply connected algebraic k-group $\mathscr{G}$, which is (quasi-)split over k, (and whose absolute rank r and exponents $m_1,\dots,m_r$ have been recalled above), so that the integer $\mathfrak{s}(\mathscr{G})$ associated with $\mathscr{G}$ in [**Pra**, §0.4] is $\mathfrak{s}(\mathscr{G}) = 0$ since $\mathscr{G} = \mathrm{Sp}_3$ splits over $k = \mathbb{Q}$, and the Tamagawa number of $\mathbf{G}$ is

$$\tau_k(\mathbf{G}) = 1$$

by page 109 of op. cit. since k is a number field, $\ell = \mathbb{Q}$ is a smallest splitting field of $\mathscr{G}$ over $\mathbb{Q}$ (since $\mathscr{G}$ is split over $\mathbb{Q}$), and the discriminants of k and ℓ over $\mathbb{Q}$ are $D_k = D_\ell = 1$,

- $S = \{\infty\}$,

which is a finite set of places of k, containing all the Archimedean ones, such that $\mathbf{G}(\mathbb{R}) = U_q$ is non compact, so that $G_S = \prod_{v\in S}\mathbf{G}(k_v) = U_q$ and $S_f = S \cap V_f$ is empty,

- $\mu_S = \mu$,

which is the Haar measure on $G_S = U_q$ normalized as in [**Pra**, §3.6], and

- Λ is the principal S-arithmetic lattice associated with the coherent family of special (maximal) parahoric subgroups $(Y_v)_{v\in V-S}$ constructed after Remark 5·3, which does satisfy the assumptions of [**Pra**, §1.2], since when $\underline{G}$ splits over k_v for $v \in V_f$, then Y_v is hyperspecial by Lemma 5·1.

With this notation, we may now state Prasad's theorem in the special case when S_f is empty, which is the case at hand. Theorem 3.7 of [**Pra**] says that

$$\mu_S(\Lambda\backslash G_S) =$$
$$D_k^{\frac{1}{2}\dim G}\Big(\frac{D_\ell}{D_k^{[\ell:k]}}\Big)^{\frac{1}{2}\mathfrak{s}(\mathscr{G})}\Big(\prod_{v\in S_\infty}\Big|\prod_{i=1}^{r}\frac{(m_i)!}{(2\pi)^{m_i+1}}\Big|_v\Big)\tau_k(\mathbf{G})\prod_{v\notin S}\frac{q_v^{(\dim \overline{M}_v+\dim\overline{\mathscr{M}}_v)/2}}{|\,\overline{M}_v(\mathfrak{f}_v)|}\ . \qquad (5·6)$$

Assume first that D_A is odd, so that $\Lambda = U_q(\mathscr{O})$. Equation (5·6) hence gives, since $\overline{\mathscr{M}}_p = \overline{M}_p$ if p does not divide D_A and by Equation (5·1) for the second equality,

$$\begin{aligned}\mu(\mathrm{U}_q(\mathscr{O})\backslash \mathrm{U}_q) &= \prod_{i=1}^{r}\frac{(m_i)!}{(2\pi)^{m_i+1}}\prod_{p\in\mathscr{P}}\frac{p^{(\dim \overline{M}_p+\dim\overline{\mathscr{M}}_p)/2}}{|\,\overline{M}_p(\mathbb{F}_p)|}\\ &= \frac{720}{(2\pi)^{12}}\prod_{p\in\mathscr{P}}\frac{p^{\dim\overline{\mathscr{M}}_p}}{|\overline{\mathscr{M}}_p(\mathbb{F}_p)|}\prod_{p|D_A}\frac{|\,\overline{\mathscr{M}}_p(\mathbb{F}_p)|}{|\,\overline{M}_p(\mathbb{F}_p)|}\,p^{(\dim\overline{M}_p-\dim\overline{\mathscr{M}}_p)/2}\ . \end{aligned}\qquad (5·7)$$

Using Euler's product formula $\zeta(s) = \prod_{p\in\mathscr{P}}\frac{1}{1-p^{-s}}$ for Riemann's zeta function, we have by Equation (5·3), since $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$ and $\zeta(6) = \frac{\pi^6}{945}$,

$$\prod_{p\in\mathscr{P}}\frac{p^{\dim\overline{\mathscr{M}}_p}}{|\,\overline{\mathscr{M}}_p(\mathbb{F}_p)|} = \zeta(2)\,\zeta(4)\,\zeta(6) = \frac{\pi^{12}}{510300}\ . \qquad (5·8)$$

Thus Theorem 1·4 in the Introduction when D_A is odd follows from Equations (5·7), (5·8), (5·4), and from Lemma 5·4.

Assume now that D_A is even. The right hand side of Equation (5·7) computes the covolume of the principal lattice Λ associated with the coherent family $(Y_p)_{p\in\mathscr{P}}$. By construction, the group $\mathrm{U}_q(\mathscr{O})$ is exactly the subgroup of elements in Λ which, when considered in $\underline{G}(\mathbb{Q}_2)$, belong to the finite index subgroup $\mathrm{U}_q(\mathscr{O}_2)$ of Y_2. Since Λ is dense in Y_2 and $\mathrm{U}_q(\mathscr{O})$ is dense in $\mathrm{U}_q(\mathscr{O}_2)$, this proves that the index m_A of $\mathrm{U}_q(\mathscr{O})$ in Λ is equal to the index of $\mathrm{U}_q(\mathscr{O}_2)$ in Y_2, which is $m_A = 24$ by Equation (5·5). This proves Theorem 1·4 of the introduction when D_A is even.

6. *Horospherical quaternionic hyperbolic geometry*

In this section, we describe the geometry of the horospheres in the quaternionic hyperbolic space $\mathbf{H}^n_{\mathbb{H}}$ (see also [**KiP**, **Phi**]). We introduce the quaternionic Heisenberg group and discuss the geometry of its quaternionic contact structure (see for instance [**Biq**]).

The *horospherical coordinates* $(\zeta, u, t) \in \mathbb{H}^{n-1} \times \operatorname{Im} \mathbb{H} \times [0, +\infty[$, that we will use from now on unless otherwise stated, of $(w_0, w) \in \mathbf{H}^n_{\mathbb{H}} \cup (\partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\})$ are

$$\begin{aligned} (\zeta, u, t) &= (w,\ 2\operatorname{Im} w_0,\ \mathtt{tr}\ w_0 - \mathtt{n}(w)) \\ \text{hence} \quad (w_0, w) &= \Big(\, \frac{\mathtt{n}(\zeta) + t + u}{2},\ \zeta \,\Big)\,, \end{aligned} \tag{6·1}$$

so that the Riemannian metric of $\mathbf{H}^n_{\mathbb{H}}$ is given by

$$ds^2_{\mathbf{H}^n_{\mathbb{H}}} = \frac{1}{4\,t^2}\big(dt^2 + \mathtt{n}(du - 2\operatorname{Im} \overline{d\zeta} \cdot \zeta) + 4\,t\ \mathtt{n}(d\zeta)\,\big)\,. \tag{6·2}$$

In horospherical coordinates, the geodesic lines from $(\zeta, u, 0) \in \partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\}$ to ∞ are, up to translations at the source, the maps $s \mapsto (\zeta, u, e^{2s})$, by the normalisation of the metric.

The *Busemann cocycle* of $\mathbf{H}^n_{\mathbb{H}}$ is the map $\beta : \partial_\infty \mathbf{H}^n_{\mathbb{H}} \times \mathbf{H}^n_{\mathbb{H}} \times \mathbf{H}^n_{\mathbb{H}} \to \mathbb{R}$ defined by

$$(\xi, x, y) \mapsto \beta_\xi(x, y) = \lim_{s \to +\infty} d(\xi_s, x) - d(\xi_s, y)\ ,$$

where $s \mapsto \xi_s$ is any geodesic ray ending at ξ. It is invariant under the diagonal action of the isometry group of $\mathbf{H}^n_{\mathbb{H}}$. The *horosphere* with centre $\xi \in \partial_\infty \mathbf{H}^n_{\mathbb{H}}$ through $x \in \mathbf{H}^n_{\mathbb{H}}$ is $\{y \in \mathbf{H}^n_{\mathbb{H}}\ :\ \beta_\xi(x, y) = 0\}$, and $\{y \in \mathbf{H}^n_{\mathbb{H}}\ :\ \beta_\xi(x, y) \geqslant 0\}$ is the (closed) *horoball* centred at ξ bounded by this horosphere.

Given two points $x = (\zeta, u, t)$ and $x' = (\zeta', u', t')$ in $\mathbf{H}^n_{\mathbb{H}}$, the maps $\xi_s : s \mapsto (\zeta, u, e^{2s})$ and $\xi'_s : s \mapsto (\zeta', u', e^{2s})$ are geodesic lines in $\mathbf{H}^n_{\mathbb{H}}$ through x and x' respectively, converging to ∞ as $s \to +\infty$. The Riemannian length of the affine path from ξ_s to ξ'_s, whose last coordinate is constant and equal to $t = e^{2s}$, is bounded by a constant times $(\frac{t}{t^2})^{\frac{1}{2}} = e^{-s}$, hence $\lim_{s \to +\infty} d(\xi_s, \xi'_s) = 0$. Thus

$$\beta_\infty(x, x') = \frac{1}{2}\ \ln \frac{t'}{t}\ . \tag{6·3}$$

The closed horoballs centred at $\infty \in \partial_\infty \mathbf{H}^n_{\mathbb{H}}$ are therefore the subsets

$$\mathscr{H}_s = \{(\zeta, u, t) \in \mathbf{H}^n_{\mathbb{H}}\ :\ t \geqslant s\} = \{(w_0, w) \in \mathbf{H}^n_{\mathbb{H}}\ :\ \mathtt{tr}\, w_0 - \mathtt{n}(w) \geqslant s\}, \tag{6·4}$$

and the horospheres centred at ∞ are their boundaries $\partial \mathscr{H}_s$, where s ranges in $]0, +\infty[$. Note that, for every $s \geqslant 1$, we have

$$d(\partial \mathscr{H}_1, \partial \mathscr{H}_s) = \frac{\ln s}{2}\ . \tag{6·5}$$

The *Cygan distance*[8] on $\mathbf{H}^n_{\mathbb{H}} \cup (\partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\})$ is (see for instance [**KiP**])

$$d_{\mathrm{Cyg}}((\zeta, u, t), (\zeta', u', t')) = \mathtt{n}\,\big(\ \mathtt{n}(\zeta - \zeta') + |t - t'| + (u - u' - 2\operatorname{Im} \overline{\zeta} \cdot \zeta')\ \big)^{1/4}\,. \tag{6·6}$$

The *quaternionic Heisenberg group* $\mathbb{H}\mathrm{eis}_{4n-1}$ of dimension $4n - 1$ is the real Lie group structure on $\mathbb{H}^{n-1} \times \operatorname{Im} \mathbb{H}$ with law

$$(\zeta, u)(\zeta', u') = (\zeta + \zeta', u + u' + 2\operatorname{Im} \overline{\zeta} \cdot \zeta')$$

and inverses $(\zeta, u)^{-1} = (-\zeta, -u)$. When we have $n = 2$, using the change of coordinates $\zeta = w$ and $u = 2\operatorname{Im} w_0$ as explained in Equation (6·1) with $t = 0$, we recover the definition given in the Introduction. The group $\mathbb{H}\mathrm{eis}_{4n-1}$ identifies with $\partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\}$ by the map

[8] It is analogous to the Euclidean distance on the closure in $\mathbb{R}^n$ of the upper halfspace model of $\mathbb{H}^n_{\mathbb{R}}$.

$(\zeta, u) \mapsto (\zeta, u, 0)$. It furthermore identifies with a subgroup of $\mathrm{PB}_q \subset \mathrm{PU}_q$ by the map $(\zeta, u) \mapsto \pm \begin{pmatrix} 1 & \zeta^* & \frac{\mathtt{n}(\zeta)+u}{2} \\ 0 & I_{n-1} & \zeta \\ 0 & 0 & 1 \end{pmatrix}$, where $\zeta \in \mathscr{M}_{n-1,\,1}(\mathbb{H})$ also denotes the column vector of $\zeta \in \mathbb{H}^{n-1}$. It acts on the space $\mathbf{H}^n_{\mathbb{H}} \cup (\partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\})$ by the *Heisenberg translations*

$$(\zeta, u)(\zeta', u', t') = (\zeta + \zeta',\ u + u' + 2\,\mathrm{Im}\,\overline{\zeta}\cdot\zeta',\ t')\ .$$

They are isometries for both the Riemannian metric and the Cygan distance, and they preserve the horospheres centred at ∞. For every $u \in \mathrm{Im}\,\mathbb{H}$, the Heisenberg translation by $(0, u)$ is called a *vertical translation*.

It is easy to see that the Cygan distance on $\mathbb{H}\mathrm{eis}_{4n-1}$ is the unique left-invariant distance on $\mathbb{H}\mathrm{eis}_{4n-1}$ with

$$d_{\mathrm{Cyg}}((\zeta, u), (0, 0)) = (\mathtt{n}(\zeta)^2 + \mathtt{n}(u))^{\frac{1}{4}}\ ,$$

or equivalently using Equation (6·1) that if $(w_0, w) \in \partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\}$, then

$$d_{\mathrm{Cyg}}((w_0, w), (0, 0)) = (4\,\mathtt{n}(w_0))^{\frac{1}{4}}\ .$$

We conclude this section with geometric lemmas that will be useful in Sections 7 and 8. See also [**Kim**, §3], with slightly different conventions, for a computation similar to Lemma 6·1. The proofs are analogous to those in the complex hyperbolic case with the added ingredient of being careful with the noncommutativity of the multiplication in the present quaternionic case.

LEMMA 6·1. *For all points $x = (\zeta, u, t)$ and $x' = (\zeta', u', t')$ in $\mathbf{H}^n_{\mathbb{H}}$, and for every $(\xi, r) \in \mathbb{H}\mathrm{eis}_{4n-1} = \partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{\infty\}$, we have*

$$\beta_{(\xi,\,r)}(x, x') = \frac{1}{2}\ln \frac{t'\, d_{\mathrm{Cyg}}(x, (\xi, r))^4}{t\, d_{\mathrm{Cyg}}(x', (\xi, r))^4}\ .$$

Proof. It is easy to check that the map $\iota : (w_0, w) \mapsto (w_0^{-1}, w w_0^{-1})$ is an isometric involution of $\mathbf{H}^n_{\mathbb{H}}$ sending $(0, 0) \in \partial_\infty \mathbf{H}^n_{\mathbb{H}}$ to ∞, induced by $\begin{pmatrix} 0 & 0 & 1 \\ 0 & I_{n-1} & 0 \\ 1 & 0 & 0 \end{pmatrix}$, which does belong to U_q. Hence, with $x = (w_0, w)$ and $x' = (w_0', w')$, using Equations (6·3) and (6·1) and the fact that $d_{\mathrm{Cyg}}(x, (0, 0))^4 = 4\,\mathtt{n}(w_0)$ and $d_{\mathrm{Cyg}}(x', (0, 0))^4 = 4\,\mathtt{n}(w_0')$, we have

$$\begin{aligned}\beta_{(0,\,0)}(x, x') = \beta_{\iota(0,\,0)}(\iota x, \iota x') &= \frac{1}{2}\ln \frac{\mathtt{tr}\,(w_0')^{-1} - \mathtt{n}(w'(w_0')^{-1})}{\mathtt{tr}\,w_0^{-1} - \mathtt{n}(w w_0^{-1})} \\ &= \frac{1}{2}\ln \frac{t'\,\mathtt{n}((w_0')^{-1})}{t\,\mathtt{n}(w_0^{-1})} = \frac{1}{2}\ln \frac{t'\, d_{\mathrm{Cyg}}(x, (0, 0))^4}{t\, d_{\mathrm{Cyg}}(x', (0, 0))^4}\ .\end{aligned}$$

The Heisenberg translation τ by (ξ, r) preserves the last horospherical coordinates and the Cygan distances. We have $\beta_{(\xi,\,r)}(x, x') = \beta_{(0,\,0)}(\tau^{-1}x, \tau^{-1}x')$, since τ is an isometry of $\mathbf{H}^n_{\mathbb{H}}$. This proves Lemma 6·1. □

LEMMA 6·2. *The orthogonal projection from $\partial_\infty \mathbf{H}^n_{\mathbb{H}} - \{(0, 0), \infty\}$ to the geodesic line in $\mathbf{H}^n_{\mathbb{H}}$ with points at infinity $(0, 0)$ and ∞ is $(w_0, w) \mapsto (2\,\mathtt{n}(w_0)^{\frac{1}{2}},\ 0)$, that is, in horospherical coordinates, $(\zeta, u, 0) \mapsto \big(0, 0, (\,\mathtt{n}(\zeta)^2 + \mathtt{n}(u)\,)^{1/2}\big)$.*

In particular, the point preimages by this orthogonal projection are the spheres of center $(0,0)$ for the Cygan distance on $\mathbb{H}\mathrm{eis}_{4n-1}$.

Proof. For every parameter a ranging in $]0,+\infty[$, consider the horosphere $\partial\mathscr{H}_a$ centred at ∞. Its image by the isometric involution $\iota:(w_0,w)\mapsto({w_0}^{-1},w{w_0}^{-1})$ is, using Equation (6·1), the horosphere $\{\,(\xi,r,t)\in\mathbf{H}^n_{\mathbb{H}}\;:\;t=\frac{a}{4}\big((\mathtt{n}(\xi)+t)^2+\mathtt{n}(r)\big)\,\}$ centred at $(0,0)$. The image of this horosphere by the Heisenberg translation by (ζ,u) is the horosphere

$$\{\,(\xi,r,t)\in\mathbf{H}^n_{\mathbb{H}}\;:\;t=\frac{a}{4}\,\big((\mathtt{n}(\xi-\zeta)+t)^2+\mathtt{n}(r-u-2\operatorname{Im}\overline{\zeta}\cdot\xi)\big)\,\}$$

centred at (ζ,u). The orthogonal projection of (ζ,u) on the geodesic line ℓ from $(0,0)$ to ∞ is attained when the parameter a gives a double point of intersection $(0,0,t)$ between this horosphere and ℓ. The quadratic equation

$$t=\frac{a}{4}\,((\mathtt{n}(\zeta)+t)^2+\mathtt{n}(u))$$

whose unknown is t has a double solution if and only if its reduced discriminant given by $\Delta'=(\mathtt{n}(\zeta)-\frac{2}{a})^2-(\mathtt{n}(\zeta)^2+\mathtt{n}(u))$ vanishes, that is, since $a>0$, if and only if

$$a=\frac{2}{(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{1/2}+\mathtt{n}(\zeta)}\;,$$

giving $t=(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{1/2}$. The result follows. □

LEMMA 6·3. *Let C be the quaternionic geodesic line $\{(w_0,w)\in\mathbf{H}^n_{\mathbb{H}}\;:\;w=0\}$. The orthogonal projection from $\mathbf{H}^n_{\mathbb{H}}$ to C is the map $(w_0,w)\mapsto(w_0,0)$. On $\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\partial_\infty C$ endowed with the horospherical coordinates, this map extends as $(\zeta,u,0)\mapsto(0,u,\mathtt{n}(\zeta))$.*

Proof. Let $(w_0,w)\in\mathbf{H}^n_{\mathbb{H}}$. It is easy to check that the distance given (see Equation (3·2)) by the formula

$$\cosh^2 d((w_0,w),(w_0',0))=\frac{\mathtt{n}(\,\overline{w_0}+w_0')}{-q(w_0,w,1)\;\mathtt{tr}\,w_0'}$$

is minimised over $(w_0',0)\in C$ exactly when $w_0'=w_0$.

Since C is totally geodesic, the closest point mapping from $\mathbf{H}^n_{\mathbb{H}}$ to C coincides with the orthogonal projection, which is hence $(w_0,w)\mapsto(w_0,0)$. The expression in horospherical coordinates of the boundary extension follows from the equations in (6·1). □

LEMMA 6·4. *For every $(w_0,w)\in\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\{\infty\}$ with $w_0\neq0$, the map from $\mathbb{R}$ to $\mathbf{H}^n_{\mathbb{H}}$ defined by $s\mapsto(w_0(1+2e^{2s}w_0)^{-1},\;w(1+2e^{2s}w_0)^{-1})\in\mathbf{H}^n_{\mathbb{H}}$ is a geodesic line from (w_0,w) to $(0,0)$.*

Proof. The image of $\begin{pmatrix}1&0&0\\ ww_0^{-1}&I_{n-1}&0\\ w_0^{-1}&(ww_0^{-1})^*&1\end{pmatrix}$ in $\mathrm{PGL}_{n+1}(\mathbb{H})$ is the conjugate by the isometric involution $\iota:(w_0',w')\mapsto({w_0'}^{-1},w'{w_0'}^{-1})$ of the Heisenberg translation by $\iota(w_0,w)$. Hence it belongs to PU_q, fixes $\iota(\infty)=(0,0)$ and maps ∞ to (w_0,w). It thus sends the geodesic line from ∞ to $(0,0)$ defined by $s\mapsto(0,0,e^{-2s+\ln 2})$ in horospherical coordinates, hence by $s\mapsto[e^{-2s}:0:1]$ in homogeneous coordinates by Equation (6·1), to a geodesic line from (w_0,w) to $(0,0)$. An easy computation gives that this geodesic line is

$$s\mapsto\big(w_0(1+e^{2s}w_0)^{-1},\;w(1+e^{2s}w_0)^{-1}\big)\;,$$

as wanted, after a time translation. $\square$

LEMMA 6·5. *For all $g \in U_q$ and $s > 0$ such that the horoballs $\mathscr{H}_s$ and $g\mathscr{H}_s$ have disjoint interiors, if c_g is the $(3,1)$-entry of the matrix g, then*

$$d(\mathscr{H}_s, g\mathscr{H}_s) = \frac{1}{2}\ln \mathtt{n}(c_g) + \ln\frac{s}{2}\,.$$

Proof. We follow [**PaP1**, Lem. 6.3]. As seen in Section 3, if we had $c_g = 0$, then g would fix ∞ and would stabilise $\mathscr{H}_s$, which contradicts the assumption. Thus $c_g \neq 0$. Multiplying g on the left and right by elements of $\mathbb{H}\mathrm{eis}_{4n-1}$ does not change c_g or $d(\mathscr{H}_s, g\mathscr{H}_s)$. We may hence assume that $g(\infty) = (0,0)$ and $g^{-1}(\infty) = (0,0)$ (in the coordinates (w_0, w)). Writing $g = \begin{pmatrix} a & \gamma^* & b \\ \alpha & A & \beta \\ c & \delta^* & d \end{pmatrix}$, the first condition implies that $a = 0$ and $\alpha = 0$, and the second one that $\beta = 0$ and $d = 0$. The first and second equations of Formula (3·3) then imply that $\gamma = \delta = 0$, the third one implies that A is unitary, and the fourth one gives $c\bar{b} = 1$. Thus,

$$g = \begin{pmatrix} 0 & 0 & \overline{c}^{-1} \\ 0 & A & 0 \\ c & 0 & 0 \end{pmatrix}$$

with $A \in \mathrm{Sp}(n-1)$. It is easy to check, using the properties of $\mathtt{tr}$ and $\mathtt{n}$, that

$$g\mathscr{H}_s = \{(w_0, w) \in \mathbb{H} \times \mathbb{H}^{n-1} : \mathtt{tr}\, w_0 - \mathtt{n}(w) \geqslant s\,\mathtt{n}(c)\,\mathtt{n}(w_0)\}\,.$$

The points of intersection of the geodesic line from $(0,0)$ to ∞ with the horospheres $\partial\mathscr{H}_s$ (centred at ∞) and $g\,\partial\mathscr{H}_s$ (centred at $(0,0)$) are $(\frac{s}{2}, 0)$ and $(\frac{2}{s\,\mathtt{n}(c)}, 0)$. The distance between them is as required by the statement. $\square$

7. *Measure computations in quaternionic hyperbolic spaces*

Let Γ be a nonelementary discrete group of isometries of $\mathbf{H}^n_{\mathbb{H}}$, let $\Lambda\Gamma$ be its limit set (the smallest closed nonempty Γ-invariant subset of $\partial_\infty \mathbf{H}^n_{\mathbb{H}}$) and let δ_Γ be its critical exponent, defined by

$$\delta_\Gamma = \lim_{n \to +\infty} \frac{1}{n} \ln \mathrm{Card}\,(\Gamma x_0) \cap B(x_0, n)$$

for any $x_0 \in \mathbf{H}^n_{\mathbb{H}}$. We refer to [**BrPP**, Chap. 1] for the background definitions and informations on the notions of this section. In this section, we give proportionality constants relating, on the one hand, Patterson, Bowen-Margulis and skinning measures associated to some convex subsets and, on the other hand, the corresponding Riemannian measures, in the quaternionic hyperbolic case. These results were announced in [**BrPP**, Chap. 7].

We start by briefly recalling the construction of these measures. Let $(\mu_x)_{x \in \mathbf{H}^n_{\mathbb{H}}}$ be a *Patterson density* for Γ, that is a family $(\mu_x)_{x \in \mathbf{H}^n_{\mathbb{H}}}$ of nonzero finite (nonnegative Borel) measures on $\partial_\infty \mathbf{H}^n_{\mathbb{H}}$ whose support is $\Lambda\Gamma$, such that $\gamma_* \mu_x = \mu_{\gamma x}$ and

$$\frac{d\mu_x}{d\mu_y}(\xi) = e^{-\delta_\Gamma \beta_\xi(x,\, y)}$$

for all $\gamma \in \Gamma$, $x, y \in \mathbf{H}^n_{\mathbb{H}}$ and (almost all) $\xi \in \partial_\infty \mathbf{H}^n_{\mathbb{H}}$.

For every $v \in T^1\mathbf{H}^n_{\mathbb{H}}$, let $\pi(v) \in \mathbf{H}^n_{\mathbb{H}}$ be its footpoint, and let v_-, v_+ be the points at $-\infty$ and $+\infty$ of the geodesic line defined by v. Let $x_0 \in \mathbf{H}^n_{\mathbb{H}}$ be a basepoint. The *Bowen-Margulis measure* $\widetilde{m}_{\rm BM}$ for Γ on $T^1\mathbf{H}^n_{\mathbb{H}}$ is defined, using Hopf's parametrisation $v \mapsto (v_-, v_+, \beta_{v_+}(x_0, \pi(v))\,)$ from $T^1\mathbf{H}^n_{\mathbb{H}}$ into $\partial_\infty\mathbf{H}^n_{\mathbb{H}} \times \partial_\infty\mathbf{H}^n_{\mathbb{H}} \times \mathbb{R}$, by

$$d\widetilde{m}_{\rm BM}(v) = e^{-\delta_\Gamma(\beta_{v_-}(\pi(v),\, x_0)+\beta_{v_+}(\pi(v),\, x_0))}\; d\mu_{x_0}(v_-)\, d\mu_{x_0}(v_+)\, dt\,. \tag{7·1}$$

Note that in the right hand side of this equation, $\pi(v)$ may be replaced by any point x' on the geodesic line defined by v, since $\beta_{v_-}(\pi(v),\, x')+\beta_{v_+}(\pi(v),\, x') = 0$. We will use this elementary observation in the proof of Lemma 7·2 *(ii)*. The measure $\widetilde{m}_{\rm BM}$ is nonzero, independent of x_0, is invariant under the geodesic flow, the antipodal map $v \mapsto -v$ and the action of Γ. Thus, it defines a nonzero measure $m_{\rm BM}$ on $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$ which is invariant under the geodesic flow of $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$ and the antipodal map, called the *Bowen-Margulis measure* on $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$.

Let D be a nonempty proper closed convex subset of $\mathbf{H}^n_{\mathbb{H}}$, with stabiliser Γ_D in Γ, such that the family $(\gamma D)_{\gamma\in\Gamma/\Gamma_D}$ is locally finite in $\mathbf{H}^n_{\mathbb{H}}$. We denote by $\partial^1_{\pm}D$ the *outer/inner unit normal bundle* of ∂D, that is, the set of $v \in T^1\mathbf{H}^n_{\mathbb{H}}$ such that $\pi(v) \in \partial D$, $v_\pm \in \partial_\infty\mathbf{H}^n_{\mathbb{H}} - \partial_\infty D$ and the closest point projection on D of $v_\pm$ is $\pi(v)$. Using the endpoint homeomorphism $v \mapsto v_\pm$ from $\partial^1_\pm D$ to $\partial_\infty\mathbf{H}^n_{\mathbb{H}} - \partial_\infty D$, we defined in [**PaP4**] (generalising the definition of Oh and Shah [**OhS**, §1.2] when D is a horoball or a totally geodesic subspace in the real hyperbolic space $\mathbb{H}^n_{\mathbb{R}}$) the outer/inner *skinning measure* $\widetilde{\sigma}^{\pm}_D$ of Γ on $\partial^1_\pm D$, by

$$d\widetilde{\sigma}^{\pm}_D(v) = e^{-\delta_\Gamma\,\beta_{v_\pm}(\pi(v),\, x_0)}\; d\mu_{x_0}(v_\pm)\,. \tag{7·2}$$

The measure $\widetilde{\sigma}^{\pm}_D$ is independent of x_0. It is nonzero if $\Lambda\Gamma$ is not contained in $\partial_\infty D$, and it satisfies $\widetilde{\sigma}^{\pm}_{\gamma D} = \gamma_*\widetilde{\sigma}^{\pm}_D$ for every $\gamma \in \Gamma$. The measure $\sum_{\gamma\in\Gamma/\Gamma_D}\ \gamma_*\widetilde{\sigma}^{\pm}_D$ is a well defined Γ-invariant locally finite measure on $T^1\mathbf{H}^n_{\mathbb{H}}$. Hence, it induces a locally finite measure $\sigma^{\pm}_D$ on $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$, called the outer/inner *skinning measure* of D in $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$. Note that if $\iota : v \mapsto -v$ is the antipodal map, then $\iota_*\widetilde{\sigma}^-_D = \widetilde{\sigma}^+_D$. In particular $\pi_*\widetilde{\sigma}^-_D = \pi_*\widetilde{\sigma}^+_D$, and the measures σ^-_D and σ^+_D have the same total mass.

We will denote the standard Lebesgue measures on the Euclidean spaces $\mathbb{H}^{n-1}$ and $\operatorname{Im}\mathbb{H}$ by $d\zeta$ and du respectively, so that the usual left Haar measure $d\lambda_{4n-1}$ on the Lie group $\mathbb{H}\mathrm{eis}_{4n-1}$ is

$$d\lambda_{4n-1}(\zeta, u) = d\zeta du\,. \tag{7·3}$$

In horospherical coordinates, the volume form of $\mathbf{H}^n_{\mathbb{H}} = \mathbb{H}\mathrm{eis}_{4n-1} \times\,]0, +\infty[$ is[9]

$$d\,\mathrm{vol}_{\mathbf{H}^n_{\mathbb{H}}}(\zeta, u, t) = \frac{1}{16\, t^{2n+2}}\; d\zeta\, du\, dt\ . \tag{7·4}$$

We begin by giving a lemma that relates the Riemannian volume of a Margulis cusp neighbourhood with the Riemannian volume of its boundary, close to [**KiP**, Lem. 3.1].

LEMMA 7·1. *Let D be a horoball in $\mathbf{H}^n_{\mathbb{H}}$ and let Γ be a discrete group of isometries of $\mathbf{H}^n_{\mathbb{H}}$ preserving ∂D (hence D). Then* $\mathrm{Vol}(\Gamma\backslash\partial D) = (4n+2)\,\mathrm{Vol}(\Gamma\backslash D)$.

Proof. Since the group of isometries of $\mathbf{H}^n_{\mathbb{H}}$ acts transitively on the set of horospheres of $\mathbf{H}^n_{\mathbb{H}}$, we may assume that $D = \mathscr{H}_1$. The horosphere centred at ∞ passing through a point $(\zeta, u, t) \in \mathscr{H}_1$ is equal to $\partial\mathscr{H}_t$ and its orthogonal geodesic line at this point is

[9] See also [**KiP**, page 301] with a different normalisation.

$s \mapsto (\zeta, u, e^{2s})$, hence

$$d\,\mathrm{vol}_{\mathbf{H}^n_{\mathbb{H}}}(\zeta,u,t) = d\,\mathrm{vol}_{\partial\mathscr{H}_t}(\zeta,u,t)\,\frac{dt}{2t}\,.$$

By Equation (7·4), we hence have

$$d\,\mathrm{vol}_{\partial\mathscr{H}_t}(\zeta,u,t) = \frac{1}{8\,t^{2n+1}}\;d\zeta\,du \tag{7·5}$$

for every $t>0$, therefore $d\,\mathrm{vol}_{\partial\mathscr{H}_t}(\zeta,u,t) = \frac{1}{t^{2n+1}}\,d\,\mathrm{vol}_{\partial\mathscr{H}_1}(\zeta,u,1)$. The homeomorphism from $\partial\mathscr{H}_t$ to $\partial\mathscr{H}_1$ defined by $(\zeta,u,t)\mapsto(\zeta,u,1)$ commutes with the action of Γ. Thus,

$$\begin{aligned}\mathrm{Vol}(\Gamma\backslash\mathscr{H}_1) &= \int_{\Gamma\backslash\mathscr{H}_1} d\,\mathrm{vol}_{\mathbf{H}^n_{\mathbb{H}}}(\zeta,u,t) = \int_{t=1}^{+\infty}\int_{\Gamma\backslash\partial\mathscr{H}_t} d\,\mathrm{vol}_{\partial\mathscr{H}_t}(\zeta,u,t)\,\frac{dt}{2t}\\ &= \int_{t=1}^{+\infty}\int_{\Gamma\backslash\partial\mathscr{H}_1} d\,\mathrm{vol}_{\partial\mathscr{H}_1}(\zeta,u,1)\,\frac{dt}{2t^{2n+2}} = \frac{1}{4n+2}\;\mathrm{Vol}(\Gamma\backslash\partial\mathscr{H}_1)\,.\end{aligned}$$

□

Let Γ be a lattice in $\mathrm{Isom}(\mathbf{H}^n_{\mathbb{H}})$, that is, a discrete group of isometries of $\mathbf{H}^n_{\mathbb{H}}$ such that the orbifold $\Gamma\backslash\mathbf{H}^n_{\mathbb{H}}$ has finite volume. Its critical exponent is

$$\delta_\Gamma = 4n+2 \tag{7·6}$$

(see for instance [**Cor**, Theo. 4.4 (i)]). The Patterson density $(\mu_x)_{x\in\mathbf{H}^n_{\mathbb{H}}}$ of Γ is uniquely defined up to a multiplicative constant, and is independent of Γ. We will choose the normalisation as follows. Let μ_∞ be the $\mathbb{H}\mathrm{eis}_{4n-1}$-invariant measure on $\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\{\infty\}$ defined (see for instance [**BrPP**, Eq. (7.5)]) by

$$\mu_\infty = \lim_{t\to+\infty} e^{\delta_\Gamma t}\,\mu_{\rho(t)} \tag{7·7}$$

where ρ is the geodesic ray starting from any point in $\partial\mathscr{H}_1$ and converging to ∞. By the uniqueness property of Haar measures on $\mathbb{H}\mathrm{eis}_{4n-1}$, we may uniquely normalise the Patterson density so that μ_∞ coincides with λ_{4n-1} on $\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\{\infty\}=\mathbb{H}\mathrm{eis}_{4n-1}$, that is

$$d\mu_\infty(\xi,r) = d\lambda_{4n-1}(\xi,r) = d\xi\,dr\,.$$

The various computations of Patterson, Bowen-Margulis and skinning measures are gathered in the following statement.

LEMMA 7·2. *Let Γ be a lattice in $\mathrm{Isom}(\mathbf{H}^n_{\mathbb{H}})$, and let $(\mu_x)_{x\in\mathbf{H}^n_{\mathbb{H}}}$ be its Patterson density, normalised as above. For all $x=(\zeta,u,t)$ and $x'=(\zeta',u',t')$ in $\mathbf{H}^n_{\mathbb{H}}$, for all (ξ,r) in $\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\{\infty\}$ and for all v in $T^1\mathbf{H}^n_{\mathbb{H}}$ such that $v_\pm\neq\infty$, we have*

(i) $d\mu_x(\xi,r) = \dfrac{t^{2n+1}}{d_{\mathrm{Cyg}}(x,(\xi,r))^{8n+4}}\,d\xi\,dr\,;$

(ii) using a Hopf parametrisation $v\mapsto(v_-,v_+,s)$,

$$d\widetilde{m}_{\mathrm{BM}}(v) = \frac{d\lambda_{4n-1}(v_-)\,d\lambda_{4n-1}(v_+)\,ds}{d_{\mathrm{Cyg}}(v_-,v_+)^{8n+4}}\,;$$

(iii)

$$\widetilde{m}_{\mathrm{BM}} = \frac{1}{2^{4n-4}}\;\mathrm{vol}_{T^1\mathbf{H}^n_{\mathbb{H}}}\,,$$

and in particular, if $M = \Gamma\backslash\mathbf{H}^n_{\mathbb{H}}$*, the total mass of the Bowen-Margulis measure of* $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$ *is*

$$\|m_{\rm BM}\| = \frac{\pi^{2n}}{2^{4n-5}\,(2n-1)!}\ \operatorname{Vol}(M)\ ;$$

(iv) using the homeomorphism $v \mapsto v_+$ *from* $\partial^1_+\mathscr{H}_1$ *to* $\partial_\infty\mathbf{H}^n_{\mathbb{H}} - \{\infty\} = \mathbb{H}{\rm eis}_{4n-1}$*, we have*

$$d\widetilde{\sigma}^+_{\mathscr{H}_1}(v) = d\lambda_{4n+1}(v_+)\,;$$

for every horoball D *in* $\mathbf{H}^n_{\mathbb{H}}$*, we have*

$$\pi_*\widetilde{\sigma}^{\pm}_D = 8\ {\rm vol}_{\partial D}\ ,$$

and the total mass of the skinning measure of D *in* $\Gamma\backslash T^1\mathbf{H}^n_{\mathbb{H}}$ *is*

$$\|\sigma^{\pm}_D\| = 16(2n+1)\ \operatorname{Vol}(\Gamma_D\backslash D)\ ;$$

(v) for every geodesic line D *in* $\mathbf{H}^n_{\mathbb{H}}$*, we have*

$$d\pi_*\widetilde{\sigma}^{\pm}_D = \frac{2n+1}{2^{4n-2}\,(4n-1)}\ d\pi_*\,{\rm vol}_{\partial^1_\pm D}$$

and, with m *the order of the pointwise stabiliser of* D *in* Γ*,*

$$\|\sigma^{\pm}_D\| = \frac{\pi^{2n-1}\,(2n+1)!}{m\,n\,(4n-1)!}\ \operatorname{Vol}(\Gamma_D\backslash D)\,;$$

(vi) for every quaternionic geodesic line D *in* $\mathbf{H}^n_{\mathbb{H}}$*, we have*

$$d\pi_*\widetilde{\sigma}^{\pm}_D = \frac{1}{2^{4n-1}}\ d\pi_*\,{\rm vol}_{\partial^1_\pm D}$$

and, with m *the order of the pointwise stabiliser of* D *in* Γ*,*

$$\|\sigma^{\pm}_D\| = \frac{\pi^{2n-2}}{m\,2^{4n-2}\,(2n-3)!}\ \operatorname{Vol}(\Gamma_D\backslash D)\,.$$

Proof. In the computations below, it is useful to note that Lemma 6·1 implies that

$$e^{-(4n+2)\,\beta_{(\xi,\,r)}(x,\,x')} = \frac{t^{2n+1}\ d_{\rm Cyg}(x',(\xi,r))^{8n+4}}{(t')^{2n+1}\ d_{\rm Cyg}(x,(\xi,r))^{8n+4}}\,. \tag{7·8}$$

(i) The geodesic line from (ξ,r) to ∞ goes through $\partial\mathscr{H}_1$ at the point $(\xi,r,1)$. For all $\eta\in\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\{\infty\}$, let $x_{\mathscr{H}_1,\,\eta}$ be the intersection point with $\partial\mathscr{H}_1$ of the geodesic line from η to ∞. By the normalisation of $d\mu_\infty$ and by the definition of μ_∞ and the Radon-Nikodym property of the Patterson density, we have

$$\frac{d\mu_x}{d\mu_\infty}(\eta) = e^{-\delta_\Gamma\beta_\eta(x,\,x_{\mathscr{H}_1,\,\eta})}\ ,$$

for all $x\in\mathbf{H}^n_{\mathbb{H}}$ and (almost all) $\eta\in\partial_\infty\mathbf{H}^n_{\mathbb{H}}-\{\infty\}$. Hence we have

$$\frac{d\mu_x}{d\xi\,dr}(\xi,r) = \frac{d\mu_x}{d\mu_\infty}(\xi,r) = e^{-\delta_\Gamma\,\beta_{(\xi,\,r)}(x,\,(\xi,\,r,\,1))}\,.$$

The result then follows from Equations (7·6), (7·8) and (6·6).

(ii) Note that if x' is on the geodesic line ℓ defined by v, since ℓ is asymptotic near v_- to the geodesic line from v_- to ∞, an easy computation using Equation (6·6) shows that

$d_{\rm Cyg}(x',v_-)^2 \sim t'$ as $x' \to v_-$. Hence, by Equation (7·1) and the comment following it, by Equations (7·6) and (7·8), by Assertion *(i)*, and by letting x' converge to v_- on the geodesic line defined by v, we have

$$\begin{aligned}
d\widetilde{m}_{\rm BM}(v) &= e^{-(4n+2)(\beta_{v_-}(x',\,x)+\beta_{v_+}(x',\,x))}\; d\mu_x(v_-)\, d\mu_x(v_+)\, ds\\
&= \Big(\frac{t'\, d_{\rm Cyg}(x,v_-)^4\, t'\, d_{\rm Cyg}(x,v_+)^4\, t^2}{t\, d_{\rm Cyg}(x',v_-)^4\, t\, d_{\rm Cyg}(x',v_+)^4\; d_{\rm Cyg}(x,v_-)^4\, d_{\rm Cyg}(x,v_+)^4}\Big)^{2n+1}\\
&\quad\; d\lambda_{4n-1}(v_-)\, d\lambda_{4n-1}(v_+)\, ds\\
&= \frac{1}{d_{\rm Cyg}(v_-,v_+)^{8n+4}}\; d\lambda_{4n-1}(v_-)\, d\lambda_{4n-1}(v_+)\, ds\,.
\end{aligned}$$

(iii) Recall that the Liouville measure $\operatorname{vol}_{T^1\mathbf{H}^n_{\mathbb{H}}}$ (which is the Riemannian measure for Sasaki's metric on $T^1\mathbf{H}^n_{\mathbb{H}}$) disintegrates under the fibration $\pi : T^1\mathbf{H}^n_{\mathbb{H}} \to \mathbf{H}^n_{\mathbb{H}}$ over the Riemannian measure $\operatorname{vol}_{\mathbf{H}^n_{\mathbb{H}}}$ of $\mathbf{H}^n_{\mathbb{H}}$, with conditional measures the spherical measures on the unit tangent spheres:

$$d\operatorname{vol}_{T^1\mathbf{H}^n_{\mathbb{H}}}(v) = \int_{x\in\mathbf{H}^n_{\mathbb{H}}} d\operatorname{vol}_{T^1_x\mathbf{H}^n_{\mathbb{H}}}(v)\; d\operatorname{vol}_{\mathbf{H}^n_{\mathbb{H}}}(x)\;.$$

Let $x=(\zeta,u,t)\in\mathbf{H}^n_{\mathbb{H}}$. Since the group I_x of isometries of $\mathbf{H}^n_{\mathbb{H}}$ fixing x acts transitively on $T^1_x\mathbf{H}^n_{\mathbb{H}}$, since both μ_x and the Riemannian measure $\operatorname{vol}_{T^1_x\mathbf{H}^n_{\mathbb{H}}}$ are invariant under I_x, using the I_x-equivariant homeomorphism $v\mapsto v_+$ from $T^1_x\mathbf{H}^n_{\mathbb{H}}$ to $\partial_\infty\mathbf{H}^n_{\mathbb{H}}$, we have, for all $v\in T^1_x\mathbf{H}^n_{\mathbb{H}}$ such that $v_+\neq\infty$, using Assertion *(i)* for the last equality,

$$d\operatorname{vol}_{T^1_x\mathbf{H}^n_{\mathbb{H}}}(v) = \frac{\operatorname{Vol}(\mathbb{S}^{4n-1})}{\|\mu_x\|}\; d\mu_x(v_+) = \frac{\operatorname{Vol}(\mathbb{S}^{4n-1})\; t^{2n+1}}{\|\mu_x\|\; d_{\rm Cyg}(x,v_+)^{8n+4}}\; d\lambda_{4n-1}(v_+)\,. \tag{7·9}$$

By homogeneity, by Assertion *(i)*, by Equation (6·6) applied with $(\zeta,u,t)=(0,0,1)$ and $(\zeta',u',t')=(\xi,r,0)$, by using the spherical coordinates in the Euclidean spaces $\mathbb{H}^{n-1}$ and $\operatorname{Im}\mathbb{H}$ of real dimensions $4n-4$ and 3 so that $d\xi = s^{4n-5}ds\, d\operatorname{vol}_{\mathbb{S}^{4n-5}}$ and $dr=\rho^2 d\rho\, d\operatorname{vol}_{\mathbb{S}^2}$, and by using the changes of variables $\rho\mapsto\frac{\rho}{s^2+1}$ and $s\mapsto s^2$, we have

$$\begin{aligned}
\|\mu_x\| = \|\mu_{(0,0,1)}\| &= \int_{\mathbb{H}^{n-1}\times\operatorname{Im}\mathbb{H}} \frac{d\xi\, dr}{((\mathtt{n}(\xi)+1)^2+\mathtt{n}(r))^{2n+1}}\\
&= \operatorname{Vol}(\mathbb{S}^{4n-5})\operatorname{Vol}(\mathbb{S}^2)\iint_0^{+\infty}\frac{s^{4n-5}\rho^2\; ds\, d\rho}{((s^2+1)^2+\rho^2)^{2n+1}}\\
&= \pi\;\operatorname{Vol}(\mathbb{S}^{4n-5})\int_{-\infty}^{+\infty}\frac{\rho^2\; d\rho}{(1+\rho^2)^{2n+1}}\;\int_0^{+\infty}\frac{s^{2n-3}\, ds}{(s+1)^{4n-1}}\;.
\end{aligned}$$

By the residue formula at a pole of order $2n+1$ and by Leibniz formula, considering the map $f: z\mapsto \frac{1}{(z+i)^{2n+1}}$ which satisfies $\frac{\partial^k f}{\partial z^k} = \frac{(-1)^k(2n+k)!}{(2n)!(z+i)^{2n+k+1}}$ for every $k\in\mathbb{N}$, we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\frac{\rho^2\; d\rho}{(\rho^2+1)^{2n+1}} = 2i\pi\operatorname{Res}_{z=i}\frac{z^2}{(z^2+1)^{2n+1}} = 2i\pi\frac{1}{(2n)!}\frac{\partial^{2n}}{\partial z^{2n}}\Big|_{z=i}(z^2 f(z))\\
&= \frac{2i\pi}{(2n)!}\Big(z^2\,\frac{\partial^{2n}f}{\partial z^{2n}} + 4n\, z\,\frac{\partial^{2n-1}f}{\partial z^{2n-1}} + 2n(2n-1)\frac{\partial^{2n-2}f}{\partial z^{2n-2}}\Big)_{z=i} = \frac{\pi\, n\,(4n-2)!}{2^{4n-2}((2n)!)^2}\;.
\end{aligned}$$

By integration by parts and by induction, we have

$$\int_0^{+\infty} \frac{s^{2n-3}\,ds}{(s+1)^{4n-1}} = \frac{(2n-3)!}{(4n-2)\dots(2n+2)} \int_0^{+\infty} \frac{ds}{(s+1)^{2n+2}} = \frac{(2n-3)!(2n)!}{(4n-2)!}\;.$$

Since $\operatorname{Vol}(\mathbb{S}^{4n-1}) = \frac{\pi^2}{(2n-1)(2n-2)} \operatorname{Vol}(\mathbb{S}^{4n-5})$, we hence have

$$\|\mu_x\| = \frac{1}{2^{4n-1}} \operatorname{Vol}(\mathbb{S}^{4n-1})\;.$$

Hence, by Equations (7·4) and (7·9), using the homeomorphism $v \mapsto (v_+, \pi(v) = (\zeta, u, t))$ from $T^1\mathbf{H}^n_{\mathbb{H}}$ to $\partial_\infty \mathbf{H}^n_{\mathbb{H}} \times \mathbf{H}^n_{\mathbb{H}}$, we have, for all $v \in T^1\mathbf{H}^n_{\mathbb{H}}$ such that $v_+ \neq \infty$,

$$d\operatorname{vol}_{T^1\mathbf{H}^n_{\mathbb{H}}}(v) = \frac{2^{4n-5}}{t\, d_{\rm Cyg}((\zeta,u,t), v_+)^{8n+4}}\, d\lambda_{4n-1}(v_+)\, d\zeta\, du\, dt\,. \tag{7·10}$$

Now, let us consider the map $F : \mathbb{H}\mathrm{eis}_{4n-1} \times \mathbb{R} \to \mathbf{H}^n_{\mathbb{H}}$ defined by

$$\begin{aligned}(\xi, r, s) \mapsto \Big(\ & \zeta = \xi\,(1 + (\mathtt{n}(\xi) + r)\, e^{2s})^{-1},\\ & u = \operatorname{Im}\big((\mathtt{n}(\xi)+r)(1+(\mathtt{n}(\xi)+r)\,e^{2s})^{-1}\big),\\ & t = \frac{(\mathtt{n}(\xi)^2 + \mathtt{n}(r))\,e^{2s}}{\mathtt{n}(1+(\mathtt{n}(\xi)+r)\,e^{2s})}\ \Big)\,.\end{aligned}$$

Note that $F(0, i, 0) = (0, \frac{i}{2}, \frac{1}{2})$. By Lemma 6·4 and Equation (6·1), the map $s \mapsto F(\xi, r, s)$ is a geodesic line in $\mathbf{H}^n_{\mathbb{H}}$ starting from $(\xi, r, 0)$ and ending at $(0, 0, 0)$. On this geodesic line, s and the time parameter in Hopf's parametrisation differ only by an additive constant, hence have the same differential.

Recall that by homogeneity, the two measures $\widetilde{m}_{\rm BM}$ and $\operatorname{vol}_{T^1\mathbf{H}^n_{\mathbb{H}}}$ are proportional. Hence we compute their (constant) Radon-Nikodym derivative at the unit tangent vector $v \in T^1\mathbf{H}^n_{\mathbb{H}}$ such that $v_- = (0, i, 0)$ and $\pi(v) = (0, \frac{i}{2}, \frac{1}{2})$, so that v is tangent to the geodesic line $s \mapsto F(0, i, s)$ at $s = 0$, hence $v_+ = (0, 0, 0)$. By Assertion *(ii)* with $(\xi, r, 0)$ parametrising v_- and by Equations (7·10) and (6·6), we have

$$\begin{aligned}\frac{d\operatorname{vol}_{T^1\mathbf{H}^n_{\mathbb{H}}}}{d\widetilde{m}_{\rm BM}} &= \frac{2^{4n-5}\, d_{\rm Cyg}((0,i,0),(0,0,0))^{8n+4}}{\frac{1}{2}\, d_{\rm Cyg}((0,\frac{i}{2},\frac{1}{2}),(0,0,0))^{8n+4}}\ \frac{d\zeta\,du\,dt}{d\xi\,dr\,ds}(0,i,0)\\ &= 2^{6n-3}\ \frac{d\zeta\,du\,dt}{d\xi\,dr\,ds}(0,i,0)\,.\end{aligned}$$

Let us compute the Jacobian at $(0, i, 0)$ of the map $F : (\xi, r, s) \mapsto (\zeta, u, t)$. Taking the derivatives at the point $(0, i, 0)$, we have, using the canonical basis i, j, k of $\operatorname{Im}\mathbb{H}$ in order to write $r = r_1 i + r_2 j + r_3 k$ and $u = u_1 i + u_2 j + u_3 k$,

$$\frac{\partial\zeta}{\partial\xi} = \frac{1}{1+i}\operatorname{Id}_{\mathbb{H}^{n-1}},\quad \frac{\partial\zeta}{\partial r} = \frac{\partial\zeta}{\partial s} = 0,\quad \frac{\partial t}{\partial s} = \frac{\partial t}{\partial\xi} = 0\,,\quad \frac{\partial u}{\partial\xi} = 0\,,$$

$$\frac{\partial u}{\partial r} = \begin{pmatrix}0&0&0\\0&\frac{1}{2}&0\\0&0&\frac{1}{2}\end{pmatrix},\quad \frac{\partial u}{\partial s} = \begin{pmatrix}-1\\0\\0\end{pmatrix},\quad \frac{\partial t}{\partial r} = \begin{pmatrix}\frac{1}{2}&0&0\end{pmatrix}\,.$$

Since the Jacobian matrix of F at $(0, i, 0)$ is block diagonal when the variables are separated into the $4(n-1)$ first ones and the last 4 ones, since the multiplication by $\frac{1}{1+i}$ in $\mathbb{H}^{n-1}$ is a Euclidean homothety of ratio $\frac{1}{\sqrt{2}}$, and since the determinant of the 4×4 matrix of the partial derivatives of u, t with respect to r, s has absolute value $\frac{1}{8}$, the Jacobian of F at $(0, i, 0)$ is equal to $\big(\frac{1}{\sqrt{2}}\big)^{4(n-1)}\frac{1}{8} = \frac{1}{2^{2n+1}}$. The first claim of Assertion *(iii)* follows.

The second claim follows from the facts that $\operatorname{Vol}(T^1M) = \operatorname{Vol}(\mathbb{S}^{4n-1})\ \operatorname{Vol}(M)$ and that $\operatorname{Vol}(\mathbb{S}^{4n-1}) = \frac{2\,\pi^{2n}}{(2n-1)!}$.

(iv) By the definition of the skinning measure $\widetilde{\sigma}^{+}_{\mathscr{H}_1}$ in Equation (7·2) and of the measure μ_∞ in Equation (7·7), we have

$$d\widetilde{\sigma}^{+}_{\mathscr{H}_1}(v) = d\mu_\infty(v_+)$$

for every $v \in \partial^1_+\mathscr{H}_1$, since $\beta_{v_+}(\pi(v), \rho(t)) = -t + \mathrm{o}(1)$ as $t \to +\infty$. The first claim of Assertion *(iv)* follows by the normalisation of the Patterson density.

By Equation (7·5), we have

$$d\operatorname{vol}_{\partial\mathscr{H}_1}(\zeta, u, 1) = \frac{1}{8}\, d\zeta\, du\,. \tag{7·11}$$

Hence $\pi_*\widetilde{\sigma}^{\pm}_{\mathscr{H}_1} = 8\ \operatorname{vol}_{\partial\mathscr{H}_1}$, and by the transitivity of the isometry group of $\mathbf{H}^n_{\mathbb{H}}$ on the set of horoballs in $\mathbf{H}^n_{\mathbb{H}}$, the second claim of Assertion *(iv)* follows. Therefore, by Lemma 7·1,

$$\|\sigma^{\pm}_D\| = \|\pi_*\sigma^{\pm}_D\| = 8\ \operatorname{Vol}(\Gamma_D\backslash\partial D) = 16(2n+1)\ \operatorname{Vol}(\Gamma_D\backslash D)\,.$$

(v) By the transitivity of the isometry group of $\mathbf{H}^n_{\mathbb{H}}$ on the set of its geodesic lines, we may assume that D is the geodesic line in $\mathbf{H}^n_{\mathbb{H}}$ with points at infinity $(0,0)$ and ∞. The map from the full-measure open subset $\big\{(\zeta, u) \in \mathbb{H}\mathrm{eis}_{4n-1}\ :\ \zeta \neq 0,\ u \neq 0\big\}$ in $\mathbb{H}\mathrm{eis}_{4n-1}$ to the product manifold $\mathbb{S}^{4n-5} \times \mathbb{S}^2 \times\,]0, +\infty[\, \times\,]0, \frac{\pi}{2}[$ defined by

$$(\zeta, u) \mapsto \Big(\sigma = \frac{\zeta}{\mathtt{n}(\zeta)^{\frac{1}{2}}},\ w = \frac{u}{\mathtt{n}(u)^{\frac{1}{2}}},\ \rho = (\mathtt{n}(\zeta)^2 + \mathtt{n}(u))^{1/2},\ \theta = \arctan\frac{\mathtt{n}(u)^{1/2}}{\mathtt{n}(\zeta)}\Big) \tag{7·12}$$

is a diffeomorphism. Since

$$\mathtt{n}(\zeta) = \rho\cos\theta \quad \text{and} \quad \mathtt{n}(u)^{1/2} = \rho\sin\theta\,, \tag{7·13}$$

we have

$$\begin{aligned} d\zeta\, du &= \frac{1}{2}\,\mathtt{n}(\zeta)^{2n-3}\, d(\mathtt{n}(\zeta))\ d\operatorname{vol}_{\mathbb{S}^{4n-5}}\Big(\frac{\zeta}{\mathtt{n}(\zeta)^{1/2}}\Big)\ \mathtt{n}(u)\, d(\mathtt{n}(u)^{1/2})\ \ d\operatorname{vol}_{\mathbb{S}^2}\Big(\frac{u}{\mathtt{n}(u)^{1/2}}\Big) \\ &= \frac{1}{2}\cos^{2n-3}\theta\ \sin^2\theta\ \rho^{2n}\ \ d\operatorname{vol}_{\mathbb{S}^{4n-5}}(\sigma)\ \ d\operatorname{vol}_{\mathbb{S}^2}(w)\ d\rho\ d\theta\,. \end{aligned} \tag{7·14}$$

Using respectively in the following sequence of equalities

- the definition of the skinning measure in Equation (7·2) with basepoint $x_0 = (0, 0, 1)$ and the homeomorphism sending $v \in \partial^1_+ D$ to $v_+ = (\zeta, u) \in \mathbb{H}\mathrm{eis}_{4n-1} - \{(0,0)\}$, Equation (7·6) and Lemma 6·2,
- Equation (7·8) and Assertion *(i)*,
- Equation (6·6), and
- Equations (7·13) and (7·14),

we have

$$\begin{aligned}
d\widetilde{\sigma}^+_D(v) &= e^{-(4n+2)\,\beta_{(\zeta,u)}(\,(0,0,(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{1/2}),\,(0,0,1)\,)}\;d\mu_{(0,0,1)}(\zeta,u)\\
&= \frac{(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{(2n+1)/2}}{d_{\rm Cyg}((0,\,0,\,(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{1/2}),(\zeta,u,0))^{8n+4}}\;d\zeta\,du\\
&= \Big(\frac{(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{1/2}}{(\mathtt{n}(\zeta)+(\mathtt{n}(\zeta)^2+\mathtt{n}(u))^{1/2})^2+\mathtt{n}(u)}\Big)^{2n+1}\;d\zeta\,du\\
&= \frac{\cos^{2n-3}\theta\;\sin^2\theta}{2^{2n+2}\,(1+\cos\theta)^{2n+1}}\;d\operatorname{vol}_{\mathbb{S}^{4n-5}}(\sigma)\;\;d\operatorname{vol}_{\mathbb{S}^2}(w)\;\frac{d\rho}{\rho}\;d\theta\,.
\end{aligned}$$

Thus,

$$d\pi_*\widetilde{\sigma}^+_D(0,0,\rho) = \frac{c'_n\;\operatorname{Vol}(\mathbb{S}^{4n-5})\operatorname{Vol}(\mathbb{S}^2)}{2^{2n+2}}\;\frac{d\rho}{\rho}\;, \tag{7·15}$$

where, using the change of variable $t=\tan\frac{\theta}{2}$,

$$c'_n=\int_0^{\frac{\pi}{2}}\frac{\cos^{2n-3}\theta\;\sin^2\theta}{(1+\cos\theta)^{2n+1}}\;d\theta=\frac{1}{2^{2n-2}}\int_0^1(1-t^2)^{2n-3}\,t^2\,(1+t^2)\;dt\,.$$

With $I_{p,q}=\int_{-1}^1 t^{2p}(1-t^2)^q\;dt$, we have by integration by parts and by induction

$$I_{p,q}=\frac{2^{2q+1}\,q!\,(2p)!\,(p+q)!}{p!\,(2p+2q+1)!}\;.$$

Hence $c'_n=\frac{1}{2^{2n-1}}\;(I_{1,\,2n-3}+I_{2,\,2n-3})=\dfrac{2^{2n-1}\,(2n-3)!\,(2n-1)!\,(2n+1)}{(4n-1)!}$.

The next step is to obtain an expression similar to Equation (7·15) for the Riemannian measure of the submanifold ∂^1_+D of $T^1\mathbf{H}^n_{\mathbb{H}}$ (endowed with Sasaki's metric). For every $x\in D$, let us denote by ν^1_xD the fiber over x of the normal bundle map $v\mapsto\pi(v)$ from ∂^1_+D to D. We endow ν^1_xD with the spherical metric induced by the scalar product of the tangent space $T_x\mathbf{H}^n_{\mathbb{H}}$ at x. The Riemannian measure of ∂^1_+D disintegrates under this fibration over the Riemannian measure of D as

$$d\operatorname{vol}_{\partial^1_+D}(v)=\int_{x\in D}d\operatorname{vol}_{\nu^1_xD}(v)\;d\operatorname{vol}_D(x)\,.$$

By looking at the expression (6·2) of the Riemannian metric of $\mathbf{H}^n_{\mathbb{H}}$ in horospherical coordinates, using the homeomorphism $\rho\mapsto x=(0,0,\rho)$ from $]0,+\infty[$ to D, we have

$$d\operatorname{vol}_D(x)=\frac{d\rho}{2\rho}\,.$$

Hence

$$d\pi_*\operatorname{vol}_{\partial^1_+D}(0,0,\rho)=\operatorname{Vol}(\mathbb{S}^{4n-2})\;\frac{d\rho}{2\rho}\,. \tag{7·16}$$

We have $\operatorname{Vol}(\mathbb{S}^{4n-2})=\frac{2^{4n-1}\,\pi^{2n-1}\,(2n-1)!}{(4n-2)!}$ and $\operatorname{Vol}(\mathbb{S}^{4n-5})=\frac{2\,\pi^{2n-2}}{(2n-3)!}$. Equations (7·15) and (7·16) give the first claim of Assertion *(v)*.

The second one follows, since pushforwards of measures preserve their total mass, and since $\operatorname{Vol}(\Gamma_{\partial^1_\pm D}\backslash\partial^1_\pm D)=\frac{\operatorname{Vol}(\mathbb{S}^{4n-2})}{m}\;\operatorname{Vol}(\Gamma_D\backslash D)$.

(vi) By the transitivity of the isometry group of $\mathbf{H}^n_{\mathbb{H}}$ on the set of its quaternionic

geodesic lines, we may assume that D is the quaternionic geodesic line

$$C = \{(w_0, w) \in \mathbf{H}^n_{\mathbb{H}} \ : \ w = 0\}$$

or, in horospherical coordinates, $C = \{(\zeta, u, t) \in \mathbf{H}^n_{\mathbb{H}} \ : \ \zeta = 0\}$.

Hence, using the homeomorphism from $\partial^1_+ C$ to $\{(\zeta, u) \in \mathbb{H}\mathrm{eis}_{4n-1} \ : \ \zeta \neq 0\}$ sending a normal unit vector v to its point at infinity $v_+ = (\zeta, u)$, by the definition of the skinning measure in Equation (7·2) with basepoint $x_0 = (0, 0, 1)$, by Equation (7·6), by Lemma 6·3, by Equations (7·8) and (6·6), and by Assertion *(i)*, we have

$$\begin{aligned} d\widetilde{\sigma}_C(v) &= e^{-(4n+2)\,\beta_{(\zeta,u)}(\,(0,\,u,\,\mathtt{n}(\zeta)),\ (0,\,0,\,1)\,)}\ d\mu_{(0,\,0,\,1)}(\zeta, u) = \frac{1}{2^{4n+2}\,\mathtt{n}(\zeta)^{2n+1}}\ d\zeta\, du \\ &= \frac{1}{2^{4n+3}\,\mathtt{n}(\zeta)^4}\ d(\mathtt{n}(\zeta))\ d\,\mathrm{vol}_{\mathbb{S}^{4n-5}}\Big(\frac{\zeta}{\mathtt{n}(\zeta)^{1/2}}\Big)\ du\,. \end{aligned}$$

In particular,

$$d\pi_*\widetilde{\sigma}_C(0, u, \mathtt{n}(\zeta)) = \frac{\mathrm{Vol}(\mathbb{S}^{4n-5})}{2^{4n+3}}\ du\ \frac{d(\mathtt{n}(\zeta))}{\mathtt{n}(\zeta)^4}\ .$$

For every $x \in C$, let us denote by $\nu^1_x C$ the fiber over x of the normal bundle map $v \mapsto \pi(v)$ from $\partial^1_+ C$ to C, endowed with the spherical metric induced by the scalar product of the tangent space $T_x\mathbf{H}^n_{\mathbb{H}}$ at x. The Riemannian measure of $\partial^1_+ C$ disintegrates under this fibration over the Riemannian measure of C as

$$d\,\mathrm{vol}_{\partial^1_+ C}(v) = \int_{x\in C} d\,\mathrm{vol}_{\nu^1_x C}(v)\ d\,\mathrm{vol}_C(x)\,.$$

Using Equation (6·2) and the homeomorphism $(u, t = \mathtt{n}(\zeta)) \mapsto x = (0, u, t)$ from $\mathrm{Im}\,\mathbb{H}\times\,]0, +\infty[$ to C, we have

$$d\,\mathrm{vol}_C(x) = \Big(\frac{1}{2\,t}\Big)^4 du\, dt = \frac{1}{2^4}\ du\ \frac{d(\mathtt{n}(\zeta))}{\mathtt{n}(\zeta)^4}\ .$$

Hence

$$d\pi_*\ \mathrm{vol}_{\partial^1_+ C}(x) = \mathrm{Vol}(\mathbb{S}^{4n-5})\ d\,\mathrm{vol}_C(x) = \frac{\mathrm{Vol}(\mathbb{S}^{4n-5})}{2^4}\ du\ \frac{d(\mathtt{n}(\zeta))}{\mathtt{n}(\zeta)^4}\ .$$

The result follows as in the end of the proof of the previous Assertion. □

8. *Equidistribution and counting in quaternionic hyperbolic geometry*

In this section, we first use the general results of [**PaP5**] (see also [**BrPP**]) and the computations of Section 7 to give explicit asymptotic counting and equidistribution results on the number of common perpendiculars that are shorter than a given bound between two properly embedded locally convex proper closed subsets of $\Gamma\backslash\mathbf{H}^n_{\mathbb{H}}$, for any lattice Γ in PU_q. Using Sections 4 and 5, we then give two arithmetic applications, generalising Theorems 1·1 and 1·2 in the introduction. We refer to [**PaP7**] for geometric applications.

Let Γ be a lattice in PU_q. Let D^- and D^+ be nonempty proper closed convex subsets of $\mathbf{H}^n_{\mathbb{H}}$, with stabilisers Γ_{D^-} and Γ_{D^+} in Γ respectively, such that the families $(\gamma D^-)_{\gamma\in\Gamma/\Gamma_{D^-}}$ and $(\gamma D^+)_{\gamma\in\Gamma/\Gamma_{D^+}}$ are locally finite in $\mathbf{H}^n_{\mathbb{H}}$. With the measures defined at the beginning of Section 7, let

$$c(D^-, D^+) = \frac{\|\sigma^+_{D^-}\|\ \|\sigma^-_{D^+}\|}{\delta_\Gamma\ \|m_{\mathrm{BM}}\|}\ .$$

For all γ, γ' in Γ, the convex sets γD^- and $\gamma' D^+$ have a common perpendicular if and only if their closures $\overline{\gamma D^-}$ and $\overline{\gamma' D^+}$ in $\mathbf{H}^n_{\mathbb{H}} \cup \partial_\infty \mathbf{H}^n_{\mathbb{H}}$ do not intersect. We denote by $\alpha_{\gamma,\,\gamma'}$ this common perpendicular, starting from γD^- at time $t = 0$, and by $\ell(\alpha_{\gamma,\,\gamma'})$ its length. The *multiplicity* of $\alpha_{\gamma,\gamma'}$ is

$$m_{\gamma,\gamma'} = \frac{1}{\operatorname{Card}(\gamma \Gamma_{D^-} \gamma^{-1} \cap \gamma' \Gamma_{D^+} {\gamma'}^{-1})}\,,$$

which equals 1 for all $\gamma, \gamma' \in \Gamma$ when Γ acts freely on $T^1 \mathbf{H}^n_{\mathbb{H}}$ (for instance when Γ is torsion-free). For all $s > 0$ and $x \in \partial D^-$, let

$$m_s(x) = \sum_{\gamma \in \Gamma/\Gamma_{D^+}\;:\; \overline{D^-} \cap \overline{\gamma D^+} = \varnothing,\; \alpha_{e,\,\gamma}(0) = x,\; \ell(\alpha_{e,\,\gamma}) \leqslant s} m_{e,\gamma}$$

be the multiplicity of x as the origin of common perpendiculars with length at most s from D^- to the elements of the Γ-orbit of D^+. For every $s > 0$, let

$$\mathscr{N}_{D^-,\,D^+}(s) = \sum_{(\gamma,\,\gamma') \in \Gamma \backslash ((\Gamma/\Gamma_{D^-}) \times (\Gamma/\Gamma_{D^+}))\;:\; \overline{\gamma D^-} \cap \overline{\gamma' D^+} = \varnothing,\; \ell(\alpha_{\gamma,\,\gamma'}) \leqslant s} m_{\gamma,\gamma'}\,,$$

where Γ acts diagonally on $\Gamma \times \Gamma$. When Γ has no torsion, $\mathscr{N}_{D^-,\,D^+}(s)$ is the number (with multiplicities coming from the fact that $\Gamma_{D^\pm} \backslash D^\pm$ is not assumed to be embedded in $\Gamma \backslash \mathbf{H}^n_{\mathbb{H}}$) of the common perpendiculars of length at most s between the images of D^- and D^+ in $\Gamma \backslash \mathbf{H}^n_{\mathbb{H}}$.

Let us determine some constants before stating Theorem 8.1 giving the asymptotics of $\mathscr{N}_{D^-,\,D^+}(s)$ as $s \to +\infty$, and its associated equidistribution claim. We assume from now on that D^- is a horoball in $\mathbf{H}^n_{\mathbb{H}}$ centred at a parabolic fixed point of Γ. We assume from now on that D^+ is one of the following three possibilities, we denote by m^+ the cardinality of the pointwise stabiliser of D^+ in Γ and we compute $c(D^-, D^+)$ using Lemma 7·2 and Equation (7·6). If D^+ is also a horoball in $\mathbf{H}^n_{\mathbb{H}}$ centred at a parabolic fixed point of Γ, then

$$c(D^-, D^+) = \frac{2^{4n+1}\,(2n+1)!}{n\,\pi^{2n}}\;\frac{\operatorname{Vol}(\Gamma_{D^-} \backslash D^-)\operatorname{Vol}(\Gamma_{D^+} \backslash D^+)}{\operatorname{Vol}(\Gamma \backslash \mathbf{H}^n_{\mathbb{H}})}\,.$$

If D^+ is a geodesic line in $\mathbf{H}^n_{\mathbb{H}}$ such that $\Gamma_{D^+} \backslash D^+$ is compact, then

$$c(D^-, D^+) = \frac{2^{4n}\,(2n-1)!(2n+1)!}{\pi\, m^+\,(4n)!}\,\frac{\operatorname{Vol}(\Gamma_{D^-} \backslash D^-)\operatorname{Vol}(\Gamma_{D^+} \backslash D^+)}{\operatorname{Vol}(\Gamma \backslash \mathbf{H}^n_{\mathbb{H}})}\,.$$

If D^+ is a quaternionic geodesic line in $\mathbf{H}^n_{\mathbb{H}}$ such that $\Gamma_{D^+} \backslash D^+$ has finite volume, then

$$c(D^-, D^+) = \frac{2\,(n-1)\,(2n-1)}{\pi^2\, m^+}\,\frac{\operatorname{Vol}(\Gamma_{D^-} \backslash D^-)\operatorname{Vol}(\Gamma_{D^+} \backslash D^+)}{\operatorname{Vol}(\Gamma \backslash \mathbf{H}^n_{\mathbb{H}})}\,.$$

Lemma 7·2 (iv) also gives that

$$\frac{1}{\|\sigma^+_{D^-}\|}\, d\,\pi_* \sigma^+_{D^-} = \frac{1}{2\,(2n+1)\,\operatorname{Vol}(\Gamma_{D^-} \backslash D^-)}\, d\operatorname{vol}_{\partial D^-}\,.$$

Recall that every lattice in PU_q is arithmetic, by the works of Margulis, Corlette, Gromov-Schoen, see [**GS**, Theo. 8.4]. The following counting and equidistribution result of common perpendiculars follows from [**PaP5**, Theo. 15 (2)] (with the remark preceding it concerning the proof by Kleinbock-Margulis and Clozel of the exponential mixing property for the Sobolev regularity of the geodesic flow), see also [**BrPP**, §12.2-3]. We denote by Δ_x the unit Dirac mass at a point x.

THEOREM 8·1. *Let Γ, D^-, D^+ be as above. There exists $\kappa > 0$ such that, as $s \to +\infty$,*

$$\mathscr{N}_{D^-,\,D^+}(s) = c(D^-, D^+)\, e^{(4n+2)\,s}\,\big(1 + \mathrm{O}(e^{-\kappa s})\big)\,.$$

Furthermore, the origins of the common perpendiculars from D^- to the images of D^+ under the elements of Γ equidistribute in ∂D^- to the induced Riemannian measure: as $s \to +\infty$,

$$\frac{2\,(2n+1)\ \mathrm{Vol}(\Gamma_{D^-}\backslash D^-)}{c(D^-, D^+)}\ e^{-(4n+2)\,s} \sum_{x \in \partial D^-} m_s(x)\ \Delta_x \ \overset{*}{\rightharpoonup}\ \mathrm{vol}_{\partial D^-}\,. \quad \Box \qquad (8{\cdot}1)$$

For smooth functions ψ with compact support on ∂D^-, there is an error term in the equidistribution claim of Theorem 8·1 when the measures on both sides are evaluated on ψ, of the form $\mathrm{O}(e^{-\kappa s}\,\|\psi\|_\ell)$ for some $\kappa > 0$, where $\|\psi\|_\ell$ is the Sobolev norm of ψ for some $\ell \in \mathbb{N}$.

We now apply Theorem 8·1 in order to prove an analog of Mertens's formula and Neville's equidistribution theorem in the quaternionic Heisenberg group. See for example the Introduction of [**PaP6**] for an explanation of the name.

Let $\mathfrak{m}$ be a nonzero bilateral ideal in $\mathscr{O}$ stable by conjugation. As defined in Equations (1·1)–(1·3) in the Introduction, the action by shears on $\mathscr{O} \times \mathscr{O} \times \mathscr{O}$ of the nilpotent group $\mathscr{N}(\mathscr{O})$ preserves $\mathscr{O} \times \mathfrak{m} \times \mathfrak{m}$. We will study the asymptotic of the counting function $\Psi_{\mathfrak{m}}$, where, for every $s \geqslant 0$, the number $\Psi_{\mathfrak{m}}(s)$ is the cardinality of

$$\mathscr{N}(\mathscr{O})\backslash\big\{(a, \alpha, c) \in \mathscr{O} \times \mathfrak{m} \times \mathfrak{m}\ :\ \mathtt{tr}(a\,\overline{c}) = \mathtt{n}(\alpha),\ {}_{\mathscr{O}}\langle a, \alpha, c\rangle = \mathscr{O},\ 0 < \mathtt{n}(c) \leqslant s\big\}\,.$$

We endow the ring $\mathscr{O}/\mathfrak{m}$ with the involution induced by the quaternionic conjugation. Let $\mathrm{U}_q(\mathscr{O}/\mathfrak{m})$ be the finite group of 3×3 matrices in $\mathscr{O}/\mathfrak{m}$, preserving the Hermitian form $-\overline{z_0}\,z_2 - \overline{z_2}\,z_0 + \overline{z_1}\,z_1$ on $(\mathscr{O}/\mathfrak{m})^3$. Let $\mathrm{B}_q(\mathscr{O}/\mathfrak{m})$ be its upper triangular subgroup.

THEOREM 8·2. *There exists $\kappa > 0$ such that, as $s \to +\infty$,*

$$\Psi_{\mathfrak{m}}(s) = \frac{2^3 \cdot 3^6 \cdot 5 \cdot 7\ D_A^{\,4}\ |\mathrm{B}_q(\mathscr{O}/\mathfrak{m})|}{\pi^8\ m_A\ |\mathscr{O}^\times|\ \prod_{p|D_A}(p-1)(p^2+1)(p^3-1)\ |\mathrm{U}_q(\mathscr{O}/\mathfrak{m})|}\ s^5\ (1 + \mathrm{O}(s^{-\kappa}))\,.$$

The particular case $\mathfrak{m} = \mathscr{O}$ gives Theorem 1·1 in the introduction. We will prove this result simultaneously with the next one. We endow the Lie group $\mathbb{H}\mathrm{eis}_7$ with its Haar measure $\mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}$ defined in the Introduction. The following result is an equidistribution result of the set of $\mathbb{Q}$-points (satisfying some congruence properties) in $\mathbb{H}\mathrm{eis}_7$, seen as the set of $\mathbb{R}$-points of a $\mathbb{Z}$-form of a $\mathbb{Q}$-algebraic group with set of $\mathbb{Q}$-points $\mathbb{H}\mathrm{eis}_7 \cap (A \times A)$ and set of $\mathbb{Z}$-points $\mathscr{N}(\mathscr{O})$. The particular case $\mathfrak{m} = \mathscr{O}$ gives Theorem 1·2 in the introduction.

THEOREM 8·3. *As $s \to +\infty$, we have*

$$\frac{\pi^8\ m_A\ |\mathscr{O}^\times|\ \prod_{p|D_A}(p-1)(p^2+1)(p^3-1)\ |\mathrm{U}_q(\mathscr{O}/\mathfrak{m})|}{2^5 \cdot 3^6 \cdot 5 \cdot 7\ D_A^2\ |\mathrm{B}_q(\mathscr{O}/\mathfrak{m})|}\ s^{-5}\ \times \sum_{\substack{(a,\,\alpha,\,c) \in \mathscr{O} \times \mathfrak{m} \times \mathfrak{m},\ 0 < \mathtt{n}(c) \leqslant s \\ \mathtt{tr}(a\,\overline{c}) = \mathtt{n}(\alpha),\ {}_{\mathscr{O}}\langle a,\,\alpha,\,c\rangle = \mathscr{O}}} \Delta_{(ac^{-1},\,\alpha c^{-1})}\ \overset{*}{\rightharpoonup}\ \mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}\,.$$

As in Theorem 8·1, for smooth functions ψ with compact support on $\mathbb{H}\mathrm{eis}_7$, there is an error term in this equidistribution result when the measures on both sides are evaluated on ψ, of the form $\mathrm{O}(s^{-\kappa}\,\|\psi\|_\ell)$ for some $\kappa > 0$, where $\|\psi\|_\ell$ is the Sobolev norm of ψ for some $\ell \in \mathbb{N}$.

Proofs of Theorem 8·2 and Theorem 8·3. We start by introducing the notation used in these proofs.

We consider the quaternionic Hermitian form q defined in Equation (2·1) with $n = 2$. For every subgroup G of U_q, we denote by $\overline{G}$ its image in PU_q, and again by g the image in PU_q of any element g of U_q.

We consider the lattice $\Gamma = \mathrm{U}_q(\mathscr{O})$ in U_q defined in Section 4, so that $\overline{\Gamma} = \mathrm{PU}_q(\mathscr{O})$. We denote by $\Gamma_{\mathfrak{m}}$ the Hecke congruence subgroup of Γ modulo $\mathfrak{m}$, that is the preimage, by the group morphism $\Gamma \to \mathrm{U}_q(\mathscr{O}/\mathfrak{m})$ of reduction modulo $\mathfrak{m}$, of the upper triangular subgroup $\mathrm{B}_q(\mathscr{O}/\mathfrak{m})$. Since $-\mathrm{id} \in \Gamma_{\mathfrak{m}}$, we have

$$[\,\overline{\Gamma}\;:\;\overline{\Gamma_{\mathfrak{m}}}\,] = [\Gamma : \Gamma_{\mathfrak{m}}] = \frac{|\,\mathrm{U}_q(\mathscr{O}/\mathfrak{m})|}{|\,\mathrm{B}_q(\mathscr{O}/\mathfrak{m})|}\;. \tag{8·2}$$

We denote by $\Gamma_{\mathscr{H}_1}$ the stabiliser in $\Gamma_{\mathfrak{m}}$ of the horoball $\mathscr{H}_1$ defined in Equation (6·4). It is equal to $\mathrm{B}_q \cap \Gamma_{\mathfrak{m}}$ where B_q has been defined in Section 3, since an element of Γ fixes ∞ if and only if it preserves $\mathscr{H}_1$. The group $\Gamma_{\mathscr{H}_1}$ is independent of $\mathfrak{m}$, by the definition of $\Gamma_{\mathfrak{m}}$.

The projection map from $\Gamma_{\mathscr{H}_1}$ to $\overline{\Gamma_{\mathscr{H}_1}}$ is 2-to-1 since $-\mathrm{id} \in \Gamma_{\mathscr{H}_1}$. We identify the lattice $\mathscr{N}(\mathscr{O})$ of $\mathbb{H}\mathrm{eis}_7$ with its image $\overline{\mathscr{N}(\mathscr{O})}$ by the embedding of $\mathbb{H}\mathrm{eis}_7$ in PU_q defined in Section 6. The description of B_q at the end of Section 3 gives

$$[\,\overline{\Gamma_{\mathscr{H}_1}} : \overline{\mathscr{N}(\mathscr{O})}\,] = \frac{1}{2}\,[\Gamma_{\mathscr{H}_1} : \mathscr{N}(\mathscr{O})] = \frac{1}{2}\,|\mathscr{O}^\times|^2\,. \tag{8·3}$$

The following result gives in particular the computation of the volume of the cusp at infinity for $\overline{\Gamma}$.

LEMMA 8·4. *The Haar measure λ_7 on $\mathbb{H}\mathrm{eis}_{4n-1}$ defined in Equation* (7·3) *coincides with the Haar measure* $\mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}$ *defined in the introduction, that is, the total mass of the measure induced by λ_7 on $\mathscr{N}(\mathscr{O})\backslash\mathbb{H}\mathrm{eis}_7$ is $\frac{D_A^2}{4}$. Furthermore*

$$\mathrm{Vol}(\,\overline{\Gamma_{\mathscr{H}_1}}\backslash\mathscr{H}_1) = \frac{D_A^2}{160\,|\,\mathscr{O}^\times|^2}\;. \tag{8·4}$$

For instance, if $D_A = 2$ and $\mathscr{O}$ is the Hurwitz order $\frac{1+i+j+k}{2}\mathbb{Z} + \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}k$, which has 24 units and $\mathrm{Im}\,\mathscr{O} = \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}k$, then $\mathrm{vol}(\overline{\Gamma_{\mathscr{H}_1}}\backslash\mathscr{H}_1) = \frac{1}{23040}$, to be compared with [**KiP**, Prop.5.8] for a related computation.

Proof. Note that $\mathtt{tr} : \mathbb{H} \to \mathbb{R}$ is a fibration, with fiber $\frac{t}{2} + \mathrm{Im}\,\mathbb{H}$ over $t \in \mathbb{R}$. The Lebesgue measure of the Euclidean space $\mathbb{H}$ disintegrates by this fibration over the Lebesgue measure of $\mathbb{R}$, with conditional measures on the fiber 1/2 the Lebesgue measure of the fiber: $dx_0dx_1dx_2dx_3 = (\frac{1}{2}dx_1dx_2dx_3)d(2x_0)$. Since the map $\mathtt{tr}$ is additive, since it maps $\mathscr{O}$ onto $\mathbb{Z}$ with kernel $\mathrm{Im}\,\mathscr{O}$ as recalled in Section 2, this implies that $\mathrm{Vol}(\mathrm{Im}\,\mathscr{O}\backslash\mathrm{Im}\,\mathbb{H}) = 2\,\mathrm{Vol}(\mathscr{O}\backslash\mathbb{H})$. Again by the surjectivity of $\mathtt{tr} : \mathscr{O} \to \mathbb{Z}$, for every $w \in \mathscr{O}$, the set $\{w_0 \in \mathscr{O} \;:\; \mathtt{tr}\,w_0 = \mathtt{n}(w)\}$ is a translate of $\mathrm{Im}\,\mathscr{O}$. Hence by Equations (6·1) and (2·4), we have

$$\lambda_7(\mathscr{N}(\mathscr{O})\backslash\mathbb{H}\mathrm{eis}_7) = 2\,\mathrm{Vol}(\mathrm{Im}\,\mathscr{O}\backslash\mathrm{Im}\,\mathbb{H})\,\mathrm{Vol}(\mathscr{O}\backslash\mathbb{H}) = 4\,\mathrm{Vol}(\mathscr{O}\backslash\mathbb{H})^2 = \frac{D_A^{\,2}}{4}\;.$$

This proves the first claim of Lemma 8·4.

Now, by Equation (8·3), by Lemma 7·1, by Equation (7·11), we have

$$\begin{aligned}\operatorname{Vol}(\,\overline{\Gamma_{\mathscr{H}_1}}\backslash\mathscr{H}_1) &= \frac{2}{|\mathscr{O}^\times|^2}\operatorname{Vol}(\,\overline{\mathscr{N}(\mathscr{O})}\backslash\mathscr{H}_1) = \frac{1}{5\,|\mathscr{O}^\times|^2}\operatorname{Vol}(\,\overline{\mathscr{N}(\mathscr{O})}\backslash\partial\mathscr{H}_1)\\ &= \frac{1}{40\,|\mathscr{O}^\times|^2}\,\lambda_7(\mathscr{N}(\mathscr{O})\backslash\mathbb{H}\mathrm{eis}_7) = \frac{D_A^{\,2}}{160\,|\mathscr{O}^\times|^2}\,.\end{aligned}$$

□

We need one more notation before giving the proof of Theorem 8·2. Consider an element $g\in\Gamma_{\mathfrak{m}}$ such that $g\mathscr{H}_1$ and $\mathscr{H}_1$ are disjoint (there are only finitely many double classes $[g]\in\overline{\Gamma_{\mathscr{H}_1}}\backslash\overline{\Gamma_{\mathfrak{m}}}/\overline{\Gamma_{\mathscr{H}_1}}$ for which this is not the case). We denote by $\ell(\delta_g)$ the length of the common perpendicular δ_g between $g\,\mathscr{H}_1$ and $\mathscr{H}_1$. If $\begin{pmatrix}a_g\\ \alpha_g\\ c_g\end{pmatrix}$ is the first column of g, then $g\cdot\infty=[a_g:\alpha_g:c_g]$.

We use the following facts in the system of equations below.

- For the first equality, note that the cardinality of each nonempty fiber of the projection map from $\big\{(a,\alpha,c)\in\mathscr{O}\times\mathfrak{m}\times\mathfrak{m}\;:\;{}_{\mathscr{O}}\langle a,\alpha,c\rangle=\mathscr{O}\big\}$ to $\mathbb{P}^2_{\mathrm{r}}(\mathbb{H})$ is $|\mathscr{O}^\times|$ and that the projection from $\mathscr{N}(\mathscr{O})$ to $\overline{\mathscr{N}(\mathscr{O})}$ is injective.
- The second and third equalities follow from Proposition 4·2 (2).
- The fourth equality follows from Lemma 6·5.
- The fifth equality follows by Equation (8·3) and by the definition of the counting function $\mathscr{N}_{\mathscr{H}_1,\mathscr{H}_1}$.
- The sixth equality follows from the first claim in Theorem 8·1 with $n=2$, $\Gamma=\overline{\Gamma_{\mathfrak{m}}}$ and $D^-=D^+=\mathscr{H}_1$.
- The last equality follows from Equations (8·4) and (8·2) and from Theorem 1·4.

We hence have, for some $\kappa>0$ and for every $s>0$,

$$\begin{aligned}\Psi_{\mathfrak{m}}(s) &= |\mathscr{O}^\times|\;\operatorname{Card}\;\overline{\mathscr{N}(\mathscr{O})}\backslash\Big\{[a:\alpha:c]\in\mathbb{P}^2_{\mathrm{r}}(\mathbb{H})\;:\;\begin{array}{c}(a,\alpha,c)\in\mathscr{O}\times\mathfrak{m}\times\mathfrak{m},\\ {}_{\mathscr{O}}\langle a,\alpha,c\rangle=\mathscr{O},\\ \mathtt{tr}(a\,\overline{c})=\mathtt{n}(\alpha),\;0<\mathtt{n}(c)\leqslant s\end{array}\Big\}\\
&= |\mathscr{O}^\times|\;\operatorname{Card}\;\overline{\mathscr{N}(\mathscr{O})}\backslash\Big\{[a:\alpha:c]\in\overline{\Gamma_{\mathfrak{m}}}\cdot\infty\;:\;\begin{array}{c}(a,\alpha,c)\in\mathscr{O}\times\mathfrak{m}\times\mathfrak{m},\\ {}_{\mathscr{O}}\langle a,\alpha,c\rangle=\mathscr{O},\;0<\mathtt{n}(c)\leqslant s\end{array}\Big\}\\
&= |\mathscr{O}^\times|\;\operatorname{Card}\big\{[g]\in\overline{\mathscr{N}(\mathscr{O})}\backslash\overline{\Gamma_{\mathfrak{m}}}/\overline{\Gamma_{\mathscr{H}_1}}\;:\;0<\mathtt{n}(c_g)\leqslant s\big\}\\
&= |\mathscr{O}^\times|\,[\,\overline{\Gamma_{\mathscr{H}_1}}:\overline{\mathscr{N}(\mathscr{O})}\,]\operatorname{Card}\big\{[g]\in\overline{\Gamma_{\mathscr{H}_1}}\backslash\overline{\Gamma_{\mathfrak{m}}}/\overline{\Gamma_{\mathscr{H}_1}}\;:\;\ell(\delta_g)\leqslant\frac{\ln s}{2}-\ln 2\big\}+\mathrm{O}(1)\\
&= \frac{1}{2}\,|\mathscr{O}^\times|^3\;\mathscr{N}_{\mathscr{H}_1,\mathscr{H}_1}\big(\frac{\ln s}{2}-\ln 2\big)+\mathrm{O}(1)\\
&= \frac{15\,|\mathscr{O}^\times|^3\,\big(\operatorname{Vol}(\,\overline{\Gamma_{\mathscr{H}_1}}\backslash\mathscr{H}_1)\big)^2}{\pi^4\;\operatorname{Vol}(\,\overline{\Gamma_{\mathfrak{m}}}\backslash\mathbf{H}^2_{\mathbb{H}})}\;s^5\,(1+\mathrm{O}(s^{-\kappa}))\\
&= \frac{2^3\cdot 3^6\cdot 5\cdot 7\,D_A^{\,4}\,|\,\mathrm{B}_q(\mathscr{O}/\mathfrak{m})|}{\pi^8\,m_A\,|\mathscr{O}^\times|\,|\,\mathrm{U}_q(\mathscr{O}/\mathfrak{m})|\,\prod_{p|D_A}(p-1)(p^2+1)(p^3-1)}\;s^5\,(1+\mathrm{O}(s^{-\kappa}))\,.\end{aligned}$$

This concludes the proof of Theorem 8·2.

Let us prove now Theorem 8·3. The orthogonal projection map $f:\partial_\infty\mathbf{H}^2_{\mathbb{H}}-\{\infty\}\to\partial\mathscr{H}_1$ is the homeomorphism defined by $[w_0:w:1]\mapsto(\zeta=w,u=2\operatorname{Im}w_0,1)$ using the homogeneous coordinates on $\partial_\infty\mathbf{H}^2_{\mathbb{H}}-\{\infty\}$ and the horospherical coordinates on $\partial\mathscr{H}_1$

(see Equation (6·1)). Let $x \in \partial\mathscr{H}_1$ be the origin of a common perpendicular of length at most s from $\mathscr{H}_1$ to a horoball $\gamma\mathscr{H}_1$ for some $\gamma \in \Gamma_{\mathfrak{m}}$ not fixing ∞. The point x is the orthogonal projection on $\mathscr{H}_1$ of the point at infinity of this horoball $\gamma\mathscr{H}_1$. This point at infinity may be written $[ac^{-1}, \alpha c^{-1} : 1]$ for some triple $(a, \alpha, c) \in \mathscr{O} \times \mathfrak{m} \times \mathfrak{m}$ with ${}_{\mathscr{O}}\langle a, \alpha, c\rangle = \mathscr{O}$, $\mathtt{tr}(a\,\overline{c}) = \mathtt{n}(\alpha)$ and $0 < \mathtt{n}(c) \leqslant 4\,e^{2s}$ (using Lemma 6·5). Such a writing is not unique, there are exactly $|\mathscr{O}^\times|$ such triples. Hence by the second claim of Theorem 8·1 with $\Gamma = \overline{\Gamma_{\mathfrak{m}}}$ and $D^- = D^+ = \mathscr{H}_1$, using the horospherical coordinates on $\partial\mathscr{H}_1$, we have, as $s \to +\infty$,

$$\frac{\pi^4\ \mathrm{Vol}(\overline{\Gamma_{\mathfrak{m}}}\backslash\mathbf{H}^2_{\mathbb{H}})}{2^{10}\cdot 3\ |\mathscr{O}^\times|\ \mathrm{Vol}(\overline{\Gamma_{\mathscr{H}_1}}\backslash\mathscr{H}_1)}\ e^{-10\,s}\times \sum_{\substack{(a,\,\alpha,\,c)\in\mathscr{O}\times\mathfrak{m}\times\mathfrak{m},\ 0<\mathtt{n}(c)\leqslant 4\,e^{2s}\\ \mathtt{tr}(a\,\overline{c})=\mathtt{n}(\alpha),\ {}_{\mathscr{O}}\langle a,\,\alpha,\,c\rangle=\mathscr{O}}} \Delta_{(ac^{-1},\,2\,\mathrm{Im}(\alpha c^{-1}),\,1)} \overset{*}{\rightharpoonup} \mathrm{vol}_{\partial\mathscr{H}_1}\ . \tag{8·5}$$

Recall that the Haar measure λ_7 on $\mathbb{H}\mathrm{eis}_7 = \partial_\infty\mathbf{H}^2_{\mathbb{H}} - \{\infty\}$, defined in Equation (7·3), coincides with the Haar measure $\mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}$ by Lemma 8·4. Its image by the above map f is, by Equation (7·5),

$$f_*\,\mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7} = f_*\,\lambda_7 = 8\,\mathrm{vol}_{\partial\mathscr{H}_1}\ .$$

Using the change of variables $s \mapsto 4\,e^{2s}$ and the continuity of the pushforward by f^{-1} of the measures on $\partial\mathscr{H}_1$ applied to Equation (8·5), we hence have, as $s \to +\infty$,

$$\frac{8\,\pi^4\ \mathrm{Vol}(\overline{\Gamma_{\mathfrak{m}}}\backslash\mathbf{H}^2_{\mathbb{H}})}{3\,|\mathscr{O}^\times|\ \mathrm{Vol}(\overline{\Gamma_{\mathscr{H}_1}}\backslash\mathscr{H}_1)}\ s^{-5} \sum_{\substack{(a,\,\alpha,\,c)\in\mathscr{O}\times\mathfrak{m}\times\mathfrak{m},\ 0<\mathtt{n}(c)\leqslant s\\ \mathtt{tr}(a\,\overline{c})=\mathtt{n}(\alpha),\ {}_{\mathscr{O}}\langle a,\,\alpha,\,c\rangle=\mathscr{O}}} \Delta_{(ac^{-1},\,\alpha c^{-1})} \overset{*}{\rightharpoonup} \mathrm{Haar}_{\mathbb{H}\mathrm{eis}_7}\ .$$

Finally, Theorem 8·3 follows from this, from Equations (8·4) and (8·2) and from Theorem 1·4. □

REMARK 8·5. Theorems 8·2 and 8·3 have generalisations in higher dimension. Theorem 8·1 (which is valid in any dimension), applied with $\Gamma = \mathrm{PU}_q(\mathscr{O})$ and with $D^- = D^+$ the horoball of points in $\mathbf{H}^n_{\mathbb{H}}$ with last horospherical coordinates at least 1, gives a counting and equidistribution result of the orbit $\Gamma\cdot\infty - \{\infty\}$ in $\partial_\infty\mathbb{H}\mathrm{eis}_{4n-1}$with error term. The volume of $\Gamma\backslash\mathbf{H}^n_{\mathbb{H}}$ could be computed using **[EmK]**, up to computing the index of Γ in a principal arithmetic subgroup containing it. The volume of the cusp corresponding to ∞ in $\Gamma\backslash\mathbf{H}^n_{\mathbb{H}}$ may also be computed by the same method as for the proof of Equation (8·4).

Other counting and equidistribution results of arithmetically defined points in the quaternionic Heisenberg group $\mathbb{H}\mathrm{eis}_{4n-1}$ may be obtained by varying the cusp (when $n = 2$ and $h_A \neq 1$, there are at least two cusps by Theorem 1·3), the integral quaternionic Hermitian form q of Witt signature $(1, n)$ and the arithmetic lattice Γ in U_q.

Acknowledgements. We thank Gaetan Chenevier for his help with Section 4. The second author thanks Hee Oh and the hospitality of Yale University for a one month stay, decisive for the writing of this paper. We thank the referee for her/his comments, which have helped a lot in order to improve the readability of this paper.

REFERENCES

[All] D. Allcock. *New complex- and quaternion-hyperbolic reflection groups.* Duke Math. J. **103** (2000) 303–333.

[BeQ] Y. Benoist and J.-F. Quint. *Stationary measures and invariant subsets of homogeneous spaces II.* J. Amer. Math. Soc. **26** (2013) 659–734.

[BGM] M. Berger, P. Gauduchon and E. Mazet. *Le spectre d'une variété riemannienne.* Lect. Notes Math. **194**, Springer Verlag, 1971.

[Biq] O. Biquard. *Quaternionic contact structures.* Proceedings of the Second Meeting on Quaternionic Structures in Mathematics and Physics, Roma 1999 (S. Marchiafava, P. Piccinni, M. Pontecorvo, eds), 23–30, World Scientific, 2001.

[Bor] A. Borel. *Reduction theory for arithmetic groups.* In "Algebraic Groups and Discontinuous Subgroups", A. Borel and G. D. Mostow eds, Proc. Sympos. Pure Math. (Boulder, 1965), pp. 20–25, Amer. Math. Soc. 1966.

[Bou] N. Bourbaki. *Groupes et algèbres de Lie : chap. 4, 5 et 6.* Masson, 1981.

[Bre] E. Breuillard. *Local limit theorems and equidistribution of random walks on the Heisenberg group.* Geom. Funct. Anal. **15** (2005) 35–82.

[BrPP] A. Broise-Alamichel, J. Parkkonen and F. Paulin. *Equidistribution and counting under equilibrium states in negative curvature and trees. Applications to non-Archimedean Diophantine approximation.* With an Appendix by J. Buzzi. Prog. Math. **329**, Birkhäuser, 2019.

[BrT1] F. Bruhat and J. Tits. *Schémas en groupes et immeubles des groupes classiques sur un corps local.* Bull. Soc. Math. France **112** (1984) 259–301.

[BrT2] F. Bruhat and J. Tits. *Schémas en groupes et immeubles des groupes classiques sur un corps local. II : groupes unitaires.* Bull. Soc. Math. France **115** (1987) 141–195.

[CaP1] W. Cao and J. Parker. *Jørgensen's inequality and collars in n-dimensional quaternionic hyperbolic space.* Quarterly J. Math. **62** (2011) 523–543.

[CaP2] W. Cao and J. Parker. *Shimizu's lemma for quaternionic hyperbolic space.* Comput. Meth. Funct. Theo. **18** (2018) 159–191.

[ChP] G. Chenevier and F. Paulin. *Sur les minima des formes hamiltoniennes binaires définies positives.* Publications Mathématiques de Besançon : Algèbre et théorie des nombres (2020) 5–25.

[Cor] K. Corlette. *Hausdorff dimensions of limit sets I.* Invent. Math. **102** (1990) 521–542.

[Eic] M. Eichler. *Über die Idealklassenzahl total definiter Quaternionenalgebren.* Math. Z. **43** (1938) 102–109.

[Eme] V. Emery. *Du volume des quotients arithmétiques de l'espace hyperbolique.* Thèse n^0 1648, Univ. Fribourg (Suisse), 2009.

[EmK] V. Emery and I. Kim. *Quaternionic hyperbolic lattices of minimal covolume.* Preprint [`arXiv:1802.07776`].

[FS] G. B. Folland and E. M. Stein. *Estimates for the $\overline{\partial}_b$ complex and analysis on the Heisenberg group.* Comm. Pure Appl. Math. **27** (1974) 429–522.

[Frö] A. Fröhlich. *Locally free modules over arithmetic orders.* J. reine angew. Math. **274** (1975) 112–124.

[Gol] W. M. Goldman. *Complex hyperbolic geometry.* Oxford Univ. Press, 1999.

[GLS] D. Gorenstein, R. Lyons and R. Solomon. *The classification of the finite simple groups.* Math. Surv. Mono. **40.1**, Ameri. Math. Soc. 1994.

[GS] M. Gromov and R. Schoen. *Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one.* Pub. Math. IHÉS **76** (1992) 165–246.

[Hel] S. Helgason. *Differential geometry, Lie groups and symmetric spaces.* Academic Press, 1978.

[Kim] I. Kim. *Counting, mixing and equidistribution of horospheres in geometrically finite rank one locally symmetric manifolds.* J. reine angew. Math. **704** (2015) 85–133.

[KiP] I. Kim and J. Parker. *Geometry of quaternionic hyperbolic manifolds.* Math. Proc. Camb. Phil. Soc. **135** (2003) 291–320.

[KO] V. Krafft and D. Osenberg. *Eisensteinreihen für einige arithmetisch definierte Untergruppen von* $\mathrm{SL}_2(\mathbb{H})$. Math. Z. **204** (1990) 425–449.

[Kra] S. Krantz. *Explorations in harmonic analysis, with applications to complex function theory and the Heisenberg group.* Appl. Num. Harm. Anal., Birkhäuser, 2009.

[Mos] G. D. Mostow. *Strong rigidity of locally symmetric spaces.* Ann. Math. Studies **78**, Princeton Univ. Press, 1973.

[OhS] H. Oh and N. Shah. *The asymptotic distribution of circles in the orbits of Kleinian* *groups*. Invent. Math. **187** (2012) 1–35.

[Ono] T. Ono. *On algebraic groups and discontinuous groups.* Nagoya Math. J. **27** (1966) 279–322.

[PaP1] J. Parkkonen and F. Paulin. *Prescribing the behaviour of geodesics in negative curvature.* Geom. & Topo. **14** (2010) 277–392.

[PaP2] J. Parkkonen and F. Paulin. *On the arithmetic and geometry of binary Hamiltonian forms.* Appendix by Vincent Emery. Algebra & Number Theory **7** (2013) 75–115.

[PaP3] J. Parkkonen and F. Paulin. *On the arithmetic of crossratios and generalised Mertens' formulas.* Special issue "Aux croisements de la géométrie hyperbolique et de l'arithmétique", F. Dal'Bo, C. Lecuire eds, Ann. Fac. Sci. Toulouse **23** (2014) 967–1022.

[PaP4] J. Parkkonen and F. Paulin. *Skinning measure in negative curvature and equidistribution of equidistant submanifolds.* Erg. Theo. Dyn. Sys. **34** (2014) 1310–1342.

[PaP5] J. Parkkonen and F. Paulin. *Counting common perpendicular arcs in negative curvature.* Erg. Theo. Dyn. Sys. **37** (2017) 900–938.

[PaP6] J. Parkkonen and F. Paulin. *Counting and equidistribution in Heisenberg groups.* Math. Annalen **367** (2017) 81–119.

[PaP7] J. Parkkonen and F. Paulin. *Counting and equidistribution of quaternionic Cartan chains.* Preprint [arXiv:2002.05130], to appear in Ann. Math. Blaise Pascal.

[Phi] Z. Philippe. *Invariants globaux des variétés hyperboliques quaternioniques.* Université de Bordeaux, Dec. 2016, https://tel.archives-ouvertes.fr/tel-01661448.

[PlR] V. Platonov and A. Rapinchuck. *Algebraic groups and number theory.* Pure Appl. Math. **139**, Academic Press, 1994.

[Pra] G. Prasad. *Volumes of S-arithmetic quotients of semi-simple groups.* Publ. Math. Inst. Hautes Études Sci. **69** (1989) 91–117.

[Rei] I. Reiner. *Maximal orders.* Academic Press, 1972.

[She] T. Shemanske. *The arithmetics and combinatorics of buidings for* Sp_n. Trans. Amer. Math. Soc **359** (2007) 3409–3423.

[Tit1] J. Tits. *Classification of algebraic semisimple groups.* In "Algebraic Groups and Discontinuous Subgroups" (Boulder, 1965) pp. 33–62, Proc. Symp. Pure Math. IX, Amer. Math. Soc. 1966.

[Tit2] J. Tits. *Reductive groups over local fields.* In "Automorphic forms, representations and L-functions" (Corvallis, 1977), Part 1, pp. 29–69, Proc. Symp. Pure Math. XXXIII, Amer. Math. Soc., 1979.

[Vig] M. F. Vignéras. *Arithmétique des algèbres de quaternions.* Lect. Notes in Math. **800**, Springer Verlag, 1980.

[Zin] T. Zink. *Über die Anzahl der Spitzen einiger arithmetischer Untergruppen unitärer Gruppen.* Math. Nachr. **89** (1979) 315–320.