BULLETIN OF THE BRAZILIAN MATHEMATICAL SOCIETY

Bull Braz Math Soc, New Series 34(1), 1-42

Arithmetics of binary quadratic forms, symmetry of their continued fractions and geometry of their de Sitter world

V. Arnold*

— Dedicated to IMPA on the occasion of its 50^{th} anniversary

Abstract. This article concerns the arithmetics of binary quadratic forms with integer coefficients, the De Sitter's world and the continued fractions.
Given a binary quadratic forms with integer coefficients, the set of values attaint at integer points is always a multiplicative "tri-group". Sometimes it is a semigroup (in such case the form is said to be *perfect*). The diagonal forms are specially studied providing sufficient conditions for their perfectness. This led to consider hyperbolic reflection groups and to find that the continued fraction of the square root of a rational number is palindromic.
The relation of these arithmetics with the geometry of the modular group action on the Lobachevski plane (for elliptic forms) and on the relativistic De Sitter's world (for the hyperbolic forms) is discussed. Finally, several estimates of the growth rate of the number of equivalence classes versus the discriminant of the form are given.

Keywords: arithmetics, quadratic forms, De Sitter's world, continued fraction, semigroup, tri-group.

Mathematical subject classification: 11A55, 11E16, 20M99, 20M99, 53A35.

Introduction

This article is a description of a long chain of numerical experiments with quadratic forms and periodic continued fractions, leading to some strange theorems, confirming the conjectures, originated from these experiments.

Received 5 June 2002.
*Partially supported by RFBR, grant 02-01-00655.

One of these theorems states that the set of values of a form is in many cases a multiplicative semigroup of integers: the product of any two values, attained at integer points, is itself a value, attained at some integer point of the plane of the arguments. I call such forms *perfect forms*.

The simplest example of a perfect form is the form $x^2 + y^2$, for which the semigroup property follows from the existence of the Gauss complex numbers multiplication. For the general forms the semigroup property is replaced by the *trigroup property*: the product of any *three* values is still a value (while the products of some pairs of values are not attained). It happens, for instance, for the form $2x^2 + 3y^2$, which attains the value 2 but does not attain the value 4 (even modulo 3).

The semigroup and trigroup properties are closely related to the hyperbolic reflection groups of symmetries of the hyperbolae in the integer plane.

These symmetries imply a strange *palindrome property* of the periods of the continued fractions of such quadratic irrational numbers, as the square roots of ordinary fractions, $\sqrt{m/n}$.

I shall also discuss below the strange relation of these arithmetical problems to the geometry of the modular group action on the relativistic de Sitter world. This world is represented by the continuation of the Klein model of the Lobachevsky plane from the interior part of a disc to its complementary domain. Quadratic binary forms of a fixed negative determinant are represented by the points of this relativistic world (or rather of its two-fold covering). The $\mathrm{SL}(2, \mathbb{Z})$-classification of the integer quadratic forms, whose invariants the present paper is studying, is the study of the action on the de Sitter world of the group, generated by the reflections in the three sides of the Lobachevsky infinite modular triangle.

The relation of the de Sitter world to the Klein model of the Lobachevsky geometry, as well as the problem of the geometric investigation of the modular group action on the de Sitter world, had been published as "Problem 1996 - 15" in the book *Arnold's Problems*, Phasis, Moscow, 2000, pp. 126 and 422. Other applications of these ideas are discussed in the recent papers [1]-[5].

1 The semigroups and trigroups of the values of quadratic forms

Let $f(x, y) = mx^2 + ny^2 + kxy$ (where the arguments (x, y) and the coefficients (m, n, k) are integers) be a binary quadratic form.

Theorem 1. *The product of any three values of such a form is also its value:*

$$f(x, y) f(z, w) f(p, q) = f(X, Y),$$

where, for instance,

$$X = ap + bq, \quad Y = cp + dq \ ,$$
$$a = m(xz) - n(yw),$$
$$b = n(yz + xw) + k(xz),$$
$$c = m(xw + yz) + k(yw),$$
$$d = n(yw) - m(xz).$$

Proof. (F. Aicardi) By the definitions of X and Y,

$$\begin{aligned} f(X,Y) &= p^2(ma^2 + nc^2 + kac) + q^2(mb^2 + nd^2 + kbd) + \\ &\qquad + pq(2mab + 2ncd + k(ad + bc)) \\ &= mp^2m^2(xz)^2 + ... \qquad \text{(there are 52 terms).} \end{aligned}$$

The product of the three values is, by the definition of f, the integer

$$f(x, y)f(z, w)f(p, q) = (mp^2 + nq^2 + kpq)$$
$$(mx^2 + ny^2 + kxy)(mz^2 + nw^2 + kzw).$$

This product consists of exactly the same 52 monomials in (x, y, z, w), as $f(X, Y)$.

Definition. A form is *perfect* if the product of any two values of the form at integer points is also the value of the form at some integer point.

Corollary 1. *Any form, representing the number* $1 = f(p, q)$, *is perfect:* $f(x, y)f(z, w) = f(X, Y)$, *where one may choose, for instance,*

$$\begin{aligned} X &= (mp + kq)xz + nq(yz + xw) - npyw, \\ Y &= -mqxz + mp(yz + xw) + (kp + nq)yw. \end{aligned}$$

Example 1. Any form $x^2 + ny^2$ is perfect.

There exist perfect forms which do not represent the number 1.

The following corollary shows, for instance, that this property holds for the form $2x^2 + 2y^2$.

Corollary 2. *If a form is representing an integer number* N, *then its product with* N *is a perfect form.*

Proof. It follows from the identity $(NA)(NB) = N(ABN)$, where A and B are values, since ABN is a value by the trigroup property, proved by the Theorem.

Example 2. The form x^2+y^2 represents 2, hence the form $2x^2+2y^2$ is perfect.

The form $2x^2 + ny^2$ represents 2, hence the form $4x^2 + 2ny^2$ is perfect (as well as is any form $m^2x^2 + mny^2$).

Remark. The Theorem defines a trilinear operation, sending three vectors $u = (x, y)$, $v = (z, w)$, $r = (p, q)$ to the vector $U = (X, Y)$. This operation depends linearly on the form f (that is, on the coefficients (m, n, k)). Moreover, this operation is natural, that is independent on the coordinates choice (while it is defined in the Theorem by the long coordinate formula).

Namely, an $\mathrm{SL}(2, \mathbb{Z})$-belonging operator A sends the form f to a new form $\tilde{f}$ and sends the 4 vectors $(u, v, r; U)$ to the 4 new vectors $(\tilde{u}, \tilde{v}, \tilde{r}; \tilde{U})$. The naturality claim means that the operation, defined by the transformed form $\tilde{f}$, sends the transformed vectors $(\tilde{u}, \tilde{v}, \tilde{r})$ to the transformed version $\tilde{U}$ of the vector U.

These remarks are perhaps sufficient to find the formula of the Theorem for the trigroup operation from the particular cases, like that of the forms mx^2+ny^2, for which I had first discovered these formulae (as a conclusion of some hundreds of numerical examples).

I had therefore tried to find for this operation an intrinsic formula, using rather the form and the 3 vectors, than the coordinates and the components. The final answer (which is strangely asymmetrical and hence provides 3 different points U, permuting the arguments) has been found by F. Aicardi:

$$U = F(u, v)r + F(v, r)u - F(r, u)v \ ,$$

where F is the symmetric bilinear form, equal to f along the diagonal.

It is an interesting question to find for which forms do the values form semigroups and for which ones they do not, for instance, what is the proportion of the perfect forms among all the forms (say, in the ball $m^2 + n^2 + k^2 \leq R^2$ of a large radius R). The tables of the forms $mx^2 + ny^2$ with bounded $|m|$ and $|n|$ are presented below in the section 3. The perfect forms fill, it seems, approximately 20 percents of the square which I had studied. The asymptotical proportion for large R is perhaps representable in terms of π and ζ.

The statistics should be also studied, counting the $\mathrm{SL}(2, \mathbb{Z})$-orbits of the forms, rather than the forms themselves. The statistics of these orbits is discussed below in section 4, where the classes of the hyperbolic forms of a fixed determinant are counted (by the integer points in an ellipse). And these statistics should be compared with the natural Lie algebra structure (of the quadratic Hamiltonians) of the space of forms.

2 Hyperbolic reflections groups and periodic continued fractions palindromy

The hyperbola's symmetries form a subgroup in $\mathrm{GL}(2,\mathbb{R})$, consisting of the hyperbolic rotations, belonging to $\mathrm{SL}(2,\mathbb{R})$, and of the hyperbolic reflections (of determinant -1). The quadratic forms f on $\mathbb{Z}^2$ and their hyperbolae $f = c$, $c \in \mathbb{Z}$ are related to the subgroups of those symmetries, which preserve the lattice $\mathbb{Z}^2$ in $\mathbb{R}^2$. We shall now provide some methods, constructing such symmetries. They are essentially some reformulations of the Picard-Lefschetz formula for cycle reflection of singularity theory, but I shall rather imitate, than use, these formulae.

Suppose first that the quadratic form f attains the value $1 = f(v)$, at an integer vector $v = (p, q)$.

Theorem 2. *The following operator $R_v : \mathbb{R}^2 \to \mathbb{R}^2$ is a reflection (of determinant -1), preserving the integer points lattice, the hyperbola $f = 1$ and its point v:*

$$R_v w = -w + \lambda v, \quad \lambda = 2F(v, w)/f(v) .$$

Here F is the symmetric bilinear form, coinciding with f along the diagonal. For the quadratic form $f(x, y) = mx^2 + ny^2 + kxy$ this bilinear form takes at the vectors $v = (p, q)$ and $w = (x, y)$ the value $F(v, w) = mpx + nqy + k(py + qx)/2$:

$$2F(v, w) = f(v + w) - f(v) - f(w) .$$

Proof of Theorem 2. The operator R_v is a linear operator, since F depends on w linearly. The vector v remains invariant: $\lambda = 2F(v, v)/f(v) = 2$, $R_v v = -v + 2v = v$. The f-orthogonal to v vectors w change their sign : $\lambda = 0$ since $F(v, w) = 0$, $R_v w = -w$. Hence the vectors tangent to the hyperbola $f = 1$ at v are reversed : at the point v the differential of f takes on any vector w the value $2F(v, w)$.

The integer points lattice is preserved by operator R_v, since 1) $2F(v, w)$ is integer-valued, 2) $f(v) = 1$ and 3) $\det R_v = -1$ (since $R_v v = v$, $R_v w = -w$ for the two vectors considered above).

The form f (and hence the hyperbola $f = 1$) is preserved by the operator R_v:

$$f(R_v w) = f(-w) + 2\lambda F(-w, v) + f(\lambda v) = f(w),$$

since $f(-w) = f(w)$, $f(\lambda v) = \lambda^2 f(v)$ and hence the increment is vanishing:

$$2\lambda F(-w, v) + f(\lambda v) = \lambda(-2F(v, w) + \lambda) = 0$$

by the definition of the coefficient λ.

Thus, the operator R_v sends the point v, the lattice $\mathbb{Z}^2$ and the hyperbola $f = 1$ to themselves, reversing the orientations of the hyperbola and of the plane.

Corollary. *The reflection operator R_v sends each hyperbola $f = const$ to itself and sends the integer points on any of these hyperbolae to the integer points on the same one.*

Remark. One might provide an interesting operator R_v for any integral vector v, whatever the integer $f(v) \neq 0$ is. To avoid the nonintegral points, we choose now

$$R_v w = -f(v)w + 2F(v, w)v.$$

In this case $f(R_v w) = f^2(v) f(w)$, hence the form f is no longer invariant, unless $f(v) = \pm 1$.

The hyperbola and the lattice are sent onto themselves only if $f(v) = \pm 1$, otherwise they are sent onto the homotetical ones. Such generalized reflections do not generate a group, but only a semigroup of linear operators, which is still interesting for the quadratic form arithmetics (and which is evidently related to the values semigroup, studied in the section 1). The semigroup, formed by the values, is the commutative version of the semigroup of the linear operators, generated by the generalized reflections R_v, corresponding to all integer points v and to a given quadratic form f.

Remark 1. It would be interesting to know whether the product of three such reflections, R_u, R_v, R_w, is itself a generalized reflection. If it is the case, the products of pairs, $R_u R_v$, do form a *semigroup of linear operators*:

$$(R_a R_b)(R_u R_v) = R_c R_v$$

for $R_c = R_a R_b R_u$, which is an interesting SL(2, $\mathbb{Z}$)-invariant of the form f.

Remark 2. For a multiplicative semigroup of integers, the products of all the semigroup elements by an element N do form a new semigroup, divisible by N:

$$(NA)(NB) = N(NAB).$$

It would be interesting which of the semigroups of the values of quadratic forms can be represented as such products of a number with a "divided" semigroup. The form $4x^2 - 2y^2$ is the first interesting example of the strange sets of values : namely, every square is a value of the form $2x^2 - y^2$, for $2^{a_2}\,3^{a_3}\,5^{a_5}\ldots = 2x^2 - y^2$, a_p is even for every prime number $p = 8q + 3$, $8q + 5$, while the prime numbers $p = 8q \pm 1$ are all representable by the form $2x^2 - y^2$, it seems.

The hyperbolic reflections are related to the palindromic structure of the continued fractions by the following construction. Let the hyperbola $f = 1$ contains two different integer points, u and v. The product of the two reflections R_u and R_v is then a hyperbolic rotation, preserving the hyperbola. It moves the points of the hyperbola from one of the infinite points (corresponding to an asymptotic direction of the hyperbola) to the other.

This symmetry explains the periodicity of the continued fraction, representing the inclination t of the asymptote $x = ty$ of the hyperbola ($f(x, y) = 0$ along this line). The continued fraction $t = [a_0, a_1, \ldots]$ has the form

$$t = a_0 + \frac{1}{a_1 + \ldots},$$

where a_k $(k > 0)$ are natural numbers. We suppose here, for simplicity, that t is positive. This can be always achieved by a convenient $\mathrm{SL}(2, \mathbb{Z})$ choice of the coordinates.

The continued fraction algorithm describes a sequence of integer vectors v_k, approximating the line $x = ty$. Namely, the points v_k are the vertices of the two boundaries of the convex hulls of the two sets of the integer points : one is formed by the points in the angle $\{x > ty\}$ and the other in the angle $\{x < ty\}$ (into which angles the line $x = ty$ divides the quadrant $x \geq 0$, $y \geq 0$).

The traditional notations for the approximating vectors are

$$v_{-1} = (0, 1), \quad v_0 = (1, 0); \quad v_{k+1} = v_{k-1} + a_k v_k.$$

These formulae provide the algorithm of the construction of the two convex hulls and of the continued fraction expansion for $t = [a_0, a_1, \ldots]$.

It starts from the choice of $a_0 = [t]$ (the integral part). This choice, as well as the next ones, describes the motion from the vector v_{k-1}, preceding the last already constructed one, v_k, adding to it that last constructed vector as many (a_k) times, as it is possible before the crossing of the line $x = ty$.

We shall denote the coordinates of the point v_k by p_k and q_k. The construction, described above, implies that the vectors v_k and v_{k+1} are situated on the different sides of the line $x = ty$. The area of the parallelogram, formed by these two vectors, is equal to one (at every step k), or, taking the $p \wedge q$ orientation into account, to

$$\det \begin{pmatrix} p_k & p_{k+1} \\ q_k & q_{k+1} \end{pmatrix} = (-1)^k.$$

Suppose now, that the quadratic form f attains the value 1 at two different points on the same branch of the hyperbola $f = 1$.

Theorem 3. *The periods of the periodic continued fraction, representing the tangents t of the inclination of an asymptote of the hyperbola, is in this case palindromic (sent to itself by a symmetry, reversing the order of the elements: $a_{r+i} = a_{r-i}$, for some r).*

Proof. The hyperbolical rotation $R_v R_w$, where v and w are the given points of value 1, sends to itself the hyperbola asymptote $x = ty$ and preserves the sets of the integer points above it and below it. Hence it preserves the convex hulls (at least far from the origin), and hence sends the vertices v_i of its boundary to the vertices v_{i+2s} of the same boundary. The reflection R_v also permutes the vertices of one of the 2 boundaries of the convex hull (namely on the one, to which belongs the reflection center $v = v_j$). These boundary vertices are the vertices v_{j+2s}, and R_{v_j} sends v_{j+2s} to v_{j-2s}.

It is easy to derive from this invariance property the continued fraction palindromy. Indeed, the number a_k being the integral length of the segment between v_{k-1} and v_{k+1}, the symmetrical segments (v_{j-2s-2}, v_{j-2s}) and (v_{j+2s}, v_{j+2s+2}) have equal integral lengths, and hence we get the equality $a_{j-2s-1} = a_{j+2s+1}$.

The palindromic property of the numbers a_{j+2s} follows from the description of these numbers as of the integer angles of the same convex hull boundary at the vertices points:

$$\det(v_k, v_{k+2}) = a_{k+1} \det(v_k, v_{k+1}) = a_{k+1} (-1)^k .$$

Whence the symmetry R_v, mapping the boundary of the convex hull to itself, permutes, reversing the order, the numbers a_{j-2s}, and so we end the palindromy proof : $a_{j-2s} = a_{j+2s}$.

Example. The form $f = x^2 - ny^2$, where n is not a square, attains the value 1 at the point (1, 0) and at some other integer point with positive coordinates (the Pell equation theory). Hence the continued fractions expansions of the irrational numbers $t = \sqrt{n}$ are palindromic.

Thus, the preceding algorithm provides the 4-periodic continued fraction

$$\sqrt{167} = 12 + \cfrac{1}{11 + \cfrac{1}{1+\cfrac{1}{24+\cfrac{1}{1+\cfrac{1}{11+\cfrac{1}{1+\dots}}}}}},$$

which we shall denote by the symbol

$$\sqrt{167} = [12; (11, 1, 24, 1), (11, 1, 24, 1), \dots] .$$

It is palindromic with respect to any of the symmetries centers a_j, $j = 4k+3$: $a_{j-s} = a_{j+s}$ (whenever both indices are positive).

The continued fraction calculation for the quadratic form's f asymptotic direction is very fast, since the intersection with the line $x = ty$ may be recognized by the change of the sign of the form. We use the notations

$$f_k = f(v_k), \qquad F_k = F(v_{k-1}, v_k)$$

and introduce new vectors, slightly crossing the line:

$$\tilde{v}_{k+1} = v_{k-1} + \tilde{a}_k v_k = v_{k+1} + v_k \qquad (\text{where } \tilde{a}_k = a_k + 1) .$$

The vector $\tilde{v}_{k+1}$ is the first vector on the ray $\{v_{k-1} + a v_k\}$, where the sign of f differs from that of $f(v_{k-1}) = f_{k-1}$. We denote this first opposite sign value by

$$\tilde{f}_{k+1} = f(\tilde{p}_{k+1}, \tilde{q}_{k+1}),$$

where $\tilde{p}_{k+1}$ and $\tilde{q}_{k+1}$ are the components of the vector $\tilde{v}_{k+1}$.

With these notations, we get from the calculation of a_k the identity

$$f_{k+1} = f_{k-1} + 2a_k F_k + a_k^2 f_k ,$$

where a_k should be the maximal integral value, for which the sign of f_{k+1} remains equal to the sign of f_{k-1} (while that of $\tilde{f}_{k+1}$, corresponding to the next value of a, $\tilde{a}_{k+1} = a_k + 1$, should differ).

To continue the calculations it is useful to observe that we have recurrently

$$F_{k+1} = F_k + a_k f_k .$$

It is also useful to represent the increment $\Delta_k = f_{k+1} - f_{k-1}$ in the form

$$\Delta_k = a_k\,(2F_k + a_k f_k)\ ,$$

making easy the control of the signs of the increments Δ_k and $\tilde{\Delta}_k = \tilde{a}_k(2F_k + \tilde{a}_k f_k)$. The recurrent calculations are represented below by the tables, similar to the following one, made for $t = \sqrt{167}$, $f = x^2 - 167y^2$:

k	-1	0	1	2	3	4	5	6	$\dots$
a_k		12	1	11	1	24	1	11	$\dots$
p_k	0	1	12	13	155	168	4 187	4 355	$\dots$
q_k	1	0	1	1	12	13	324	337	$\dots$
f_k		+1	−23	+2	−23	+1	−23	+2	$\dots$
F_k		0	12	−11	11	−12	12	−11	$\dots$
$\tilde{p}_k$			13	25	168	323	4 355	8 542	$\dots$
$\tilde{q}_k$			1	2	13	25	337	661	$\dots$
$\tilde{f}_k$			+2	−43	+1	−46	+2	−43	$\dots$

The palindromic properties are here the identities:

$$a_{4+s} = a_{4-s} \quad (-3 \le s \le 3), \qquad a_{6+s} = a_{6-s} \quad (-5 \le s \le 5), \dots ;$$
$$f_{4+s} = f_{4-s} \quad (-4 \le s \le 4), \qquad f_{6+s} = f_{6-s} \quad (-6 \le s \le 6), \dots ;$$
$$F_i = -F_j \quad (i + j = 5;\ i, j > 0), \qquad (i + j = 9;\ i, j > 0)\ , \dots .$$

The periodicity holds for the four lines of the table:

$$a_{i+4} = a_i, (i > 0)\ , \quad f_{i+4} = f_i, (i \ge 0)\ ,$$
$$F_{i+4} = F_i, (i > 0), \quad \tilde{f}_{i+4} = \tilde{f}_i, (i > 0)\ .$$

The hyperbola $f = 1$ contains two integer points v_0 and v_4. The reflection operator R_{v_0} has the matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

The resulting hyperbolic rotation is

$$T = R_{v_4} R_{v_0} = \begin{pmatrix} 56\,447 & 729\,456 \\ 4\,368 & 56\,447 \end{pmatrix},$$

acting on the vertices of the convex hull as the shift, $T v_i = v_{i+8}$.

Now we shall prove the palindromic property on the continued fractions of the square roots of rational numbers. Let m/n be a rational number, the integer m and n having no common nontrivial divisor (different from 1), and m having no (nontrivial) square factors (m may be 6, but may not be 12).

Theorem 4. *The continued fraction of the square root $\sqrt{m/n}$ is palindromic.*

Proof. The value $f(v)$ of the form $f = mx^2 - ny^2$ at the point $v = (1, 0)$ is m. The value of the bilinear form $F(v, w)$ at any vector $w = (x, y)$, being mx, it is divisible by m. Hence the reflection operator, acting on w as

$$R_v w = -w + \left(\frac{2F(v, w)}{f(v)}\right) v \, ,$$

preserves the integral lattice $\mathbb{Z}^2$. It preserves also the hyperbola $f = m$.

To find a second integral point at this hyperbola, we have to solve the equation $mx^2 - ny^2 = m$. This equation implies that $y = mz$ (since m and n have no common divisors and m has no square divisors). We obtain for the integers x and z the Pell equation $x^2 - mnz^2 = 1$, which has, as it is well known, a nontrivial solution (where $x^2 \neq 1$). Thus we get the second integer point, $u = (x, mz)$, at the hyperbola $f = m$.

The value of the bilinear form $F(u, w)$, where $w = (a, b)$, is equal to $mxa - nmzb$. This number is divisible by m. Hence the reflection R_u is defined by an integral elements matrix. It is of determinant -1 and hence it preserves the lattice $\mathbb{Z}^2$.

Thus we had constructed two symmetries R_u, R_v of the quadratic form f. These symmetries act on the continued fractions of the two numbers $t = x/y$, defining the inclinations of the asymptotes $f(x, y) = 0$ of the hyperbola $f(x, y) = m$.

These symmetries provide the palindromic structure of the periodic continued fractions of the numbers t, as it has been explained in the proof of Theorem 3 above.

In many cases one can deduce from the palindromic structure of the period of a continued fraction its relation to the situation of Theorem 4.

Theorem 5. *Let x be the number, whose continued fraction has an odd period and is palindromic with the period $(b, \dots, d, d, \dots, b, 2a)$. Then x is the square root of a rational number.*

Denote the vectors like $(x, 1)$ by the capitals like X, and denote by M_a and R the matrices

$$M_a = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}, \qquad R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The representation $y = a + \frac{1}{x}$ means that $M_a X$ is parallel to Y. Hence the continued fraction

$$y = a + \cfrac{1}{b + \cdots + \frac{1}{d + \frac{1}{x}}}$$

means that the vector Y is parallel to the result of the application to X of the product operator $M_{[a,b,\dots,d]} = M_a M_b \cdots M_d$.

Palindromic Lemma. *The inverse to the product operator is conjugate to the "palindromic product" operator by the projective line involution R:*

$$(M_{[a,b,\dots,d]})^{-1} = \pm\, R\, M_{[d,\dots,b,a]}\, R.$$

Proof of the lemma. The relation $RM_aR = M_a^{-1}$ is obvious, since the equation $y = a + \frac{1}{x}$ is equivalent to the equation

$$-\frac{1}{x} = a + \frac{1}{(-\frac{1}{y})} .$$

Hence we represent the long inverse product in the form

$$(M_a M_b \cdots M_d)^{-1} = M_d^{-1} \cdots M_a^{-1} = (RM_dR) \cdots (RM_aR) = \pm RM_d \cdots M_b M_a R ,$$

(since $R^2 = -1$) as required.

The Lemma implies the inverse continued fraction formula:

$$\left(-\frac{1}{x}\right) = d + \frac{1}{c + \cdots + \frac{1}{b + \frac{1}{a + (-1/y)}}} .$$

Proof of Theorem 5. The palindromic property of the continued fraction of x means (in the above notations) that for $z = 1/x$ one has the parallelisms

$$X \,||\, (M_{[a,\dots,d]}\, Y) , \quad Y \,||\, (M_{[d,\dots,a]}\, Z) .$$

According to the Palindromic Lemma, we can write this condition in the form

$$(RY) \,||\, M\,(RX) , \quad Y \,||\, (MZ) ,$$

where $M = M_{[d,\dots,a]}$ is the product linear operator; we shall denote its matrix by $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$.

We thus get the expressions for the right part vectors of the above parallelisms,

$$RX = (-1, x), \quad M\,(RX) = (-\alpha + \beta x, -\gamma + \delta x),$$
$$RMRX = (\gamma - \delta x, -\alpha + \beta x), \quad (MZ) \,||\, (\alpha + \beta x, \gamma + \delta x).$$

The required parallelism condition, $(RMRX)||(MZ)$, means therefore the vanishing of the determinant

$$\det \begin{pmatrix} \alpha+\beta x & \gamma-\delta x \\ \gamma+\delta x & -\alpha+\beta x \end{pmatrix} = \beta^2 x^2 - \alpha^2 - \gamma^2 + \delta^2 x^2,$$

whence x^2 is the rational number $(\alpha^2+\gamma^2)/(\beta^2+\delta^2)$. Theorem 5 is thus proved.

Remark. The golden ratio, $(\sqrt{5}+1)/2 = [1, 1, 1, \dots]$, is not the square root of a rational number. The proof of Theorem 5 is however applicable to the palindromes of even periods of the form $[b, \dots, c, d, c, \dots, b, 2a]$. Several examples are discussed at the end of the next section.

3 Statistics of diagonal forms

Forms $f = mx^2 + ny^2$ provide interesting examples for many properties of their arithmetics and geometry. I shall present below the tables, showing the places occupied on the plane with coordinates (m, n) by these (perfect) forms, whose values set are multiplicative semigroups.

The form mx^2 is perfect if and only if the number m is a square. Indeed, if $m = n^2$, we get

$$(mx^2)(my^2) = m(nxy)^2.$$

For mx^2 to be perfect, m^2 should be attainable, since $m1^2 = m$ is. Thus we get for a perfect form the equality $m^2 = mx^2$ for some integer $x = n$, and hence $m = n^2$.

Consider now the forms $x^2 + ny^2$.

Theorem 6. *All these forms are perfect.*

Proof. The value $f(x, y) = 1$ is attainable (at $x = 1$, $y = 0$).

Hence the values form a semigroup (by the Corollary 1 of the Theorem 1).

Turn now to the forms $-x^2 + ny^2$.

Theorem 7. *No such form, where n is negative, is perfect. For the positive values of n between 1 and 100, the form is perfect if and only if n has one of the following 21 values :*

$$n = 1, 2, 5, 10, 13, 17, 26, 29, 37, 41, 50, 53, 58,$$
$$61, 65, 73, 74, 82, 85, 89, 97 .$$

The periods of the corresponding continued fractions $\sqrt{n}$ are odd numbers.

Remark. The proofs of this Theorem and of the next similar Theorems are based on the studies of some infinite series of values of n, for which we prove either the semigroup property of the set of the values of the form or its absence. The restriction of the smallness of n is only used to check that our series do cover all the values of n (till the required limit).

Proof. First one should consider the residues mod u. Since $f = -1$ is attainable (at $x = 1$, $y = 0$), if values of f form a semigroup, the equation of the representability of $f = 1$ (implying that $-x^2 = 1 \bmod u$) should be solvable for any u.

For $u = 3, 4, 7, 11, 19, 23, 31, 43, 47$, this congruence has no solutions. Hence the form $-x^2 + ny^2$ is not perfect, provided that n is divisible by one of these 8 numbers.

A similar argument shows that no $n = 3 \bmod 4$ is possible for a perfect form $-x^2 + ny^2$. Indeed, $x^2 = 0$ or $1 \bmod 4$, hence $-x^2 + ny^2$ is congruent mod 4 to (0 or -1) +3 (0 or 1), which is not congruent to 1 mod 4. Thus, the value 1 is not attainable by the form, while the value -1 is, and hence the form is not perfect.

For $n = 25$ the value 1 is also unattainable since the equality $25y^2 - x^2 = (5y - x)(5y + x) = 1$ implies, that $5y - x = 5y + x = \pm 1$, and hence $x = 0$.

The remaining values of n, smaller than 100, are all of the form $a^2 + b^2$. Among them $n = 34$ does not generate a semigroup, since $-x^2 + 34y^2$ does not attain the value 1. To prove this it suffice to calculate the continued fraction of $\sqrt{34}$, as it is explained in the section 2, and to see whether $f_k = -1$ is attained for $f = x^2 - 34y^2$. We get, from the algorithm of the section 2, the table

k	-1	0	1	2	3	4	5	6	7	8	9	...
a_k		5	1	4	1	10	1	4	1	10	1	...
f_k	-34	+1	-9	+2	-9	+1	-9	+2	-9	+1	-9	...

proving that the value -1 is never attained by the form $x^2 - 34y^2$. Hence the form $-x^2 + 34y^2$ is not perfect: the number -1 is a value and +1 is not.

The remaining 21 values of n (smaller than 100) are listed above. The corresponding form $-x^2 + ny^2$ is perfect by corollary 1 of Theorem 1, since the value 1 is attained by the form at the following place:

n	1	2	5	10	13	17	26	29	37	41	50
x	0	1	2	3	18	4	5	70	6	32	7
y	1	1	1	1	5	1	1	13	1	5	1

n	53	58	61	65	73	74	82	85	89	97
x	182	99	29 718	8	1 068	43	9	378	500	5 604
y	25	13	3 805	1	125	5	1	41	53	569

These places are easily calculated by the above algorithm. But some series of them might be obtained with no calculations. For instance, if $n = a^2 + 1$, it suffice to take $x = a,\ \ y = 1$ (cases $n = 1, 2, 5, 10, 17, 26, 37, 50, 65, 82$).

The restriction $n < 100$ is not used in these studies of the series.

For the series $n = a^2+4$, where a is odd ($n = 5, 13, 29, 53, 85, \dots$), the value 1 of the quadratic form $-x^2+ny^2$ is attained at $x = a(a^2+3)/2,\ \ y = (a^2+1)/2$.

The value set of the form is a semigroup (accordingly to Corollary 1 of Theorem 1). Theorem 7 is thus proved.

I do not know whether similar methods work for $n = a^2 + b^2$. At least for $n = 45(= 36 + 9)$ and $n = 34(= 25 + 9)$ the value sets of the quadratic forms $-x^2 + ny^2$ do not contain 1 and hence do not form a semigroup.

It is interesting that, every time when the Diophantine equation $mx^2+ny^2 = N$ was solvable mod p (for sufficiently many p's), it had been solvable in the integers. I do not know whether this observation might be proved as a general theorem (either for our quadratic forms representation equations (where the mod p^q version had been proved by Hasse), or for the general Diophantine systems, and either provided that the existence of a solution modulo any prime number p is given, or even modulo any integer, which might be virtually non equivalent to the mod p solvability).

This difficulty is similar to the calculus convergence problem situations, where the existence of a formal Taylor series (or of a solution modulo any degree of the maximal ideal) does not imply the genuine existence of a holomorphic solution of a differential equation.

Consider the quadratic forms $\pm 2x^2 + ny^2$. If both signs are negative, the form can't be perfect, since -2 is attained (at $x = 1, y = 0$) while 4 is not (being positive).

If both signs are positive and n is at least 2, the value 4 can be attained only when $4 = 2x^2+ny^2 \geq 2(x^2+y^2)$, that is at the places where $x^2 \leq 1$ and $y^2 \leq 1$. We get thus only two perfectness candidates cases ($x = 0, n = 4, y^2 = 1$) and ($x^2 = 1, n = 2, y^2 = 1$). These two forms, $2x^2 + 4y^2 = 2(x^2 + 2y^2)$ and $2x^2 + 2y^2 = 2(x^2 + y^2)$, are perfect, accordingly to Corollary 2 of Theorem 1, since $N = 2$ is attained by the form $x^2 + 2y^2$ (at (0, 1)) and by $x^2 + y^2$ (at (1,1)).

The remaining nonnegative forms (with $n < 2$) $2x^2$ and $2x^2 + y^2$, define values sets, the first of which does not form a semigroup ($2x^2$ does not attain the value 4), the second form being perfect (by Corollary 1 of Theorem 1), since the second form takes the value 1 at (0, 1).

The study of the hyperbolic forms $2x^2 - ny^2$ and $-2x^2 + ny^2$ ($n > 0$) leads to the following conclusions.

Theorem 8. *The quadratic form $f = -2x^2 + ny^2$ ($0 < n < 100$) is perfect if and only if n has one of the following 27 values:*

$$n = 1, 3, 4, 6, 9, 11, 12, 19, 22, 27, 33, 36, 38, 43, 44,$$
$$51, 54, 57, 59, 67, 73, 76, 81, 83, 86, 89, 99.$$

Proof. The direct calculation of the residues of the squares of integers mod u shows, that 4 is not congruent to $-2x^2$ (mod u) for the following 12 values of u:

$$u = 5, 7, 13, 23, 29, 31, 37, 47, 53, 61, 71, 79\ .$$

Since $f = -2$ for ($x = 1$, $y = 0$), the form f is not perfect, if the equation $-2x^2 + ny^2 = 4$ has no integral solution. Thus, the form $-2x^2 + ny^2$, corresponding to an integer n, divisible by any of the 12 factors u listed above, is not perfect.

This is also true for any n, congruent to 0 or to 2 mod 8 (since x^2 is congruent to 0, 1 or 4 mod 8 and hence $-2x^2 + ny^2$ is not congruent to 4 mod 8, as it should be if $-2x^2 + ny^2 = 4$). The condition $n < 100$ is not used here. For $n < 100$ the above congruences leave not so many candidates for the perfect forms $-2x^2 + ny^2$. The values $n = 2a^2 + 1$ (like 1, 3, 9, 19, 33, 51, 73, 99) do define perfect forms, accordingly to the Corollary 1 of Theorem 1, since $f = 1$ for ($x = a$, $y = 1$).

Another infinite series of the perfect forms is provided by the choice of $n = 2a^2 + 4$, (like $n = 4, 6, 12, 22, 36, 54, 76$).

Indeed, these quadratic forms are divisible by $2 : -2x^2 + (2a^2 + 4)y^2 = -2(x^2 - (a^2+2)y^2)$, and $x^2 - (a^2+2)y^2 = -2$ for ($x = a$, $y = 1$). Therefore, the form $-2x^2 + (2a^2+4)y^2$ is perfect, accordingly to the Corollary 2 of Theorem 1 (where $N = 2$). We had not used the restriction here. Taking this restriction into account the remaining numbers n (candidates to perfectness) are only the 16 values, 11, 17, 27, 38, 41, 43, 44, 57, 59, 67, 68, 81, 83, 86, 89, 97.

For many of these values of n the form $-2x^2 + ny^2$ attains the value 1 and hence it is perfect, accordingly to Corollary 1 of Theorem 1. These 9 numbers n of the preceding list (and those (x, y) where $f = 1$) are listed in the following table:

n	11	27	43	57	59	67	81	83	89
x	7	11	51	16	277	191	70	20 621	20
y	3	3	11	3	51	33	11	3 201	3

To prove the perfectness of the form $-2x^2+ny^2$ for the even number $n=2m$, it suffices to solve the equation $-x^2+my^2=2$: the Corollary 2 of Theorem 1 (for $N=2$) implies that the set $\{-2x^2+2my^2\}$ is then a semigroup.

The corresponding numbers n of our list (and their x and y) are listed in the following table:

n	38	44	86
x	13	14	59
y	3	3	9

We have thus proved the perfectness for all the 27 cases of Theorem 8. It only remains to prove the nonperfectness in the few remaining cases, which are $n=17, 41, 68, 97$.

Lemma. *The form $-2x^2+ny^2$ does not attain the value 4 for any of these 4 values of n.*

Proof. Applying the (quadratic) continued fractions algorithm, described in the section 2, we find the vertices v_k of the boundaries of the convex hulls and the values f_k of the form $f=2x^2-ny^2$ at these vertices. The absence of the value -4 in these tables proves its unattainability (accordingly to the convexity arguments and to the easy calculation of the values of f on the segments, joining the neighbouring vertices of the same convex hull).

Continued fractions of $\sqrt{n/2}$ $(f=2x^2-ny^2)$.

Case $n=17$: $\sqrt{17/2}=[2, (1, 10, 1, 4), (1, \dots]$.

k	-1	0	1	2	3	4	5	...
a_k		2	1	10	1	4	1	...
p_k	0	1	2	3	32	35	172	...
q_k	1	0	1	1	11	12	59	...
f_k	-17	$+2$	-9	$+1$	-9	$+2$	-9	...
F_k		0	$+4$	-5	$+5$	-4	$+4$	...
$\tilde{p}_k$			3	5	35	67	207	...
$\tilde{q}_k$			1	2	12	23	71	...
$\tilde{f}_k$			$+1$	-18	$+2$	-15	$+1$	...

This table implies that the negative values of $2x^2 - 17y^2$ are smaller (or equal) than -9.

Case $n = 41 : \sqrt{41/2} = [4, (1, 1, 8), (1, 1, 8), \dots]$.

The table below shows that the negative values of $2x^2 - 41y^2$ are either equal to -2 or are smaller (or equal) than -9. Indeed, $a_2 = 1$, hence there are no integer points inside the segment, joining v_1 to v_3. Similarly, $a_4 = 1$, hence there are no integer points inside the segment joining v_3 to v_5.

Inside the segment joining v_5 to v_7, there are $7 = a_6 - 1$ integer points, but the values of f at these points are smaller than the -9 value, attained at both ends of the segment.

k	-1	0	1	2	3	4	5	6	7
a_k		4	1	1	8	1	1	8	1
p_k	0	1	4	5	9	77	86	163	1 390
q_k	1	0	1	1	2	17	19	36	307
f_k	-41	$+2$	-9	$+9$	-2	$+9$	-9	$+2$	-9
F_k		0	$+8$	-1	$+8$	-8	$+1$	-8	$+8$
$\tilde{p}_k$			5	9	14	86	163	249	1 553
$\tilde{q}_k$			1	2	3	19	36	55	343
$\tilde{f}_k$			$+9$	-2	$+23$	-9	$+2$	-23	$+9$

Case $n = 68 : \sqrt{68/2} = [5, (1, 4, 1, 10), (1, \dots]$.

The table below shows that the negative values of $2x^2 - 68y^2$ are smaller (or equal) than the value -18 (attained, for instance, at $(x = 5, y = 1)$). In this case no large value is attained at the vertices of the convex hull boundary, and the values along a segment of the boundary is smaller than at its ends, since the quadratic function, that we restrict to the boundary, is convex along this segment.

k	-1	0	1	2	3	4	5	6
a_k		5	1	4	1	10	1	4
p_k	0	1	5	6	29	35	379	414
q_k	1	0	1	1	5	6	65	71
f_k	-68	$+2$	-18	$+4$	-18	$+2$	-18	$+4$
F_k		0	10	-8	$+8$	-10	$+10$	-8
$\tilde{p}_k$			6	11	35	64	414	793
$\tilde{q}_k$			1	2	6	11	71	136
$\tilde{f}_k$			$+4$	-30	$+2$	-36	$+4$	-30

Thus, the values set of the form $-2x^2 + 68y^2$ does not contain the value 4, while the value -2 is attained (at $x = 1$, $y = 0$). Therefore, this form is not perfect.

Case $n = 97$: $\sqrt{97/2} = [6, (1, 26, 1, 12), \dots]$.

The table below shows that the set of the values of the form $-2x^2 + 97y^2$ is not a semigroup, since it contains -2 and does not contain 4. In this table the form $f = 2x^2 - 97y^2$ is considered, and the value $f = -4$ is not attained, since the negative values of f are smaller (or equal) than the value -25 (attained at $x = 6$, $y = 1$).

k	-1	0	1	2	3	4	5	6
a_k		6	1	26	1	12	1	26
p_k	0	1	6	7	188	195	2 528	2 723
q_k	1	0	1	1	27	28	363	391
f_k	-97	$+2$	-25	$+1$	-25	$+2$	-25	$+1$
F_k		0	$+12$	-13	$+13$	-12	$+12$	-13
$\tilde{p}_k$			7	13	195	383	3 723	5 251
$\tilde{q}_k$			1	2	28	55	391	754
$\tilde{f}_k$			$+1$	-50	$+2$	-47	$+1$	-50

We had thus proved the completeness of the list of the perfect forms $-2x^2+ny^2$, provided by Theorem 8 (for $0 < n < 100$).

Theorem 9. *The quadratic form* $f = 2x^2 - ny^2$ $(0 < n < 100)$ *is perfect if and only if* n *has one of the following* 18 *values:*

$$n = 1, 4, 7, 14, 17, 23, 28, 31, 46, 47, 49, 62, 68, 71, 79, 92, 94, 97 .$$

Proof. As in the proof of Theorem 8, we start with some quadratic residues calculations, showing that for some values of n the number $2x^2 - ny^2$ is not congruent to 4 mod u, while it should be congruent (and even equal) to 4 if f is perfect, since $f = 2$ is attained (for $x = 1$, $y = 0$).

These 18 perfectness restrictions are listed in the following table, presenting the values of the modulo u and the forbidden values of the residues r of n mod u (obstructing the perfectness of the form $2x^2 - ny^2$) :

u	3	5	8	8	11	13	16	19	29
r	0	0	0	2	0	0	6	0	0
u	32	32	37	43	53	59	61	67	83
r	12	20	0	0	0	0	0	0	0

As in the other cases, the condition $n < 100$ of Theorem 9 is of no importance for these restrictions.

Next we prove the perfectness of some infinite series of forms $2x^2 - ny^2$.

Series $n = 2a^2 - 1$ (containing, for instance, the 7 values $n = 1, 7, 17, 31, 49, 71, 97$, smaller than 100).

For $(x = a, y = 1)$ we get $2x^2 - ny^2 = 1$, and hence the form is perfect, accordingly to Corollary 1 of Theorem 1.

Series $n = 2a^2 - 4$ (containing, for instance, the 6 values $n = 4$, 14, 28,46, 68, 94, smaller than 100).

For $(x = a, y = 1)$ we get $x^2 - (a^2 - 2)y^2 = 2$, and hence the form $2(x^2 - (a^2 - 2)y^2)$ is perfect accordingly to Corollary 2 of Theorem 1.

Series $n = a^2 - 2$, where a is odd (containing for instance the values $n = 7$, 23, 47, 79, 119, 167, the first 4 being smaller than 100).

For these 6 members of this infinite series I had computed (using the quadratic continued fractions algorithm of the section 2) the explicit representations of 1 by the forms $2x^2 - ny^2$, given in the table below. This table implies that these 6 forms are perfect.

a	3	5	7	9	11	13
n	7	23	47	79	119	167
x	2	78	732	44	54	3 993 882
y	1	23	151	7	7	437 071

Unfortunately, I was unable to find any formula for these experimental results, and the conjecture that the equation $2x^2 - (a^2 - 2)y^2 = 1$ is solvable for any odd value of a remains unproved.

Series $n = 2(a^2 - 2)/b^2$, that is $2a^2 - nb^2 = 4$.

The cases $b^2 \neq 1$ of this series are not immediately evident, but they do exist:

n	4	14	28	46	62	92
a	2	3	4	5	39	156
b	1	1	1	1	7	23

When $b = 2c$ is even, $2a^2 - 4nc^2 = 4$ hence $a = 2d$ is even, and $2d^2 - nc^2 = 1$. In this case the form $2x^2 - ny^2$ is perfect, accordingly to Corollary 1 of Theorem 1.

When b is odd (as in our examples), $n = 2m$ is even, and $a^2 - mb^2 = 2$. In this case the form $x^2 - my^2$ attains the value 2, and hence the doubled form $2x^2 - ny^2$ is perfect, accordingly to Corollary 2 of Theorem 1.

Therefore all the forms of our series are perfect (with no smallness restriction on n).

Returning now to the case $n < 100$ (of Theorem 9), we see that the preceding statements of the perfectness and imperfectness decide the perfectness questions for all the values of n, except the following 3 values: $n = 41, 73, 89$.

The continued fractions, proving the nonperfectness of these 3 forms $2x^2 - ny^2$, are presented in the following 3 tables.

Case $n = 41 : \sqrt{41/2} = [4, (1, 1, 8), (1, 1, 8), \ldots]$.

The table (presented in the proof of Theorem 8 above, in the Lemma) shows that $2x^2 - 41y^2$ does not attain the value 1 (since it is at least $+2$ at the vertices of the boundary of the convex hull, where the form is positive).

This fact implies that the value +4 is not attained too. Indeed, if it where $2x^2 - 41y^2$, the value y should be even : $y = 2z$. Therefore, one should have $x^2 - 82z^2 = 2$, and $x = 2t$ should be even. Thus, we would obtain $2t^2 - 41z^2 = 1$ and the form $2x^2 - 41y^2$ would attain the value 1.

We have thus proved the nonperfectness of the form $2x^2 - 41y^2$, which attains the value 2 but does not attain the value 4.

Case $n = 73 : \sqrt{73/2} = [6, (24, 12), (24, 12), \ldots]$.

Applying the algorithm of the section 2 to the form $f = 2x^2 - 73y^2$, we get the following table.

k	-1	0	1	2	3	4	...
a_k		6	24	12	24	12	...
p_k	0	1	6	145	1 746	42 049	...
q_k	1	0	1	24	289	6 960	...
f_k	-73	$+2$	-1	$+2$	-1	$+2$	...
F_k		0	$+12$	-12	$+12$	-12	...
$\tilde{p}_k$			7	151	1 891	43 795	...
$\tilde{q}_k$			1	25	313	7 249	...
$\tilde{f}_k$			$+25$	-23	$+25$	-23	...

This table implies that the value $f = +1$ is never attained by the form f. We deduce that form f is not perfect: it attains 2, but does not attain the value 4, since otherwise we would have $(2x^2 - 73y^2 = 4) \Rightarrow (y = 2z,\ x^2 - 2 \cdot 73z^2 = 2) \Rightarrow (x = 2t,\ 2t^2 - 73z^2 = 1) \Rightarrow (f = +1$ would be attained).

Thus the form $2x^2 - 73y^2$ is not perfect.

Case $n = 89 : \sqrt{89/2} = [6, (1, 2, 26, 2, 1, 12), \dots]$.
The table for the form $f = 2x^2 - 89y^2$ is:

k	-1	0	1	2	3	4	5	6	7
a_k		6	1	2	26	2	1	12	1
p_k	0	1	6	7	20	527	1 074	1 601	20 286
q_k	1	0	1	1	3	79	161	240	3 041
f_k	-89	+2	-17	+9	-1	+9	-17	+2	-17
F_k		0	+12	-5	+13	-13	+5	-12	+12
$\tilde{p}_k$			7	13	27	547	1 601	2 675	21 887
$\tilde{q}_k$			1	2	4	82	240	401	3 281
$\tilde{f}_k$			+9	-18	+34	-18	+2	-39	+9

It is clear from the table that the value $f = +1$ is never attained by the form. This fact implies that the value 4 is not attained, too. Indeed, if it were attained, we would deduce

$$(2x^2 - 89y^2 = 4) \Rightarrow (y = 2z,\ x^2 - 2\cdot 89z^2 = 2) \Rightarrow (x = 2t,\ 2t^2 - 89z^2 = 1)$$

and the value $f = 1$ would be attained. The unattainability of the value 4 proves that the form $2x^2 - 89y^2$ is not perfect, since the value $f = 2$ is attained (at $x = 1$, $y = 0$) by the form. Therefore, Theorem 9 is proved.

Remark. The series of the forms $f = 2x^2 - ny^2$, $n = a^2 - 2$, which we had studied in the proof, has interesting relations to the series of the forms $X^2 - NY^2$, where $N = 2n$.

Theorem 10. *Let* $2p^2 - nq^2 = 1$. *Then the vector* $(P = 4p^2 - 1,\ Q = 2pq)$ *satisfies the equation* $P^2 - NQ^2 = 1$. *Moreover, the unimodular operators defined by the matrices*

$$(M) = \begin{pmatrix} P & nQ \\ 2Q & P \end{pmatrix}, \qquad (\tilde{M}) = \begin{pmatrix} P & NQ \\ Q & P \end{pmatrix},$$

preserve the forms $f = 2x^2 - ny^2$ *and* $\tilde{f} = x^2 - Ny^2$ *respectfully, whenever* $P^2 - NQ^2 = 1$.

Proof. By the definitions, we have the relations $P = 2(2p^2)-1,\ \ Q^2 = 4p^2q^2$. Substituting $2p^2 = nq^2+1$ in these relations, we get the equalities

$$P = 2nq^2+1, \qquad Q^2 = 2q^2(nq^2+1),$$
$$P^2 = 4n^2q^4+4nq^2+1, \qquad NQ^2 = 4nq^2(nq^2+1),$$

and thus $P^2 - NQ^2 = 1$.

The forms preservation means the identities

$$2(Px+nQy)^2 - n(2Qx+Py)^2 \equiv 2x^2 - ny^2\ ,$$
$$(Px+NQy)^2 - N(Qx+Py)^2 \equiv x^2 - Ny^2\ ,$$

which can be written in the way

$$2P^2-4nQ^2 = 2, \quad 2n^2Q^2 - nP^2 = -n\ ,$$
$$P^2 - NQ^2 = 1, \quad N^2Q^2 - NP^2 = -N\ .$$

All these identities follow from the equation $P^2 - NQ^2 = 1$, proved above.

Remark. For $N = a^2-2,\ \ P = a^2-1,\ \ Q = a$, we get $P^2 - NQ^2 = 1$:

$$(a^2-1)^2 - (a^2-2)a^2 = a^4 - 2a^2 + 1 - a^4 + 2a^2 = 1\ .$$

Therefore, the operator defined by the matrix

$$\begin{pmatrix} a^2-1 & a(a^2-2) \\ a & a^2-1 \end{pmatrix}\ ,$$

preserves the form $f = x^2 - Ny^2,\ \ N = a^2-2$.

For $a = 2, \dots, 13$ we obtain the useful symmetries matrices for the quadratic form f:

$$\begin{matrix} N=2 & N=7 & N=14 & N=23 & N=34 \\ a=2 & a=3 & a=4 & a=5 & a=6 \\ \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}, & \begin{pmatrix} 8 & 21 \\ 3 & 8 \end{pmatrix}, & \begin{pmatrix} 15 & 56 \\ 4 & 15 \end{pmatrix}, & \begin{pmatrix} 24 & 115 \\ 5 & 24 \end{pmatrix}, & \begin{pmatrix} 35 & 204 \\ 6 & 35 \end{pmatrix}, \end{matrix}$$

$$\begin{matrix} N=47 & N=62 & N=79 & N=98 \\ a=7 & a=8 & a=9 & a=10 \\ \begin{pmatrix} 48 & 329 \\ 7 & 48 \end{pmatrix}, & \begin{pmatrix} 63 & 496 \\ 8 & 63 \end{pmatrix}, & \begin{pmatrix} 80 & 711 \\ 9 & 80 \end{pmatrix}, & \begin{pmatrix} 99 & 980 \\ 10 & 99 \end{pmatrix}, \end{matrix}$$

$$\begin{array}{ccc} N = 119 & N = 142 & N = 167 \\ a = 11 & a = 12 & a = 13 \\ \begin{pmatrix} 120 & 1\,309 \\ 11 & 120 \end{pmatrix}, & \begin{pmatrix} 143 & 1\,704 \\ 12 & 143 \end{pmatrix}, & \begin{pmatrix} 168 & 2\,171 \\ 13 & 168 \end{pmatrix}. \end{array}$$

The weak point of the preceding theory is that it reduces the solution of the equation $P^2 - NQ^2 = 1$ to the solution of a more difficult one, $2p^2 - nq^2 = 1$, while the inverse reduction to the Pell equation would be more useful: the existence of the solution of the equation $2p^2 - (a^2 - 2)q^2 = 1$, for the large odd integer values of the parameter a, is still a conjecture.

Turn now to the quadratic forms $\pm 3x^2 + ny^2$. If the signs of both terms are the same, there are few perfect forms. Indeed, the relation $3x^2 + ny^2 = 9$ for a positive n implies that $x^2 \leq 1$, $y^2 \leq 9$. If $x = 0$, the relation $ny^2 = 9$ implies that either ($n = 1$, $y = \pm 3$) or ($n = 9$, $y = \pm 1$). The form $3x^2 + y^2$ is perfect, since it attains the value 1. The form $3x^2 + 9y^2$ is also perfect, since $x^2 + 3y^2$ attains the value 3 (Corollary 2 of Theorem 1, $N = 3$). Thus the contribution of the case $x = 0$ to the list of perfect forms $3x^2 + ny^2$ (for positive n's) consists of only two forms $3x^2 + y^2$ and $3x^2 + 9y^2$.

In the case $x^2 = 1$ the equation $ny^2 = 6$ has the only solution ($n = 6$, $y^2 = 1$), and the form $3x^2 + 6y^2 = 3(x^2 + 2y^2)$ is perfect, since $x^2 + 2y^2 = 3$ at the point (1,1).

Therefore, the complete list of the positive perfect forms $3x^2 + ny^2$, where $n > 0$, consists of the 3 forms:

$$3x^2 + y^2, \quad 3x^2 + 6y^2, \quad 3x^2 + 9y^2.$$

The class of the negative definite forms, including $-3x^2 - ny^2$ (for positive n), does not contain any perfect form, since the squares of the values (like +9), are not attained, all the nonzero values of the form being negative.

Theorem 11. *The quadratic form* $f = 3x^2 - ny^2$ *(where* $0 < n < 100$*) is perfect if and only if n has one of the following* 14 *values:*

$$n = 2, 3, 11, 18, 23, 26, 39, 47, 59, 66, 71, 74, 83, 99\ .$$

Proof. The quadratic residues mod u show that the equation of the representation of the value 9, which should be attained, if the form is perfect, has no solutions, provided that the residue r of the coefficient n mod u has the value shown in the following table:

u	3	4	5	7	9	17	19	27	27	29	31	41	53
r	1	0	0	0	6	0	0	0	9	0	0	0	0

Example. $3x^2 - ny^2 = 9$, where $n \equiv 1\,(3)$, $y = 3z$, $x^2 - 3nz^2 = 3$, $x = 3t$, $3t^2 - nz^2 = 1$ impossible mod 3.

The forms $3x^2 - ny^2$ *defined by such values of* n *are not perfect*, and the list of the remaining possibilities (for $0 < n < 100$) is not large.

The forms, corresponding to $n = 3a^2 - 1$ *are perfect*. Indeed, $f(a, 1) = 1$ in this case and the perfectness is implied by the Corollary 1 of Theorem 1 ($n = 2, 11, 26, 47, 74, \dots$).

Another *series of perfect forms is defined by* $n = 3a^2 - 9$. Indeed, $x^2 - my^2$ attains the value 3 (at $x = a$, $y = 1$) in the case $n = 3m$.

Therefore, the tripled form, f, is perfect, accordingly to Corollary 1 of Theorem 1 ($n = 3, 18, 39, 66, 99$).

The forms $f = 3x^2 - ny^2$ *are perfect for* $n = 12m - 1$, *where* $m = 1, 2, 4, 5, 6, 7$ (excluding 3 and 9, and I do not know what is this series continuation). This perfectness follows from the fact that these forms attain the value 1, at the points, shown in the table (including the values (23, 59, 71, 83), missing in the preceding series):

m	1	2	4	5	6	7
n	11	23	47	59	71	83
x	2	36	732	102	180	1194
y	1	13	151	23	73	227

Remark. These values provide the solutions of many problems, as it is explained above (Theorem 1, Theorem 10). For instance, in the case $n = 47$, it follows from Theorem 10 that $P = 2\,143\,295$, $Q = 221\,064$, $N = 94$, and that *the operator, defined by the matrix*

$$(M) = \begin{pmatrix} 2\,143\,295 & 10\,390\,008 \\ 442\,128 & 2\,143\,295 \end{pmatrix},$$

preserves the form $f_2 = 2x^2 - 47y^2$ (providing the symmetry of the continued fraction $\sqrt{47/2}$).

Theorem 1 provides in this case the f_2*-perfectness proving operation*

$$\big((x, y), (z, w)\big) \mapsto (X, Y),$$

(such that $f_2(x, y) f_2(z, w) = f_2(X, Y)$) :

$$X = 1\,464xz - 7\,097(xw + yz) + 34\,404yw\ ,$$
$$Y = -302xz + 1\,464(xw + yz) - 7\,097yw\ .$$

I have no general formula for the series $n = 12m - 1$.

The set of all the values of n, smaller than 100, is covered by the above series, and so Theorem 11 is proved.

Theorem 12. *The quadratic form* $f = -3x^2 + ny^2$ *(where* $0 < n < 100$*) is perfect if and only if n has one of the following* 25 *values:*

$$n = 1, 3, 4, 7, 9, 12, 13, 19, 21, 28, 31, 36, 37, 39,$$
$$43, 49, 57, 61, 63, 67, 76, 79, 84, 91, 93\ .$$

Proof. The residues of $x^2 \mod u$ show the *nonsolvability of the equation* $-3x^2 + ny^2 = 9$ for the following values of u and of the residue r of the coefficient $n \mod u$:

u	3	5	8	9	17	22	23	27	27	29	41	43
r	2	0	0	6	0	0	0	0	18	0	0	0

Example. If $n = 9m+6$ we find : $(-3x^2+ny^2 = 9) \Rightarrow (-x^2+(3m+2)y^2 = 3) \Rightarrow (-x^2 - y^2 = 0(\mod 3)) \Rightarrow (x = 3z, y = 3w) \Rightarrow (3 = 9(-z^2 + (3m+2)w^2))$, which is impossible. Hence, the equation $-3x^2 + ny^2 = 9$ is not solvable, and the form $-3x^2 + ny^2$ is not perfect, since it attains the value -3.

The table above eliminates most of the candidates n for the perfect forms f. Perfect forms are provided by the following series of forms $f = -3x^2 + ny^2$.

Series $n = 3a^2 + 1$ (including, for instance, $n = 1, 4, 13, 28, 49, 76$). For this choice of n the form f attains the value 1 (at $(x = a,\ y = 1)$). Hence it is perfect, accordingly to Corollary 1 of Theorem 1.

Series $n = 3a^2 + 9$ (including, for instance, $n = 9, 12, 21, 36, 57, 84$). For this choice of n the form $f/(-3) = x^2 - my^2$, (where $m = a^2 + 3$) attains the value -3 (at $(x = a,\ y = 1)$). Hence the form f is perfect, accordingly to Corollary 2 of Theorem 1.

Series $n = 3b^2 + 3b + 3$ (including, for instance, $n = 3, 9, 21, 39, 63, 93$). For this choice of n the form $f/(-3) = x^2 - my^2$, where $m = b^2 + b + 1$, attains the value -3 (at the point $(2b + 1, 2)$). Hence the form f is perfect, accordingly to Corollary 2 of Theorem 1.

Remark. The continued fraction corresponding to the value $n = 3b^2 + 3b + 1$ is, according to F. Aicardi,

$$\sqrt{n/3} = [b, (1, 1, 6b + 2, 1, 1, 2b), (1, 1, \dots, 2b), \dots] .$$

The list of difficult values of the coefficient n, for which the preceding series do not claim neither the perfectness, nor the imperfectness, is rather small, if $0 < n < 100$. Calculating the continued fractions for these difficult cases by the algorithm of the section 2, we find the following results.

Series $n = 12a - 5$. The following 7 values of n define the perfect forms $f = -3x^2 + ny^2$, attaining the value 1 at the following places:

n	7	19	31	43	67	79	91
x	3	5	45	53	293	195	11
y	2	2	14	14	62	38	2

For $n = 61$ (not entering in the list above) the form is also perfect, because in this case $f(9, 2) = 1$.

The 3 remaining difficult values of n define imperfect forms $f = -3x^2 + ny^2$, where the continued fractions, proving the imperfectness, are provided by the following tables (whose notations had been explained in the section 2).

Continued fraction $\sqrt{52/3} = [4, (6, 8), (6, 8), \dots]$.

k	-1	0	1	2	3	4
a_k		4	6	8	6	8
p_k	0	1	4	25	204	1 249
q_k	1	0	1	6	49	300
f_k	+52	-3	+4	-3	+4	-3
F_k		0	-12	+12	-12	+12
$\tilde{p}_k$			5	29	229	1 453
$\tilde{q}_k$			1	7	55	349
$\tilde{f}_k$			-23	+25	-23	+25

This table shows that *the form $f = -3x^2 + 52y^2$ does not attain the value* $+1$ (the minimal value at the vertices of the boundary of the convex hull, corresponding to the positive f, being equal to 4).

Hence f *is imperfect*. Indeed, the value 9 is not attained, since if it were attained, we would have : $(-3x^2 + 59y^2 = 9) \Rightarrow (y = 3z, -x^2 + 3 \cdot 59x^2 = 3) \Rightarrow (x = 3t, -3t^2 + 59z^2 = 1)$, while the value $f = 1$ is not attained, as we have seen from the continued fraction.

Continued fraction $\sqrt{73/3} = [4, (1, 13, 1, 8), (1, 13, 1, 8), \dots]$.

k	-1	0	1	2	3	4	5	6
a_k		4	1	13	1	8	1	13
p_k	0	1	4	5	69	74	661	735
q_k	1	0	1	1	14	15	134	149
f_k	$+73$	-3	$+25$	-2	$+25$	-3	$+25$	-2
F_k		0	-12	$+13$	-13	$+12$	-12	$+13$
$\tilde{p}_k$			5	9	74	143	735	1 396
$\tilde{q}_k$			1	2	15	29	149	283
$\tilde{f}_k$			-2	$+49$	-3	$+46$	-2	$+49$

This table proves *the imperfectness of the form* $f = -3x^2 + 73y^2$, since this form does not attain the value $+9$, its minimal value at the positive vertices being $+25$ (while $f(1, 0) = -3$).

Continued fraction $\sqrt{97/3} = [5, (1, 2, 5, 1, 2, 10), (1, 2, \dots), \dots]$.

k	-1	0	1	2	3	4	5	6	7
a_k		5	1	2	5	2	1	10	1
p_k	0	1	5	6	17	91	199	290	3 099
q_k	1	0	1	1	3	16	35	51	545
f_k	$+97$	-3	$+22$	-11	$+6$	-11	$+22$	-3	$+22$
F_k			-15	$+7$	-15	$+15$	-7	$+15$	-15
$\tilde{p}_k$			6	11	23	108	290	489	3 389
$\tilde{q}_k$			1	2	4	19	51	86	596
$\tilde{f}_k$			-11	$+25$	-35	$+25$	-3	$+49$	-11

It follows that *the form* $f = -3x^2 + 97y^2$ *is imperfect*, since it does not attain the value $+9$. This unattainability can be seen from the easy calculations of the values of f along the segments of the boundary of the convex hull (between the vertices v_1 et v_3 and between the vertices v_3 and v_5), but one can also immediately observe that the value $+1$ is not attained, while if 9 were attainable, one would have : $(-3x^2 + 73y^2 = +9) \Rightarrow (y = 3z, -x^2 + 3 \cdot 73z^2 = +3) \Rightarrow$

($x = 3t$, $-3t^2 + 73z^2 = +1$), and so the value $f = +1$ would be attained (contradicting the inequalities $+6 > +1$, $+22 > +1$, making $+1$ unattainable).

Theorem 12 is now proved, since the series, for which we had proved either the perfectness or the imperfectness of the form $-3x^2 + ny^2$, do cover all the interval $0 < n < 100$.

Remark. Putting together the information on the perfect forms $f = mx^2 + ny^2$ for $|m| \le 16$, $|n| \le 16$, we get the following table of the small perfect diagonal forms:

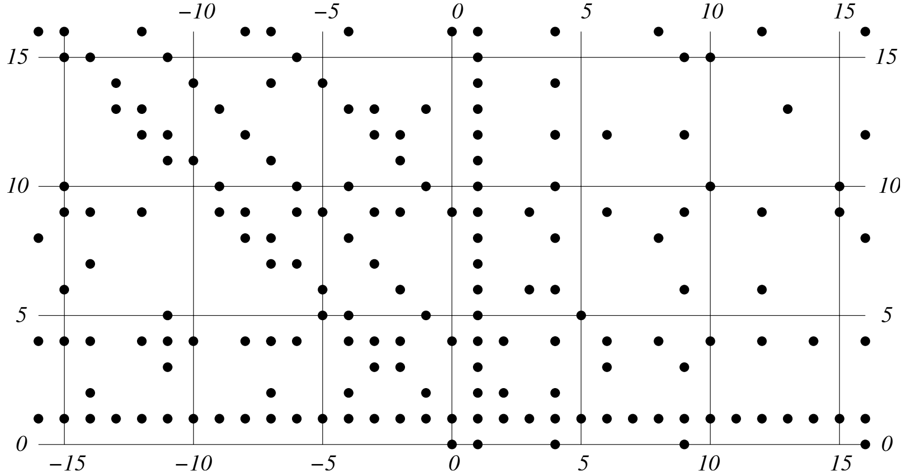

Figure 1: Table of perfect forms.

There are 561 forms in the table, and the statistics of the table is ($-16 \le m \le 16,\ 0 \le n \le 16$) :

forms type	elliptic	hyperbolic	parabolic	all
total number	256	256	49	561
number of perfect forms	71	93	9	173

Therefore, for all the forms in the square $|m| \le 16$, $|n| \le 16$, the statistics is

forms type	positive	hyperbolic	negative	parabolic	total
number	256	512	256	65	1 089
perfect forms	71	186	0	9	266

To compute this table, I had used all the methods, explained above. These particular series studies, having no restrictions on the coefficients (m, n) smallness, might be useful also for other purposes, so I list here a small part of the series studies results.

Series $aux^2 + ny^2$. To be perfect, such a form should represent the number $n^2 \pmod u$. The quadratic residues show that in the cases, mentioned in the following table, it is impossible, and hence *the following series contain no perfect form* (in the table r means the residue of $n \mod u$) :

u	3	4	5	5	6	6	7	7	7	8	8	8	8	8
r	2	3	2	3	2	5	3	5	6	2	3	5	6	7

Series $apx^2 + aqy^2$. To be perfect, such a form should represent the number a^2p^2, so one should have the representation

$$px^2 + qy^2 = ap^2.$$

Suppose, that p is a prime number and that q is not divisible by it. Then we get $y = pz$, and get the representations

$$x^2 + pqz^2 = ap\ , \qquad x = pt\ , \qquad pt^2 + qz^2 = a\ .$$

If this is impossible (say, due to the quadratic residues, or, in the case of a definite form, to the checking of the small finite list of candidates (t, z)), then *the initial form were not perfect*.

Some of the many examples of this kind are listed in the following *table of the pairs* $(m = ap, n = aq)$*, defining imperfect forms*:

m	-4	-4	-4	6	6	6	6	6	6	-6	-6
n	22	24	26	-21	-9	-8	10	15	18	21	22

m	5	-5	7	7	-9	-9	10	12	13	14	21
n	-10	10	-21	14	12	15	26	22	26	22	18

Among the strange series of the perfect forms, discovered due to the computation of the small forms table, I have no theory, explaining the perfectness of the following series, $f = 4x^2 - ny^2$, $n = 4s + 3$, *whose forms attain the value* 1 *at the following strange places*:

n	3	7	11	15	19	23	27	31	35
x	1	4	5	2	85	12	13	760	3
y	1	3	3	1	39	5	5	273	1

I do not know, how to continue this series, but $n = 39$ should not be included. An easy case is *the perfect forms sub-series* $n = 4a^2 - 1$: in this case $f(x, 1) = 1$, and the perfectness follows from Corollary 1 of Theorem 1.

Perfect forms series $f = (2a + 1)x^2 - (2a - 3)y^2$. Its perfectness follows from Corollary 1 of Theorem 1, since $f(a - 1, a) = 1$.

Perfect forms series $f = (3a + 9)x^2 - 3ay^2$. Its perfectness follows from Corollary 2 of Theorem 1, since $f = 3[(a+3)x^2 - ay^2]$, while $(a+3)x^2 - ay^2 = 3$ at the point $(x = 1, y = 1)$.

Series $f = 4x^2 - a^2y^2$, *where a is odd, contains no perfect forms*. Indeed, a^4 is not representable by f, since the representation

$$4x^2 - a^2y^2 = a^4 \qquad 4x^2 = a^2(a^2 + y^2) \ ,$$

would imply, that y is odd, that $a^2 = 1 \pmod 4$, $y^2 = 1 \pmod 4$, $4x^2 = 2 \pmod 4$, which is impossible. Hence, the forms $4x^2 - 9y^2$ and $4x^2 - 25y^2$ are imperfect.

Perfect forms series $f = 4a(x^2 - y^2)$. Such a form is always perfect, accordingly to Corollary 2 of Theorem 1, since $x^2 - y^2 = 4a$ at the point $(x = a + 1, y = a - 1)$.

Perfect forms series $f = a^2x^2 + aby^2$. Such a form is always perfect, accordingly to Corollary 2 of Theorem 1, since $ax^2 + by^2 = a$ at the point $(x = 1, y = 0)$. For instance, *every form* $4x^2 + ny^2$ *is perfect, if n is even.*

4 De Sitter relativistic world and statistics of quadratic forms classes

The *Klein model* of the Lobachevsky plane inside the unit disc might be extended outside the disc, providing a pseudoriemannian structure of Lorentzian signature, which I shall call “de Sitter 2-world”. A similar construction might be used in n dimensions. Its definition is described below (in a more general situation of hyperbolic geometry).

A projective algebraic hypersurface of degree d in $\mathbb{R}P^n$ is called *hyperbolic (with respect to a “time-like” point)*, if every real straight line, containing this point, intersects the hypersurface at d real points (counted with their multiplicities). In the smooth hypersurface case no multiplicities are needed, and the hypersurface is called *strictly hyperbolic*.

A useful example is provided (*in the space* $\mathbb{R}^N$, $N = n(n+1)/2$) *by the cone of the degenerate forms*, which is hyperbolic with respect to any positive-definite direction (since the eigenvalues of symmetric matrices are real). Similar hyperbolic hypersurfaces are provided by the cones of the degenerate Hermitian (or HyperHermitian) complex (or quaternionic) matrices.

Starting from a homogeneous function, defining the hyperbolic hypersurface as the set where the function value is zero (for instance, from the determinant of the quadratic forms), one might imitate the Klein – de Sitter world construction (explained below), defining interesting pseudoriemannian structures on the hypersurface of constant determinant forms (or on a non zero level hypersurface of the cone defining homogeneous function). These structures might be interpreted also as pseudoriemannian structures (of different signatures) on the components of the complement to the hyperbolic hypersurface in the projective space.

These pseudoriemanian structures signatures, for the quadratic forms space case, may be calculated (due to the transitivity of the $\mathrm{SL}(n, \mathbb{R})$ group action on the space of the forms of fixed signature in $\mathbb{R}^n$), using the following fact.

The signature of the second quadratic form of the {zero determinant quadratic forms} cone hypersurface at its smooth point is equal to the signature of the very quadratic form, which is this point. (This fact is "Theorem 4" in the paper: V. I. Arnold, *Ramified covering* $\mathbb{C}P^2 \to S^4$*, hyperbolicity and projective topology*, Siberian Mathematical Journal, 1988, vol. **29**, N. 5, pp. 36-47).

These natural "de Sitter type symmetric manifolds", generalizing the E. Cartan symmetric riemannian manifolds of positive definite forms, had not been studied (neither by the geometers, nor by the relativists), and I shall discuss below the case of the real binary forms ($n = 2$, $N = 3$).

The degenerate forms in $\{mx^2 + ny^2 + kxy\}$ form the quadratic cone $\{D = 0\}$ in this 3-space, where the determinant is $D = 4mn - k^2$ (the matrix of the form is multiplied by 2 to avoid the fractional coefficients).

The *pseudoriemannian metric of the generalized Klein model* is constructed using a homogeneous polynomial D (we shall only consider the above determinant of the binary quadratic forms, but the same construction works in many other cases).

Consider the hypersurface $\{D = 1\}$ in the vector space $\mathbb{R}^N$, where the cone $\{D = 0\}$ lives. The *Klein type metric* on this hypersurface is provided by the following construction. Consider a neighboring homotetical hypersurface, which is defined by the slightly deformed equation, $\{D = 1 + \epsilon\}$. The hyperplane, tangent to the initial hypersurface, $\{D = 1\}$, is intersected by the deformed hypersurface along a "quadric", defined by the equation $g(\xi) = \epsilon$, where g is a quadratic form of the tangent vector ξ. The field of quadratic forms, defined along the hypersurface $\{D = 1\}$ by this construction (sometimes with a minus sign, if one wishes obtain a genuine riemannian metric) is the pseudoriemannian metric on the hypersurface $\{D = 1\}$, which I propose to call the "generalized Klein model".

The resulting pseudoriemannian metric is evidently invariant under the action of the group of those linear transformations of the ambient vector space $\mathbb{R}^N$ of the cone, which preserve the cone (and the function D), since nothing else had been used in the above coordinates-free definition. To reduce the dependence of the metric on the scale (that is, on the choice of D), one usually projects the hypersurface $\{D = 1\}$ to the projective space $\mathbb{R}P^{N-1}$ (whose points are the rays, connecting the nonzero points of our vector space $\mathbb{R}^N$ to 0).

In the case of the forms determinant, $D = 4mn - k^2$, considered above, the two-sheeted hyperboloid $\{D = 1\}$ is projected onto the interior disc of the "absolute" circle (to which is projected the degenerate forms cone). This identification of the interior part of the disc with the hyperboloid of the binary quadratic forms of determinant $D = 1$ (or, to be exact, with any of the two connected components of this hyperboloid – say, of that formed by the positive definite forms) may be extended to the case of any other fixed nonzero value of the determinant.

The binary forms of fixed negative determinant (which are hyperbolic, having the $(+, -)$ signature) are projected by the rays from the origin to the *exterior* domain of the disc in the projective plane. This exterior domain is topologically the Möbius band. The surface of the binary forms of a fixed negative determinant is a one-sheeted hyperboloid. The projection by the rays, sending it to the projective plane, is a two-fold covering of the Möbius band by the hyperboloid surface, which is diffeomorphic to the cylinder. Thus, *the exterior domain of the disc of the Klein model of the Lobachevsky plane can be considered as the manifold of the hyperbolic binary quadratic forms of a fixed determinant* (the forms, to be strict, should be considered up to the sign, since every image point of the projection represents two opposite preimage points on the hyperboloid).

Unlike the interior disk, where the pseudoriemannian metric defined above is the Lobachevsky riemannian metric of the usual Klein model, in the exterior domain it is Lorentzian, that is, it has signature $(+, -)$ and has two "light directions" at each point. The light directions on the hyperboloid are the directions of its generating straight lines, and on the projective plane exterior domain of the Lobachevsky disc they are the directions of the two tangents to the absolute circle, bounding the disc.

The projective transformations of the projective plane, preserving the disc, act both inside it (as the Lobachevsky plane isometries) and outside it (as the de Sitter world isometries). The de Sitter 2-world is naturally projectively equivalent to the space of the projective lines in the Lobachevsky disc, which space is diffeomorphic to the Lobachevsky disc of the Klein model, where one point (the center) is resolved by a sigma process (replacing it by a real projective line, which is the infinity line in the standard affine plane Klein model).

The group $\mathrm{SL}(2, \mathbb{R})$ of the area-preserving linear transformations of the (x, y)-plane, on which our binary quadratic forms are defined (as well as its modular subgroup $\mathrm{SL}(2, \mathbb{Z})$, preserving the integer points sublattice $\mathbb{Z}^2$) acts on the space $\mathbb{R}^3$ of the quadratic forms $f(x, y)$ (as the linear transformations), preserving the cone of the degenerate forms $\{D = 0\}$ and even preserving the determinant function D and the hyperboloids $\{D = c\}$ in $\mathbb{R}^3$.

The *modular group* is the reflection group of the Lobachevsky infinite *modular triangle*, represented in the Klein model by an ordinary Euclidean equilateral triangle, inscribed into the “absolute” circle. It acts also on the Klein model with one point resolved, which is the de Sitter world.

The geodesics of the (pseudo)riemannian metric of the model are simply the usual projective lines in the projective plane (both inside the Lobachevsky disc and outside it), since the hyperboloid has symmetries, preserving the planes, containing the origin.

Therefore, *the study of the classification of the hyperbolic forms* $mx^2 + ny^2 + kxy$ *of a fixed determinant means geometrically the investigation of the orbits of the points of the de Sitter world under the action of the symmetries of the modular triangle, described above*. This study might be done by the usual geometrical methods of the discrete groups theory, starting from the fundamental domains in the de Sitter world (similar to the covering of the Lobachevsky plane by the images of the modular triangle, reflected many times in its sides straight lines).

I shall report below the results on the invariants of integer coefficients binary quadratic forms of fixed determinant and on these forms orbits under the action of the modular group on the de Sitter world. However I shall not represent the initial geometric reasonings which had lead me to these results, to avoid the drawings of the too numerous Klein-Fricket type diagrams.

One might eliminate the de Sitter world point of view, representing the exterior points in the Klein model by the polar dual interior straight lines of the usual Lobachevsky plane. In the projective plane, containing a circle, the line, polar to an exterior point, joints the two tangency points of the circle with the two straight lines, containing this exterior point. The Lobachevsky line might be characterized by its point, closest to the origin (except the absolute diameters line) – this is the diffeomorphism interpretating the de Sitter world as the one point resolution of the Lobachevsky plane, quoted above.

The results, described below, might be interpreted as statements on the action of the group, generated by the reflections in the sides of the modular triangle, on Lobachevsky lines (representing the integer coefficients quadratic forms of a given negative value of the determinant).

Some invariants of this $\mathrm{SL}(2,\mathbb{Z})$-classification were already introduced above : a perfect form remains perfect, whatever $\mathrm{SL}(2,\mathbb{Z})$-frame in the (x, y)-plane is used to define the coefficients (m, n, k).The set of the values of the form is itself an invariant, be it a semigroup or not. One might distinguish in this set the subset of those values, which are attained at the indivisible points (x and y having no nontrivial common divisor). And the geometry of the continued fraction (that is of the boundaries of the convex hulls of the sets of the integer points, separated by the asymptotes $f = 0$ of the form f) is an $\mathrm{SL}(2,\mathbb{Z})$-invariant characteristics of the form. One might also consider as the invariants of the form the relation of its values set with different prime numbers (similar to the classical description of the $x^2 + y^2$ values) : such descriptions are missing even for the perfect forms, discussed above.

Theorem 13. *The set of the* $\mathrm{SL}(2,\mathbb{Z})$*-equivalence classes of the binary forms* $mx^2 + ny^2 + kxy$ *with integer coefficients* m, n, k *having a fixed negative value of the determinant* $D = 4mn - k^2$*, is finite. Moreover, the number of the classes is smaller than*

$$\frac{\pi}{2}|D| + 4\sqrt{|D|}\ .$$

Proof. We shall distinguish the "irrational" forms, whose lines $f = 0$ do contain no nontrivial integer points (different from zero), and the "rational" ones, where there is such a nonzero point. In the second case integer points do exist on both straight lines $f = 0$, since the quadratic equation, defining these lines inclination, has integer coefficients, and hence if one of its roots is rational, the other is rational too, accordingly to the Vieta formula.

Lemma. *The number of the* $\mathrm{SL}(2,\mathbb{Z})$*-classes of binary irrational forms of negative determinant* D *with integer coefficients does not exceed the number of the integer points in the ellipse,* $k^2 + 4r^2 \leq |D|$ *(where the points on the axis*

$r = 0$ *are not counted and the others points on the boundary curve are counted with multiplicity* $1/2$ *each).*

Proof. For an irrational form one can choose such a frame of two integer vectors, that the value of the form on the first of them is positive, on the second negative, while the oriented area of the parallelogram, defined by these vectors, is equal to 1 (this is impossible for some "rational" forms, like $f = xy$, for instance).

To choose the frame in the irrational case, it suffices to apply the continued fraction algorithm: two consecutive vectors (v_k, v_{k+1}) have the required property.

This reasoning is the only place where we have to use the irrationality. In the coordinates, defined by the frame, chosen above, two of the 3 form's coordinates will have fixed signs, $m > 0$, $n < 0$. Let a, b the natural numbers $|m| = a$, $|n| = b$. The ordered pair (a, b) defines, together with the coefficient k, the "normal form" of the class of the quadratic form, $ax^2 - by^2 + kxy$, verifying the determinant condition $4ab + k^2 = |D|$.

We shall now count the number of the solutions of this equation. Observe first, that the equality $ab = r^2$ implies, that either $1 \le a \le r$, or $1 \le b < r$, since otherwise $(a > r, b \ge r) \Rightarrow (ab > r^2)$.

Therefore, the number of the natural solutions (a, b) of the equation $ab = r^2$ does not exceed the sum of the numbers of the solutions of the two inequalities $1 \le a \le r$ and $1 \le b < r$.

The number of solutions of the first one is the integer part $[r]$ of r, while the number $]r[$ of solutions of the second inequality is equal to $[r]$ if r is not an integer and to $[r] - 1$ if it is a positive integer.

Thus, the number of the solutions (a, b) of the equation $ab = r^2$ does not exceed the sum $[r]+]r[$.

On the other side, the number of the integer points, u, on the segment $-r \le u \le r$ is equal to the same sum (provided, that $u = 0$ is not counted and that $u = +r$ and $u = -r$ are counted together as 1 point).

Applying this result to the equation $4ab = |D| - k^2$, where k is fixed, we see that the number of its solutions does not exceed the number of the integer points in the interval $-r \le u \le r$, where r is defined by the ellipse equation, $4r^2 + k^2 = |D|$. As in the preceding reasoning, the point $u = 0$ is not counted and the boundary points $u = r$ and $u = -r$ are counted together as one point (in the case where r is a positive integer).

The lemma is thus proved. To prove the theorem, we evaluate the number of the integer points (r, k) in the above ellipse, for which $k \ne 0$, associating to any such point a unit coordinate square domain, of which the point is a vertex, the square being directed from this vertex toward the origin.

Since the open squares do not intersect and lie inside the ellipse, their number does not exceed the area inside the ellipse, which is $\pi\sqrt{|D|}(\sqrt{|D|}/2)$ since the axes are $k^2 \leq |D|$, $4r^2 \leq |D|$. The number of the integer points, for which $k = 0$, does not exceed $2[\sqrt{|D|}/2]$. Thus, the total number of the classes of the irrational forms of negative determinant D does not exceed $(\pi/2)|D| + 2\sqrt{|D|}$.

To count the rational forms classes, choose the closest integer point on the line $f = 0$ as the first basic vector. In such a coordinate system, the form will be $f(x, y) = ny^2 + kxy$. The determinant being $D = -k^2$, we have only to normalize the coefficient n. To do this, it suffices to add y to the coordinate x. This $\mathrm{SL}(2, \mathbb{Z})$-change of coordinates shifts the coefficient n by k. It follows that the residue of $n \mod k$ is an invariant of the form, and that any such form is equivalent to one, for which $1 \leq n \leq |k|$. Hence, the number of classes does not exceed $2|k| = 2\sqrt{|D|}$ (one might observe that, essentially, we had classified the integer lattice parallelograms of a given area k).

Adding the number of the rational forms classes to the previous upper bound of the number of the irrational forms, we obtain the upper bound claimed by Theorem 13.

Remark. This bound is larger than the actual number of the classes, and I shall provide below some more realistic arguments, suggesting the smaller growth rate of the number of classes of forms of a fixed determinant D (at least one hope it should be correct for the total number of classes of those forms which determinants lie between 0 and $-|D|$).

To evaluate the number of the natural solutions of the equation $ab = c$ (and of the inequality $ab \leq c$), one may use the hyperbola $b = c/a$ and the area below it (taking into account the restriction $a \leq c$, b being natural). Adding the rectangle $(0 \leq a \leq 1, 0 \leq b \leq c)$ to take into account the points with $a = 1$, while counting the unit squares, issued from the integer points below the hyperbola to the direction of smaller coordinates, we get for the number N of the integer points below the hyperbola the "entropie type" formula

$$N \sim c + c \ln c\ .$$

It suggests that the number $M(c)$ of points on one hyperbola should behave like $\frac{dN}{dc} = 2 + \ln c$ (neglecting the fluctuations, making some hyperbolae, like $c = n!$, more populated then the others like those, corresponding to prime c, where $M = 2$).

In this sense we shall call $2 + \ln c$ the "averaged upper bound".

Example. For $c = 24$ the genuine number $N = 84$ of the integer points below the hyperbola is smaller than the upper bound $25 + 25 \ln 25 \simeq 105$, provided by our reasoning. However, the number $M = 8$ of the integer points exactly on our hyperbola is larger, than our "averaged upper bound", which is $2 + \ln 24 \simeq 5, 2$.

In all these upper bounds reasonings, we have not taken into account the arithmetical restriction on the value of $4ab = |D| - k^2$: this integer should be divisible by 4. To eliminate 3/4 of the hyperbolae, we first write the total number of classes as the sum of the numbers of points on the particular hyperbolae,

$$\sum_{k, k^2 \leq |D|} M(C_k) , \qquad \text{where } 4C_k = |D| - k^2 .$$

Replacing next the sum by the integral and the actual integers M by the above averaged upper bound, we shall then divide the integrand by 4, to take the arithmetical restriction into account, and we shall multiply the positive k integral by 2, to take into account the negative values of k as well. The resulting "euristical" integral formula for the number $A(D)$ of the SL(2, $\mathbb{Z}$)-classes of binary quadratic forms with integral coefficients of fixed negative determinant D is:

$$A(D) \text{ "} \sim \text{" } \int_{0 \leq k \leq \sqrt{|D|}} \frac{2 + \ln C}{2} \, dk , \qquad \text{where } 4C = |D| - k^2.$$

Taking into account the relations $4\, dC = -2k\, dk$, $dk = -2dC/k$, we rewrite this integral in the independent variable C form,

$$A(D) \text{ "} \sim \text{" } \int_1^{|D|/4} \frac{2 + \ln C}{\sqrt{|D| - 4C}} \, dC .$$

The main contribution to this convergent integral is provided by the right boundary point, $C = |D|/4$. It is of the order of magnitude, provided by the formula

$$A(D) \text{ "} \sim \text{" } \big(2 + \ln(|D|/4)\big)\sqrt{|D|}/2 .$$

This "euristical upper bound" is much smaller, than the bound of the order $|D|$, rigorously proved in Theorem 13. The euristical formula asymptotics is closer to the actual numbers observed in the examples than the rigorous bound, proved in Theorem 13. The actual numbers of hyperbolic forms classes of mild determinants $-|D|$ are, accordingly to the numeric calculations by F. Aicardi, those, presented in the following table:

$\lvert D\rvert$	1	4	5	8	9	12	13	16	17	20	21	24	25	28	29
A	1	2	1	1	3	2	1	4	1	2	2	2	5	2	1

$\|D\|$	32	33	36	37	40	41	44	45	48	49	52	53	56	57
A	3	2	6	1	2	1	2	3	4	7	2	1	2	2

$\|D\|$	60	61	64	65	68	69	72	73	76	77	80	81	84	85
A	4	1	8	2	2	2	3	1	2	2	4	9	2	2

$\|D\|$	88	89	92	93	96	97	100	101	104	105
A	2	1	2	2	6	1	10	1	2	4

This table suggests for the total number of the hyperbolic forms classes of determinants between 0 and $-|D|$ the approximate value $1.4\cdot|D|$ (for $|D| \leq 100$).

The symbol "$\sim$" means the absence of a rigorous proofs of the corresponding relations : they represent rather the *averaged* asymptotics, providing more reliable estimations for the sums $\sum_{d\leq|D|} A(d)$, than for the values of the numbers $A(d)$ themselves, which might fluctuate around the mean values estimated below.

For the positive definite forms the numbers A of the classes of forms corresponding to a given value of the determinant $D = 4mn - k^2$ is, accordingly to F. Aicardi, for $0 < D \leq 104$, given by the table

$\|D\|$	3	4	7	8	11	12	15	16	19	20	23	24	27	28
A	1	1	1	1	1	2	2	2	1	2	3	2	2	2

$\|D\|$	31	32	35	36	39	40	43	44	47	48	51	52	55	56
A	3	3	2	3	4	2	1	4	5	4	2	2	4	4

$\|D\|$	59	60	63	64	67	68	71	72	75	76	79	80	83	84
A	3	4	5	4	1	4	7	3	3	4	5	6	3	4

$\|D\|$	87	88	91	92	95	96	99	100	103	104
A	6	2	2	6	8	6	3	3	5	6

Denote by $\mu(D)$ the number of the classes of determinant between 0 and D:

$$\mu(D) = \sum A(d), \quad 0 < d \leq D.$$

The table provides the approximate formula $\mu(D) \approx 7D/4$ for $D \leq 104$. Aicardi suggested the empirical growth rate $\mu(D) \approx 0.16 D^{1.5}$ (the theoretical "euristical" arguments below provide for the upper bound the suggestion of order $D^{7/4}$).

For the hyperbolic case (where $D < 0, \ -|D| \leq d < 0$) the empirical growth rate suggestion is

$$\mu(D) \approx 0.48\,|D|^{1.23},$$

while the theoretical "euristical" arguments below provide the upper bound of the order of $D^{3/2}$.

The difference between the theoretical and the empirical growth rates might be explained by the fact, that the theoretical bound is an attempt to evaluate the *maximal* fluctuations, while in the empirical study only the *sum* over d is evaluated.

The fluctuations sum might be smaller, than the sum of the maximal fluctuations, the maximum being attained only for a small part of the values of d.

Theorem 14. *The number of the* $\mathrm{SL}(2,\mathbb{Z})$*-equivalence classes of the positive definite forms* $f = mx^2 + ny^2 + kxy$ *with integer coefficients* (m, n, k)*, having a fixed positive value of the determinant* $D = 4mn - k^2$*, is smaller or equal to the number* $8D/\pi^2$*.*

We start from a geometric observation (due to Minkowsky):

Lemma. *There exists an indivisible integer point at which the value* $\tilde{A}$ *of the form is of order of* $\sqrt{D}$*:*

$$0 < \tilde{A} < 2\sqrt{D}/\pi \ .$$

Proof. The area of the ellipse $f \leq C$ equals $C(2\pi/\sqrt{D})$. If this area is greater or equal to 1, the ellipse intersects its version translated parallely to some nonzero integer center point. Hence the doubled ellipse, $f \leq 4C$, contains in this case a nonzero integer point.

For $C = \sqrt{D}/(2\pi)$ the area of the ellipse is 1, and hence we find a nonzero integer point, where the form value is at most equal to $4C = 2\sqrt{D}/\pi$, as required in the Lemma. To make this point indivisible it suffices to divide it.

Proof of Theorem 14. Choosing an $\mathrm{SL}(2,\mathbb{Z})$-coordinate system (X, Y), we make $(X = 1,\ Y = 0)$ the coordinates of the point of the Lemma. The form and the determinant have now the expressions

$$f = AX^2 + NY^2 + KXY\ , \quad D = 4AN - K^2\ .$$

When D and A are fixed, the last equation defines a parabola in the plane of (N, K).

The choice of new coordinates $(\tilde{X}, Y)$, where $X = \tilde{X} + pY$, moves the point (N, K) along this parabola, replacing the coefficient K by its new value, $\tilde{K} = K + 2Ap$. Choosing the integer p, we attain the parabola segment where $0 \leq \tilde{K} < 2A$.

This segment contains at most $2A$ integer points. We have proved that one of these points does represent the $\mathrm{SL}(2,\mathbb{Z})$-equivalence class of the form f. Hence *the number of the equivalence classes of forms of determinant D is at most equal to the product of the numbers of the possible values of the integers A and* $\tilde{K}$, that is to $2A^2 \leq 8D/\pi^2$, which proves the Theorem.

Remark. The upper bound, provided by Theorem 14, seems to be higher than the genuine asymptotic of the number of the classes, when D is large.

Indeed, the number of the integer points on the $0 \leq K < 2A$ segment of the parabola $\{N = (D + K^2)/(4A)\}$ seems to grow with A rather as $\sqrt{A}$ than as $2A$, which would provide for the number of the classes an upper bound of the order of $A^{3/2}$ (that is of $D^{3/4}$ instead of D).

To explain this growing rate of the number of the integer points on the parabola segment, consider the case $D = 0$. In this case the number $4NA$ on the parabola should be a square of an integer. Denote $A = Q^2R$, where the integer R has no squares of primes among its divisors. Then one should have $N = RS^2$, for some integer S. The bound $K < 2A$ implies the inequalities $N < A$, $S < \sqrt{A}$, and hence the number of the integer points on the segment of the parabola, for which $D = 0$, is at most $\sqrt{A}$.

Acknowledgments. I would like to thank A. Ortiz-Rodríguez and G. Capitanio who were kind enough to type this text for me.

References

[1] V.I. Arnold, *Euler groups and arithmetics of geometric progression*, Moscow, MCCME, (2003), 40pp.

[2] V.I. Arnold, *Fermat-Euler dynamical system and statistics of the arithmetics of geometrical progressions*, Funct. Anal. and its Appl., **37** (2003), N1, pp.1–20.

[3] V.I. Arnold, *Ergodic arithmetical properties of the dynamics of geometrical progressions*, Moscow Math. Journal, (2003).

[4] V.I. Arnold, *Topology of algebra: the combinatorics of the squaring operation.* Funct. Anal. and its Appl., **37** (2003), N2, pp.1–24.

[5] V.I. Arnold, *Arithmetics of binary quadratic forms, symmetry of their continued fractions and geometry of their de Sitter world*, Moscow, Dubna, MCCME, 2002, pp.1-40 (Bull. of Braz. Math. Soc. Vol. **34** No. 1, (2003), p.1–41).

V. Arnold
Autumn address:
Steklov Mathematical Institute
Gubkina str. 8, Moscow V-333, GSP-1, 117966
RUSSIA

Spring address:
CEREMADE, Universite Paris 9 - Dauphine
Place du Marechal de Lattre de Tassigny
75775 PARIS Cedex 16-e,
FRANCE