BULLETIN OF THE BRAZILIAN MATHEMATICAL SOCIETY

Bull Braz Math Soc, New Series 40(3), 301-340
© 2009, Sociedade Brasileira de Matemática

Symmetries of quadratic form classes and of quadratic surd continued fractions. Part I: A Poincaré tiling of the de Sitter world

Francesca Aicardi

Abstract. The problem of classifying the indefinite binary quadratic forms with integer coefficients is solved by introducing a special partition of the de Sitter world, where the coefficients of the forms lie, into separate domains. Under the action of the special linear group acting on the integer plane lattice, each class of indefinite forms has a well-defined finite number of representatives inside each such domain.
In the second part, we will show how to obtain the symmetry type of a class and also the number of its points in all domains from a single representative of that class.

Keywords: binary quadratic forms, reduction theory, tilings.

Mathematical subject classification: Primary: 11H55, 20M30; Secondary: 20H05, 52C20.

Introduction

In this paper, by *form*, we mean a binary quadratic form:

$$f = mx^2 + ny^2 + kxy, \tag{1}$$

where m, n, and k are integers and (x, y) ranges the integer plane lattice.

Definition. The *discriminant* of the form (1) is the integer number

$$k^2 - 4mn.$$

We denote it by Δ.

Following [1], we say that a form is *elliptic* if $\Delta < 0$, *hyperbolic* if $\Delta > 0$, and *parabolic* if $\Delta = 0$.

Received 30 August 2007.

In the usual terminology, the elliptic forms are said to be *definite* and the hyperbolic forms are said to be *indefinite*.

Given a topological space and a group acting on it, the images of a single point under the group action form an *orbit* of the group action. A *fundamental domain* is a connected subset of the space which contains exactly one point from each orbit.

The problem of classifying and counting the orbits of binary quadratic forms under the action of $\mathrm{SL}(2,\mathbb{Z})$ on the xy plane dates back to Gauss and Lagrange ([5], [6]).

The description of the orbits of the positive definite forms under the action of the modular group on the Poincaré model of the Lobachevsky disc is well known: in this model, there is a special tiling of the disc into fundamental domains. In this way, the tiles (or fundamental domains) are in one-to-one correspondence with the group elements, namely, each tile corresponds to the element that sends a chosen domain (called *principal fundamental domain*) to it. The upper sheet of the two-sheeted hyperboloid that contains the points (m, n, k) determining the positive definite forms with a given discriminant, is represented by the Lobachevsky disc in such a way that each class of forms has exactly one representative in each domain.

The complement to the Lobachevsky disc in the projective plane containing it, to which the hyperbolic forms are projected, is not tiled by the same net of lines (for instance, the straight lines of the Klein model, separating the domains of the Lobachevsky disc) into domains of finite area.

In this article, we show that the one-sheeted hyperboloid where the coefficients of the forms with a given discriminant lie can be specially partitioned into separate domains: in each such domain, each orbit has a finite (well-defined) number of points.[1]

This situation is intrinsically different from that of the Lobachevsky disc, where all domains of the partition can be chosen as principal. In our tiling of the de Sitter world, *there are two special domains*, which we call *principal* domains. An $\mathrm{SL}(2,\mathbb{Z})$ change of coordinates in the xy plane (and, consequently, on the hyperboloid) changes the shape of *only a finite set of tiles of the partition* (including the principal domains) but preserves all the peculiar properties of the tiling:

1. The complement to the principal domains of the hyperboloid is separated by the principal domains into four regions: two of them (called

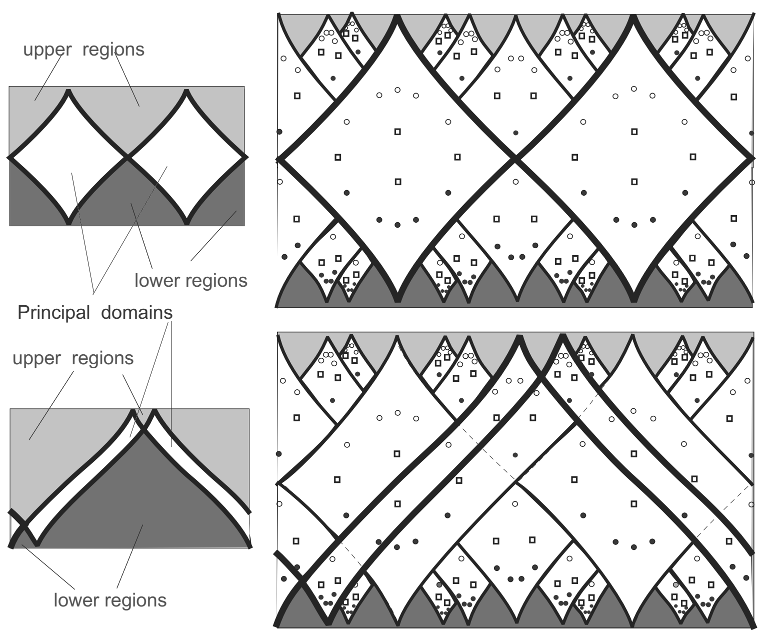
[1] This is surprising. Indeed, the orbit of a generic point (i.e., with irrational coordinates) on the de Sitter world is dense, as Arnold proved [2]. Our results imply only that the number of points of such an orbit is unbounded in each domain.

upper regions) are bounded from the circle at $+\infty$ of the hyperboloid, and the other two regions (*lower regions*) are bounded from the circle at $-\infty$. The circles at infinity are invariant under the action of the group.

2. Each of the upper and lower regions are partitioned into a countable set of domains in one-to-one correspondence with the elements of the semigroup of $\mathrm{SL}(2,\mathbb{Z})$ which are generated by $A = \binom{1\ 1}{0\ 1}$ and $B = \binom{1\ 0}{1\ 1}$.

3. Every orbit has a fixed finite number of points, N_u for example, in each domain of the upper regions, and a fixed finite number of points, N_d for example, in each domain of the lower regions. The principal domains contain $N = N_u + N_d$ points of that orbit.

To understand this unusual situation in which the partition changes without changing the number of integer points in the corresponding domains, we give an example where the properties 1 and 3 above are illustrated on a finite set of domains (also see Fig. 15 for more points of the orbits and more domains).



This figure shows some tiles of two different partitions related by a change of coordinates (namely, by the operator BA) on the hyperboloid projected onto an open cylinder and the points of the three different classes of integer quadratic forms with $k^2 - 4mn = 32$, lying in these tiles.

The principal domains are marked by a thick black boundary: they contain five points of the first orbit (circles), five points of the second orbit (black discs), and four points of the third orbit (squares).

Each domain in the upper regions contains four points of the first orbit, one point of the second orbit, and two points of the third orbit. Each domain in the lower regions contains one point of the first orbit, four points of the second orbit, and two points of the third orbit.

The classical reduction theory introduced by Lagrange for indefinite forms states that there is a finite number of forms such that m and n are positive and $m+n$ is less than k. The reduction procedure, which allows finding these forms, can be described in terms of the tiling introduced in this work. We will see this relation in more detail in Part II.

The reduction theory that follows directly from our tiling is closer to that expounded in [4] because the 'reduced' forms here are those with $mn < 0$, as in our definition. The number of reduced forms by Lagrange is equal to the number n_u of forms in each domain of the upper regions in our partition, while the number of reduced forms by our definition is the number $n_u + n_d$ of reduced forms in the principal domains.

The essential new element with respect to the known theories is the geometric standpoint, which allows seeing the action of the group in the space of forms, exactly as for the modular group action on the Lobachevsky disc.

Here we also introduce a classification of the types of symmetries of the classes of forms. This classification allows us to classify the symmetries of the periods of the continued fractions of the quadratic irrationalities (or *surds*), answering more recent questions posed by Arnold [3], as we will show in Part II, where we will also see how to calculate the number of points in each domain for every class of hyperbolic forms in terms of the coefficients of a form belonging to that class.

I am deeply grateful to Arnold, who posed the problem of the missing geometrical model for hyperbolic forms in [1]. A special thank to Ricardo Uribe Vargas, for his genuine interest in this work and to the referee for his accurate reading to correct the style of the manuscript.

1 The space of forms and their classes

Besides the coordinates m, n, k in the space of forms, we will systematically use also the following coordinates:

$$\begin{aligned} K &= k, \\ D &= m - n, \\ S &= m + n. \end{aligned}$$

Remark. A point having integer coordinates K, D, S represents a form if and only if $D \equiv S \mod 2$. In such coordinates, the discriminants is

$$\Delta = K^2 + D^2 - S^2.$$

Definition. A point having integer coordinates (m, n, k) or integer coordinates $[K, D, S]$ such that $D \equiv S \mod 2$ is called a *good point* and is denoted by a bold letter.

Notation. To avoid confusion, the $[K, D, S]$ coordinates of a good point will be indicated in square brackets whereas the coordinates (m, n, k) in round brackets.

1.1 Action of $\mathrm{SL}(2, \mathbb{Z})$ on the form coefficients

Let $\mathbf{f}$ be the triple (m, n, k) of the coefficients of the form (1), and let $\mathbf{f}'$ be the triple (m', n', k') corresponding to the form f' obtained from f by the action of an operator $\mathbf{L}$ of $\mathrm{SL}(2, \mathbb{Z})$ on $\mathbb{Z}^2$. That is, if $\mathbf{v} = (x, y)$, then we define $f'(\mathbf{v}) = f(\mathbf{L}(\mathbf{v}))$.

With $\mathbf{L}$, we thus associate the operator L acting on $\mathbb{Z}^3$ as

$$\mathbf{f}' = L\mathbf{f}. \tag{2}$$

This defines an homomorphism: $\mathbf{L} \mapsto L$ from $\mathrm{SL}(2, \mathbb{Z})$ to $\mathrm{SL}(3, \mathbb{Z})$. Let $\mathcal{T}$ denote the image of this homomorphism. The subgroup $\mathcal{T}$ is isomorphic to $\mathrm{PSL}(2, \mathbb{Z})$ because $\mathbf{L}$ and $-\mathbf{L}$ have the same image.

Definition. The *orbit* or *class* of a good point $\mathbf{f}$ is the set of points obtained by applying all elements of the group $\mathcal{T}$ to $\mathbf{f}$. The class of $\mathbf{f} = (m, n, k)$ is denoted by $C(\mathbf{f})$ or by $C(m, n, k)$.

The following statements are obvious or easy to prove:

– All points of the orbit of a good point are good.

– All forms of one orbit have the same discriminant, say Δ, that is, they belong to the hyperboloid

$$K^2 + D^2 - S^2 = \Delta.$$

Moreover, in the elliptic case, the orbit lies entirely either on the upper or the lower sheet of the hyperboloid; in the parabolic case, it lies entirely either on the upper or the lower cone.

1.2 The semigroups $\mathcal{T}^+$ and $\mathcal{T}^-$

We consider the generators of the group $\mathrm{SL}(2,\mathbb{Z})$

$$\mathbf{A}=\begin{pmatrix}1&1\\0&1\end{pmatrix},\quad \mathbf{B}=\begin{pmatrix}1&0\\1&1\end{pmatrix},\quad \mathbf{R}=\begin{pmatrix}0&1\\-1&0\end{pmatrix} \tag{3}$$

and their inverse operators denoted by $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$, and $\mathbf{R}^{-1}$.

Note that

$$\mathbf{R}=\mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}=\mathbf{A}\mathbf{B}^{-1}\mathbf{A} \quad \text{and} \quad \mathbf{R}^{-1}=\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}=\mathbf{B}\mathbf{A}^{-1}\mathbf{B}. \tag{4}$$

Let A, B, and R denote the corresponding operators of $\mathcal{T}$ obtained from Eq. (2) and A^{-1}, B^{-1}, and R^{-1} denote their inverses.

Observe that $\mathbf{A}$ and $\mathbf{R}$ are sufficient to generate $\mathrm{SL}(2,\mathbb{Z})$ and constitute the standard basis of this group.

Remark. In the coordinates (m,n,k) the matrices of the generators A, B, and R of the group $\mathcal{T}$, are

$$A=\begin{pmatrix}1&0&0\\1&1&1\\2&0&1\end{pmatrix},\quad B=\begin{pmatrix}1&1&1\\0&1&0\\0&2&1\end{pmatrix},\quad R=\begin{pmatrix}0&1&0\\1&0&0\\0&0&-1\end{pmatrix}.$$

The matrices of the same generators in the coordinates $[K,D,S]$ are

$$A=\begin{pmatrix}1&1&1\\-1&1/2&-1/2\\1&1/2&3/2\end{pmatrix},\quad B=\begin{pmatrix}1&-1&1\\1&1/2&1/2\\1&-1/2&3/2\end{pmatrix},$$

$$R=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix}.$$

The matrices of $\mathbf{A}$ and $\mathbf{B}$ are the transposes of each other, and the same holds for A and B, while the transpose of $\mathbf{R}$ is equal to $\mathbf{R}^{-1}$. Since the transpose and the inverse of R are both equal to R, relations (4) become

$$R=B^{-1}AB^{-1}=AB^{-1}A=A^{-1}BA^{-1}=BA^{-1}B. \tag{5}$$

Let $\mathcal{T}^+$ ($\mathcal{T}^-$) denote the multiplicative semigroup of the elements of $\mathcal{T}$ generated by the identity and by the operators A and B (respectively by A^{-1} and B^{-1}).

Definition. A product of n operators $\prod_{i=1}^{n} T_i$, where either $T_i = A$ or $T_i = B$, is called *word of length n* in A and B.

Lemma 1.1.

(a) *Each operator $T \in \mathcal{T}^+$ ($T \in \mathcal{T}^-$) is written uniquely as word in A and B (in A^{-1} and B^{-1}).*

(b) *Each operator $T \in \mathcal{T}$ can be written as the product VSU, where S belongs to $\mathcal{T}^+$ and the operators U and V belong to the set $\{E, R\}$, where E is the identity. Statement (b) also holds with $\mathcal{T}^+$ replaced with $\mathcal{T}^-$.*

Proof.

(a) There is no relation involving only the operators $\mathbf{A}$ and $\mathbf{B}$ in $\mathrm{SL}(2, \mathbb{Z})$, and hence no relation involving only A and B.

(b) Relations (5) allow transforming any word in A, B, R, and their inverse operators into a word of type VSU. $\square$

Figure 3 shows the one-to-one correspondence between the tiles of the Lobachevsky disc and the elements of the group $\mathcal{T} \simeq \mathrm{PSL}(2, \mathbb{Z})$: the domain corresponding to a given element of the group is the image of the principal fundamental domain (I) by that element. Thus, by Lemma 1.1, any domain is the image of I by an element of $\mathcal{T}$ of the form VSU.

Indeed, any domain in the right half-disc is obtained from I by an operator of the form SU (using the notation of Lemma 1.1). The same holds for the domains in the left half-disc replacing $\mathcal{T}^+$ with $\mathcal{T}^-$. The multiplication by R from the left acts as a reflection with respect to the center. Hence, each domain in one half-disc can be obtained from I by the operator corresponding to the domain symmetric to it with respect to the center, multiplied from the left by R.

1.3 Symmetries of the form classes

We present some different types of symmetries that the classes of forms may have.

To each form $\mathbf{f} = (m, n, k)$, there correspond eight forms, obtained from $\mathbf{f}$ by combining three involutions (see Fig. 1):

$$\begin{aligned} \mathbf{f}_c &= (n, m, -k), \\ \bar{\mathbf{f}} &= (m, n, -k), \\ \mathbf{f}^* &= (-n, -m, k). \end{aligned} \tag{6}$$

All these involutions commute and preserve the discriminant because they correspond to changes of sign of some of the coordinates K, D, S. In these coordinates,

$$\mathbf{f}_c = [-K, -D, S], \quad \bar{\mathbf{f}} = [-K, D, S], \quad \mathbf{f}^* = [K, D, -S].$$

Thus, the eight forms defined by these involutions on the form $\mathbf{f}$ lie on the same hyperboloid as $\mathbf{f}$.

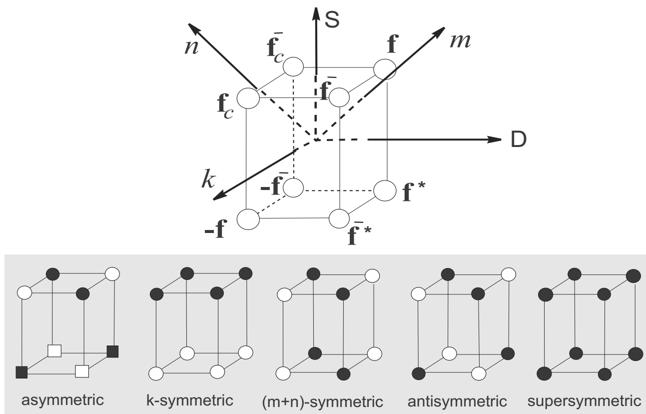


Figure 1: For every symmetry type, the forms denoted by the same symbol and the same color belong to the same class.

The *complementary* form $\mathbf{f}_c$ always belongs to the class of $\mathbf{f}$ because $\mathbf{f}_c = R\mathbf{f}$ and $R \in \mathcal{T}$. We note that the complementary form $\mathbf{f}_c$ of the form $\mathbf{f}$ satisfies $f_c(x, y) = f(y, -x) = f(-y, x)$ in the xy plane, and the corresponding $\mathrm{PSL}(2, \mathbb{Z})$ change of coordinates is a rotation by $\pi/2$.

The complementary of the *conjugate* form $\bar{\mathbf{f}}_c = (n, m, k)$ of the form $\mathbf{f} = (m, n, k)$ is obtained by the reflection of the xy plane with respect to the diagonal, whose operator $\left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$ does not belong to $\mathrm{SL}(2, \mathbb{Z})$.

The opposite form $-\mathbf{f} = (-m, -n, -k) = [-K, -D, -S]$ is the complementary of the adjoint of $\mathbf{f}$, i.e., $-\mathbf{f} = \mathbf{f}_c^*$.

The forms obtained from a form $\mathbf{f}$ by conjugation and/or adjunction may or may not belong to $C(\mathbf{f})$. But if a class contains a pair of forms related by one of the above involutions or a form that is invariant under such an involution, then the entire class is invariant under that involution.

Proposition 1.2. *Let σ be one of the involutions: $\sigma(\mathbf{f}) = \bar{\mathbf{f}}$, $\sigma(\mathbf{f}) = \mathbf{f}^*$ or $\sigma(\mathbf{f}) = \bar{\mathbf{f}}^*$. If, for some $\mathbf{f}$, $\sigma(\mathbf{f}) \in C(\mathbf{f})$, then any $\mathbf{g} \in C(\mathbf{f})$ satisfies $\sigma(\mathbf{g}) \in C(\mathbf{f})$.*

Proof. We must first prove the following lemma.

Lemma 1.3. *The following identities hold:*

$$\begin{aligned} &1.\quad (A\mathbf{f})^* = B^{-1}\mathbf{f}^*; && (B\mathbf{f})^* = A^{-1}\mathbf{f}^*;\\ &2.\quad \overline{A\mathbf{f}} = A^{-1}\,\overline{\mathbf{f}}; && \overline{B\mathbf{f}} = B^{-1}\,\overline{\mathbf{f}};\\ &3.\quad (\overline{A\mathbf{f}})^* = B\overline{\mathbf{f}}^*; && (\overline{B\mathbf{f}})^* = A\overline{\mathbf{f}}^*. \end{aligned} \tag{7}$$

Proof of the lemma. Let $\mathbf{f} = (m, n, k)$. We have $A\mathbf{f} = (m, m+n+k, k+2m)$ and $B\mathbf{f} = (m + n + k, n, k + 2n)$.

1. Since $\mathbf{f}^* = (-n, -m, k)$, we have $(A\mathbf{f})^* = (-m-n-k, -m, k+2m)$ and $B^{-1}\mathbf{f}^* = (-n-m-k, -m, k+2m)$; $(B\mathbf{f})^* = (-n, -n-m-k, k+2n)$ and $A^{-1}(\mathbf{f}^*) = (-n, -m - n - k, k + 2n)$.
2. Since $\overline{\mathbf{f}} = (m, n, -k)$, we have $\overline{A\mathbf{f}} = (m, m + n + k, -k - 2m)$ and $A^{-1}\,\overline{\mathbf{f}} = (m, m + n + k, -k - 2m)$; $\overline{B\mathbf{f}} = (m + n + k, n, -k - 2m)$ and $B^{-1}\,\overline{\mathbf{f}} = (m, m + n + k, -k - 2m)$.
3. Since $\overline{\mathbf{f}}^* = (-n, -m, -k)$, we have $\overline{A\mathbf{f}} = (m, m + n + k, -k - 2m)$, $(\overline{A\mathbf{f}})^* = (-n-m-k, -m, -k-2m)$, and $B\overline{\mathbf{f}}^* = (-n-m-k, -m, -k-2m)$; $\overline{B\mathbf{f}} = (m+n+k, n, -k-2n)$, $(\overline{B\mathbf{f}})^* = (-n, -n-m-k, -k-2n)$, and $A\overline{\mathbf{f}}^* = (-n, -m - n - k, -k - 2n)$. □

Proof of Proposition 1.2. If $\mathbf{g} \in C(\mathbf{f})$ then $\mathbf{g} = T\mathbf{f}$ for some operator $T \in \mathcal{T}$. But any operator $T \in \mathcal{T}$ can be written as a word in A and B and their inverses. Then Lemma 1.3 implies that $\sigma(T\mathbf{f}) = T'\sigma(\mathbf{f})$ for some $T' \in \mathcal{T}$. Therefore, if $\sigma(\mathbf{f}) \in C(\mathbf{f})$, then $\sigma(\mathbf{g}) = T'\sigma(\mathbf{f}) \in C(\mathbf{f})$. □

Definition. A class of forms is said to be (see Fig. 1)

1. *asymmetric* if it is only invariant under reflection with respect to the axis of the coordinate S ($K = 0$, $D = 0$). (It contains only pairs of complementary forms);
2. *supersymmetric* if it contains all eight forms obtained by combining the three involutions.
3. *k-symmetric* if it is not supersymmetric but is invariant under reflection with respect to the plane $k = 0$ ($K = 0$). (It contains the *conjugate* form $\overline{\mathbf{f}}$ for each $\mathbf{f}$);

4. $(m+n)$-*symmetric* if it is not supersymmetric but is invariant under reflection with respect to the plane $m+n=0$ $(S=0)$. (It contains the *adjoint* form $\mathbf{f}^*$ for each $\mathbf{f}$);

5. *antisymmetric* if it is not supersymmetric but is invariant under reflection with respect to the planes $m=0$ and $n=0$ $(|S|+|D|=0)$. (It contains the *antipodal* form $\bar{\mathbf{f}}^* = (-n,-m,-k) = [-K, D, -S]$ for each $\mathbf{f}$).

Remarks.

1. Each of the above types of classes is invariant under reflection with respect to some plane or some axis through the origin of the coordinate system, a plane or axis that is noninvariant under the action of the group $\mathcal{T}$. Hence, these symmetries a priori no longer hold in another system of coordinates. But we proved (Proposition 1.2) that the action of the group $\mathcal{T}$ preserves each of the symmetries, and the same symmetry definitions hence hold in any system of coordinates obtained by a $\mathcal{T}$ coordinate transformation. This is equivalent to saying that a symmetry of a class of forms is a symmetry with respect to all infinite planes (or axes) that are the images under $\mathcal{T}$ of one of such symmetry planes (or axes).

2. The opposite form, $-\mathbf{f}$, belongs to the class of $\mathbf{f}$ only if the class is $(m+n)$-symmetric or supersymmetric (see Fig. 1).

2 Elliptic forms

In this section, we treat the classification of positive definite forms to introduce some notions and terms that are used in Secs. 3 and 4.

We define a map from one sheet of the two-sheeted hyperboloid to the open unitary disc, which gives the explicit one-to-one correspondence between the integer points of an orbit on the hyperboloid and the domains of the classical Poincaré tiling of the Lobachevsky disc.

Let $\mathcal{P}$ be the following *normalized* projection from the upper sheet of the hyperboloid $K^2 + D^2 - S^2 = \Delta$ $(\Delta < 0)$ to the disc of unit radius. Let $\mathbf{p} = [K, D, S]$ be a point on the hyperboloid (see Fig. 6, left), $\mathbf{p}'$ be its projection from the point $O' = [0, 0, -\rho]$ $(\rho = \sqrt{-\Delta})$ to the disc of radius ρ in the plane $S = 0$. The image of the normalized projection $\mathcal{P}\mathbf{p}$ is defined by

$$\mathcal{P}\mathbf{p} = \begin{cases} \widetilde{K} = \frac{K}{\rho+S} \\ \widetilde{D} = \frac{D}{\rho+S}. \end{cases} \tag{8}$$

Let $\mathbf{L} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be an operator of $\mathrm{SL}(2, \mathbb{Z})$ and L be its corresponding operator of $\mathcal{T}$ defined by (2).

Let $\widetilde{L}$ denote the operator acting on the disc of radius 1 by the rule

$$\widetilde{L}(\mathcal{P}\mathbf{p}) = \mathcal{P}(L\mathbf{p}). \tag{9}$$

Besides $\widetilde{L}$, the operator $\mathbf{L} \in \mathrm{SL}(2, \mathbb{Z})$ determines another map from the disc to itself. Let H_L be the homographic operator acting on the upper complex half-plane $\{z \in \mathbb{C} : \mathrm{Im}(z) \geq 0\}$:

$$H_L z = \frac{az + b}{cz + d}. \tag{10}$$

The following map $\pi : z \to w$ sends the upper complex half-plane to the unitary complex disc $\{w \in \mathbb{C} : |z| \leq 1\}$ (Fig. 2):

$$w = \pi z = \frac{1 + iz}{1 - iz}. \tag{11}$$

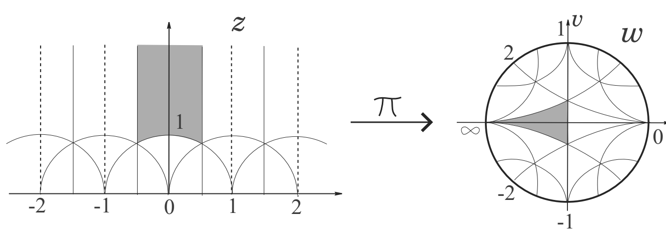


Figure 2: The map π from the upper-half plane to the Lobachevsky disc. The standard principal fundamental domain is shown in gray color.

We define the operator $\widehat{L}$ acting on the complex unit disc by

$$\widehat{L}(w) = \pi(H_L \circ \pi^{-1}(w)). \tag{12}$$

Proposition 2.1. *The actions of the operators $\widehat{L}$ and $\widetilde{L}$ on the unitary disc coincide under the identification*

$$\widetilde{D} = \mathrm{Re}(w), \quad \widetilde{K} = \mathrm{Im}(w).$$

We prove it by writing the operators corresponding to the $\mathrm{SL}(2, \mathbb{Z})$ generators explicitly and by comparing their actions on the coordinates of a point of the disc.

Remark. The group of operators defined by Eq. (9) is isomorphic to the homographic group of the operators H_L and to the group $\mathcal{T}$, i.e., to $\mathrm{PSL}(2, \mathbb{Z})$. We designate its generators by the same letters as the corresponding generators of $\mathcal{T}$.

We take the pair $(\widetilde{K}, \widetilde{D})$ as coordinates in the Lobachevsky disc. Hence, our Lobachevsky disc is obtained from the unitary complex disc with the coordinate $w = u + iv$ (see Figs. 2 and 3) by reflection with respect to the diagonal $v = u$.

The Lobachevsky disc is shown in Figure 3. The principal fundamental domain is indicated by the letter I (the bold line at the boundary belongs to it and the dotted line does not). The other domains are obtained by applying some elements of $\mathrm{PSL}(2, \mathbb{Z})$, written in terms of R, A, B, and their inverses to I.

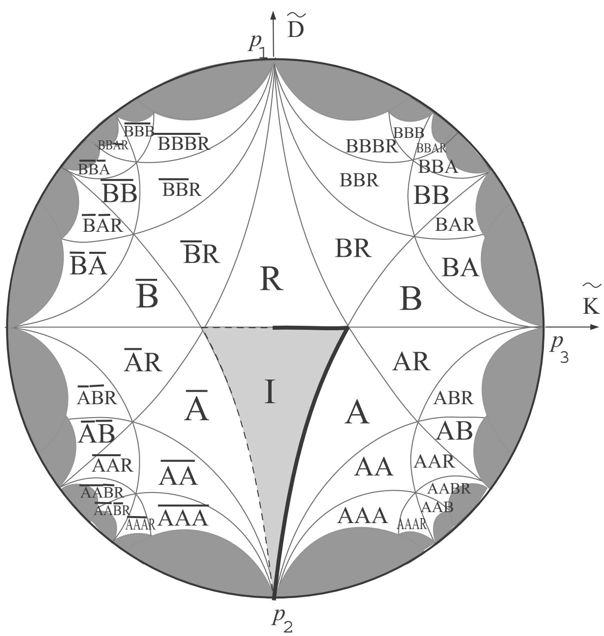


Figure 3: A finite set of domains in the Lobachevsky disc with coordinates $\widetilde{K}$, $\widetilde{D}$.

The expressions are not unique because of relations (5) involving these generators. We have chosen this representation in order to see the meaning of Lemma 1.1, which is decisive in Sec. 4.

Remark. The choice of a fundamental domain as principal is arbitrary, as well as the choice of a coordinate system in the plane of forms related to the canonical one by an element of $\mathrm{PSL}(2, \mathbb{Z})$.

In the figures, the inverse operators A^{-1} and B^{-1} are denoted by $\overline{A}$ and $\overline{B}$.

Each orbit has exactly one point in each domain. In Figure 4, the Lobachevsky disc with a finite subset of domains is shown together with a finite part of the three distinct orbits in the case $\Delta = -31$.

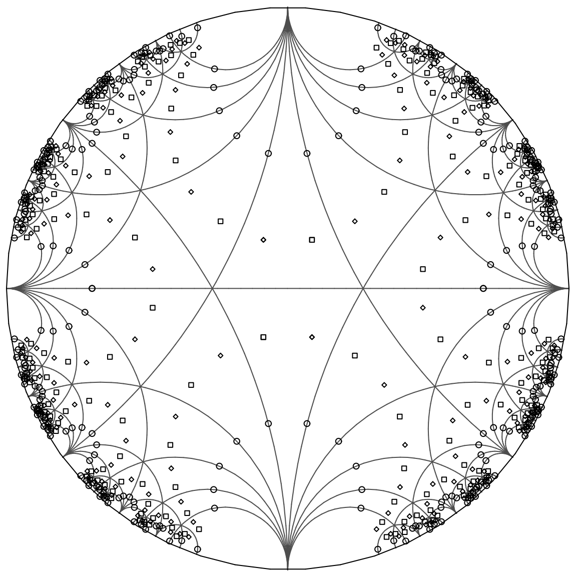


Figure 4: Finite subsets of the three distinct orbits in the case $\Delta = -31$, projected to the Lobachevsky disc, with coordinates $(\widetilde{K}, \widetilde{D})$. The projection of the representative point (m, n, k) of each orbit lies in the principal fundamental domain. Two asymmetric orbits are shown by boxes (2,4,−1) and rhombi (2,4,1), and one k-symmetric orbit is shown by circles (1,8,1).

Remark. The opposite, the antipodal, and the adjoint of a positive-definite or negative-definite quadratic form f are respectively negative-definite or positive-definite quadratic forms; hence, they cannot belong to the same class of f. Therefore, a class of elliptic forms can have only two types of symmetries: it is either k-symmetric or asymmetric.

2.1 The hierarchy of the points at infinity

This section is important for our study of hyperbolic forms.

Let C denote the circle at infinity bounding the Lobachevsky disc. By the $\mathrm{PSL}(2, \mathbb{Z})$ action, the points of C with rational coordinates inherit a hierarchy (explained below), on which our partition of the de Sitter word is based.

Let π' denote the composition of the map π (see eq. (11)) with the reflection of the disc-image ($|w| \leq 1$) with respect to the diagonal ($\mathrm{Im}(w) = \mathrm{Re}(w)$).

We consider only the right half of the circle C because of the symmetry of the picture. This semicircle ($\widetilde{K} \geq 0$) is the image under π' of the real half-line (together with the infinite) $x \equiv \mathrm{Re}(z) > 0$ of the half-plane where the homographic operators act.

The endpoints of this semicircle, $p_1 \colon (\widetilde{K}, \widetilde{D}) = (0, 1)$ and $p_2 \colon (\widetilde{K}, \widetilde{D}) = (0, -1)$, are the points of the zeroth generation, being the images under π' of the points $x_1 = 0$ and $x_2 = \infty$.

The rational points x_i of the real half-line are written as fractions: $0 \equiv 0/1$, $\infty \equiv 1/0$, $q \equiv q/1$ if $q \in \mathbb{Z}$, etc., and the points p_i are their images under π' on C.

Here, A and B are the generators of the homographic group associated with the generators $\mathbf{A}$ and $\mathbf{B}$ of $\mathrm{SL}(2, \mathbb{Z})$:

$$\mathrm{A} : x \to \frac{x+1}{1}; \quad \mathrm{B} : x \to \frac{x}{x+1}. \tag{13}$$

We consider the iterated action of such generators on the points $x_1 = 0$ and $x_2 = \infty$. We first have

$$\mathrm{A}x_1 = x_3, \quad \mathrm{B}x_1 = x_1, \quad \mathrm{A}x_2 = x_2, \quad \mathrm{B}x_2 = x_3, \tag{14}$$

where $x_3 = 1/1$ is the preimage under π' of the point p_3 with coordinates $\widetilde{K} = 1$, $\widetilde{D} = 0$.

Definition. The *points of the nth generation* ($n > 1$) in $\mathbb{R}^+$ are obtained from the point $x_3 = 1/1$ of first generation by applying all the 2^n words of length n in the generators A and B to it.

The hierarchy and the order of these points is shown in the following scheme, where T denotes any word of length $n - 1$ in the generators A and B:

generation				
n		T(1/1)		
n+1	TB(1/1)		TA(1/1)	
n+2	TBB(1/1)	TBA(1/1)	TAB(1/1)	TAA(1/1)

The points of all generations have a nice algebraic property that we recall.

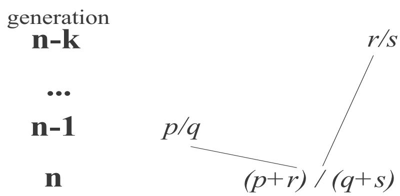
Definition. We call the points TA(1/1) and TB(1/1) of the (n+1)th generation *sons* (respectively, the A-son and the B-son) of the point T(1/1) of the nth generation. Thus, T(1/1) is the *father* of his sons. In the scheme above, the segments indicate the father-son relations.

Remark. The following order relations hold: TB(1/1) < T(1/1) < TA(1/1). Moreover,

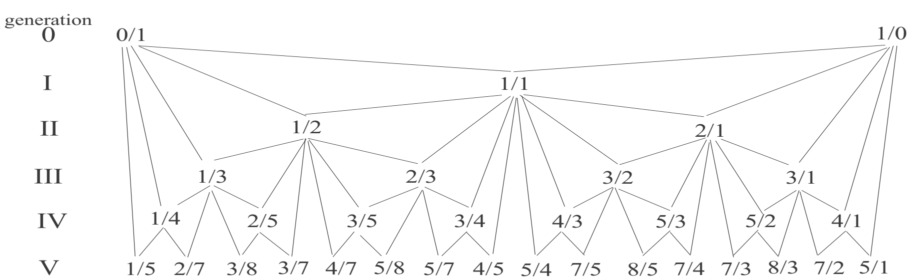
$$\mathrm{TBU}(1/1) < \mathrm{T}(1/1) < \mathrm{TAV}(1/1),$$

where U and V are arbitrary words in the generators A and B. That is, all descendants from the B-son of a number x_i are less than all descendants from the A-son of x_i.

Farey rule. *The coordinate of a point x_i of the nth generation can be calculated directly from those of his father and his nearest ancestor (i.e., the point, among its ancestors, which is the closest to it in $\mathbb{R}^+$), by the rule shown in the following scheme:*



In the following scheme of the hierarchy, each point x_i is connected by segments to its two sons, to its father, and to its closest ancestor. Note that the descendants of B(1/1) are the inverse fractions of the descendants of A(1/1).



Remark. All positive rational numbers are covered by this procedure.

The point $x_i = p/q$ is sent by π' to the point of C with the coordinates $(\widetilde{K}, \widetilde{D})$:

$$\widetilde{K} = \frac{2pq}{p^2 + q^2}, \qquad \widetilde{D} = \frac{p^2 - q^2}{p^2 + q^2}. \tag{15}$$

The map π', restricted to the right half line, induces an ordering and a hierarchy on the rational points of the right half-circle, from those of the positive rational points.

We will define a similar ordering and hierarchy on the rational points of the left half-circle. Observe that the point $(\widetilde{K}, \widetilde{D}) = (-1, 0)$ of C is the image under π' of the point $-1/1$. The above construction can be repeated by the iterated action of the inverse generators A^{-1} and B^{-1} on the point $(-1/1)$ by regarding the point $-1/0 = -\infty$ as the preimage of the point p_2. The rational points in the real half-line $x < 0$ are thus endowed with an ordering and a hierarchy which are inherited by the points with rational coordinates on the left half of the circle C.

3 Parabolic forms

The map π' sends the rational numbers to the Pythagorean triples $\{(a, b, c) \in \mathbb{Z}^3 : a^2 + b^2 = c^2\}$:

$$\frac{p}{q} \to \left(2pq, \quad q^2 - p^2, \quad p^2 + q^2\right). \tag{16}$$

Definition. A good point $[K, D, S]$ is said to be *Pythagorean* if $K^2 + D^2 = S^2$. If K, D, and S have no common divisors, then the triple is said to be *simple*. If $[K, D, S]$ is Pythagorean, then the set of points $\{[\lambda K, \lambda D, \lambda S], \ \lambda \in \mathbb{Z}\}$, all Pythagorean, is called a *Pythagorean line*.

The Pythagorean points belong to the cone $\Delta = 0$, and any good point belonging to the cone is Pythagorean.

Lemma 3.1. *There exists a one-to-one correspondence between the Pythagorean lines and the points with rational coordinates p_i on the circle C.*

Proof. By formula (16) we associate a Pythagorean triple with each point p_i on the circle C. This triple represents a good point because $b - c \equiv 0 \mod 2$. On the other hand, given a simple good point $[K, D, S]$, the equations

$$K = 2pq, \quad D = p^2 - q^2, \quad S = p^2 + q^2$$

have the solution

$$p = \sqrt{\frac{(S + D)}{2}}, \quad q = \sqrt{\frac{(S - D)}{2}}, \quad \text{i.e.,} \quad p = \sqrt{m}, \quad q = \sqrt{n}.$$

Since $S = m + n$ and $D = m - n$ and in this case $K = 2\sqrt{mn}$, m and n have no common divisors; otherwise, the triple $[K, D, S]$ would not be simple. But the equality $K^2 = 4mn$ with m and n relatively prime implies that $m = p^2$

and $n = q^2$ for some integers p and q. Hence, we associate a point p_i on the circle C with each simple Pythagorean triple, and vice versa. The Pythagorean line corresponding to $p_i = (\tilde{K}, \tilde{D})$ is the line through 0 and $[\tilde{K}, \tilde{D}, 1]$ on the cone of parabolic forms. $\square$

Theorem 3.2. *The set of classes of forms with* $\Delta = 0$ *is parametrized by the forms of type* ax^2, *with* $a \in \mathbb{Z}$.

Proof. We prove that for all $a \in \mathbb{Z}$, the orbits containing the points $[K, D, S] = [0, a, a]$ are distinct. Let us suppose that the point $\mathbf{r} := [0, b, b]$ belongs to the same orbit of the point $\mathbf{p} = [0, a, a]$. The form $f = ax^2$ is therefore in the same class as the form $f' = bx^2$. This means that there exists an operator $\mathbf{L} = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ of $\mathrm{SL}(2, \mathbb{Z})$ such that $a(\alpha x + \beta y)^2 = bx^2$. This can be satisfied only by $\beta = 0$, and by $\alpha = \delta = 1$ because $\alpha\delta - \gamma\beta = 1$. Hence, $a = b$. On the other hand, any parabolic form $f = mx^2 + ny^2 + kxy$, satisfying $k^2 - 4mn = 0$, can be written as $a(\alpha x + \beta y)^2$, where a is the greatest common divisor of (m, n, k). The integers α and β are relatively prime because they are the elements of a row of an $\mathrm{SL}(2, \mathbb{Z})$ matrix. For every pair of relatively prime integers α and β, there exist two integers γ and δ such that $\alpha\delta - \gamma\beta = 1$. Hence, the inverse of the operator L is the operator of $\mathrm{SL}(2, Z)$ transforming f into ax^2. $\square$

4 Hyperbolic forms

Let X be the set of planes through the origin in the three-dimensional space with coordinates K, D, S obtained from the plane $D = 0$ by the action of group $\mathcal{T}$. These planes subdivide the interior of the cone $(K^2 + D^2 < S^2)$ into domains (some of these planes are shown in Fig. 5). These planes intersect both sheets of the two-sheeted hyperboloid and subdivide them into domains; the domains that belong to the upper sheet are projected by $\mathcal{P}$ (see Sec. 2) to the domains of the Lobachevsky disc.

The closure $\overline{X}$ of X contains the planes tangent to the cone along all Pythagorean lines.

The intersections of the planes of $\overline{X}$ with the one-sheeted hyperboloid H $(K^2 + D^2 - S^2 = 1)$ form a net of lines that is dense in H.

In the interior of the unit disc, the intersections of the plane $S = 1$ with the planes of the set $\overline{X}$ are the lines of the Klein model of the Lobachevsky disc. The arcs of circles joining pairs of points of the circle at infinity of the Poincaré model are substituted by the chords connecting these points. We are interested in the prolongations of these chords outside the disc. The description of the de

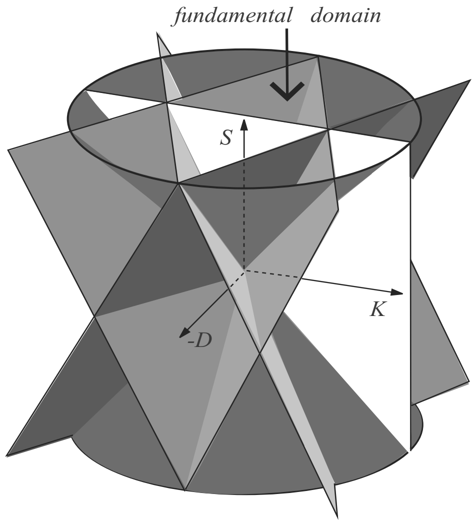


Figure 5: The principal fundamental domain in the space of form coefficients.

Sitter world is based on the "limit" chords, i.e., the tangents to the circle at all rational points of it.

4.1 The Poincaré tiling of the de Sitter world

Let H_Δ denote the one sheeted hyperboloid with equation $K^2 + D^2 - S^2 = \Delta$, $\Delta > 0$, in the coordinates $[K, D, S]$.

By analogy with the standard projection $\mathcal{P}$ from the upper sheet of the two-sheeted hyperboloid to the Lobachevsky disc, we have chosen the following projection $\mathcal{Q}$ from H_Δ to the open cylinder C_H:

$$C_H = \Big\{[K, D, S] \,:\, K^2 + D^2 = 1, \;\; |S| < 1\Big\}.$$

The coordinates (s, ϕ) of the cylinder C_H are obtained from the coordinates K, D, S of H_Δ by:

$$s = \frac{S}{r + \rho}, \qquad \text{where} \qquad r = \sqrt{K^2 + D^2} \quad \text{and} \qquad \rho = \sqrt{\Delta}, \qquad (17)$$

and ϕ is the angle defined by the relations $K = r\cos\phi$ and $D = r\sin\phi$ (see Fig. 6, right).

Remark. Two points $\mathbf{f}$ and $\mathbf{f}'$ belonging to different hyperboloids H_Δ and $H_{\Delta'}$ have the same projection in C_H iff $\mathbf{f}' = \alpha\mathbf{f}$, where $\alpha = \sqrt{\Delta'/\Delta}$.

The border of the cylinder consists of two circles, denoted by c_1 ($s = 1$) and c_2 ($s = -1$).

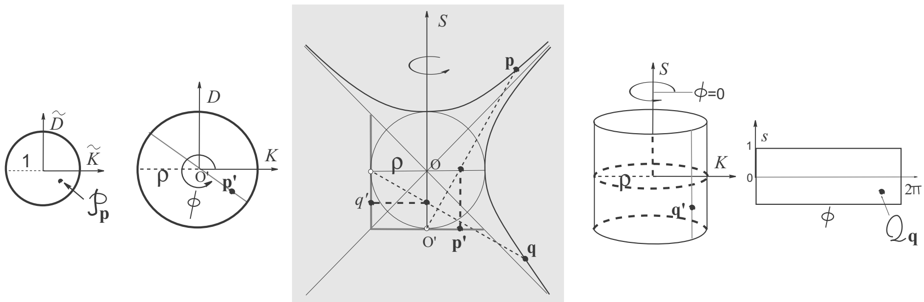


Figure 6: Projection $\mathcal{P}$, left, and $\mathcal{Q}$, right.

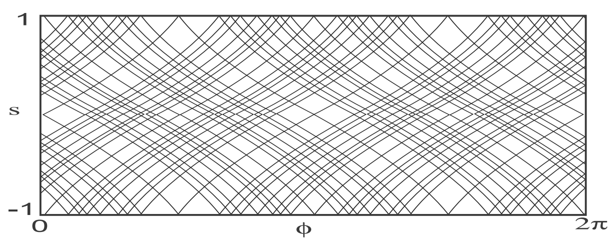


Figure 7: Projection on the cylinder C_H of some lines, intersection of H with the planes tangent to the cone along Pythagorean lines.

Let H^0 and H^0_R denote the open domains

$$\begin{aligned} H^0 &= \{[K, D, S] \in H : |S| < |D|, D > 0\}; \\ H^0_R &= \{[K, D, S] \in H : |S| < |D|, D < 0\}. \end{aligned}$$

Observe that $H^0_R = RH^0$.

Remark. Since the planes tangent to the cone intersect the hyperboloid H along two of its generatrices, the boundaries of H^0 and H^0_R are straight half-lines.

For simplicity, we let the same letters denote the domains on H and their images under $\mathcal{Q}$ on the cylinder C_H.

The circles c_1 and c_2 in the respective planes $S = 1$ and $S = -1$ coincide with the circle C at infinity of the Lobachevsky disc. Hence, the points with rational coordinates $(\widetilde{K}, \widetilde{D})$ on c_1 and c_2 are also mapped, according to (15), into the

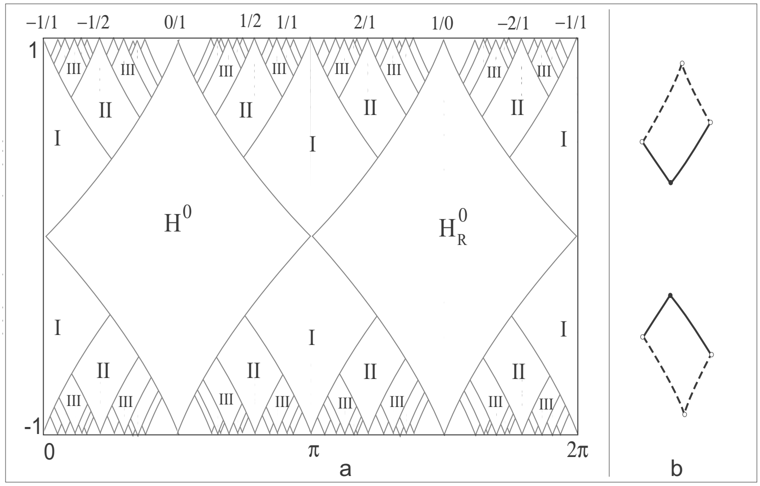


Figure 8: a) The Poincaré tiling of the de Sitter world: domains of the first, second, and third generations are indicated by I, II, and III. b) The segments at the border of a connected component are solid if they belong to it, dashed otherwise. The vertices are represented by black discs if they belong to the connected component, by white discs otherwise.

points p_i of the first, second, third, ... generations with the hierarchy explained in Sec. 2.1.

On the upper circle c_1 in Figure 8, the points p_i up to the second generation are denoted by the corresponding rational numbers x_i.

Note that the upper and lower vertices of the domains H^0 and H^0_R have the coordinates $\phi = \pi/2$ and $\phi = 3\pi/2$, and correspond to the points $x_1 = 0/1$ and $x_2 = 1/0 = \infty$.

Definition. Let H^{x_i} and H^{-1/x_i} denote the domains obtained from H^0 and H^0_R by a rigid rotation of C_H about the S-axis such that the upper vertex of H^0 transfers to the point $\pi'(x_i)$ and the upper vertex of H^0_R transfers to the point $\pi'(-1/x_i)$. The domains H^{x_i} thus inherit the hierarchy of the points x_i. We call $H^0 = H^{0/1}$ and $H^0_R = H^{1/0}$ *rhombi*[2] *of the zeroth generation*, $H^{1/1}$ and $H^{-1/1}$ *rhombi of the first generation*, $H^{1/2}$, $H^{-2/1}$, $H^{-1/2}$, and $H^{2/1}$ *rhombi of the second generation*, and so on.

Let H^O, H^I, H^{II}, H^{III}, ... denote the unions of the rhombi of the zeroth, first, second, third, ... generations.

[2]Evidently, they are not exactly rhombi neither in H nor in C_H.

Definition. We call *Poincaré tiling* of the de Sitter world the cylinder C_H provided with the subdivision into domains obtained by the following procedure. Let $\mathcal{H}^0 = H^0 \cup H^0_R$ be the *domain of the zeroth generation.* Let $\mathcal{H}^I = H^I \setminus H^O$ be the domain of the first generation, $\mathcal{H}^{II} = H^{II} \setminus (H^O \cup H^I)$ be the domain of the second generation, and so on; the domain of the nth generation is thus obtained as

$$\mathcal{H}^{\mathbf{n}} = H^{\mathbf{n}} \setminus \big(H^O \cup H^I \cup H^{II} \cup \cdots \cup H^{\mathbf{n-1}}\big).$$

The partition of C_H we have introduced is the projection of the partition of H by planes through the origin. Therefore, the Poincaré tiling is a universal model for the hyperboloid H ($\Delta = 1$) as well as for H_Δ, for every positive Δ.

Figure 8 shows the domains of different generations. For $n > 0$, the domain of the nth generation, $\mathcal{H}^{\mathbf{n}}$, has 2^{n+1} connected components.

Each connected component of $\mathcal{H}^{\mathbf{n}}$, $n > 0$, has the form of a rhomboid in C_H. Observe that the two segments bordering the bottom of a connected component belong to this connected component iff $s \geq 0$ and the two segments bordering the top of a connected component belong to this connected component iff $s \leq 0$ (see Fig. 8,b).

The action of A, B, A^{-1} and B^{-1} on the domain H^0 is shown in Figure 9, where ϕ varies from $-\pi/2$ to $3\pi/2$ and H^0 is in the center.

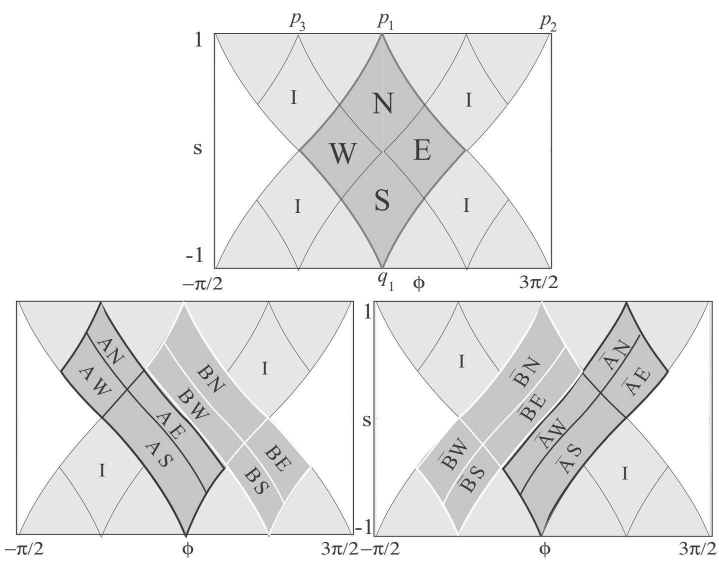


Figure 9: Images under A, B, A^{-1}, and B^{-1} of H^0. The letter I indicates the four connected components of the domain of the first generation $\mathcal{H}^I$.

In the figures and in the subscript indices, the inverse operators A^{-1} and B^{-1} are denoted by $\overline{A}$ and $\overline{B}$ for short.

We have subdivided H^0 into subdomains, denoted by N, S, E, and W. The operators A and B and their inverses map H^0 partly to itself and partly outside H^0 (see Fig. 9).

Remark. Since $H^0_R = RH^0$, the corresponding actions on H^0_R are obtained from the following relations coming from (5):

$$A = RB^{-1}R; \quad B = RA^{-1}R; \quad A^{-1} = RBR; \quad B^{-1} = RAR. \tag{18}$$

Lemma 4.1. *The parts of the images of H^0 (resp. of H^0_R) under A, B, A^{-1} and B^{-1} that are not in H^0 (resp. in H^0_R) are disjoint and they form the first-generation domain $\mathcal{H}^I$.*

Proof. The boundary of each domain is formed by segments of straight lines and/or by half-straight lines in $\mathbb{R}^3$, by construction. Therefore it suffices to calculate the action of the generators on the vertices of H^0 and on the intersections of its boundary lines with the boundaries of the rhombi of the first generation in C_H. The actions of A and B on the points on the circles c_1 and c_2 are obtained using equations (14). We need only observe that any point q_i on the circle c_2 symmetric to the point p_i on c_1, regarded as the opposite vertex of the same rhomboidal domain, is opposite ($q_i = -p_i'$) to the point p_i' on the circle c_1 at the distance π from p_i. For instance, according to (14), the image under A of the extreme north p_1 of H^0 is p_3, while the image under A of the extreme south, q_1, is q_1, because $q_1 = -p_2$ and $Ap_2 = p_2$ (see Fig. 9). To conclude the proof, we observe that the interiors of the segments separating H^0 from the connected components of $\mathcal{H}^I$ do not belong to H^0, which is open. But they belong to the images of H^0 under A and B and their inverses, and therefore coincide with the interiors of the segments at the boundary of the first generation domain. The extremes of these segments that have coordinate $s = 0$ do no belong to H^0 but belong to its image under A and B and their inverses, and therefore they coincide with the vertices of the connected components of $\mathcal{H}^I$ belonging to $\mathcal{H}^I$. The extremes of these segments that have coordinate $s \neq 0$ belong neither to H^0 nor to its images under A and B and their inverses, and indeed they do not belong to $\mathcal{H}^I$. $\square$

Let the four connected components of $\mathcal{H}^I$, where $|S| < |K|$ and $|S| \geq |D|$, be denoted, together with their images under $\mathcal{Q}$, by (see Fig. 10, left):

$$\begin{aligned} H_A &= \{[K,D,S] \in \mathcal{H}^I : S \geq 0, K > 0\}; \\ H_{\bar{A}} &= \{[K,D,S] \in \mathcal{H}^I : S \geq 0, K < 0\}; \\ H_B &= \{[K,D,S] \in \mathcal{H}^I : S \leq 0, K < 0\}; \\ H_{\bar{B}} &= \{[K,D,S] \in \mathcal{H}^I : S \leq 0, K > 0\}. \end{aligned} \tag{19}$$

Remark. The subscript of a connected component of the first-generation domain $\mathcal{H}^I$ indicates the operator that sends one part of H^0 onto this connected component. Moreover,

$$H_{\bar{A}} = RH_A; \quad H_{\bar{B}} = RH_B.$$

Figure 10, left, shows the actions of the operators A and B on H^0, H^0_R and the four connected components of $\mathcal{H}^I$.

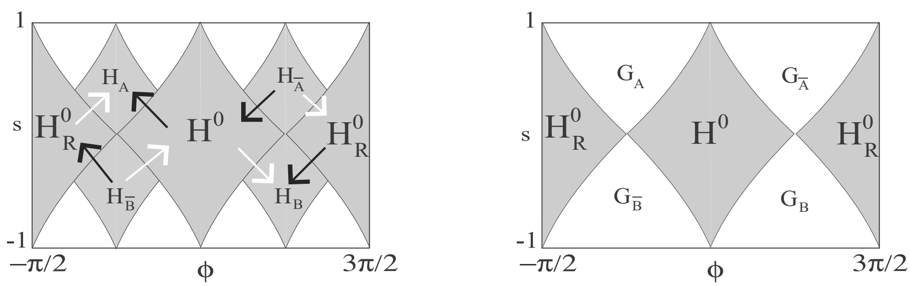


Figure 10: Left: The domains H^0, H^0_R and the domains of first generation. Black arrows indicate the operator A, white arrows the operator B. Right: the four parts of G^0.

Let G^0 denote the domain $H \setminus (H^0 \cup H^0_R)$, and its projection to C_H; it consists of four parts (see Fig. 10, right), denoted by

$$\begin{aligned} G_A &= \{[K,D,S] : S \geq |D|, K > 0\}; \\ G_{\bar{A}} &= \{[K,D,S] : S \geq |D|, K < 0\}; \\ G_B &= \{[K,D,S] : S \leq -|D|, K < 0\}; \\ G_{\bar{B}} &= \{[K,D,S] : S \leq -|D|, K > 0\}. \end{aligned}$$

Observe that $G_{\bar{A}} = RG_A$ and $G_{\bar{B}} = RG_B$.

Theorem 4.2. *Every one of the* 2^n *connected components of the domain of the* $(n+1)$*th generation,* $\mathcal{H}^{\mathbf{n+1}}$ $(n > 1)$*, that lies in* G_A *(in* G_B*) is obtained as* TH_A *(respectively, as* TH_B*), where* $T \in \mathcal{T}^+$ *is a word of length* n *in the generators* A *and* B.

Every one of the 2^n *connected components of the domain of the* $(n+1)$*th generation,* $\mathcal{H}^{\mathbf{n+1}}$ $(n > 1)$*, that lies in* $G_{\bar{A}}$ *(in* $G_{\bar{B}}$*) is obtained as* $TH_{\bar{A}}$ *(respectively, as* $TH_{\bar{B}}$*), where* $T \in \mathcal{T}^-$ *is a word of length* n *in the operators* A^{-1} *and* B^{-1}.

The proof of Theorem 4.2 consists of a computation.

In the sequel, the simple world *domain* indicates a connected component of the domain of a given generation; i.e., a domain is a tile of our model.

The correspondence between the operators of $\mathcal{T}^+$ and the domains up to to fifth generation in G_A is shown in Figure 11.

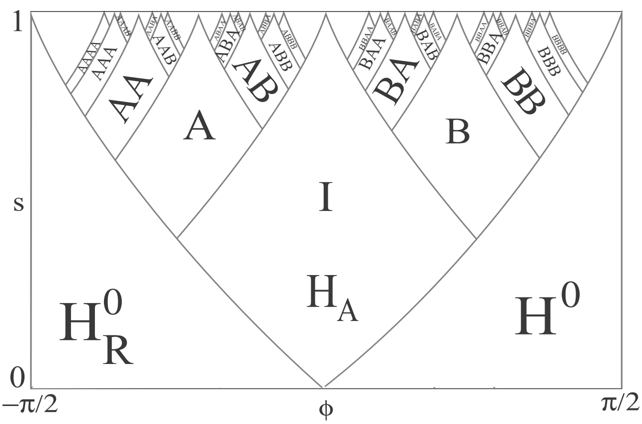


Figure 11: The domains of second, third, fourth, fifth generations lying in G_A are the image of H_A under the operators of $\mathcal{T}^+$, written as words of length one, two, three, and four in A and B.

Remark. The domains H_A and H_B behave as the respective principal "fundamental domains" for the action of the semigroup $\mathcal{T}^+$ in G_A and G_B (their images do not overlap). Similarly for the domains $H_{\bar{A}}$ and $H_{\bar{B}}$ and the action of $\mathcal{T}^-$ in $G_{\bar{A}}$ and $G_{\bar{B}}$.

4.2 Coordinate changes

Here, we see how the Poincaré tiling of the de Sitter world changes under a change of coordinates in the space of the forms (m, n, k) obtained from a $\mathrm{SL}(2, \mathbb{Z})$ coordinate change in the plane.

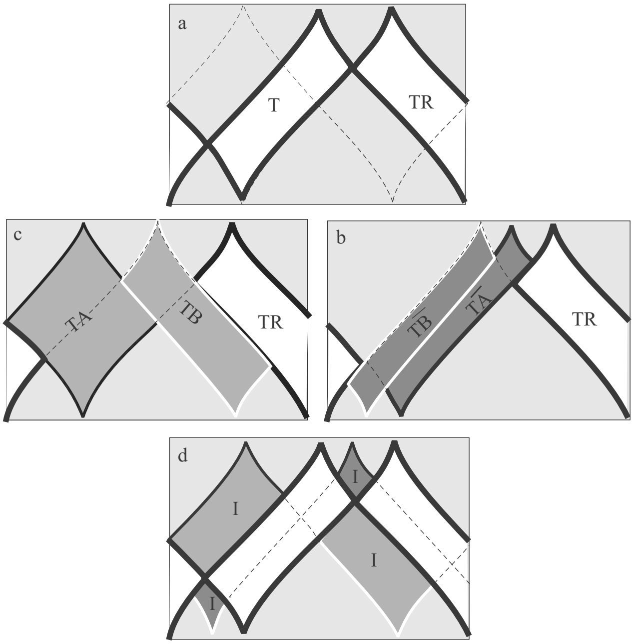


Figure 12: The domains of first generation after a coordinate change.

In the Poincaré model of the Lobachevsky plane, changing coordinates by an operator $\mathbf{L} \in \mathrm{SL}(2, \mathbb{Z})$ in the plane corresponds to replacing the principal fundamental domain with its image under L, which is another domain of the tiling.

In the Poincaré tiling of the de Sitter world, there is no fundamental domain, because the images of a domain under $\mathcal{T}$ overlap each other. The key idea to construct the tiling is to take the nonoverlapping parts of the images of the principal domains under the action of the semigroups $\mathcal{T}^+$ or $\mathcal{T}^-$.

We explain how the tiling of the de Sitter world is affected by a coordinate change. We consider the element $T \in \mathcal{T}$ corresponding to our change of coordinates. The images of H^0 and H^0_R under T are the principal domains of the new tiling. In Figure 12, $T = A^{-1}$.

In part (a), the domains marked by T and TR are the images of H^0 and $H^0_R = RH^0$ under T. They represent the new *rhombi of the zeroth generation*. The dotted lines show the boundaries of H^0 and H^0_R.

Part (b) of the figure shows the images under TA^{-1} and under TB^{-1} of the principal domain H^0 (respectively marked by $T\bar{A}$ and $T\bar{B}$), and part (c) of the figure shows the images of the principal domain H^0 under TA and TB (respectively marked by TA and TB). The union of these images form the *rhombi of the first generation*.

To obtain the *domain of the first generation* (d), we must exclude the parts of these rhombi that overlap with the rhombi of the zeroth generation, which become the principal domains H^0 and H_R^0. We thus obtain four disjoint components (denoted by I), which are the images under T of the domains H_A, $H_{\bar{A}}$, H_B, and $H_{\bar{B}}$ in the tiling introduced in the preceding section. The procedure for building the tiling continues analogously to that already explained.

4.3 Hyperbolic orbits

Figure 13 shows the action of $\mathcal{T}$ on a two-sheeted hyperboloid and on a one-sheeted hyperboloid, as we now explain.

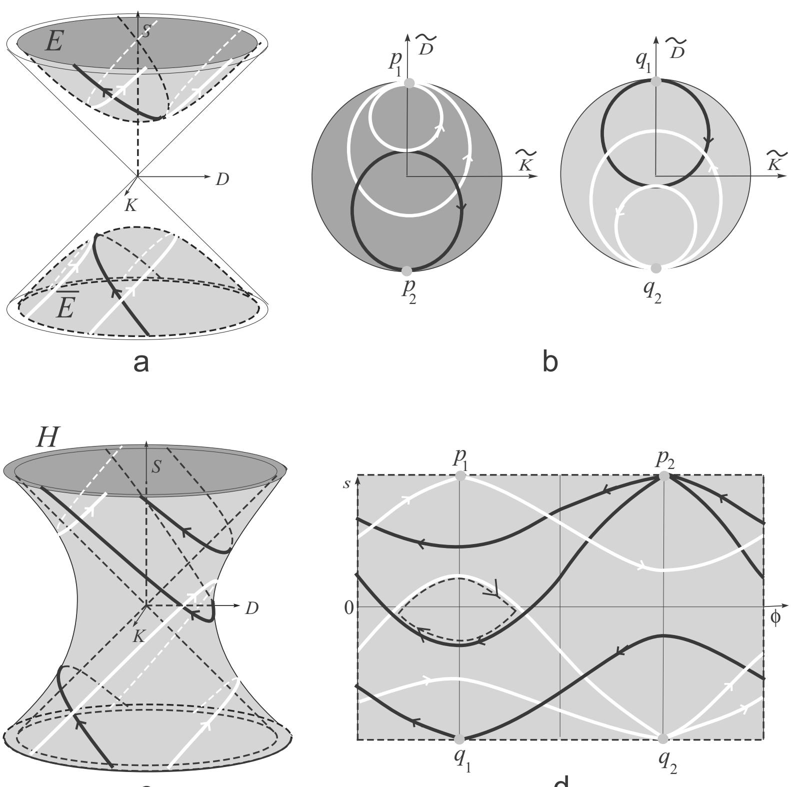


Figure 13: (a) Necklaces on a two-sheeted hyperboloid; (b) their projection on the Lobachevsky disc; (c) Necklaces on a one-sheeted hyperboloid; (d) their projection on the cylinder C_H.

Definition. Given a good point $\mathbf{f}$, the sequences of points $\{A^j\mathbf{f}\}$ and $\{B^j\mathbf{f}\}$, where $j \in \mathbb{Z}$, are called *necklaces* and are respectively denoted by $\omega_{\mathbf{f}}(A)$ and $\omega_{\mathbf{f}}(B)$.

The ordering of $\mathbb{Z}$ orders the necklaces. This order, from lower values to higher values of $j \in \mathbb{Z}$, is indicated by an arrow in Figure 13.

Proposition 4.3. *Every necklace $\omega_{\mathbf{f}}(A)$ (every necklace $\omega_{\mathbf{f}}(B)$) lies on the intersection of the hyperboloid containing it with a plane of the family $S = -D + \alpha$ (respectively $S = D + \alpha$), $\alpha \in \mathbb{Z}$.*

Proof. For every $\mathbf{f} = [K, D, S]$, $A\mathbf{f} := [K', D', S']$ satisfies $D' + S' = D + S$, and $B\mathbf{f} := [K'', D'', S'']$ satisfies $D'' - S'' = D - S$. $\square$

Remark. The notion of necklace is independent of the type of orbits (elliptic or hyperbolic). Figure 13 shows some lines where the necklaces lie on both two-sheeted and one-sheeted hyperboloids. We use black color for necklaces of type $\omega_{\mathbf{f}}(A)$ and white color for necklaces of type $\omega_{\mathbf{f}}(B)$.

In the elliptic case, the projections of the necklaces lie on horocycles tangent to C at the points $\widetilde{K} = 0$ (p_1 and p_2 in Fig. 13**b**).

For each $\mathbf{f}$ in E, we consider the set of points reached by the action of the respective semigroups $\mathcal{T}^+$ and $\mathcal{T}^-$ excluding the identity. This set never contains $\mathbf{f}$. This can be seen starting by any point in E and trying to reach it by a path composed of pieces of necklaces always in the same direction of the arrow (i.e., by an operator of either $\mathcal{T}^+$ or $\mathcal{T}^-$). The situation is different in H: a sequence of good points obtained one from the preceding one by consecutively applying either the operator A or the operator B (i.e., forming a path composed of pieces of necklaces in the positive direction) can lie on a cycle (for instance, see the dotted line in Fig. 13**d**).

Remark. Any orbit in C_H is invariant under a shift by π, because for every points $\mathbf{f}$ the images of $\mathbf{f}$ and $R\mathbf{f}$ in C_H are obtained one from the other by a shift by π.

We separately consider the case where Δ is a square number.

4.4 The case of Δ different from a square number

Theorem 4.4. *For every integer $\Delta > 0$ that is not a square number and satisfies $\Delta \equiv 0$ or $\Delta \equiv 1 \mod 4$, the hyperboloid H_Δ contains at least a good* *point* $\mathbf{f}$ *in such that* $T\mathbf{f} = \mathbf{f}$ *for some operator* T *of* $\mathcal{T}^+$ ($\mathcal{T}^-$) *different from the identity. Such a point belongs to* H^0 *or to* H^0_R.

Proof. The discriminant of an integer quadratic form $\mathbf{f} := (m, n, k)$ is $\Delta = k^2 - 4mn$ and is therefore congruent to 0 or 1 modulo 4. Theorem 4.4 is proved by the following Lemmas 4.5–4.9.

Lemma 4.5. *The domain* H^0 (H^0_R) *contains a finite number of good points.*

Proof. Since

$$mn = \frac{(S^2 - D^2)}{4},$$

the domain H^0 contains all forms where $m > 0$ and $n < 0$, while H^0_R contains all forms where $m < 0$ and $n > 0$. From the definition of Δ, we obtain

$$4mn = k^2 - \Delta.$$

This equality is satisfied by a finite number of values of k, because the product mn is negative in H^0. For each of these values, the set of pairs (m, n) such that $|4mn| = \Delta - k^2$ is finite. $\square$

The condition $\sqrt{\Delta} \notin \mathbb{Z}$ implies that m and n cannot vanish; we hence have $|D| \neq |S|$. Therefore, G^0 ($G^0 := H \setminus (H^0 \cup H^0_R)$) contains all hyperbolic forms where m and n have the same sign.

Lemma 4.6. *Every good point* $\mathbf{f}$ *in* H^0 *satisfies* $A\mathbf{f} \in H^0$ *iff* $B\mathbf{f} \in H_B$ *and* $B\mathbf{f} \in H^0$ *iff* $A\mathbf{f} \in H_A$. *Moreover,* $A^{-1}\mathbf{f} \in H^0$ *iff* $B^{-1}\mathbf{f} \in H_{\bar{B}}$ *and* $B^{-1}\mathbf{f} \in H^0$ *iff* $A^{-1}\mathbf{f} \in H_{\bar{A}}$. *Analogous statements hold with* H^0 *replaced with* H^0_R.

Proof. For the point $\mathbf{f} = (m, n, k)$, $A\mathbf{f} = (m, m + n + k, 2m + k)$ and $B\mathbf{f} = (m + n + k, n, 2n + k)$. We have $m > 0$ and $n < 0$ because $\mathbf{f}$ belongs to H^0. Now, if $A\mathbf{f}$ belongs to H^0, then $m + n + k < 0$ and therefore $B\mathbf{f} \in G^0$. If $A\mathbf{f}$ belongs to G^0, then $m + n + k > 0$ and therefore $B\mathbf{f} \in H^0$. Similar inequalities hold for the inverse generators. By Lemma 4.1, if the image under A, B, A^{-1}, and B^{-1} of a point in H^0 is in G^0, then it respectively belongs to H_A, H_B, $H_{\bar{A}}$, and $H_{\bar{B}}$. Analogous arguments hold for H^0_R because of relations (18). $\square$

Definition. A *cycle of length* t ($t > 1$) is a cyclic sequence of distinct points $[\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_t]$ such that $\mathbf{f}_i = T_{i-1}\mathbf{f}_{i-1}$ ($i = 2, \ldots, t$) and $\mathbf{f}_1 = T_t\mathbf{f}_t$, where each of the operators $T_1, T_2, \ldots, T_t$ is A or B.

A cycle of length t is denoted by $\gamma_{\mathbf{f}}(T_1, \ldots, T_t)$, where $\mathbf{f} = \mathbf{f}_1$. An equivalent notation for the cycle $\gamma_{\mathbf{f}}(T_1, \ldots, T_t)$ is obviously $\gamma_{\mathbf{g}}(T_j, \ldots, T_t, T_1, \ldots T_{j-1})$, where $\mathbf{g} = \mathbf{f}_j$.

Lemma 4.7. *For every good point* $\mathbf{f}$ *of* H^0 (H_R^0)*, there exists an integer* $t > 1$ *such that* $\mathbf{f}$ *belongs to a unique cycle* $\gamma_{\mathbf{f}}(T_1, \ldots, T_t)$*. This cycle is contained in* H^0 *(resp. in* H_R^0*).*

Proof. Let $\mathbf{f}_1 = \mathbf{f}$. By Lemma 4.6, either $A\mathbf{f}_1$ or $B\mathbf{f}_1$ belongs to H^0. Let $\mathbf{f}_2 = T_1\mathbf{f}_1$ be in H^0, where $T_1 = A$ or $T_1 = B$. Now let $\mathbf{f}_3$ be the point that belongs to H^0, and so on. We thus find a sequence of points $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \ldots$ in H^0. Since H^0 contains only a finite number of good points by Lemma 4.5, the sequence of points $\mathbf{f}_1, \mathbf{f}_2, \ldots$, starting from some index j, must be periodic, i.e., we finally find a cycle $\gamma_{\mathbf{f_j}}(T_j, \ldots, T_{j-1+t})$ for some $j \geq 1$. We now prove that $j = 1$, i.e., $\mathbf{f}$ itself belongs to the cycle. Let us suppose that $\mathbf{f}_j \neq \mathbf{f}$. In this case, $\mathbf{f}_{j-1}$ does not belong to the cycle. We have $A\mathbf{f}_{j-1} = \mathbf{f}_j$ or $B\mathbf{f}_{j-1} = \mathbf{f}_j$. But we also have $A\mathbf{f}_{j-1+t} = \mathbf{f}_k$ or $B\mathbf{f}_{j-1+t} = \mathbf{f}_k$. This contradicts Lemma 4.6, by which the images of f_j under A^{-1} and B^{-1} cannot both lie in H^0. Therefore, $\mathbf{f}_j = \mathbf{f}$.

The proof for H_R^0 is analogous. □

Lemma 4.8. *Different cycles are disjoint.*

Proof. By Lemma 4.6, any point of a cycle determines the others. Hence, if two cycles have a common point, then they coincide. □

Lemma 4.9. *If* $\sqrt{\Delta} \notin \mathbb{Z}$*, then* H^0 *and* H_R^0 *both contain at least one good point.*

Proof. Either the discriminant Δ is divisible by 4, or $\Delta = 4d + 1$. If $\Delta = 4d$, then the point $(d, -1, 0)$ belongs to H^0 and the point $(-1, d, 0)$ belongs to H_R^0. If $\Delta = 4d + 1$, then the point $(d, -1, 1)$ belongs to H^0 and the point $(-1, d, 0)$ belongs to H_R^0. □

Proof of Theorem 4.4.[3] Let $\mathbf{f}$ be a good point of H^0 (Lemma 4.9). By Lemma 4.7, it belongs to a cycle $\gamma_{\mathbf{f}}(T_1, \ldots, T_t)$. Hence, $T = T_tT_{t-1}\cdots T_2T_1 \in \mathcal{T}^+$ satisfies $T\mathbf{f} = \mathbf{f}$. Observe now that if $\mathbf{f}$ satisfies $T\mathbf{f} = \mathbf{f}$, for some $T \in \mathcal{T}^+$, then $\mathbf{f}$ satisfies also $\mathbf{f} = T^{-1}\mathbf{f}$, with $T^{-1} \in \mathcal{T}^-$. Suppose that $\mathbf{f}$ belongs to a domain of n-th generation ($n > 0$). If $\mathbf{f} \in G_A$ or $\mathbf{f} \in G_B$, then, as a consequence of Theorem 4.2, every $T \in \mathcal{T}^+$, different from the identity, sends $\mathbf{f}$ into a domain of the $(n + j)$-th generation, where j is the length of the word T.

[3] An alternative proof of this theorem is to show that the set of eigenvectors corresponding to the eigenvalue $\lambda = 1$ of the operators of $\mathcal{T}^+$ contains an integer good vector (v_1, v_2, v_3) such that $|v_3| < |v_2|$, $v_1^2 + v_2^2 - v_3^2 = 4d + e$, for every $d \in \mathbb{N}$ and $e \in \{0, 1\}$ whenever $4d + e$ is not a square number.

Similarly, if $\mathbf{f} \in G_{\bar{A}}$ or $\mathbf{f} \in G_{\bar{B}}$, then $\mathbf{f}$ is sent into a domain of $(n + j)$-th generation by any word $T \in \mathcal{T}^-$ in A^{-1} and B^{-1} of length j. Therefore if $\mathbf{f} = T\mathbf{f}$ for some operator $T \in \mathcal{T}^+$ (or $\mathcal{T}^-$) different from the identity, then either $\mathbf{f} \in H^0$ or $\mathbf{f} \in H^0_R$. □

Theorem 4.10. *If Δ is not a square number, then the $\mathcal{T}$-orbit of every good point of H_Δ contains exactly one cycle in H^0.*

Proof. We must prove that: (a) every orbit contains a cycle in H^0, and (b) this cycle is unique. We consider a point $\mathbf{f} \in G^0$ and suppose that $\mathbf{f} \in G_A$ (the proof is analogous in the other cases).

(a) By Theorem 4.2, there exists a unique operator T in $\mathcal{T}^+$ such that $\mathbf{g} := T^{-1}\mathbf{f}$ belongs to H_A. Hence, $\mathbf{h} := A^{-1}\mathbf{g} = A^{-1}T^{-1}\mathbf{f}$ is inside H^0 and thus belongs to a cycle $\gamma_{\mathbf{h}}$ by Lemma 4.7. Consequently, the orbit of $\mathbf{f}$ contains a cycle.

(b) We must prove that a point of H^0 that does not belong to the cycle $\gamma_{\mathbf{h}}$ cannot be obtained from $\mathbf{f}$ by an operator in $\mathcal{T}$. By Lemma 1.1, every operator of $\mathcal{T}$ can be written as USV, where S belongs to $\mathcal{T}^+$ or $\mathcal{T}^-$ and V and U are equal to the identity or to the operator R. Hence, we try to reach H^0 from $\mathbf{f}$ by such operators. We must start by R, reaching a point $\mathbf{p}$ of $G_{\bar{A}}$. As before, there is only one operator of $\mathcal{T}^-$ such that $\mathbf{p}$ is the image under it of a point in $H_{\bar{A}}$. Indeed, since $\mathbf{p} = R\mathbf{f}$ and $\mathbf{f} = T\mathbf{g}$,

$$\mathbf{p} = R\ T\mathbf{g} = \hat{T}R\mathbf{g}\ ,$$

where $\hat{T}$ is the operator obtained from T by replacing each A with B^{-1} and each B with A^{-1}. The operator $\hat{T}$ belongs to $\mathcal{T}^-$, and the point $\mathbf{j} := R\mathbf{g}$ is in $H_{\bar{A}}$ because $\mathbf{g} \in H_A$. We can now reach a point either in H^0 (by A) or in H^0_R (by B). Since $\mathbf{g} = A\mathbf{h}$, we obtain

$$B\mathbf{j} = RR\ B\mathbf{j} = R\ A^{-1}R\ R\mathbf{g} = R\ A^{-1}A\mathbf{h} = R\mathbf{h}.$$

Therefore, the point in H^0_R reached from $\mathbf{p}$ in $G_{\bar{A}}$ in this case is exactly $R\mathbf{h}$. On the other hand, $A\mathbf{j} \in H^0$ is equal to

$$A\mathbf{j} = RR\ A\mathbf{j} = R\ B^{-1}R\ R\mathbf{g} = R\ B^{-1}\ A\mathbf{h} = AR\ A\mathbf{h} = B\mathbf{h}.$$

Since $A\mathbf{h} \in H_A$, $B\mathbf{h} \in H^0$ (by Lemma 4.6) and hence $B\mathbf{h}$ belongs to the same cycle of $\mathbf{h}$. The proof is complete. □

Figure 15 shows examples of orbits projected to the cylinder C_H.

For every $T \in \mathcal{T}^+$, let us denote by $t_A(T)$ and $t_B(T)$ respectively the number of times that A and B appear in T.

Theorem 4.11. *Let T be an operator in $\mathcal{T}^+$ satisfying $T\mathbf{f} = \mathbf{f}$ for some good point $\mathbf{f} \in H^0$, being not the power of an operator $W \in \mathcal{T}^+$ which satisfies $W\mathbf{f} = \mathbf{f}$. The number of points of $C(\mathbf{f})$ in H_A and H_B is respectively equal to $t_B(T)$ and $t_A(T)$.*

Proof. By Theorem 4.4, the set of points of every cycle in H^0 is subdivided into two disjoint subsets: the set of points whose image under A belongs to H_A and the set of points whose image under B belongs to H_B. By Lemma 4.6, the image under A of a point is inside H_A if and only if its image under B is inside H^0. The number of such points is evidently equal to $t_B(T)$. (see Figs. 14 and 10). Similarly, the image under B of a point is inside H_B if and only if the image under A is inside H^0. The number such points is equal to $t_A(T)$. $\square$

As a consequence of Theorem 4.2 we obtain the following

Corollary 4.12. *The orbit of a form $\mathbf{f}$ has the same number of points in H_A, in $H_{\bar{A}}$, and in every domain in G_A and $G_{\bar{A}}$; similarly, it it has the same number of points in H_B, in $H_{\bar{B}}$, and in every domain in G_B and $G_{\bar{B}}$.*

Proof. The fact that the number of points in H_A and in $H_{\bar{A}}$ coincide, is due to the fact that the forms in H_A and $H_{\bar{A}}$ belonging to to the same orbit constitute pairs of complementary forms, being $H_A = RH_{\bar{A}}$. Similarly for the forms in H_B and in $H_{\bar{B}}$. Alternatively, we may proceed with the same reasoning proving Theorem 4.4, applied to the operator T^{-1}, which is a word containing $t_A(T)$ times A^{-1} and $t_B(T)$ times B^{-1}, to obtain the number of points of $C(\mathbf{f})$ in the interior of $H_{\bar{A}}$ and of $H_{\bar{B}}$ (see Fig. 14). The conclusion of the proof follows from Theorem 4.2. $\square$

4.5 The case where Δ is a square number

Definition. A form *represents zero* if it vanishes at least at an integer point of the plane different from the origin. More generally, a form *represents z* if it takes the value z at an integer point of the plane..

If Δ is a square number, then H_Δ contains forms where the coefficient m or n vanishes.

A form with vanishing m or n represents zero, as well as any form in its class.

For the forms where m or n vanish, either $S = D$ or $S = -D$ and the corresponding good points therefore lie on the boundaries of H^0 and H^0_R.

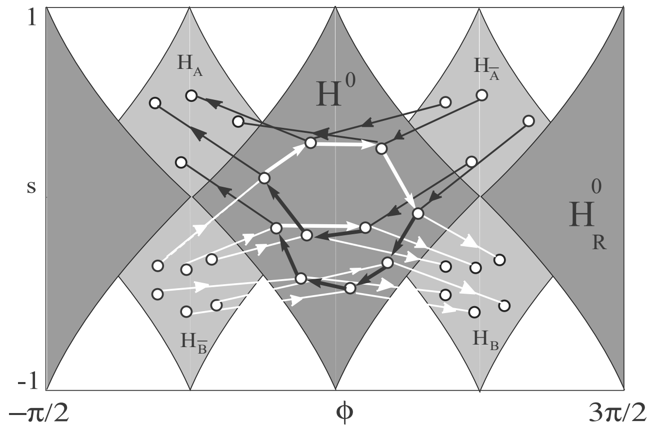


Figure 14: The cycle of the orbit of [18, 20, -10] in H^0, and the points of the orbit in H_A, H_B, $H_{\bar{A}}$ and $H_{\bar{B}}$. Black and white arrows show the action respectively of A and of B. The orbit is asymmetric.

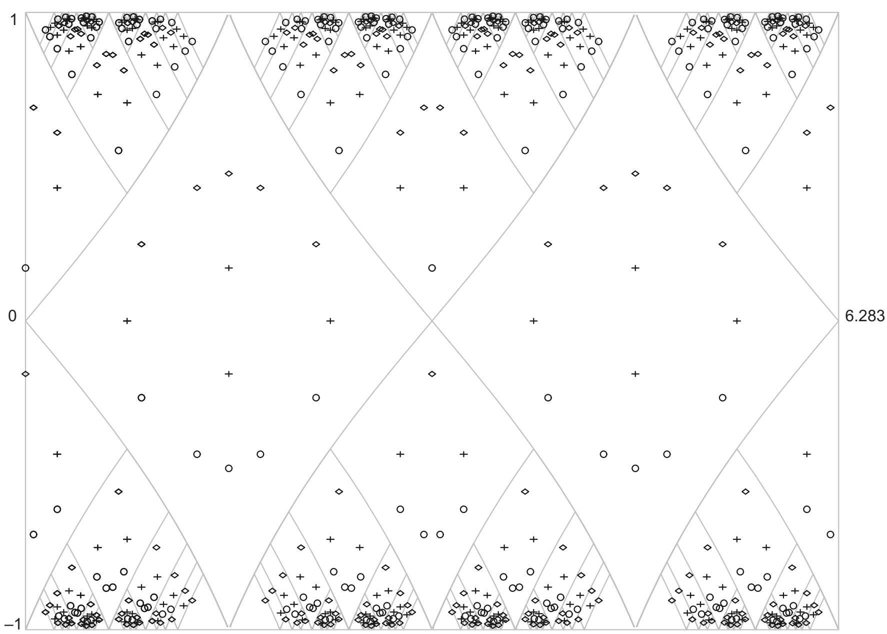


Figure 15: The parts of the three distinct orbits with $\Delta = 32$ contained in the domains of the first five generations. The representative points of these orbits are in coordinates [K, D, S]: [-6,0,2], circles, k-symmetric orbit; [6,0,2], diamonds, k-symmetric orbit; [-4,4,0], crosses, supersymmetric orbit.

Theorem 4.13. *For each square number $\Delta = \rho^2$, $\rho \in \mathbb{N}$, the hyperboloid H_Δ contains exactly ρ orbits. They are the orbits of the points $[K, D, S] = [\rho, r, r]$, $r = 0, \ldots, \rho - 1$.*

The proof follows from Lemmas 4.14 and 4.15.

In H_Δ, the segments that belong to the respective common boundaries of H^0 with H_A, $H_{\bar{A}}$, H_B, and $H_{\bar{B}}$, and belong respectively to H_A, $H_{\bar{A}}$, H_B, and $H_{\bar{B}}$, are denoted by

$$\begin{aligned} F_A &= \left\{[\ \rho,\ r,\ r],\ \ 0 \le r < \rho\right\};\\ F_{\bar{A}} &= \left\{[-\rho,\ r,\ r],\ \ 0 \le r < \rho\right\};\\ F_{\bar{B}} &= \left\{[\ \rho,\ r, -r],\ \ 0 \le r < \rho\right\};\\ F_B &= \left\{[-\rho,\ r, -r],\ \ 0 \le r < \rho\right\}. \end{aligned}$$

We will use the same notations as well as for their images in C_H.

Observe that F_A is the lower-right side of the boundary of H_A, $F_{\bar{A}}$ is the lower-left side of the boundary of $H_{\bar{A}}$, F_B is the upper-left side of the boundary of H_B, and $F_{\bar{B}}$ is the upper-right side of the boundary of $H_{\bar{B}}$ (see Fig. 16).

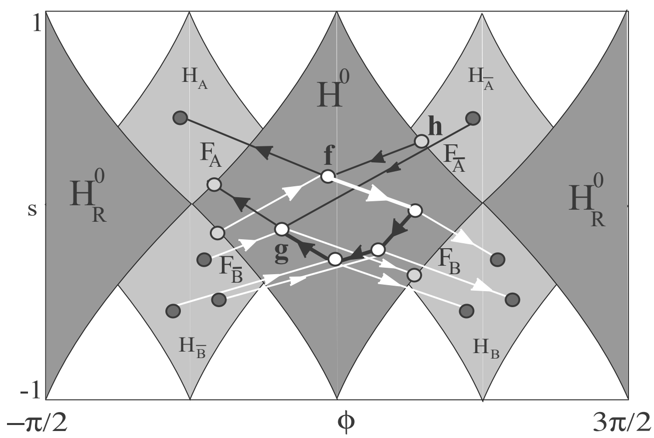


Figure 16: The chain of length 5 in H^0 of the orbit of [9, 2, 2]. Black and white arrows show the action respectively of A and of B. The orbit is asymmetric.

The action of $\mathcal{T}^+$ on F_A and on F_B is deduced from that of $\mathcal{T}^+$ on H_A and H_B, and also the action of $\mathcal{T}^-$ on $F_{\bar{A}}$ and on $F_{\bar{B}}$ is deduced from that of $\mathcal{T}^-$ on H_A and H_B (see Theorem 4.2). The following lemma is indeed a corollary of Theorem 4.2.

Lemma 4.14. *Every good point* $\mathbf{f}$ *in* $H^0 \subset H_\Delta$, *where* Δ *is a square number, satisfies* $A\mathbf{f} \in H^0$ *iff* $B\mathbf{f} \in H_B$, $B\mathbf{f} \in H^0$ *iff* $A\mathbf{f} \in H_A$, $A^{-1}\mathbf{f} \in H^0$ *iff* $B^{-1}\mathbf{f} \in H_{\bar{B}}$, *and* $B^{-1}\mathbf{f} \in H_0$ *iff* $A^{-1}\mathbf{f} \in H_{\bar{A}}$.

Moreover, $A\mathbf{f} \in F_A$ *iff* $B\mathbf{f} \in F_B$ *and* $A^{-1}\mathbf{f} \in F_{\bar{A}}$ *iff* $B^{-1}\mathbf{f} \in F_{\bar{B}}$.

Proof. This lemma is the version of Lemma 4.6 where Δ is equal to a square number. Indeed, if $\mathbf{f} = (m, n, k)$ and $\mathbf{g} = A\mathbf{f} \in F_A$, then $(m + n + k = 0)$. This implies that $\mathbf{g}' = B\mathbf{f} = (m + n + k, n, k + 2n)$ belongs to F_B. The cases of the inverse generators are similar. $\square$

Lemma 4.15. *For every* $\Delta = \rho^2$*, the orbit of the point* $[K, D, S] = [\rho, 0, 0]$ *consists of all the lower vertices of all domains in* G_A *and* $G_{\bar{A}}$ *and all upper vertices of all domains in* G_B *and in* $G_{\bar{B}}$.

Proof. The lemma follows from Theorem 4.2 because the points $[\pm\rho, 0, 0]$ are the lower points of the domains H_A and $H_{\bar{A}}$ and the upper points of the domains H_B and $H_{\bar{B}}$. $\square$

Remark. The orbit of the points $[\pm\rho, 0, 0]$ is supersymmetric.

Proof of Theorem 4.13. By Lemma 4.1, all good points in the interior of the domain H_Z (where $Z = A, B, A^{-1}$, or B^{-1}) are the images under Z of good points at the interior of H^0. We are now interested in the images under Z^{-1} of the points of F_Z that are inside H^0. For instance, we consider a point $\mathbf{h} \in F_{\bar{A}}$ (see Fig. 16) which is sent to $\mathbf{f} \in H^0$ by A. We then sequentially apply either A or B to remain inside H^0 until we reach the point $\mathbf{g}$ such that both $A\mathbf{g}$ and $B\mathbf{g}$ belong to the boundary of H^0 (namely, to F_A and F_B) by Lemma 4.14. Lemma 4.14 also states that $B^{-1}\mathbf{f}$ belongs to $F_{\bar{B}}$ because $A^{-1}\mathbf{f}$ belongs to $F_{\bar{A}}$. By a similar procedure, with each point $\mathbf{h}$ of any one of the sets F_Z, we associate an ordered sequence of points in H^0 and three other points of the orbit of $\mathbf{f}$, one in each of the other sets F_Y, $Y \neq Z$ (see Fig. 16). With any point in H^0, we thus associate a unique ordered sequence of points inside H^0, whose initial point is respectively mapped by A^{-1} and B^{-1} to two points of $F_{\bar{A}}$ and $F_{\bar{B}}$ and whose final point is respectively mapped by A and B to two points of F_A and F_B (see Fig. 16). Since different sequences cannot have common elements also in this case, we have $\rho - 1$ distinct orbits corresponding to all integer points in $F_{\bar{A}}$ ($[\rho, r, r]$ for $r = 1 \ldots \rho - 1$) plus the orbit of the point $[\rho, 0, 0]$, given by Lemma 4.15. $\square$

Definition. We call *chain of length t* The ordered sequence of points $\mathbf{p}_1, \ldots,$ $\mathbf{p}_t$ in H^0 satisfying $\mathbf{p}_{i+1} = T_i\mathbf{p}_i$, $i = 1, \ldots, t - 1$, where T_i is A or B and $A\mathbf{p}_t \in F_A$, $B\mathbf{p}_t \in F_B$, $A^{-1}\mathbf{p}_1 \in F_{\bar{A}}$, and $B^{-1}\mathbf{p}_1 \in F_{\bar{B}}$.

The proof of the above theorem implies the following corollary.

Corollary 4.16. *The good points in $H^0 \in H_\Delta$, where Δ is a square number, are partitioned into disjoint chains. Every orbit different from that of* $[\sqrt{\Delta}, 0, 0]$ *contains exactly one chain in such an H_Δ.*

Figure 17 shows the case where $\Delta = 25$.

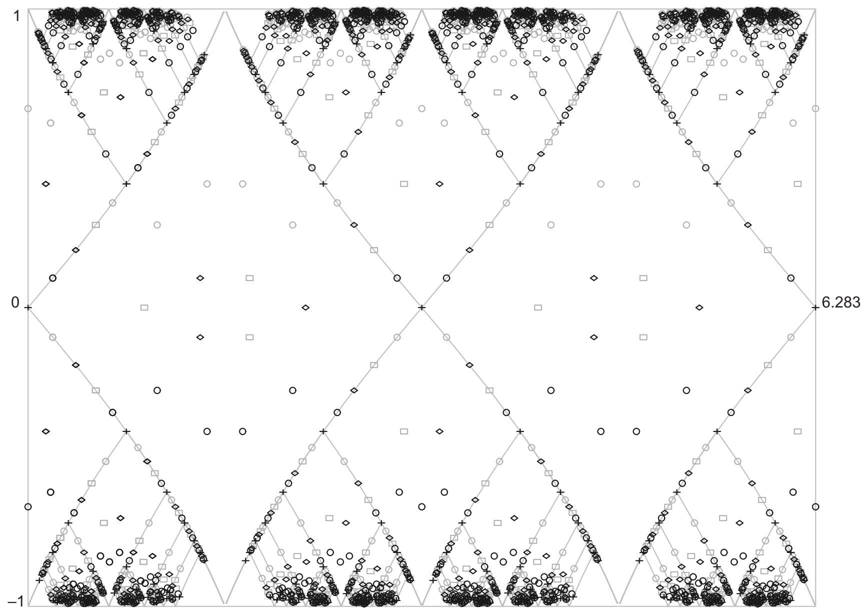


Figure 17: The points of 5 distinct orbits for $\Delta = 25$ belonging to domains of the first five generations. The representant points of the orbits are, in coordinates $[K, D, S]$: $[5, 0, 0]$, crosses, supersymmetric orbit; $[5, 1, 1]$, black circles, k-symmetric orbit; $[5, 2, 2]$, black diamonds, $(m + n)$-symmetric orbit; $[5, 3, 3]$, gray boxes, $(m + n)$-symmetric orbit; $[5, 4, 4]$, gray circles, k-symmetric orbit.

Theorem 4.17. *Let t be the length of a chain inside H^0 with initial point $\mathbf{f}$ and final point $\mathbf{g}$. Let T be the operator in $\mathcal{T}^+$ satisfying $T\mathbf{f} = \mathbf{g}$. Then T is a word of length $t - 1$ in A and B, and the orbit of $\mathbf{f}$ contains exactly $t_B(T)$ points in the interior of H_A and exactly $t_A(T)$ points in the interior of H_B.*

Proof. The points $\mathbf{f} = \mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_t = \mathbf{g}$ of the chain satisfy $\mathbf{f}_{i+1} = T_i \mathbf{f}_i$, $i = 1, \ldots, t-1$, where T_i is A or B. Hence, $g = Tf$, where $T = T_{t-1}T_{t-2} \cdots T_1$. By Lemma 4.14, if the image under A (under B) of a point of the chain belongs to the chain, then the image under B (resp., under A) belongs to H_B (resp., to H_A). Therefore, the orbit of $\mathbf{f}$ has $t_B(T)$ points in the interior of H_A and $t_A(T)$

points in the interior of H_B. That there are no other points inside these domains follows from Corollary 4.16 and from the fact that no domains other than H^0 are mapped by A to H_A or by B to H_B. □

We obtain the following corollary from Theorem 4.2.

Corollary 4.18. *The orbit of a form* **f** *representing zero has the same number of points in the interior of H_A, of $H_{\bar{A}}$, and of every domain in G_A and $G_{\bar{A}}$; similarly, it it has the same number of points in the interior of H_B, of $H_{\bar{B}}$, and of every domain in G_B and $G_{\bar{B}}$.*

Proof. The proof is analogous to that of Corollary 4.12. □

Example. In Figure 17, the orbit of [5, 1, 1] (black circles) has a chain in H_0 composed of four points. The three operators between them are all of type A, i.e. $T = A^3$. Hence $t_A(T) = 3$ and $t_B(T) = 0$: indeed, there are no black circles in H_A and $H_{\bar{A}}$. Conversely, for the orbit of [5, 4, 4] (gray circles), we have $T = B^3$, $t_A(T) = 0$ and $t_B(T) = 3$. The orbits of [5, 2, 2] and [5, 3, 3] (respectively diamonds and boxes) in H^0 contain a chain of three points, with $t_A(T) = t_B(T) = 1$ for both orbits: indeed, there is a diamond and a box inside each domain in the regions G_A, $G_{\bar{A}}$, G_B, and $G_{\bar{B}}$.

5 Supplementary remarks

5.1 Remark on the behavior of the nonrational points under $\mathcal{T}$

Note that on the hyperboloid $K^2 + D^2 - S^2 = \Delta$, the orbit of any point with the fractional coordinates $[K, D, S]$ with a common denominator μ is obtained (by a scale reduction) from the orbit of the good point with the integer coordinates $[2\mu K, 2\mu D, 2\mu S]$ on the hyperboloid with the discriminant $4\mu^2\Delta$ and hence has a finite number of points in H^0, in H^0_R, and also in each domain of any generation.

In the hyperbolic case, the situation is completely different from the elliptic case, where the orbit under $\mathcal{T}$ of an irrational point is described exactly as that of an integer point (close points in the Lobachevsky disc have close orbits under $\mathrm{PSL}(2, \mathbb{Z})$, that is, the images of two close points under any element of the group are close). Indeed, Theorems 4.2 and 4.4 imply that two close points in H have close orbits under the semigroup $\mathcal{T}^+$ or $\mathcal{T}^-$ only if these points belong to the complement of H^0 and H^0_R. But the orbits under $\mathcal{T}$ (and even under $\mathcal{T}^+$ or $\mathcal{T}^-$) of two close points in H^0 or in H^0_R are not close. This follows from the fact that a statement analogous to Lemma 4.6 holds for all points in H^0 and in H^0_R and not only for the good points.

Moreover, from Theorems 4.2 and 4.4, we obtain the following corollary for the irrational points on the one-sheeted hyperboloid.

Corollary 5.1. *The orbit of any point having at least one irrational coordinate contains an infinite number of points in H^0 and in H^0_R (and hence in each connected component of the domains of all generations of the hyperboloid).*

5.2 The solar eclipse model of the de Sitter world

In this section, we see the Poincaré tiling of the de Sitter world under an alternative projection.

We consider the domains of different generations directly on the hyperboloid H_Δ. The lines bounding such domains belong to the straight-line generatrices of the hyperboloid.

The segment joining the point p_i and its opposite point on the circle c_1, upper vertices of a pair of opposite rhombi of some generation in C_H, defines a direction, ℓ_i, in the plane $S = 1$. In the plane $S = 0$, consider the circle of radius $\rho = \sqrt{\Delta}$, intersection of H_Δ with this plane, and the two straight lines l_i and l'_i tangent to this circle, in the direction ℓ_i. The four generatrices of the hyperboloid that bound the domains projected by $\mathcal{Q}$ to the two rhombi, lie on two planes parallel to the S axis and passing through l_i and l'_i respectively.

The hierarchy of the points p_i is inherited by the pairs of parallel lines (l_i, l'_i) on the plane $S = 0$ and also by the regions bounded by such pairs of lines and by the circle $K^2 + D^2 = \rho^2$. The regions of the nth generation *lie behind* those of all preceding generations. The view of the domains on the hyperboloid projected to the plane $S = 0$ is shown in Figure 18. Observe that we map only one half of H_Δ by this projection.

To introduce the "solar eclipse model," we consider a further projection of the hyperboloid H_Δ.

Let $\mathcal{P}'$ be the projection of H_Δ to the plane $S = \rho = \sqrt{-\Delta}$ along the direction of the S axis, and let $\mathbf{r} = \mathcal{P}'\mathbf{f}$. Let $\mathcal{P}''$ be the projection of the plane $S = \rho$ to the sphere of radius ρ from the coordinate origin, and let $\mathbf{s} = \mathcal{P}''\mathbf{r}$. Let $\mathcal{P}'''$ be the stereographic projection of the upper half-sphere to the disc of radius ρ in the plane $S = 0$ from the point O' ($K = D = 0$, $S = -\rho$), and let $\mathbf{g} = \mathcal{P}'''\mathbf{s}$ (see Fig. 19, right).

Remark. If the point $\mathbf{p}$ belongs to the upper sheet E of the two-sheeted hyperboloid ($K^2 + D^2 = S^2 + \Delta$, $\Delta < 0$; see Fig. 19, left), then a short

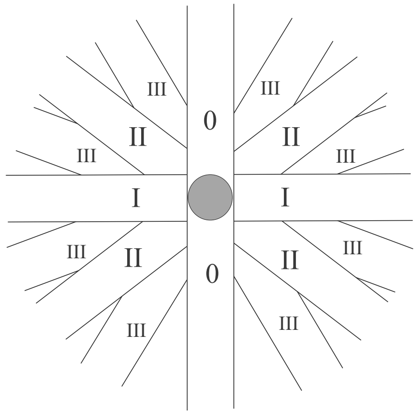


Figure 18: Projection to the plane $S = 0$, with coordinates (K, D), of the tiling of H_Δ.

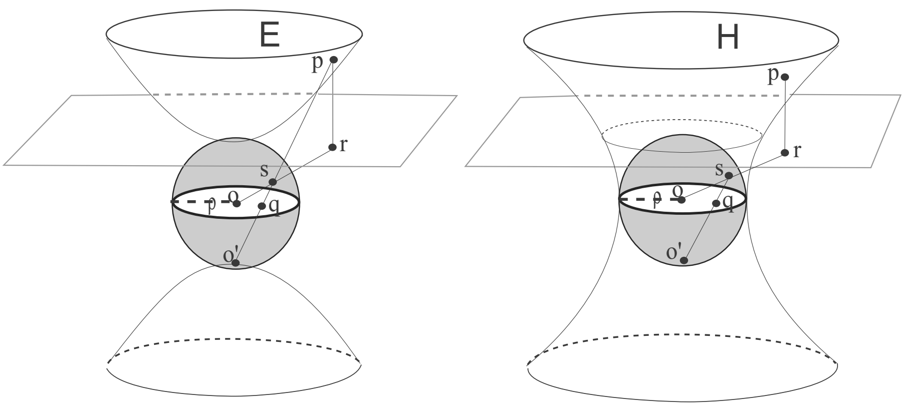


Figure 19: The point p on the hyperboloid is sent to q on the disc by a projection that results from the composition of three projections.

calculation shows that the point $\mathbf{q} = \mathcal{P}'''\mathcal{P}''\mathcal{P}'\mathbf{p}$ coincides with the image of $\mathbf{p}$ by the projection $\mathcal{P}$.

We project the upper half-hyperboloid H_Δ by $\mathcal{P}'''\mathcal{P}''\mathcal{P}'$ to the disc of radius ρ in the plane $S = 0$. The image is contained in the ring $\alpha\rho \leq \sqrt{K^2 + D^2} < \rho$, where $\alpha = \sqrt{2} - 1$. The pairs of straight lines tangent to the circle $K^2 + D^2 = \rho^2$ in the plane $S = \rho$ are projected by $\mathcal{P}''$ to half-meridian circles of the sphere of radius ρ and hence by the stereographic projection $\mathcal{P}'''$ to arcs of circles tangent to the circle $\sqrt{K^2 + D^2} = \alpha\rho$. The final disc of unit radius is obtained by rescaling. On the boundary C of this disc, the same points p_i considered at the boundary of the Lobachevsky disc are the extreme points of the domains of all generations (forming the solar corona). The empty (black) disc of radius α is the moon in the solar eclipse model (see Fig. 20).

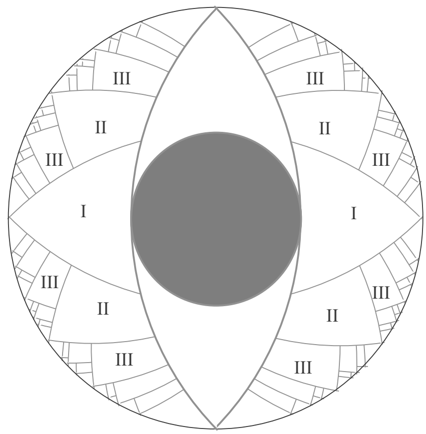


Figure 20: Solar eclipse model of the de Sitter world.

Remark. The complementary forms in the solar eclipse model are symmetric with respect to the center of the disc. Hence, the picture of any orbit has this symmetry. The k-symmetric orbits are symmetric with respect to the vertical axis of the disc.

A complete representation of an orbit requires two copies of this model, one for the upper and the other for the lower half-hyperboloid, merged along the circle bounding the sun.

Identifying these two copies by identifying the images of points of the hyperboloid H_Δ symmetric with respect to the origin of the coordinates, we obtain a Möbius band, tiled in the same way. We observe that opposite points of the circle bounding the sun are identified. Here the number N of points of an orbit in each domain is the same in all the non principal domains, as well as in both half parts (upper and lower in Fig. 20) of the unique principal domain, which contains therefore $2N$ points of that orbit. The number N is obtained from the numbers $t^\uparrow$ and $t_\downarrow$ of points of that orbit in each domain of the respective upper and lower regions in the Poincaré tiling as follows:

– supersymmetric and $(m+n)$-symmetric orbits: $N = t^\uparrow = t_\downarrow$ (the two points of the orbit with opposite coordinates have the same image).

– asymmetric, k-symmetric and antisymmetric orbits: $N = t^\uparrow + t_\downarrow$ (there are no pairs of points with opposite coordinates in such orbits). In particular, for antisymmetric orbits $N = 2t^\uparrow = 2t_\downarrow$.

References

[1] V.I. Arnold. *Arithmetics of binary quadratic forms, symmetry of their continued fractions and geometry of their de Sitter world.* Bull. of Braz. Math. Soc., **34**(1) (2003), 1–41.

[2] V.I. Arnold. Letter to the seminar (2005).

[3] V.I. Arnold. Questions to the seminar (2007).

[4] J.H. Conway. *The sensual (quadratic) forms.* The Mathematical Association of America, Princeton University (1997).

[5] C.F. Gauss. *Disquisitiones Arithmeticae.* Werke, V, Fleischer, Leipzig (1801).

[6] J.L. De Lagrange. *Recherches d'Arithmétique.* Œuvres, tome 3, Gauthier-Villar, Paris (1899).

Francesca Aicardi
Sistiana 56 PR, 34011 Trieste
ITALY

E-mail: aicardi@sissa.it