Fermat Dynamics of Matrices, Finite Circles and Finite Lobachevsky Planes

V.I. ARNOLD *

Introduction

The unusual version of Lobachevsky geometry that we shall use for number theory, had been first introduced in 1965 in my Paris lectures [1] on dynamical systems.

The Fermat dynamical system analysis (studied from the modern point of view in the recent article [2]) investigates the geometric progressions of the powers of a residue modulo n , $\{a^t\}$, $t = 1, 2, 3, \dots$, including their turbulent chaoticity and pseudo-randomness.

The present article describes some matrix versions of the small Fermat theorem, claiming the periodicity of the residues of geometric progressions.

Starting from the article [2], we relate the studies of these geometric progressions to the topology of some graphs and trees, describing the combinatorics of the squaring operation of matrices or residues.

In the present article, continuing these extensions of the small Fermat theorem and of the primitive residues theory, we shall consider the deep relations of these subjects to the Lobachevsky and to the relativistic de Sitter geometries (see [3]) over the finite fields, involved in the study of the equations $A^m = 1$ in the finite groups of matrices (whose solutions form finite Lobachevsky and finite de Sitter planes).

The relations of these geometries and of the corresponding Riemann surfaces to the Kepler's cubes inscribed into the dodecahedron, had been started to be explored in [2]. The present article contains more number-theoretic applications of these ideas, the positivity condition, defining upper half-plane of the Lobachevsky geometry model, being replaced by the quadraticity condition for the residues.

I would like to thank R. Uribe-Vargas who was kind enough to edit this article from two manuscripts that I give to him.

*Partially supported by RFBR, grant 02-01-00655.

1 Dynamics of traces

Let A be a matrix of order r with integral elements and let $\text{tr}A$ be its trace. Fix a prime number p .

Studying the traces of the powers of this matrix and their residues modulo p , we get an extension of the small Fermat theorem.

Theorem 1. *The following congruence modulo p holds:*

$$\text{tr}(A^p) \equiv (\text{tr}A)^p.$$

Moreover, the difference of the left and the right hand sides is equal to an integer coefficient polynomial of the coefficients of the characteristic polynomial of A , the integral coefficients being divisible by p :

$$(\text{tr}A)^p - \text{tr}(A^p) = pF_p(\sigma_1, \dots, \sigma_r), \quad F_p \in \mathbb{Z}[\sigma],$$

where $\sigma = (\sigma_1, \dots, \sigma_r)$, σ_k denoting the basic symmetric function of degree k of the roots of the characteristic polynomial (coinciding, up to signs, to its coefficients):

$$\sigma_1 = \text{tr}A = \sum_j \lambda_j, \quad \sigma_2 = \sum_{j < m} \lambda_j \lambda_m, \quad \dots, \quad \sigma_r = \det A = \prod_j \lambda_j.$$

Theorem 1 is essentially equivalent to the following polynomial identity:

Theorem 1'. *The following identity holds*

$$(\lambda_1 + \dots + \lambda_r)^p - (\lambda_1^p + \dots + \lambda_r^p) = pF_p(\sigma(\lambda)), \quad F_p \in \mathbb{Z}[\sigma_1, \dots, \sigma_p],$$

the weighted homogeneous Newton-Girard polynomial pF_p (of degree p for weights $\deg \sigma_k = k$) being independent on the number r of the arguments λ_j .

Example 1. The identity $(\lambda_1 + \lambda_2)^3 - (\lambda_1^3 + \lambda_2^3) = 3(\lambda_1 + \lambda_2)(\lambda_1 \lambda_2)$.

Our theorems extend this identity $s_3 = \sigma_1^3 - 3\sigma_1\sigma_2$ (corresponding to $r = 2$, $p = 3$) and some similar divisibility theorems of the elements of Pacal's triangle to the cases of higher values of r and of p (the "Newton symmetric function" s_n is defined as the sum of powers $\lambda_1^n + \dots + \lambda_r^n$).

Example 2. For the identity matrix (or for the values $\lambda_j \equiv 1$) our theorems have the form of the congruence

$$r^p \equiv r \pmod{p}$$

of the small Fermat theorem. In this sense *our Theorem 1 is a matricial version of the small Fermat theorem.*

Of course, the divisibility by p of the value of the difference described in Theorem 1', at any integer point λ follows directly from the small Fermat theorem. However, the divisibility of the *values* of a polynomial by p does not imply the divisibility by p of its *coefficients* (for instance, the values of the polynomial $x^2 + x$ at integer points are even, its coefficients being odd). Our Theorem 1' shows that the image of the set $\mathbb{Z}^r$ under the Vieta mapping $\sigma : \mathbb{C}^r \rightarrow \mathbb{C}^r$ (sending λ to $\sigma(\lambda)$) has the special property that the vanishing of the values (mod p) of our polynomial at the points $\sigma(\lambda)$, $\lambda \in \mathbb{Z}^r$, of this image implies the vanishing of the values (mod p) at all the points $\sigma \in \mathbb{Z}^r$. I do not know whether this property of the values of our polynomial on the image $\sigma(\mathbb{Z}^r)$ holds for the values of other polynomials, neither which properties of the image of the Vieta mapping are important for the divisibility propagation that holds in our case.

Note that our theorems describe extremely simple and classical identities.

Example 3. The explicit formulae for the Newton-Girard ([4], [5]) polynomials F_p of Theorems 1 and 1' are strangely simple:

$$\begin{aligned} (\lambda_1 + \dots + \lambda_r)^2 - (\lambda_1^2 + \dots + \lambda_r^2) &= 2\sigma_2(\lambda), \\ (\lambda_1 + \dots + \lambda_r)^3 - (\lambda_1^3 + \dots + \lambda_r^3) &= 3(\sigma_1\sigma_2 - \sigma_3), \\ (\lambda_1 + \dots + \lambda_r)^5 - (\lambda_1^5 + \dots + \lambda_r^5) &= 5(\sigma_1^3\sigma_2 - \sigma_1^2\sigma_3 + \sigma_1\sigma_4 - \sigma_1\sigma_2^2 + \sigma_2\sigma_3 - \sigma_5), \\ (\lambda_1 + \dots + \lambda_r)^7 - (\lambda_1^7 + \dots + \lambda_r^7) &= 7(\sigma_1^5\sigma_2 - \sigma_1^4\sigma_3 + \sigma_1^3\sigma_4 - 2\sigma_1^3\sigma_2^2 - \sigma_1^2\sigma_5 + \\ &+ 3\sigma_1^2\sigma_2\sigma_3 + \sigma_1\sigma_6 - 2\sigma_1\sigma_2\sigma_4 - \sigma_1\sigma_3^2 + \sigma_1\sigma_2^3 - \sigma_7 + \sigma_2\sigma_5 + \sigma_3\sigma_4 - \sigma_2^2\sigma_3). \end{aligned}$$

The general formula for the coefficients of the right side polynomials will be written below.

In these formulae, many of the coefficients of the polynomials F_p are equal to ± 1 , the sum of the polynomial coefficients being 0, and all coefficients being rather small (for $p \leq 7$).

These coefficients may be considered as forming a function of the area p Young diagrams, associated to the partitions of p : $p = \sum a_k$ (these partitions correspond to the monomials $\prod \sigma_{a_k}$). The explicit formula for this function is given below (virtually replacing the prime number p by any natural number n).

The “+” sign of a monomial in the formulae above occurs exactly when the number of even degrees a_k in the monomial is odd (making odd the permutation corresponding to the partition $p = \sum a_k$): this number is 3 for the monomial $+\sigma_1\sigma_2^3$, being 2 for the monomial $-2\sigma_1\sigma_2\sigma_4$.

Proof of Theorems 1 and 1'.

The product of p equal multipliers $(\lambda_1 + \dots + \lambda_r)$ of σ_1^p is the sum of r^p monomials of degree p , each of these monomials being the product of the powers of different arguments,

$$\mu^a = \mu_1^{a_1} \cdots \mu_s^{a_s}, \text{ where } \mu_k = \lambda_{j(k)}, \text{ while } \sum a_k = p, s \leq p, \mu_k \neq \mu_p.$$

The same monomial μ^a is repeated in the sum several times, since, say, the product $\mu_1^{a_1}$ is the contribution of any subset formed by a_1 multipliers μ_1 among the p multipliers of σ_1^p ; the product $\mu_2^{a_2}$ being the contribution of a different subset (disjoint with the first one), and so on. The choices of these disjoint subsets are arbitrary.

Therefore the total number of repetitions of the chosen monomial μ^a in our sum of r^p summands is equal to the product of the numbers of consecutive combinations, counting the consecutive choices of the subsets,

$$N(a) = C_{p_1}^{a_1} C_{p_2}^{a_2} \dots C_{p_s}^{a_s},$$

where $p_1 = a_1 + \dots + a_s = p$, $p_2 = a_2 + \dots + a_s = p - a_1$, $\dots$, $p_s = a_s$.

Indeed we have $C_{p_1}^{a_1}$ possible choices from the first subset of a_1 elements (from the total set of p_1 multipliers), then $C_{p_2}^{a_2}$ possible choices from the second subset of a_2 elements (from the set of p_2 remaining multipliers), and so on.

The binomial coefficient (equal to the number of combinations)

$$C_{p_1}^{a_1} = \frac{p(p-1) \dots (p-a_1+1)}{a_1!}$$

is divisible by the prime number p , since the denominator $a_1!$ is not divisible by p for $a_1 < p$.

Therefore, the product N of the binomial coefficients is divisible by the prime number p .

A monomial $\tilde{\mu}^a$, similar to μ^a , is reducible to μ^a by the replacement of the pairwise different arguments μ_k by some pairwise different arguments $\tilde{\mu}_k$ (chosen among the same $\lambda_j - s$). It is repeated in our sum the same number of times as μ^a (for instance, the multiplicity of the monomial $\lambda_1^2 \lambda_2^2 \lambda_3$ in our sum is equal to that of the monomial $\lambda_1^2 \lambda_4^2 \lambda_2$).

Therefore, all our sum is the sum of some symmetric polynomials of the variables λ , the coefficients $N(a)$ (depending on the area p Young diagram of the partition $p = \sum a_k$) being divisible by p .

Representing the symmetric polynomials (divided by p) in terms of the basic symmetric polynomials σ , we represent our sum in the form described by Theorem 1' (proving at the same time Theorem 1, which follows from the same polynomial identity). $\square$

Remark 1. The binomial coefficient C_m^k , where m and k are relatively prime, is divisible by m ; the binomial coefficient C_{m-1}^{k-1} being in this case divisible by k .

proof. The first statement follows from the identification of the binomial coefficient with the cardinality of the set Y of k -elements subsets X of $\mathbb{Z}_m$. The m translations $x \mapsto x + a$ (where $a = 1, \dots, m$) send X to m different k -elements subsets. Indeed, otherwise there would exist a non-trivial translation, sending X to itself. The period T of this mapping would be a divisor of m (the m -th iteration of the translation being the identity mapping). The period T would also be a divisor of

k (since the k -elements subset X consists of the disjoint orbits of this translation, each orbit containing T elements). The numbers m and k being relatively prime, T should be 1, which is impossible for $k < m$, proving the divisibility of C_m^k by m (the m equally long images of X being different elements of Y).

The second statement follows from the proven one, since

$$C_m^k = \frac{m}{k} C_{m-1}^{k-1}. \quad \square$$

2 Congruences for the binomial coefficients

Working on the matricial extensions of the Euler's version of the small Fermat theorem, I have observed several hundreds of peculiar divisibility properties of the binomial coefficients (and of their multinomial coefficients versions $K[\{m_j\}]$ and $C[\{m_j\}]$, defined and discussed in the next section).

Let p be a prime number, and a and b non-negative integers.

Theorem 2. *The following congruence modulo p^2 holds:*

$$C_{pa}^{pb} - C_a^b = 0 \pmod{p^2}.$$

Thus for $p = 2$, $a = 4$, $b = 2$ our difference is $C_8^4 - C_4^2 = 70 - 6 = 64 = 2^6$, therefore $x(4, 2) = 6$.

Example 4. The power of p dividing the left hand side of the congruence of Theorem 2 is denoted by $x(a, b)$ and it is defined by the equation

$$C_{pa}^{pb} - C_a^b = p^{x(a,b)} z(a, b), \quad \text{where } (p, z) = 1.$$

The 4 tables below (of the type of Pascal's triangle), corresponding to the primes $p = 2, 3, 5$ and 7 , are formed by the values of $x(a, b)$, where the rows start from $a = 2, 3, \dots$, and b is running from $b = 1$ to $b = a - 1$ on each row a (the values $x(a, b) = \infty$ at $b = 0$ and at $b = a$ being not shown).

For $p = 2$ and the rows $a = 2, 3, \dots, 9$, with $0 < b < a$, the numbers $x(a, b)$ are given by

$$\begin{array}{ccccccc} & & & & 2 & & & \\ & & & & 2 & 2 & & \\ & & & 3 & 6 & 3 & & \\ & & 3 & 3 & 3 & 3 & & \\ & 2 & 5 & 3 & 5 & 2 & & \\ & 2 & 2 & 3 & 3 & 2 & 2 & \\ & 4 & 8 & 4 & 9 & 4 & 8 & 4 \\ 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4. \end{array}$$

For $p = 3$, the similar table of the values of $x(a, b)$ (with the rows $a = 2, 3, \dots, 10$) is given by

2
 4 4
 3 3 3
 2 3 3 2
 4 4 5 4 4
 3 3 3 3 3 3
 2 3 3 2 3 3 2
 6 6 7 6 6 7 6 6
 4 4 4 4 4 4 4 4 4.

For $p = 5$, the rows ($a = 2, 3, \dots, 11$) of the values $x(a, b)$ are:

3
 3 3
 3 3 3
 5 5 5 5
 4 4 4 4 4
 3 4 4 4 4 3
 3 3 4 4 4 3 3
 3 3 3 4 4 3 3 3
 5 5 5 5 6 5 5 5 5
 4 4 4 4 4 4 4 4 4 4;

For $p = 7$, the rows ($a = 2, 3, \dots, 15$) of the values $x(a, b)$ are:

3
 3 3
 3 3 3
 3 3 3 3
 3 3 3 3 3
 5 5 5 5 5 5
 4 4 4 4 4 4 4
 3 4 4 4 4 4 4 3
 3 3 4 4 4 4 4 3 3
 3 3 3 4 4 4 4 3 3 3
 3 3 3 3 4 4 4 3 3 3 3
 3 3 3 3 3 4 4 3 3 3 3 3
 5 5 5 5 5 5 6 5 5 5 5 5 5
 4 4 4 4 4 4 4 4 4 4 4 4 4 4.

The fact that these numbers $x(a, b)$ exceed 2 is important for the extensions of Euler's theorem, but I have not proved neither the unboundness, nor the boundness

of $x(a, b)$ (neither fixing the prime number p , nor varying it). Many empirical peculiarities of these numbers remain unproved. Say the values $x(mp^a + 1, b)$ seem to be independent of b , as well as the values $x(mp + 1, b)$ do not depend on b for small m . These b -independent values of x are peculiar functions of a and of m , and I leave to the reader the pleasure of find them, at least empirically, continuing the Fermat experimental way to discover mathematical facts.

Theorem 2 states that $x(a, b) \geq 2$ everywhere, but I do not know whether this function is increasing.

The quantities $x(a, b)$ have many peculiar properties, of which only a small part has been proven (a discussion of related conjectures is presented in [7]).

Lemma 2.1. *If an integer c is relatively prime to p , then the binomial coefficient $C_{p^r a}^c$ is divisible by p^r .*

Proof. One uses the obvious identity

$$C_{p^r a}^c = \frac{p^r a}{c} C_{p^r a - 1}^{c-1}.$$

Since the denominator c is relatively prime to p , the whole product is divisible by p^r . $\square$

Lemma 2.2. *The following congruence modulo p^2 holds:*

$$C_{pa}^{pb} \equiv C_{p(a-1)}^{p(b-1)} + C_{p(a-1)}^{pb}.$$

Proof. By the Pascal triangle description of the binomial coefficients, the number C_{pa}^{pb} is equal to the number of the non-strictly monotonic lattice paths on the integer plane from the point $(0, 0)$ to the point (pb) of the line (pa) . We get therefore the recurrent relation

$$C_{pa}^{pb} = \sum_{0 \leq c \leq p} \left(C_{p(a-1)}^{pb-c} C_p^c \right)$$

since any path should cross at some c the level $\tilde{a} = p(a - 1)$.

If c is different from 0 and from p then $(c, p) = 1$ and $(pb - c, p) = 1$, and therefore each of the two factors is divisible by p , their product being divisible by p^2 . Therefore, the whole sum is reduced modulo p^2 to the two summands with $c = 0$ and $c = p$, providing the congruence of Lemma 2.2. $\square$

Lemma 2.3. *The quantities $\Delta(a, b) = C_{pa}^{pb} - C_a^b$ satisfy the following congruences modulo p^2 :*

$$\Delta(a, b) \equiv \Delta(a - 1, b - 1) + \Delta(a - 1, b).$$

Proof. This follows from Lemma 2.2, by the subtraction of the equality

$$C_a^b = C_{a-1}^{b-1} + C_{a-1}^b. \quad \square$$

The congruence of Lemma 2.3 is similar to the Pascal triangle definition, but it is not an equality – just a congruence.

Proof of Theorem 2. Modulo p^2 , the numbers $\Delta(a, b)$ satisfy the recurrent equations of Lemma 3 (similar to the Pascal triangle definition). The successive application of this formula leads to a system which defines a unique solution in terms of “initial conditions” of the type $\Delta(c, 0)$ and $\Delta(d, d)$. These numbers vanish:

$$\Delta(c, 0) = C_{pc}^0 - C_c^0 = 1 - 1 = 0, \quad \Delta(d, d) = C_{pd}^{pd} - C_d^d = 1 - 1 = 0.$$

Therefore, the whole solution is vanishing everywhere, and hence all the values $\Delta(a, b)$ are zero (modulo p^2), providing the proof of Theorem 2.

3 Congruences for the multinomial coefficients

The *multinomial coefficient* of a partition of a natural number $m = \sum m_k$ into non-negative integers m_k is defined as

$$C(\{m_k\}) = \frac{m!}{\Pi(m_k!)}.$$

This number counts the words of length m consisting of m_k occurrences of each character k .

For instance, the binomial coefficients are obtainable in the particular case of just two summands:

$$C(\{m_1, m_2\}) = C_a^b,$$

where $a = m_1 + m_2$, $b = m_1$ (or m_2).

The multinomial coefficients are basic for information theory, since for $m_k \rightarrow \infty$ one gets

$$\lim_{m \rightarrow \infty} \frac{\ln C(\{m_k\})}{m} = - \sum p_k \ln p_k,$$

for $\lim m_k/m = p_k$, which explains the origin of the notion of entropy.

Theorem 3. *The following congruence modulo p^2 holds:*

$$C(\{pm_k\}) - C(\{m_k\}) \equiv 0 \pmod{p^2}.$$

Lemma 3.1. *If at least one of the numbers m_k is relatively prime to p and the sum $m = \sum m_k$ is divisible by p^r , then the multinomial coefficient $C(\{m_k\})$ is divisible by p^r .*

Proof. Use the obvious formula

$$C(\{m_k\}) = \frac{m}{m_r} C(\{m_k - 1_r\}),$$

where $(p, m_r) = 1$ and where $1_r = \delta_{k,r}$ means 1 for $k = r$ and 0 otherwise.

Since the denominator m_r is not divisible by p , the multinomial coefficient, along with m , is divisible by p^r . $\square$

Lemma 3.2. *The following congruence modulo p^2 holds:*

$$C(\{pm_k\}) \equiv \sum_r C(\{p(m_k - 1_r)\}).$$

Proof. The multinomial coefficients satisfy the recurrent relations similar to those defining the Pascal triangle:

$$C(\{pm_k\}) = \sum_c (C(\{pm_k - c_k\})C(\{c_k\})),$$

where $|c| = \sum c_k = p$ and the coordinates c_k are non-negative.

This property follows from the fact that each (non-strictly monotonic) lattice path b , connecting the lattice points $(\{0\})$ and $(\{pm_k\})$, must pass the level $\sum b_k = p(m - 1)$, where $m = \sum m_k$.

If at least one of the coordinates c_k of the vector c is relatively prime to p , then each of the two factors of the corresponding term is divisible by p and their product is divisible by p^2 .

Therefore, modulo p^2 , the whole sum above is reduced to the sum of the terms for which all the components of c are divisible by p .

However, $|c| = p$, therefore such a vector c has only one non-zero component, say $c_r = p$, yielding the congruence of Lemma 3.2 (modulo p^2). $\square$

Lemma 3.3. *The quantities*

$$\Delta(\{m_k\}) = C(\{pm_k\}) - C(\{m_k\})$$

satisfy the following congruences modulo p^2 :

$$\Delta(\{m_k\}) \equiv \sum_r \Delta(\{m_k - 1_r\}).$$

Proof. The congruence follows from Lemma 3.2 by subtracting the sum

$$C(\{m_k\}) = \sum_r C(\{m_k - 1_r\}). \quad \square$$

Proof of Theorem 3. Modulo p^2 , the quantities $\Delta(\{m_k\})$ satisfy the recurrent system of Lemma 3.3 (similar to the Pascal triangle definition).

To check the uniqueness of the solution satisfying the known initial conditions

$$\Delta(\{a, 1_r\}) = C(\{pa, 1_r\}) - C(\{a, 1_r\}) = 1 - 1 = 0,$$

one should start from vectors with only one non-zero component (for which we have just proved the vanishing of the value of Δ), going then to the vectors with only two non-vanishing components.

The vanishing of Δ on such a plane follows from the boundary conditions on the two rays already examined, due to the uniqueness of the Pascal relations solutions.

Similarly, Lemma 3.3 deduces the vanishing of the values of Δ on any coordinate plane of dimension s provided it is already known on its coordinate faces of dimension $s - 1$.

Therefore, by induction on s , we come to the conclusion that $\Delta(\{m_k\})$ is congruent to zero modulo p^2 for any partition $\{m_k\}$, which proves Theorem 3. $\square$

4 Congruences for the Newton-Girard coefficients

The *Newton-Girard formula* is the representation of the sum of the n th powers

$$s_n = \lambda_1^n + \dots + \lambda_r^n$$

in terms of the basic symmetric functions of the r arguments $\lambda_1, \dots, \lambda_r$:

$$\sigma_1 = \lambda_1 + \dots + \lambda_r, \quad \sigma_2 = \sum_{i < j} \lambda_i \lambda_j, \quad \dots, \quad \sigma_r = \prod \lambda_j.$$

We write this formula in the form of a polynomial, called *Newton-Girard polynomial*, which is the sum of the monomials Π with coefficients K :

$$s_n = \sum (K[\{m_k\}] \Pi[\{m_k\}]),$$

where Π means the product of the powers of the basic functions,

$$\Pi(\{m_k\}) = \Pi(\sigma_{u_k}^{m_k}).$$

The coefficients K are provided by the following simple explicit expression, in terms of the multinomial coefficients (strangely missing in [4], [5] and in the later textbooks):

$$K[\{m_k\}] = (-1)^{n-m} \frac{n}{m} C(\{m_k\}), \quad (*)$$

where $n = \sum u_k m_k$ is the degree of σ_n in the variables λ , (that is, its weighted quasi-homogeneous degree considered as a polynomial in the variables σ_j of weights j), and $m = \sum m_k$ is the ordinary degree of s_n , considered as a polynomial in the variables σ .

Example 5. The above formulae for $s_2 = \sigma_1^2 - 2\sigma_2$, $\sigma_3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$, σ_5 and σ_7 follow from formula (*).

The proof of the formula (*) for the coefficients K depends on the evident Girard-Newton recurrent relation

$$s_n = \sigma_1 s_{n-1} - \sigma_2 s_{n-2} + \dots + (-1)^{n+1} \sigma_n s_0,$$

which is an immediate corollary of the Vieta formula

$$\lambda^n = \sigma_1 \lambda^{n-1} - \sigma_2 \lambda^{n-2} \pm \dots + (-1)^{n+1} \sigma_n,$$

(λ being a root of the polynomial with coefficients $\pm\sigma_j$).

The coefficients K verify an identity similar to the one defining the Pascal triangle,

$$K[\{m_k\}] = \sum (-1)^{u_k+1} K[\{m_1, \dots, m_{k-1}, m_k - 1, m_{k+1}, \dots, m_s\}]. \quad (**)$$

This Pascal-type identity follows from the Newton-Girard recurrent relations above (where one should replace the quantities s_l by their expressions $s_l = \sum (K\Pi)$).

The Pascal-type recurrent relations (**) for the coefficients K imply the expression (*) of these coefficients in terms of the multinomial coefficients C .

Indeed, replacing K in (**) by the expressions (*) and identifying the coefficients of the same products of basic functions, we get true Pascal-type relations for the multinomial coefficients C , which can be proved easily (either by a high-dimensional Pascal ways-counting argument, or by direct sum computation, in terms of the factorials). Therefore, the Pascal-type identity (**) is verified by the numbers (*).

This identity implies the correctness of the expressions (*), since the Pascal-type recurrent system (**) has (evidently) only one solution, and the numbers (*) verify all the relations (**).

Example 6. The coefficients K of some monomials in the expressions (*) of s_n are:

$$\text{Coeff}[\sigma_1^2 \sigma_2^2 \sigma_3] = 54 \text{ (where } m = 5, n = 9\text{);}$$

$$\text{Coeff}[\sigma_1^a \sigma_2 \sigma_3] = -n(n-4) \text{ (where } m = a+5, n = a+2\text{);}$$

$$\text{Coeff}[\sigma_u^a \sigma_v^b] = \pm \frac{n}{a+b} C_{a+b}^a \text{ (where } u \neq v\text{).}$$

Theorem 4. If $n = p^2$, the following congruence modulo p^2 holds:

$$K(\{pm_k\}) - K(\{m_k\}) \equiv 0 \pmod{p^2}.$$

Proof. Let first p be odd. The fraction $\frac{n}{m}$ preserves its value, when all the degrees m_k are multiplied by a constant p , therefore it suffices to prove the divisibility by p^2 of the quantity $\frac{n}{m} \Delta(\{m_k\})$.

The second factor, $\Delta(\{m_k\})$, is divisible by p^2 (Theorem 3). We have therefore only to prove that our denominator m can't be divisible by a higher degree of p than the numerator $n = p^2$.

But the formulae $m = \sum m_k$, $n = \sum u_k m_k$ imply the inequality $m \leq n$. Therefore, m can't be divisible by p^3 , and Theorem 4 is thus deduced from Theorem 3 for the case $p \neq 2$.

In the case $p = 2$ there are 5 partitions of $n = p^2 = 4$, the two partitions into even parts pm_k providing the contributions:

$\Pi(\{m_k\})$	σ_2	σ_1^2
$K(\{m_k\})$	-2	+1
$\Pi(p\{m_k\})$	σ_2^2	σ_1^4
$K(p\{m_k\})$	+2	+1
$K(p\{m_k\}) - K(\{m_k\})$	+4	0

Since both differences are divisible by $p^2 = 4$, Theorem 4 is proved for $p = 2$. $\square$

5 The matricial Euler Theorem

The Euler theorem extends the small Fermat theorem, $a^p \equiv a \pmod{p}$ for a prime modulus p , to the following congruence for any integer modulus n :

$$a^n \equiv a^{n-\varphi(n)} \pmod{n},$$

where φ denotes the Euler function, whose value at an integer n is equal to the number of the residues modulo n , which are relatively prime to n . Thus, for instance,

$$\varphi(p) = p - 1, \quad \varphi(p^m) = (p - 1)p^{m-1},$$

and therefore for $n = p^m$ the value of $n - \varphi(n)$ is p^{m-1} .

Here we prove a matricial version of the Euler theorem, similar to our matricial version of the small Fermat theorem (Theorem 1), for the cases $n = p^2$ and $n = p^3$:

$$\text{tr}(A^{p^2}) \equiv \text{tr} A^p \pmod{p^2},$$

$$\text{tr}(A^{p^3}) \equiv \text{tr} A^{p^2} \pmod{p^3},$$

for matrices A of any order with integer entries.

For $n = 6$, where $\varphi(6) = 2$, the congruence $\text{tr}(A^6) \equiv \text{tr} A^4 \pmod{6}$ sometimes fails (for example, for the matrices with eigenvalues $(0, 0, \pm 1, \pm i)$).

Theorem 5. *Let A be an order r matrix with integer entries, p a prime number and $n = p^2$. The following congruence modulo n holds:*

$$\text{tr}(A^n) \equiv \text{tr}(A^{n-\varphi(n)}) \pmod{n}.$$

For $n = p^2$ we have $n - \varphi(n) = p$ and thus Theorem 5 says that

$$\text{tr}(A^{p^2}) \equiv \text{tr}(A^p) \pmod{p^2},$$

which is essentially equivalent to the following theorem on symmetric functions.

Theorem 6. For $n = p^2$ the following congruence holds:

$$s_n - s_{n-\varphi(n)} \equiv 0 \pmod{n}.$$

Remark 2. While the “matrix small Fermat” Theorem 1’ claims the divisibility by $n = p$ of the *coefficients* of the difference, where $n = p$ is a prime; the “matrix Euler” Theorem 6 says nothing on the divisibility by n of the coefficients of the polynomial $s_n - s_{n-\varphi(n)}$: it only claims the divisibility by $n = p^2$ of the *values* of this integer coefficients polynomial of the variables σ at each point $\sigma \in \mathbb{Z}^r$.

Proofs of theorems 5 and 6. Let p be odd. Writing the expression (*) of section 4, for the polynomials s_n and $s_{n-\varphi(n)} = s_p$, we may neglect terms divisible by p^2 to obtain the modulo p^2 congruence.

Namely, this remark is true for all the summands, containing a power m_k , relatively prime to p . Therefore, we get, modulo p^2 , the congruence

$$s_n - s_p \equiv \sum [K(\{pa_k\})\Pi(\{pa_k\}) - K(\{a_k\})\Pi(\{a_k\})].$$

Noting that by Theorem 4

$$K(\{pa_k\}) \equiv K(\{a_k\}) \pmod{p^2},$$

we get modulo $n = p^2$ the congruence

$$s_n - s_p \equiv \sum [K(\{a_k\}) (\Pi(\{pa_k\}) - \Pi(\{a_k\}))]. \quad (**)$$

But $\Pi(\{pa_k\}) = (\Pi(\{a_k\}))^p$, therefore the difference $\Pi(\{pa_k\}) - \Pi(\{a_k\}) = \Pi^p - \Pi$ is divisible by p (by the numerical small Fermat Theorem).

The factor $K(\{a_k\})$ is also divisible by p since

$$K(\{a_k\}) = (-1)^{n-m} \frac{n}{m} C(\{a_k\}),$$

where the multinomial coefficient C is divisible by p according to Lemma 3.1 (indeed, at least one of the numbers a_k is relatively prime to p , otherwise the number $\sum pa_k = m < n = p^2$ would be divisible by p^2), while in the fraction $\frac{n}{m}$ the power of p in the denominator is smaller, than in the numerator, as we have already shown in section 4.

Thus, we have deduced from the congruence (**) the congruence

$$s_n - s_p \equiv 0 \pmod{p^2},$$

proving Theorem 6 and hence Theorem 5 for the case where p is odd.

In the case $p = 2$ the congruences modulo 4 are elementary: arguing as above, we get

$$\begin{aligned} s_4 - s_2 &\equiv 4\sigma_4 + 4\sigma_1\sigma_3 + 2\sigma_2^2 - 4\sigma_1^2\sigma_2 + \sigma_1^4 - (-2\sigma_2 + \sigma_1^2) \\ &\equiv 2(\sigma_2^2 + \sigma_2) + (\sigma_1^4 - \sigma_1^2). \end{aligned}$$

Since the numbers $\sigma_1^4 - \sigma_1^2$ and $2(\sigma_2^2 + \sigma_2)$ are divisible by 4, Theorems 5 and 6 are proved also for $p = 2$.

Remark 3. The proof of a more general matrix Euler Theorem,

$$\mathrm{tr}(A^n) \equiv \mathrm{tr}(A^{n-\varphi(n)}) \pmod{n = p^h},$$

seems to be similar (as I have verified it for many particular cases), but I had not written all the details for all values of the degree h (see subsection 5.1).

It seems to me that the methods of the present paper provide also stronger congruences (in particular, one may replace the transition p units up from the level pm used above by the transition by p^2 units: $|c| = p^2$). The existence of such congruences is suggested, for instance, by the tables of section 2 above.

5.1 The matricial Euler Theorem for $n = p^3$

Theorem 7. *The following congruence modulo $n = p^3$ holds:*

$$\mathrm{tr}(A^n) \equiv \mathrm{tr}(A^{n-\varphi(n)}) \pmod{n}.$$

Taking into account the value of the Euler function φ at $n = p^3$, we obtain $n - \varphi(n) = p^2$, therefore Theorem 7 claims that:

$$\mathrm{tr}(A^{p^3}) \equiv \mathrm{tr}(A^{p^2}) \pmod{p^3},$$

which is equivalent to the following theorem on symmetric functions:

Theorem 8. *The following congruence modulo $n = p^3$ holds:*

$$s_n - s_{n-\varphi(n)} \equiv 0 \pmod{n}.$$

Remark 4. Theorem 8 (as Theorem 6 above) says nothing on the *coefficients* divisibility of the difference (see Remark 2). Only the *values* of the polynomial $s_n - s_{n-\varphi(n)}$ are divisible by $n = p^3$ (at any integer point $\sigma \in \mathbb{Z}^r$).

I don't know whether other polynomials, whose values at the points of the set $\sigma(\mathbb{Z}^r)$ are divisible by n , are forced to be divisible by n at all integer points of $\mathbb{Z}^r$.

It would be interesting to understand which properties of a set $M \subset \mathbb{Z}^r$ imply divisibility by n of values at all points from $\mathbb{Z}^r$ of all the polynomials, whose values are divisible by n at all the points of the set M .

Otherwise, if this divisibility property depends rather on the peculiarities of our polynomial $s_n - s_{n-\varphi(n)}$ than of the characteristics of the set $M = \sigma(\mathbb{Z}^r)$, it would be interesting to understand, which peculiarities of the polynomial $s_n - s_{n-\varphi(n)}$ are relevant here.

Proofs of Theorems 7 and 8. Decompose all the degree sequences $\{m_k\}$ into several groups according to their divisibility by the powers of the prime p :

The group α contains $m_k = p^\alpha a_k$, where $(a_k, p) = 1$ for at least one k .

The terms with $\alpha = 0$ produce in s_n only the monomials, whose coefficients are divisible by p^a , since the coefficient

$$K[\{m_k\}] = \pm \frac{(n = p^a)}{m} \frac{m}{m_k} C(\{m_k - 1_k\})$$

is divisible by p^a , provided that $(m_k, p) = 1$. This group may be neglected in the proof of the theorem. Here, as above, $n = \sum (p^{\alpha+1} u_k a_k)$, and therefore we shall write later $\underline{n} = n/p$ for its value divided by p .

The contributions of the terms of the other groups in $s_n - s_{n-\varphi(n)}$ can be rewritten in the form of two summands $I + II$, where

$$I = (K[\{p^{\alpha+1} a_k\}] - K[\{p^\alpha a_k\}]) \Pi [\{p^{\alpha+1} a_k\}],$$

$$II = K[\{p^\alpha a_k\}] (\pi^{p^{\alpha+1}} - \pi^{p^\alpha}),$$

here $\pi = \Pi(\{a_k\}) = \Pi(\sigma_{u_k}^{a_k})$.

Evaluate the divisibility by the powers of the prime number p of both summands I and II .

Lemma 5.1. *The summand II is divisible by $n = p^a$.*

Proof. The number K defined in section 4 by the formula (*),

$$K = \pm \frac{(\underline{n} = p^{a-1})}{m} \frac{m}{(m_k = p^\alpha a_k)} C(\{m_k - 1_k\})$$

is obviously divisible by $p^{a-\alpha-1}$. The difference $\pi^{p^{\alpha+1}} - \pi^{p^\alpha}$ is divisible by $p^{\alpha+1}$ (by the numerical Euler theorem).

Therefore the product is divisible by $p^{(a-\alpha-1)+(\alpha+1)} = p^a$. $\square$

To evaluate the contribution of the term I , note that in the right hand side product

$$K[\{p^{\alpha+1} a_k\}] - K[\{p^\alpha a_k\}] = (\underline{n}/m) [C(\{p^{\alpha+1} a_k\}) - C(\{p^\alpha a_k\})]$$

the term $\underline{n}$ provides the contribution $p^{a-1} = p^2$, while the second factor, representing the difference of the multinomial coefficients, is divisible by p^2 (by Theorem 3 of section 3).

Therefore, to prove the divisibility of the product by n , it is sufficient to prove that the denominator m is not divisible by p^2 . It is rather easy to describe the cases of divisibility of m by p^2 , since we would have

$$(m = \sum (p^\alpha a_k)) \leq \left(n = \sum (p^\alpha a_k u_k) = p^2 \right),$$

and the divisibility of m (which is not 0, since $(a_k, p) = 1$ for some k) is only possible for $m = p^2$ and $\sum (u_k - 1) a_k = 0$. Thus only one of the coefficients a_k

is different from 0, namely the one for which $u_k = 1$, yielding $\{p^\alpha a_k\} = p^2$ and $\{p^{\alpha+1} a_k\} = p^3$. We conclude that the contribution of this term provides

$$K[\{p^{\alpha+1} a_k\}] = K[\{p^\alpha a_k\}] = 1,$$

and the difference contribution vanishes. Thus, the summand I is divisible by p^3 , and Theorem 8 is thus proved for odd primes p .

Theorems 7 and 8 are also true for $p = 2$, and the proof follows the same lines, as for the odd primes. The only difference is the need to replace the multinomial coefficients difference

$$C(\{p^{\alpha+1} a_k\}) - C(\{p^\alpha a_k\})$$

by the sum of the same multinomial coefficients (for certain degrees $\{a_k\}$).

The divisibility of the resulting sums by $p^2 = 4$ is easily provable, but it even suffices simply to write explicitly all the summands of the difference $s_8 - s_4$, whose number is rather small, and to check the divisibility directly, using the explicit expressions:

$$\begin{aligned} s_8 = & \sigma_1^8 - 8\sigma_8 + 8\sigma_1\sigma_7 - 8\sigma_1^2\sigma_6 + 8\sigma_1^3\sigma_5 - 8\sigma_1^4\sigma_4 + 8\sigma_1^5\sigma_3 - 8\sigma_1^6\sigma_2 + 20\sigma_1^4\sigma_2^2 \\ & - 32\sigma_1^3\sigma_2\sigma_3 - 16\sigma_1^2\sigma_2^3 + 24\sigma_1^2\sigma_2\sigma_4 + 12\sigma_1^2\sigma_3^2 + 24\sigma_1\sigma_2^2\sigma_3 - 16\sigma_1\sigma_2\sigma_5 - 16\sigma_1\sigma_3\sigma_4 \\ & + 2\sigma_2^4 - 8\sigma_2^2\sigma_4 - 8\sigma_2\sigma_3^2 + 8\sigma_2\sigma_6 + 8\sigma_3\sigma_5 + 4\sigma_4^2. \end{aligned}$$

Therefore we get the following modulo 8 congruence:

$$s_8 \equiv \sigma_1^8 + 20\sigma_1^4\sigma_2^2 + 12\sigma_1^2\sigma_3^2 + 2\sigma_2^4 + 4\sigma_4^2.$$

Similarly, the explicit expression for s_4 is

$$s_4 = \sigma_1^4 - 4\sigma_4 + 4\sigma_1\sigma_3 - 4\sigma_1^2\sigma_2 + 2\sigma_2^2.$$

Thus we get the modulo 8 congruence

$$\begin{aligned} s_8 - s_4 \equiv & \sigma_1^4(\sigma_1^4 - 1) + 4\sigma_4(\sigma_4 + 1) + 2\sigma_2^2(\sigma_2^2 - 1) \\ & + 4\sigma_1^2\sigma_2(\sigma_1^2\sigma_2 + 1) + 4\sigma_1\sigma_3(3\sigma_1\sigma_3 + 1). \end{aligned}$$

Obviously, each of the five summands is divisible by 8, which proves Theorems 7 and 8 for $p = 2$.

6 Matrices of order $r = 2$

Till now the order r of the matrices of Theorem 1 had been arbitrary. For matrices of order $r = 2$ Theorem 1 may be rewritten in the following way. Denote the trace of the matrix A^m by

$$t_m = \text{tr}(A^m).$$

Proposition 1. *For a matrix A of determinant 1, the trace t_m is the value at the point $t_1 = \text{tr} A$ of some polynomial of variable t with integer coefficients.*

Indeed any matrix A (with $\text{tr} A = t$ and $\det A = 1$) verifies its characteristic equation

$$A^2 = At - 1, \quad A^{m+1} = A^m t - A^{m-1}.$$

Therefore, the relation $\text{tr} A^m = t_m(t)$ implies the recurrent relation

$$t_{m+1}(t) = t t_m(t) - t_{m-1}(t),$$

whence the initial conditions $t_0(t) = \text{tr}(1) = 2$ and $t_1(t) = t$ provide a useful sequence of polynomials of one variable t :

$$\begin{aligned} t_0 &= 1, & t_1 &= t, & t_2 &= t^2 - 2, & t_3 &= t^3 - 3t, & t_4 &= t^4 - 4t^2 + 2, \\ t_5 &= t^5 - 5t^3 + 5t, & t_6 &= t^6 - 6t^4 + 9t^2 - 2, & t_7 &= t^7 - 7t^5 + 14t^3 - 7t. \end{aligned}$$

The general term of this trigonometric useful sequence is the polynomial

$$t_m(t) = \sum A_{k,l} t^k, \quad \text{where } k - 2l - 2 = m \quad \text{and} \quad 0 \leq k \leq m$$

(of course, the sum runs here along the values k whose difference with m is even).

The coefficients A in this formula are equal to

$$A_{k,l} = \frac{l(l+1) \cdots (l+k-2)(2l+k-2)}{k!} (-1)^{l+1}.$$

These coefficients form a Pascal-type “triangle”. The beginning ($k = 0, 1, 2, 3, 4$) of this triangle (neglecting the sign factor $(-1)^{l+1}$) has the form of the following table:

6	1	0	0	0	1	6	20	50	105
-5	-1	0	0	1	5	14	30	55	91
	4	1	0	1	4	9	16	25	36
		-3	-1	1	3	5	7	9	11
			2	2	2	2	2	2	2

The 5 presented lines of this “triangle” are defined by the formulae

$$\begin{aligned} \pm A_{0,l} &= 2, & \pm A_{1,l} &= 2l - 1, & \pm A_{2,l} &= l^2, \\ \pm A_{3,l} &= (2l^3 + 3l^2 + l)/6, & \pm A_{4,l} &= l(l+1)(l+2)(2l+4)/24. \end{aligned}$$

Our divisibility theorems describe the vertical segments of the above table, situated below the diagonal of 1's, over the odd primes lying on the line $\{A_{1,l}\}$ having only odd numbers. These segments are $\{3\}$ for $p = 3$, $\{5, 5\}$ for $p = 5$, $\{7, 14, 7\}$ for $p = 7$; further, one observes that 55 is divisible by 11 and 91 by 13, as it should be by Theorem 1.

Our sequence provides also the polynomials p_m and q_m , defined by the conditions

$$A^m = p_m A + q_m 1$$

(implied by the characteristic equation of the matrix A). This equation can be written as two recurrent relations:

$$p_{m+1} = tp_m + q_m, \quad q_{m+1} = -p_m.$$

The first terms of these sequences are:

$$\begin{array}{ll} p_1 = 1, & q_1 = 0; \\ p_2 = t, & q_2 = -1; \\ p_3 = t^2 - 1, & q_3 = -t; \\ p_4 = t^3 - 2t, & q_4 = -t^2 + 1; \\ p_5 = t^4 - 3t^2 + 1, & q_5 = -t^3 + 2t; \\ p_6 = t^5 - 4t^3 + 3t, & q_6 = -t^4 + 3t^2 - 1; \\ p_7 = t^6 - 5t^4 + 6t^2 - 1, & q_7 = -t^5 + 4t^3 - 3t. \end{array}$$

The polynomials p_m , q_m and t_m are related by the evident relation (of additivity of traces)

$$t_m = tp_m + 2q_m \quad (\text{where } 2 = \text{tr}1).$$

7 The circular group C_p on the finite plane

Consider the finite unit circle

$$C_p = \{(x, y) : x^2 + y^2 = 1\},$$

on the finite plane $\mathbb{Z}_p^2$, consisting of p^2 points (where $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$, p being a prime number). A well known fact of Babylonian mathematics is

Lemma 7.1. *The number of points on C_p is equal to*

$$|C_p| = p + 1 \text{ for } p = 4c - 1, \text{ and to}$$

$$|C_p| = p - 1 \text{ for } p = 4c + 1.$$

Example 7. For $p = 5$, C_p has 4 points: $(0, \pm 1)$, $(\pm 1, 0)$. For $p = 7$ there are 4 additional points on C_p : $(2, \pm 2)$ and $(-2, \pm 2)$.

Proof of Lemma 7.1. Parametrise the rational curve C_p by the “tangent of the half angle”, that is by the inclination of the straight line, connecting the point (x, y) of the curve with its point $(-1, 0)$:

$$[y = t(1 + x)] \implies \left[x = \frac{1 - t^2}{1 + t^2}, \quad y = \frac{2t}{1 + t^2} \right].$$

The number of points on the curve equals the number of possible values of the projective parameter $t \in \{1, 2, \dots, p, \infty\}$. A value is admissible if $1 + t^2 \neq 0$ (the point $(-1, 0)$ corresponding to the parameter value $t = \infty$).

To solve the equation $1 + t^2 = 0$, defining the asymptotic directions, write t in terms of a primitive residue ρ , expressing it as $t = \rho^\lambda$, $0 < \lambda < p$. Taking into account that $-1 = \rho^{\frac{p-1}{2}}$, the value of λ corresponding to an asymptotic direction should verify the congruence

$$2\lambda \equiv \frac{p-1}{2} \pmod{(p-1)},$$

that is

$$4\lambda = (p-1)(1+2a), \quad a \in \mathbb{Z}.$$

The number $2a+1$ is odd, therefore $p-1 = 4c$ (for some integer c).

We have thus proved

Lemma 7.2. *If p has the form $4c-1$ then the equation $t^2+1=0$ has no solutions in $\mathbb{Z}_p$. If p has the form $4c+1$ then the equation $t^2+1=0$ has two solutions (provided by the choice of $\lambda = c$ above, for one of them).*

Therefore, the number of asymptotic points (to be excluded from the set of values of t in the proof of Lemma 7.1) is 0 or 2 for $p \equiv -1$ or $1 \pmod{4}$. We have thus ended the proof of Lemma 7.1.

Definition 1. One calls *circular group* C_p the set of matrices of order 2 which are the matrices of the multiplication by the complex numbers, with determinant 1 $\pmod{p}$ and whose elements are residues modulo p .

Such matrices have the form

$$A = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

A matrix of such form has determinant 1 if and only if the point (x, y) belongs to the circle C_p . These matrices form (evidently) a commutative group that we call “circle C_p ”. This circular group is a subgroup (of order $p \pm 1$ for $p = 4c \mp 1$) of the group of unimodular matrices $G = SL(2, \mathbb{Z}_p)$.

Theorem 9. *The circular group C_p is cyclic.*

Proof. As every finite commutative group, the circular group is the direct product of some cyclic groups of primary orders $q_k^{a_k}$, where the numbers q_k are primes (virtually coinciding). We shall first prove that, for the circular group, all the primes corresponding to the factors of such direct product are pairwise different (making impossible the products like $\mathbb{Z}_3^2$ or $\mathbb{Z}_3\mathbb{Z}_9$ for a circular group).

Lemma 7.3. *For any positive integer m the number of solutions A of the equation $A^m = B$ on the circle C_p does not exceed its degree m .*

Proof. According to the formula for the tangent of a sum, the parameter value t_m at the point A^m is related to the parameter value t at the point A by the formula $t_m = f_m(t)/g_m(t)$, where the polynomials f_m and g_m have degrees m and $m - 1$, respectively, provided that m is odd (and degrees $m - 1$ and m , if m is even).

In both cases the equation $A^m = B$ can be written as an algebraic equation of degree m for the value t of the parameter at the point A :

$$t_m g_m(t) = f_m(t),$$

whence the number of roots t can't exceed m . Lemma 7.3 is proved. $\square$

The group $\mathbb{Z}_{q^a}$ contains exactly q elements of degree q . Therefore, the product of h primary groups $\mathbb{Z}_{q^{a_k}}$ (with the same prime number q) has q^h elements of degree q . For $h > 1$ this number is greater than q (which is impossible for the circular group, by Lemma 3). Therefore, $h = 1$ for any q , that is, all primes occurring in different primary factors are different. As the product of cyclic groups of relatively prime orders, the group C_p is also cyclic. We have thus proved Theorem 2:

$$C_p \simeq \mathbb{Z}_{p+1} \text{ for } p = 4c - 1 \text{ and } C_p \simeq \mathbb{Z}_{p-1} \text{ for } p = 4c + 1. \quad \square$$

The order of the cyclic group C_p with $p \neq 2$ is thus divisible by 4 (as the circle symmetry requires).

Remark 5. We have excluded above the group $C_2 = \mathbb{Z}_2$, which consists of two matrices of permutations of coordinates, corresponding to the points $(0, 1)$ and $(1, 0)$ of the $\mathbb{Z}_2$ -circle C_2 .

8 The dynamics of squaring matrices

Consider the squaring operation, $A \mapsto A^2$, acting on the matrices of order 2 with determinant 1, whose elements are residues modulo a prime number p :

$$A \in G = \text{SL}(2, \mathbb{Z}_p).$$

Iterating this squaring operation, we construct from a matrix A a sequence of squares belonging to G :

$A_s = A_{s-1}^2 = A^{2^s}$, $A_0 = A$, $A_1 = A^2$, $\dots$, A_r —the orbit of the “squaring dynamics”.

Example 8. For $p = 7$ we get, for instance, the chain of squares

$$A_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The number r of squarings of a chain will be called the *length* of that chain ($r = 3$ in the example of 4 matrices above).

Theorem 10. *The group $G = \text{SL}(2, \mathbb{Z}_p)$ has a squarings chain of length r , ending with $A_r = 1$ ($A_{r-1} \neq 1$) if and only if the prime number p has the form $p = 2^r c \pm 1$.*

Example 9. For $p = 7$ and 17 there are chains of length $r = 3$, and for $p = 5$ and 11 they do not exist (while chains of length $r = 2$ exist).

The number r is the maximal length of a chain of squares in G if and only if $p = 2^r c \pm 1$, c being odd.

The proof of Theorem 10 consists of several steps.

Lemma 8.1. *The order of the group $G = \text{SL}(2, \mathbb{Z}_p)$ (p being a prime number) is equal to $|\text{SL}(2, \mathbb{Z}_p)| = p(p^2 - 1)$.*

Example 10. $|\text{SL}(2, \mathbb{Z}_3)| = 24$, $|\text{SL}(2, \mathbb{Z}_5)| = 120$, $|\text{SL}(2, \mathbb{Z}_7)| = 336$.

The case $p = 5$ is related to the dodecahedron's rotation group ($|\text{PSL}(2, \mathbb{Z}_5)| = 60$) and the case $p = 7$ to the 168 rotations of the regular polyhedron of genus 3 (see [2]).

Proof of Lemma 8.1. A matrix of G , $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, satisfies $ad - bc = 1$. Denote the product ad by $u \in \mathbb{Z}_p$. The condition $\det A = 1$ takes the form of two decompositions:

$$ad = u, \quad bc = u - 1.$$

The number of ordered decompositions of the residue u , representing it as the product ad , is equal to $p - 1$ if $u \neq 0$, and equal to $2p - 1$ if $u = 0$. Counting separately the contributions of the two special values ($u = 0$, $u = 1$) and the contributions of the $p - 2$ remaining values of the product u , and then adding the products of the number of decompositions for u and for $u - 1$, we get the following expression for the number of elements of G :

$$\begin{aligned} |G| &= 2(p - 1)(2p - 1) + (p - 2)(p - 1)^2 \\ &= (p - 1)[(4p - 2) + (p^2 - 3p + 2)] \\ &= (p - 1)(p^2 + p), \end{aligned}$$

proving Lemma 8.1. □

The chain of squares $\{A_s\}$, ending with $A_r = 1$, defines in the group G the cyclic subgroup of powers A_0^k ($k = 1, 2, 3, \dots, 2^r$) of the element A_0 .

Lemma 8.2. *All these 2^r elements of the group G are pairwise different, provided that $A_r = 1$ and $A_{r-1} \neq 1$.*

Proof. Otherwise one would have $A_0^m = 1$ for some m with $0 < m < 2^r$. The largest common divisor q of the numbers m and 2^r can be written in the usual form $q = 2^u = Um + V2^r$, where u is smaller than r .

One would then have $A_0^q = 1$. Therefore, the squarings chain would contain $A_u = 1$, with $u < r$, the chain ending earlier than at A_r . Since A_{r-1} is supposed to be different from 1, this is impossible. Lemma 8.2 is proved. □

The group G , of order $p(p^2 - 1)$ contains thus a (cyclic) subgroup of order 2^r . It follows (for odd p) that either $p + 1$ or $p - 1$ is divisible by 2^{r-1} . We shall prove that in fact it is divisible even by 2^r .

Lemma 8.3. *If the group $G = \text{SL}(2, \mathbb{Z}_p)$ contains a squarings chain of length r ending at $A_r = 1$, with $A_{r-1} \neq 1$, then such squarings chain is also contained in the circular subgroup C_p of the group G .*

Proof. Consider the sequence of traces of the matrices of the squarings chain, $t_s = \text{tr} A_s$. One has the following squaring formula for the matrices with determinant $ad - bc = 1$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & d^2 + bc \end{pmatrix} = \begin{pmatrix} at - 1 & bt \\ ct & dt - 1 \end{pmatrix}, \quad (*)$$

where $t = a + d = \text{tr} A$.

Squaring formula $(*)$ implies that

$$t_2 = \text{tr}(A^2) = t^2 - 2, \quad t_s = t_{s-1}^2 - 2$$

(for $s = 1, \dots, r$). Therefore, the chain $\{t_s\}$ of traces of squared matrices ends with the number $t_r = \text{tr} 1 = 2$, preceded by a solution t_{r-1} of the equation $t_{r-1}^2 = 4$.

The only solutions are $t_{r-1} = 2$ and -2 .

For odd p , the squaring formula $(*)$ implies the vanishing of the elements b and c of the matrix A_{r-1} (since $bt = ct = 0$, the matrix A_r being 1).

Therefore the matrix A_{r-1} is a diagonal matrix:

$$A_{r-1} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad ad = 1, \quad a + d = \pm 2.$$

Solving the quadratic equation $ad = 1$ with respect to a , we find $a = d = 1$ for $t = 2$ and $a = d = -1$ for $t = -2$. Therefore, if $A_{r-1} \neq 1$ then it should be $A_{r-1} = -1$ with trace $t_{r-1} = -2$.

The omitted case, $p = 2$, is easily considered separately (in this case there are only 6 matrices in G).

To end the proof of Lemma 8.3, p being odd, we shall study the term A_{r-2} of the squaring chain (its square being -1).

Definition 2. A matrix in G whose square equals -1 is called a *G-complex structure* (*p-complex* for $G = \text{SL}(2, \mathbb{Z}_p)$).

Example 11. The *standard complex structure*

$$I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in G$$

is convenient for any p .

Lemma 8.4. *The number of possible G -complex structures for the group $G = \text{SL}(2, \mathbb{Z}_p)$ is equal to $p(p \pm 1)$ for $p = 4c \pm 1$. All these complex structures are conjugated to the standard one:*

$$A_{r-2} = gIg^{-1} \quad \text{for some } g \in G.$$

Proof of Lemma 8.4. The trace t of a complex structure verifies, by the squaring formula (*), the relation $t^2 - 2 = -2$, and is therefore vanishing. Reciproquely, if $t = \text{tr}A = 0$, the squaring formula (*) shows that A is a complex structure ($A^2 = -1$).

Calculation of the number of matrices with zero trace in the group G .

These matrices have the form

$$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \quad a^2 + bc = -1.$$

The number of (ordered) representations of the residue $u = bc$, in this product form, equals $p - 1$ for $u \neq 0$, being $2p - 1$ for $u = 0$. The number of solutions of the equation $a^2 + 1 = 0$ in G is 0 for $p = 4c - 1$, being 2 for $p = 4c + 1$ (see Lemma 7.2).

Counting separately the virtual contribution of the value $u = 0$ and of the values $u \neq 0$, we get that the number of matrices with trace zero is, for $p = 4c - 1$ (where $u = 0$ is not attained), equal to

$$\#(A \in \text{SL}(2, \mathbb{Z}_p) : \text{tr}A = 0) = p(p - 1),$$

while for $p = 4c + 1$, the 2 special values of a contributing $2(2p - 1)$, the number of matrices with trace zero equals

$$\#(A) = (p - 2)(p - 1) + 2(2p - 1) = p(p + 1).$$

Counting the number of complex structures conjugated to the standard one.

The element $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ of G belongs to the isotropy group H of the standard structure I (preserving I , while conjugating G), if

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \begin{pmatrix} -\beta & \alpha \\ -\delta & \gamma \end{pmatrix} = \begin{pmatrix} \gamma & \delta \\ -\alpha & -\beta \end{pmatrix},$$

which means that $(\alpha = \delta, \gamma = -\beta)$ with the “determinant= 1” condition $\alpha^2 + \beta^2 = 1$. By Lemma 7.1, for $p = 4c \pm 1$, the circular group H has $p + 1$ elements.

Therefore, *the number of complex structures conjugated to the standard one is equal to*

$$\frac{|G|}{|H|} = \frac{p(p^2 - 1)}{p \pm 1} = p(p \mp 1) \quad \text{for } p = 4c \mp 1.$$

This number coincides with the number of matrices in G of trace zero, calculated above.

Hence, *all matrices of trace zero in G are conjugated in G to the standard structure I* . Lemma 8.4 is proved. $\square$

We shall end the proof of Lemma 8.3. Consider a squarings chain of length r such that $A_r = 1$ and $A_{r-1} \neq 0$. We have proved that it ends with $A_{r-1} = -1$ and $A_r = 1$. Moreover, if $r \geq 2$, these matrices are preceded by a complex structure A_{r-2} , conjugated to the standard one: $A_{r-2} = g^{-1}Ig$, for some $g \in G$.

Conjugating all elements of this squarings chain $\{A_s\}$ by the same element g of G , we obtain a new squarings chain, $\tilde{A}_s = gA_sg^{-1}$ of the same length, ending with the standard triple

$$\tilde{A}_{r-2} = I, \quad \tilde{A}_{r-1} = -1, \quad \tilde{A}_r = 1.$$

Note that no matrix $\tilde{A}_s$ (or A_s), with $s < r - 2$, may have trace zero (since its vanishing would imply $A_s^2 = -1$, $A_{s+2} = 1$, implying the earlier ending of the chain that at $A_r = 1$).

The squaring formula (*) for $t \neq 0$ shows that the square anti-symmetry condition, $bt + ct = 0$, implies the anti-symmetry of the original matrix, $b + c = 0$. Therefore, all matrices A_s of the squaring chain, constructed above, verify the anti-symmetry condition $b + c = 0$ (for $s < r - 2$).

Similarly, the same squaring formula (*) shows that the symmetry relation for the squares, $(at - 1) - (dt - 1) = 0$ (where $t \neq 0$) implies the symmetry relation for the preceding matrix of the squarings chain.

Thus all the elements $\tilde{A}_s$ with $s \leq r - 2$ belong to the circular group

$$C_p = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \right\}.$$

We have thus proved Lemma 8.3. $\square$

Proof of Theorem 3. By Lemma 7.1, the circular group C_p has order $p \mp 1$, for $p = 4c \pm 1$ and, by Lemma 8.3, it contains a squarings chain of length r ($A_r = 1$), in the cases $p = 2^r c \pm 1$. The cyclic subgroup of G , $\{A_0^k\}$, of order 2^r constructed above, is also a subgroup of the circular group C_p , since the matrix $A = A_0$ belongs to C_p . Therefore, $p \pm 1$ is divisible by 2^r .

Depending on the residue of p modulo 4, we obtain the following conclusions:

$$p + 1 = 2^r q \quad \text{for} \quad p = 4c - 1 \quad (\text{and then} \quad p = 2^r q - 1),$$

$$p - 1 = 2^r q \quad \text{for} \quad p = 4c + 1 \quad (\text{and then} \quad p = 2^r q + 1).$$

Reciproquely, for such values of p we have constructed above a squarings chain of length r . Theorem 3 is proved. $\square$

9 The equation $A^p = 1$ in the group $G = \text{SL}(2, \mathbb{Z}_p)$ and Lobachevsky geometry

For a matrix A in $G = \text{SL}(2, \mathbb{Z}_p)$, the matricial Fermat Theorem of section 1 provides a congruence modulo p :

$$\text{tr}(A^p) \equiv (\text{tr}A)^p.$$

The usual small Fermat Theorem congruence modulo p claims that $t^p \equiv t$ for $t = \text{tr}A$.

Combining the two congruences, we get a new matrix version of the small Fermat Theorem in the form of a congruence

$$\text{tr}(A^p) \equiv \text{tr}A.$$

Therefore, to solve the equation $A^p = 1$ in G , we start from a study of the matrices A with trace

$$\text{tr}A (= \text{tr}A^p) = \text{tr}1 = 2.$$

Lemma 9.1. *The set M of matrices of trace 2 in $G = \text{SL}(2, \mathbb{Z}_p)$ consist of p^2 elements and is closed under the squaring operation.*

Proof. By the squaring formula (*) of section 3, the equation $\text{tr}A = 2$ implies $\text{tr}(A^2) = (\text{tr}A)^2 - 2 = 4 - 2 = 2$.

To count the matrices of trace 2, we will write them in the form

$$A = 1 + B, \quad \text{tr}B = 0, \quad B = \begin{pmatrix} x & y \\ z & -x \end{pmatrix}.$$

The condition $\det A = 1$ takes the form

$$1 - x^2 - yz = 1,$$

that is $yz = -x^2$. If $x \neq 0$, we get $p - 1$ (ordered) decompositions $yz = u$ of the residue $u = -x^2$. If $x = 0$, the number of decomposition is $2p - 1$. The total number of elements of the set equals the sum of the products of the numbers of choices of the residue x and of the decompositions, that is

$$(p - 1)^2 + 1 \cdot (2p - 1) = p^2.$$

□

Of course, the set M contains the identity and the Jordan matrix $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, since $J^b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$.

The two eigenvalues of any matrix of M verify the characteristic equation $\lambda^2 - 2\lambda + 1 = 0$ and hence both are equal to 1.

Theorem 11. *The set M of matrices of trace 2 in $G = \text{SL}(2, \mathbb{Z}_p)$, p being an odd prime number, consists of 3 classes of conjugated elements:*

$$M = \{1\} \sqcup M^+ \sqcup M^-,$$

where the $(p^2 - 1)/2$ elements belonging to M^+ are conjugated to the Jordan matrix J , the elements of M^- being conjugated to J^b for any non quadratic residue b (modulo p).

Thus the $p - 1$ elements of $M \setminus \{1\}$ (being the order p solutions of the equation $A^p = 1$, $A \neq 1$) are subdivided into two equally large equivalence classes, M^+ and M^- , corresponding to the division of the $p - 1$ non-zero residues (modulo p) into the two equally large classes of quadratic and of non quadratic residues.

This decomposition is also the version (modulo p) of the subdivision of the real plane into the upper half-plane (of the Lobachevsky geometry model) and the lower one, or of the decomposition of the real projective plane into a Lobachevsky geometry disc and the complementary Möbius band (corresponding to the relativistic de Sitter world, as it is explained in [3]).

Proof of Theorem 11. The isotropy group of the standard Jordan matrix J consists of the matrices of determinant 1 $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, for which

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \begin{pmatrix} \alpha & \alpha + \beta \\ \gamma & \gamma + \delta \end{pmatrix} = \begin{pmatrix} \alpha + \gamma & \beta + \delta \\ \gamma & \delta \end{pmatrix},$$

which means ($\gamma = 0$, $\alpha = \delta$). The condition “determinant = 1” implies $\alpha^2 = 1$. For odd p , the number of solutions is $2p$ (since $\alpha = \pm 1$, $\beta \in \mathbb{Z}_p$): $|H| = 2p$.

Therefore the number of elements of G conjugated with J equals

$$\frac{|G|}{|H|} = \frac{p(p^2 - 1)}{2p} = \frac{p^2 - 1}{2}.$$

It is easy to see that J^b is conjugated to J if and only if $b \neq 0$ is a quadratic residue. The fact that all matrices of trace 2, different from 1, are conjugated to J or to J^b (with b non quadratic residue) is proved by a calculation similar to the calculation of the number of conjugated elements. Theorem 11 is thus proved. $\square$

Remark 6. One can naturally associate to a matrix $A \neq 1$ of M its only invariant eigenline, πA containing the point 0 of the p -vector space $\mathbb{Z}_p^2$. This line can be considered as a point of the finite projective line, $P = P^1(\mathbb{Z}_p)$.

Reciprocally, one can reconstruct the matrix A belonging to $M \setminus \{1\}$ from its projection πA to P , with the following ambiguity: the preimage of each of the $p + 1$ points of P in M^+ consists of $(p - 1)/2$ matrices with the same eigenvector, including, for instance, J^b together with J (b being a quadratic residue).

Each element g of G acts on $p + 1$ points of P , permuting them (by some special even permutation g_* , the subgroup of *projective permutations* of the even permutations's group of $p + 1$ elements). This subgroup has $p(p^2 - 1)/2$ elements, the even permutations's group of P having $(p + 1)!/2$ elements.

This action of the projective permutations is homomorphic to the action of G (by conjugation with its elements) on M^+ in the sense that

$$\pi(gAg^{-1}) = g_*(\pi A).$$

For a matrix g of M^+ the projective permutation g_* leaves fixed the point πg of the projective line P .

The projective transformations of a projective line, preserving the same point, form a group L (finite, for the finite projective line P).

Choosing on the projective line P a coordinate system for which the fixed point has affine coordinate $u = \infty$, we write the transformations forming L as affine transformations $\{u \mapsto au + b, \ a \neq 0\}$.

Therefore the number of such transformations equals $|L| = p(p - 1)$. However, only those affine transformations may be lifted from the projective group to $G = \text{SL}(2, \mathbb{Z}_p)$, for which the residue a is quadratic.

Thus the image of the group G under the above homomorphism to the group of projective permutations is only the index 2 subgroup L^+ of affine transformations group L , of order $|L^+| = p(p - 1)/2$, consisting of those affine transformations $a \mapsto au + b$, for which $a = c^2 \neq 0$ is a nontrivial quadratic residue.

The finite group L^+ is the (mod p) *finite version of the Lobachevsky plane*. Indeed, one can describe the usual real Lobachevsky plane as *the group of orientation preserving ($a > 0$) affine real transformations of the line ($u \mapsto au + b$), equipped with (any) left-invariant Riemannian metric* (see the details of this description in [1]).

This real Lie group is the upper half-plane $\{(a, b) : a > 0\}$. The above reasoning shows that the p -version of the Lobachevsky plane's definition might follow the same plan, provided that one would replace the real *positivity* condition $a > 0$ by its number-theoretic (mod p) version of *being a quadratic residue*, $a = c^2$.

Remark 7. One might hope that the replacement of the positivity conditions by the condition of being a square may help to extend different branches of real calculus to algebraic geometry and to number theory. Working on the 16-th Hilbert problem, I had successfully used this trick in [6], where the understanding of the complexification of real manifolds with boundary ($f(x) > 0$) is provided. The complexified version of this manifold with boundary is the two-fold ramified covering ($f(x) = y^2$), as it is explained in [6].

The resulting possibility to apply the results of 4-dimensional differential topology to the study of real algebraic curves provides very strong results in real algebraic geometry (starting with the Gudkov congruence conjecture, discussed in [6]).

I hope that the number-theoretic applications of the same idea might be equally useful. The present article describes only the first naive steps in this direction.

For instance, one might hope to use the finite Lobachevsky geometry relations of the equation $A^p = 1$, described above, to the combinatorial study of the matrix squaring operations in G and to the corresponding Riemann surfaces (described in [2], for $p = 5$ and for $p = 7$).

The *modulo p Lobachevsky plane* consists of $p(p-1)/2$ points. The lines of the Lobachevsky plane might be considered as points of a “dual” (de Sitter) relativistic surface (described in [3]). Its modulo p version contains $p(p+1)/2$ lines (connecting pairwise the $p+1$ points of the Lobachevsky absolute, separating the Lobachevsky disc from its de Sitter complement). I hope that these two dual geometries might be useful for the study of the equations $A^p = 1$ and $A^3 = 1$ in the finite groups $G = \text{SL}(2, \mathbb{Z}_p)$ (and in their projective versions $PG = G/\{\pm 1\}$).

10 The equation $A^3 = 1$ in the group $G = \text{SL}(2, \mathbb{Z}_p)$ and finite Lobachevsky geometry

To solve the cubic equation $A^3 = 1$, we write it in the ‘quadratic equation’ form $A^2 = A^{-1}$, leading, for the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, to the system of 4 equations

$$at - 1 = d, \quad bt = -b, \quad ct = -c, \quad dt - 1 = a,$$

$t = a + d$ being the trace of A (we use that $A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, for the order 2 matrices of determinant 1).

Therefore, $b(t+1) = c(t+1) = 0$, and thus either $t = -1$, or $b = c = 0$.

In the first case we have $A^2 = -A$. In the second case we find $(a+d)t - 2 = t$, $t^2 - t - 2 = 0$, $(t+1)(t-2) = 0$, therefore $t = 2$.

From the relations $(a+d=2, ad=1)$ we deduce $a = d = 1$, therefore the second case above occurs only for the identity matrix $A = 1$. We thus proved the third order elements classification theorem:

Lemma 10.1. *The solutions of the equation $A^3 = 1$ in G , $A \neq 1$, are exactly the trace -1 matrices of determinant 1:*

$$A = \begin{pmatrix} a & b \\ c & -1-a \end{pmatrix}, \quad -bc = a^2 + a + 1.$$

Example 12. The matrix $F = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$, has order 3 for any p : $F^3 = 1$.

To count all the third order matrices in G , we consider separately the cases $u = 0$ and $u \neq 0$, where u means $a^2 + a + 1$.

Lemma 10.2. *The number of solutions a of the equation $a^2 + a + 1 = 0$ in $\mathbb{Z}_p$ equals 2 for $p = 3c + 1$ (like for $p = 7$ or 13), being 0 for $p = 3c - 1$ (like for $p = 5$ or 11), while for $p = 3$ the only solution is $a = 1$.*

Proof. Our equation implies $a^3 = 1$. Denoting by ρ a primitive residue (modulo p), we get $a = \rho^\lambda$, $a^3 = \rho^{3\lambda}$. The residue of $\lambda \pmod{p-1}$ should verify the condition $3\lambda = k(p-1)$, $k \in \mathbb{Z}$.

If the integer k is divisible by 3, we get the (parasite) root $a = 1$ (convenient only for $p = 3$).

If the integer k is not divisible by 3, we get $p-1 = 3c$, $p = 3c + 1$. Choosing $k = 1$, $\lambda = c$, we obtain one solution a_1 of our quadratic equation (the second solution being $a_2 = 1/a_1 = -a_1 - 1$). Lemma 10.2 is proved. $\square$

Theorem 12. *All solutions of the equation $A^3 = 1$ in $G = \text{SL}(2, \mathbb{Z}_p)$, different from 1, are conjugated to the special solution F . Depending on the residue of p modulo 3, their number is equal to*

$$\begin{aligned} p(p-1) & \quad (\text{for } p = 3c - 1), \quad \text{or to} \\ p(p+1) & \quad (\text{for } p = 3c + 1). \end{aligned}$$

Remark 8. These numbers (and probably, these solutions) are related to the points and to the lines of the finite Lobachevsky plane and of the finite de Sitter world (or to their two-fold coverings, having $(p \mp 1)$ elements).

Proof of Theorem 12. From Lemma 10.2 we obtain the number of the trace -1 matrices with determinant 1, depending on the residue of p modulo 3:

$$\begin{aligned} p(p-1) & \quad \text{for } p = 3c - 1, \\ (p-2)(p-1) + 2(2p-1) & = p(p+1) \quad \text{for } p = 3c + 1. \end{aligned}$$

To find the isotropy group H of the standard matrix F in G , we have to solve the equation

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \begin{pmatrix} -\beta & \alpha - \beta \\ -\delta & \gamma - \delta \end{pmatrix} = \begin{pmatrix} \gamma & \delta \\ -\alpha - \gamma & -\beta - \delta \end{pmatrix},$$

that is $(\gamma = -\beta, \delta = \alpha - \beta)$. From the “determinant= 1” condition we get a rational quadratic curve $\alpha^2 - \alpha\beta + \beta^2 = 1$. Putting $(\alpha = x - 1, \beta = tx)$, we find the second intersection point of the line $\beta = tx$ with the above quadratic rational curve (the first intersection point being $(\alpha = -1, \beta = 0)$).

Namely, we get

$$\begin{aligned} x^2(1 - t + t^2) + x(t - 2) & = 0, \\ x & = \frac{2 - t}{1 - t + t^2}, \quad \alpha = \frac{1 - t^2}{1 - t + t^2}, \quad \beta = \frac{2t - t^2}{1 - t + t^2}. \end{aligned}$$

To eliminate the (virtual) infinitely far asymptotic points from the $p + 1$ points corresponding to $t = 1, 2, \dots; \infty$, we have to solve the quadratic equation $1 - t + t^2 = 0$ (implying an easier cubic equation $t^3 = -1$).

Denoting $t = \rho^\lambda$ for a primitive residue $\rho \pmod{p}$, we obtain for λ the congruence

$$3\lambda \equiv (p-1)/2 \pmod{p-1}, \quad 6\lambda = (p-1)(2k+1).$$

If the integer $2k+1$ is divisible by 3, we get $t = \pm 1$. In this case $1 - t + t^2 = 0$ only for $p = 3$, $t = +1$.

If the integer $2k+1$ is not divisible by 3, we get $p-1 = 3c$, $p = 3c+1$. In this case our expression for λ provides two solutions t for our quadratic equation in $\mathbb{Z}_p$ (while in the case $p = 3c-1$ there are no such solutions).

Finally, we obtain for the number of the elements of the isotropy group the quantity, depending on the residue of p modulo 3:

$$|H| = p+1 \quad \text{for } p = 3c-1,$$

$$|H| = p-1 \quad \text{for } p = 3c+1.$$

Therefore, the number of solutions of the equation $A^3 = 1$ in G , which are conjugated to F , is

$$|G|/|H| = p(p \mp 1) \quad \text{for } p = 3c \mp 1.$$

It is thus equal to the number of trace -1 matrices in G . Therefore, all trace -1 matrices of G are conjugated to F . Theorem 12 is thus proved (for $p \neq 3$).

In the case $p = 3$ there are only 8 matrices with $A^3 = 1$, $A \neq 1$ (these are all the trace -1 matrices in $G = \text{SL}(2, \mathbb{Z}_3)$), 4 of them being conjugated to F and 4 to F^2 .

All this follows immediately from the known topological structure of the graph of the squaring operation. for this group G , the graph has 5 components

$$[G] = T_{1,1,6} \sqcup 4A_2,$$

in the notations of [2]. In this case the tree-component (the connected component containing the unity) is a subgroup, isomorphic to the group of quaternionic units $\{\pm 1, \pm i, \pm j, \pm k\}$, while the cycles (of length 2) of the 4 components A_2 are formed by the 8 elements of third order in G .

References

- [1] V.I. ARNOLD, A. AVEZ, *Problèmes ergodiques de la mécanique classique*. Paris, Gautier-Villars, 1967 (appendice 20).
- [2] V.I. ARNOLD, *Topology and statistics of arithmetic and algebraic formulae*, Russian Math. Surveys, 2003, **58**:4 (352), 3-28.

- [3] 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 (New Series) 1–41.
- [4] I. NEWTON, *Arithmetica Universalis*, Cambridge, 1707.
- [5] A. GIRARD, *Sur des découvertes nouvelles en algèbre*, Amsterdam, 1629.
- [6] V.I. ARNOLD, *On the arrangements of ovals of real plane algebraic curves, on involutions of four-dimensional smooth manifolds and on the arithmetics of integer quadratic forms*, Funct. Anal. and its Appl. (1971) **5**:3, 169–176.
- [7] V.I. ARNOLD AND SEVERAL AUTHORS, *Arnold Problems*. Springer-Phasis (2004).