Problems for the seminar, ICTP, 2007-2008

V. Arnold

January 21, 2008

Contents

1 Problem 1. Topological classification of polynomials

1. Consider the smooth mapping $f : \mathbb{R}^2 \to \mathbb{R}$, defined by a degree n polynomial in 2 variables. Such mappings are topologically different for different polynomials of degree n: $f = (x^2 + y^2 - 2)(2x^2 + y^2 - 1)$ is not equivalent to $g = (x^2 + y^2 - 2)(2x^2 + 4y^2 - 1)$.

The problem is to evaluate the growth rate of the number N of the topologically different types as a function of n: is it smaller than some polynomial an^b or greater than some exponent c^n ?.

2. One might restrict the class of polynomials, considering only the Morse polynomials, (whose critical points are all non degenerate and whose critical values are different), or only to Morse polynomials f, having $(n-1)^2$ real critical points. What is the growth rate of the number $N_2(n)$ of their topological types?

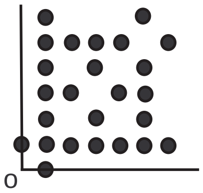
3. In the case, when the highest degree homogeneous part of f is positive definite (like $f_n = x^n + y^n$ for $n = 2m$), the polynomial mapping $f : \mathbb{R}^2 \to \mathbb{R}$ defines naturally a class of smooth mappings $\widehat{f} : S^2 \to \mathbb{R}$, having a Morse critical point $\infty \in S^2$ and attaining there a maximal value (one constructs the sphere S^2 adding to $\mathbb{R}^2$ one point ∞).

The number of topological classes of such smooth very good Morse functions on S^2, having T saddles, is growing asymptotically (for $T \to \infty$) as T^{2T} (for $T = 4$ there are 17746 classes).

It would be interesting to understand which part N_3 of these T^{2T} classes is realized by the above polynomials (of degree $n = 2m$ for $T = 2m(m-1)$). For $n = 4$ ($T = 4$) the conjectured answer N_3 is smaller than 1000.

2 Problem 2. Equipartition of indivisible integer vectors

A plane vector $(u, v) \in \mathbb{Z}^2$ is *divisible*, if the integers u and v have a common integer divisor $d > 1$ (in which case the vector, divided by d, belongs to the lattice $\mathbb{Z}^2$). Consider the set M of all indivisible vectors.



Let K be an angle at vertex 0 of the plane. Consider the set $M \cap K$ of the indivisible vectors, belonging to angle K. Denote by $N_M(R,K)$ the number of the points of set $M \cap K$ belonging to the disc $u^2+v^2 \leq R^2$ of radius R. Denote by $S(R,K)$ the total number of the integer points (divisible and not) of angle K inside this disc.

The (conjectural) equipartition property of the set M is the asymptotical independence of density ρ of M on the angle K:

$$\lim_{R\to\infty} \frac{N_M(R,K)}{S(R,K)} \to \rho,$$

the constant ρ being independent of the angle K.

Example. If $K=\mathbb{R}^2$ is the whole plane, the limiting density ρ exists and is known to be $\rho = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}$ (Euler discovered it also for $K=\mathbb{R}^m$). For the indivisible integer points in the n-dimensional space $\mathbb{Z}^m$ the asymptotical density (for $K=\mathbb{R}^m$) is $\rho = 1/\zeta(m)$, where ζ is the zeta-function

$$\zeta(m) = \sum_{n=1}^{\infty} \frac{1}{n^m} = \prod_{p=2}^{\infty} \frac{1}{1-\frac{1}{p^m}}$$

(product over all the primes, $p = 2,3,5,7,11,13,17,\dots$).

3 Problem 3. Geometrical progressions' fractional parts' equipartitions

Let $a>1$ be a real number. Consider the geometrical progression $(a, a^2, a^3, \dots)$ and replace each term t of it by its fractional part $\{t\}$ (where $t=[t]+\{t\}$, the integer part $[t]$ being an integer and $\{t\}$ belonging to $[0,1)$: $0 \leq \{t\} < 1$).

The equipartition of the resulting sequence $(\{a\}, \{a^2\}, \{a^3\}, \dots)$ means that for any interval $A \subset (0,1)$ the number $n_A(N)$ of members of this sequence $\{a^x\}$, for which $1 \leq x \leq N$, belonging to A, is growing with N the following way:

$$\lim_{N\to\infty} \frac{n_A(N)}{N} = (\text{length of } A),$$

whatever be the interval A.

The problem is to prove (or to disprove) the conjecture that *this equipartition property occurs for almost every real number a* (the exceptional values forming a set of measure zero).

The equipartition property might even be true for those real numbers a, all whose integer powers are irrational.

The (non rigorous) arguments for the geometric progressions' equipartition property, based on the physical theory of adiabatic invariants, is described in the book: V.Arnold, *Galois fields, their dynamics, statistics and projective geometry*, MCCME, 2005.

Remark. As an arithmetical model of the above equipartition conjecture one might consider the following (different) conjecture on the geometrical progressions of the residues $(a^x(\text{mod} M))$, where $x = 1, 2, 3, \dots, N$, numbers a and M being positive integers with no common divisor (greater than 1). Suppose, for simplicity, that M is a prime number.

Example: For $M = 97$, $N = 15$, $a = 3$ the resulting geometric progression of 15 residues (modulo 97) with base 3 is

$$3, 9, 27, 81, 49, 50, 53, 62, 89, 73, 25, 75, 31, 93, 85.$$

Conjecture. Such arithmetical progressions of the residues become asymptotically equipartitioned, provided that $M \to \infty$, the number N of terms of the progression being, say

$$aM < N < bM$$

for some constants a and b, $0 < a < b < 1$.

The asymptotical equipartition statement here is the statement that, for any interval $A \subset (0, 1)$ the number $n_A(N)$ among the N fractional parts

$$\left\{\frac{a^x}{M}\right\} \quad (\text{where } x = 1, \dots, N)$$

of the fractional parts, belonging to A, grows with N the following way:

$$\lim_{M \to \infty} \frac{n_A(N)}{N} = (\text{length of } A).$$

Example. Among the $N = 15$ residues modulo $M = 97$ above, 5 are smaller than $\frac{97}{3}$, and 6 bigger than $\frac{2}{3}97$; $5 = \frac{15}{3}$ and $6 \approx (1 - \frac{2}{3})15$.

4 Problem 4. Statistics of continued fractions of eigenvalues of matrices

Consider the matrices of order 2×2, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with integer elements and determinant $+1$, $A \in \mathrm{SL}(2, \mathbb{Z})$. Suppose that the eigenvalues are irrational real numbers. Their continued fractions are then periodic. Consider the periods for all such matrices A in the ball $a^2 + b^2 + c^2 + d^2 \leq R^2$ of radius R. Among the $N(R)$ elements of all these periods there are some elements equal to k; let N_k be their number:

$$N(R) = N_1(R) + N_2(R) + N_3(R) + \dots.$$

The statistics, required in this problem, is the calculation of the asymptotical frequency f_k of element k,

$$\lim_{R \to \infty} \frac{N_k(R)}{N(R)} = f_k.$$

Remark. The Gauss-Kuz'min frequency of element k for a random real number continued fraction statistics is

$$g_k = \frac{1}{\ln 2} \ln \left(1 + \frac{1}{k(k+2)} \right).$$

The frequencies f_k are perhaps different, and the problem is to evaluate how different they are. For instance, g_k decline as C/k^2 for $k \to \infty$. Is it also true for f_k?

Similar problems are also interesting for the eigenvalues of other types of matrices. For instance, one might consider the case $A \in \mathrm{End}(\mathbb{Z}^2)$ (of arbitrary 2×2 integer matrices) and the cases $A \in \mathrm{SL(n}, \mathbb{Z})$ or $A \in \mathrm{End}(\mathbb{Z}^n)$ of $n \times n$, where the periodicity fails.

Example. The mean lengths $\widehat{T}(R)$ of the periods of the continued fractions of the eigenvalues of the integer matrices $A \in \mathrm{End}(\mathbb{Z}^2)$ is growing with the radius R of the ball $a^2 + b^2 + c^2 + d^2 \leq R^2$ in the space of the matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer elements (the average growth rate being, it seems, $\widehat{T}(R) = CR$).

For the unimodular matrices $A \in \mathrm{SL}(2, \mathbb{Z})$ (for which $\det A = 1$) there is no such growth of the lengths of the periods of the continued fractions of the eigenvalues:

$$\lim_{R \to \infty} \widehat{T}_{\mathrm{SL}(2,\mathbb{Z})}(R) = 2.$$

This fact implies the difference of the frequencies F_k of the elements k of the continued fractions of the eigenvalues of these matrices $A \in \mathrm{SL}(2, \mathbb{Z})$ from the Gauss-Kuz'min frequencies g_k.

For instance, $g_1 = \frac{\ln(4/3)}{\ln 2} \approx 0,415$, while $F_1 = 0,5 > g_1$. It follows from the fact that the periods (for the continued fractions of the eigenvalues for $A \in \mathrm{SL}(2, \mathbb{Z})$) have mostly the form $[1, a]$, at least one half of their elements being 1.

The statistical problem 4 for these matrices is to find (at least empirically, for $R \sim 100$) the frequencies F_k (or at least the averaged frequencies $F_k(R)$ for $a^2 + b^2 + c^2 + d^2 \leq R^2$) of the elements of the (periods of) continued fractions for the eigenvalues of the matrices $A \in \mathrm{SL}(2, \mathbb{Z})$.

5 Problem 5. Growth rate of elements of periodic continued fractions

Consider the quadratic equation

$$x^2 + px + q = 0$$

with integer coefficients p and q, having real roots ($p^2 \geq 4q$).

The continued fractions of these roots are periodic. Consider the mean value of the elements of the period, consisting of elements $(a_1, \dots, a_T)$:

$$\widehat{a} = \frac{a_1 + a_2 + \dots a_T}{T}.$$

Numerical experiments show that $\widehat{a}$ is generally growing, when the coefficients p and q are growing.

To formulate the problem of the of study of this growth rigorously, construct the averaged mean values

$$A(R) = \frac{\left(\sum_{p^2+q^2 \le R^2} \widehat{a}(p,q)\right)}{N(R)}$$

where $N(R)$ is the number of the summands, that is of the integer points (p,q), for which $p^2 \ge 4q$, in the ball $p^2 + q^2 \le R^2$.

The problem is to evaluate (at least empirically) the growth rate of $A(R)$: is A greater than CR^α for some positive C, α ? Or is it smaller than some $C(\ln R)^\alpha$? Numerical experiments for $R \sim 1000$ are already quite interesting.

Similar problems are also interesting for other families of the integer coefficients quadratic equations.

Example. $rx^2 + px + q = 0$ (with $(r,p,q) \in \mathbb{Z}^3$), $x^2 + px + 1 = 0$ ($p \in \mathbb{Z}$), $rx^2 + q = 0$ ($(r,q) \in \mathbb{Z}^2$).

Example. The continued fractions of the roots of the equations $x^2 + px + 1 = 0$ (which are the characteristic equations of the matrices $A \in \mathrm{SL}(2,\mathbb{Z})$ discussed in Problem 4) have, for $|p| \ge 4$ the periods with $T(p,1) = 2$ elements, $[1,a]$, $a \sim |p|$.

It follows that the averaged element of the continued fractions behave for $|p| \le R$, like $R/4$, which seems to be a much faster growth than the (unknown, rigorously speaking), growth for the general 2×2 integer elements matrices $A \in \mathrm{End}(\mathbb{Z}^2)$.

6 Problem 6. Periods of geometrical progressions of residues

The geometrical progression of residues modulo an integer n $(1, a, a^2, \dots (\mathrm{mod}\ n))$ with integer base n is periodic (Fermat's little theorem for prime values of n extended by Euler to arbitrary integers n).

The period's length $T(a,n)$ is a peculiar function, and the problem is to find its asymptotical behaviour for $n \to \infty$.

Example 1. For the prime values $n = 17$ and base $a = 2$ the progression starts from

$$(1, 2, 4, 8, 16, 15, 13, 9, 1, 2, \dots)$$

and hence has the period's length $T(2,17) = 8$.

For $n = 100$ and base $a = 2$ the progression starts from

$$(1, 2, 4, 8, 16, 32, 64, 28, 12, 24, 48, 96, 92, 84, 68, 36, 72, 44, 88, 76, 52, 4, 8, \dots)$$

and hence has the period's length $T(2,100) = 20$.

Euler proved that $T(a,n)$ is a divisor of the number $\varphi(n)$ of the residues modulo n, which are mutually prime to n. The values of $\varphi(17)$ and $\varphi(100)$ are 16 and 40 (since for a prime p one has $\varphi(p) = p - 1$ and $\varphi(p^a) = (p-1)p^{a-1}$, while $\varphi(uv) = \varphi(u)\varphi(v)$ for mutually primes u and v).

To find the asymptotical behaviour of T, one starts with the study of the asymptotical behaviour of Euler's function φ. Euler calculated the Cesaro mean value $\widehat{\varphi}(n) = \frac{1}{n}\sum_{m=1}^{n} \varphi(m)$ to grow asymptotically rather regularly as cn, where the constant $c = 1/\zeta(2) = 6/\pi^2$ is approximatively $2/3$ (while the particular values of φ oscillate chaotically):

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
φ	1	2	2	4	2	6	4	6	4	10	4	12	6	8	8

Next one tried to evaluate the growth rate of the divisors of large numbers.

The number $\tau(n)$ of the integer divisors of n grows chaotically:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
τ	1	2	2	3	2	4	2	4	3	4	2	6	2	4	4	5

Dirichlet calculated the simple asymptotical behaviour of the Cesaro means:

$$\widehat{\tau}(n) = \frac{1}{n}\sum_{m=1}^{n} \tau(m) \sim \ln n$$

(the sign $\sim$ means that the ratio of the left hand side to the right hand side tends to 1 for $n \to \infty$).

Similarly, the Cesaro means $\widehat{\sigma}$ of the sums $\sigma(n)$ of the integer divisors of n regularize the chaotical oscillations of σ:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
σ	1	3	4	7	6	12	8	15	13	18	12	28	14	24	24	31

while $\widehat{\sigma}(n) = \frac{1}{n}\sum_{m=1}^{n} \sigma(n) \sim c_1 n$, the coefficient being $c_1 = \zeta(2) = \pi^2/6 \approx \frac{3}{2}$.

The mean divisor of integer n is defined as $d(n) = \sigma(n)/\tau(n)$. One is tempted to guess that the Cesaro means of the ratio behave like the ratio of the Cesaro means of the numerator and the denominator,

$$\frac{\widehat{\sigma}(n)}{\widehat{\tau}(n)} \approx \frac{c_1 n}{\ln n}.$$

However the true asymptotical behaviour of the Cesaro averaged mean divisors, $\widehat{d}(n) = \frac{1}{n}\sum_{m=1}^{n} d(n)$ does not coincide with the above ratio: it is much higher, namely

$$\widehat{d}(n) \sim \frac{c_2 n}{\sqrt{\ln n}}.$$

Of course, the fact that the mean value of the ratio differs from the ratio of the mean values is not too astonishing and the difference is even not too big for reasonable values of n ($< 10^6$).

Uniting the preceding information, one is tempted to predict the following mean behaviour of the periods $T(a, n)$ (say, for some fixed a, even for $a = 2$): T is expected to be the mean divisor of the mean Euler function value

$$(*) \qquad \widehat{d}(\widehat{\varphi}(n)) \sim \frac{c_2\widehat{\varphi}(n)}{\sqrt{\ln \widehat{\varphi}(n)}} \sim \frac{c_2 cn}{\sqrt{\ln cn}} \sim \frac{c_3 n}{\sqrt{\ln n}}$$

however, the empirical studies (for $n \approx 10^{10}$) showed a different behaviour, conjecturally

$$\widehat{T}(a, n) \sim \frac{c_4 n}{\ln n}$$

(the multipliers like $\ln\ln n$ being empirically constant even for $n \sim 10^{10}$).

These empirical observations show that something is wrong in the naive reasoning $(*)$. The problem is to discover (at least empirically) what is practically happening. Theoretically the following 3 main events are possible:

1) The asymptotics of the mean divisors $d(\varphi(n))$ of the values of the Euler function might differ from the asymptotics of the mean divisors of arbitrary integers m at places $m = \varphi(n)$, due to the difference of the arithmetics of the numbers

$$n = \prod p_s^{a_s} \quad \text{and} \quad m = \prod \left((p_s - 1)p_s^{a_s - 1}\right).$$

2) The divisor $T(a, n)$ could be not at all the mean divisor $d(\varphi(n))$, but one of the other divisors, making wrong the evaluation of $T(a, n)$ in terms of $\widehat{d(\varphi(n))}$).

3) This Cesaro mean values might also behave differently than $\widehat{d}(\widehat{\varphi}(n)))$.

The problem 6 requires both the computation of the true asymptotics of the Cesaro means $\widehat{T}(a, n)$ of the periods' lengths $T(a, n)$, and the evaluation of the differences between the terms of reasoning $(*)$, explained above in the description of 3 events.

7 Problem 7. Kolmogorov's distributions

The *Kolmogorov's distribution function* Φ equals

$$\Phi(\Lambda) = \sum_{k=-\infty}^{k=+\infty} (-1)^k \mathrm{e}^{-2k^2\Lambda^2} \quad \text{(where } \Lambda > 0\text{).} \tag{1}$$

It grows from $\Phi(0) = 0$ to $\Phi(+\infty) = 1$. Kolmogorov claimed that for $\Lambda \to 0$ it has the flat asymptotics (all the derivatives vanishing at 0)

$$\Phi(\Lambda) \sim \frac{\sqrt{2\pi}}{\Lambda} \mathrm{e}^{-\frac{\pi^2}{8\Lambda^2}}.$$

The **first question** is to explain this asymptotical formula, at least to prove it (but also to relate it to the Brownian motions behaviour).

The **second question** is to extend the Kolmogorov's formula (1) to the following more general situation than the theory of empirical observations of a real random variable x having a continuous distribution function

$$S(X) = (\text{probability of the appearance of value } x \leq X).$$

Namely, for n observed values (ordered by the growth order) $x_1 \leq x_2 \leq \cdots \leq x_n$ Kolmogorov compared their *empirical counting function* $C_n(X)$ =(number of the values $x_j \leq X$) with the *theoretical counting function* $C_0(X) = nS(X)$. He defined their deviation as the uniform convergence norm

$$F = \sup_X |C_n(X) - C_0(X)| \tag{2}$$

and defined his *stochasticity parameter value* λ_n as

$$\lambda_n = F/\sqrt{n}.$$

If $(x_1, \ldots, x_n)$ are mutually independent observed values of random variable x, the quantity λ_n is itself a random number. Kolmogorov proved (in his paper in Italian *Sulla determinazione empirica di una legge di distribuzione*, in Giorn. Ist. Ital. Attuar,(1933),Vol. **4**,1, 83-91) that the random variable λ_n has the probability distribution converging for $n \to \infty$ to the distribution, whose distribution function is his function Φ (the distribution function of λ_n converging to Φ uniformly).

The second question of Problem 7 requires the following generalizations of this Kolmogorov's theorem: is it possible to replace the real random variable $x \in \mathbb{R}$ in the Kolmogorov's theorem by the random variables with other values, for instance in the following 4 cases:

$$a)\ x \in \mathbb{Z}; \quad b)\ x \in \{1, 2, \ldots, N\}; \quad c)\ x \in S^1; \quad d)\ x \in \mathbb{Z}_N (= \mathbb{Z}/N\mathbb{Z}).$$

One might conjecture that such extensions are possible, and even that the resulting distribution function would approach the Kolmogorov's expression (1) for $N \to \infty$ (in cases *b*) and *d*)).

The **third question** on Φ is to compare the distribution of λ_n (say, for the case $x \in \mathbb{Z}_N$) of n random vertices of regular N-gon) to the statistics of the distances between the neighbours

$$\Sigma_n = |x_1 - x_2|^2 + |x_2 - x_3|^2 + \cdots + |x_n - x_1|^2$$

(where $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n < N$ and the distance $|x - y|$ is the length of the shortest path between points x and y of the finite $\mathbb{Z}_N$).

Example: for the $n = 3$ points 1,3,6 of $\mathbb{Z}_7$ the quantity Σ has the value $|1-3|^2 + |6-3|^2 + |1-6|^2 = 17$.

One expects some inequalities between the values of λ_n and Σ_n (and with its mathematical expectation). Even the experimental studies of these interrelations (say, with $N = 100$ or 1000) would provide interesting information (and conjectures).

The **fourth question** on the Kolmogorov's distribution Φ is related to the difficulty of the determination of the theoretical counting function C_0 (see, for instance, its discussion for the geometrical progressions residues case in Problem 6).

If C_0 is unknown, one might use a similar device the following way. Consider, instead of one sample $(x_1 \leq x_2 \leq \cdots \leq x_n)$ of n observations of random variable x, two (independent) samples. Denote the second sample $(x'_1 \leq x'_2 \leq \cdots \leq x'_n)$. The pair of samples define the pair of the counting functions C_n and C'_n.

To evaluate the stochasticity parameter value one might replace the distance (2) by the similar distance

$$F^* = \sup_X |C_n(X) - C'_n(X)|, \qquad \lambda^*_n = F^*/\sqrt{n}.$$

The problem is to relate the distribution Φ^*_n of this random variable λ^*_n with the Kolmogorov's distribution Φ: are Φ^*_n converging to some universal limit Φ^* for $n \to \infty$?

The natural conjecture would be the similarity of the random variable λ^*_n to $C\lambda_n$ for some constant coefficient C (which should obviously belong to the interval $1 \leq C \leq 2$, the Euclidean guess $\sqrt{2}$ being even possible).

Similar questions are also natural for the random variables, whose values belong to $\mathbb{Z}$, to $\{1, \ldots, N\}$, to S^1 and to $\mathbb{Z}_N$ (as in question 2).

Some information on the preceding question (on the relations of Φ^*_n to Φ) might, it seems, be found in the old article:

N.V. Smirnov, *On the difference between the empirical curves of a distribution for two independent series*, Moscow University Bulletin, 1939, N.2, 3-14 (in Russian).

It would be nice at least to report to the Seminar the results of this paper (whose author was highly appreciated by Kolmogorov).

8 Problem 8. Stochasticity degree of arithmetical progressions of fractional parts.

Consider the arithmetical progressions of the fractional parts of the n numbers $(a, 2a, 3a, 4a, \ldots, na)$ (each real number t is the sum of its integer part $[t] \in \mathbb{Z}$ and its fractional party $\{t\}$, which belong to the interval $0 \le \{t\} < 1$).

The n numbers $\{xa\}$ $(x = 1, 2, \ldots, n)$ form a subset of the interval $[0, 1)$, and their Kolmogorov's stochasticity parameter value λ_n is well defined (in Problem 7 above).

The question is to understand the typical behaviour of the values λ_n for $n \to \infty$.

Example 1. For all rational values of the step length a of the arithmetical progression the stochasticity parameter values tend to 0: $\lim_{n\to\infty} \lambda_n = 0$.

Example 2. There exist irrational values of the step's length a such that λ_n attains (for suitable time moments n) arbitrarily large values ($\lambda_n > K$) (and hence do not tend to 0 for $n \to \infty$).

The problem is to decide which behaviour is typical. For instance, the set of those step's lengths a, for which $\lambda_n \to 0$, is either of Lebesgue measure 0, or of full Lebesgue measure (its complement being of Lebesgue measure 0).

Similarly, for any asymptotical behaviour of the sequence of the values λ_n for $n \to \infty$ (like: $\lambda_n \nrightarrow 0$, λ_n are unbounded, $\lambda_n \to \infty$ and so on) the set of those lengths a of the step of the arithmetical progression of the fractional parts, for which this behaviour takes place, contains either almost all values $a \in \mathbb{R}$ or almost none. It means that either the Lebesgue measure of the complement of the set or the Lebesgue measure of the set itself is equal to zero: in the first case the set contains almost all the values and in the second almost none.

The examples 1 and 2, discussed above, are proved to take place at least for some sets of Lebesgue measure 0 of values $a \in \mathbb{R}$ (the subset of all the rational numbers $\mathbb{Q} \subset \mathbb{R}$, being countable, has the zero Lebesgue measure).

But whether a typical arithmetical progression of fractional parts is random or not is not clear, and even the possibility of the behaviour "$\lambda_n \to 0$ for almost every value of a" is not excluded (while I would rather expect that the behaviour "$\lambda_n \nrightarrow 0$" is typical, or even "there exists such a subsequence n_j, such that $\lambda_{n_j} \to \infty$".

9 Problem 9. Is a generic geometrical progression of fractional parts random?

Starting from a real number $a > 1$, consider the geometrical progression of the fractional parts of its powers, $\{a\}, \{a^2\}, \{a^3\}, \ldots, \{a^n\}$ (like in problem 3).

For this finite set of n elements of interval $[0, 1)$ construct the Kolmogorov's stochasticity parameter value λ_n (as in problem 7).

One might presuppose here the uniform distribution conjecture to be true, to choose the "theoretical distribution", needed to define λ_n.

Or, otherwise, one might replace it by the version λ_n^* (defined by a pair of independent progressions, say, studying

$$\{aA\}, \{a^2A\}, \{a^3A\}, \ldots, \{a^nA\} \tag{3}$$

for two different choices of A).

The problem is to study the typical behaviour of the values λ_n for $n \to \infty$.

Example. For any rational base value a the Kolmogorov's stochasticity parameters λ_n of the geometrical progressions of fractional parts (3) tend to 0 for $n \to \infty$. This follows from the Fermat and Euler periodicity theorem (see Problem 6).

However, the experiments with the shorter geometrical progressions of residues modulo N

$$(aA, a^2A, a^3A, \ldots, a^nA)(\text{mod } N) \tag{4}$$

showed rather the randomness of these sets of n residues modulo N (measuring randomness by the values of the Kolmogorov's stochasticity parameter λ_n, discussed in Problem 7).

Namely, denote by T $(= T(a, N))$ the (shortest) period of sequence (2) of residues. The "shortness" condition is the inequality

$$pT \leq n \leq qT \tag{5}$$

for some fixed constants $0 < p < q < 1$.

Example. For $n = 31$ and $a = 3$ the period is $T(3, 31) = 30$, and the 15 terms of geometrical progression (of resides mod 31) are

$$(3, 9, 27, 19, 16, 17, 20, 29, 25, 13, 8, 24, 10, 30, 28).$$

The Kolmogorov stochasticity parameter is $\lambda_{15} = 3/\sqrt{15} \approx 0,76$. (reasonably close to the mean value $\Lambda \approx 0,87$ for the Kolmogorov's distribution Φ (of Problem 7).

The conjecture is that in the shortness situation (5) the values of the stochasticity parameter λ_n of the "short" geometrical progression of residues (4) with different initial points $A \in \mathbb{Z}_N$ attain such values $\lambda_n(A)$ that their distribution tends to the Kolmogorov's distribution Φ (of Problem 7) for $n \to \infty$, $N \to \infty$.

This is the arithmetical version of the conjectural generic randomness of the geometrical progressions (3) of fractional parts (for almost every base number a): for instance, the conjecture claims that the cases $\lambda_n \to 0$ for $n \to \infty$ and $\lambda_n \to \infty$ for $n \to \infty$ are the exceptions, realized only for the base numbers a forming a Lebesgue measure 0 set.

10 Problem 10. Prime numbers distribution's randomness

The residues modulo N of the n successive prime numbers form strange sets, which look randomly. Say, for the 21 prime numbers $100 < p < 200$ we obtain the following $n = 21$ residues modulo $N = 100$:

$$1, 3, 7, 9, 13, 27, 31, 37, 39, 49, 51, 57, 63, 67, 73, 79, 81, 91, 93, 97, 99$$

(corresponding to the primes $101, 103, 107, \ldots, 199$).

The Kolmogorov's stochasticity parameter λ_n, calculated from this sequence (for the Legendre distribution of the prime numbers, whose density at n is inversely proportional to $\ln n$) is approximatively $\lambda_{21} \approx 0,5$.

The Kolmogorov's distribution Φ (of Problem 7) provides the small probability $\Phi(0,5) \approx 0,07$ of the randomness for this sequence of 21 residues.

The Problem is to understand the behaviour of the stochasticity parameter λ_n for the other sequences of successive primes' residues.

For instance, one might consider the primes in the intervals $100a < p < 100(a+1)$ (modulo $N = 100$), or in the intervals $\{p_{k+1} < \cdots < p_{k+n}\}(\text{mod } N)$ containing the modulo N residues of n successive primes (say for $n = 20$ and $N = 100$, $k \to \infty$, or for $N = 2n$, $k = 2n$, $n \to \infty$): what would be the behaviour of the stochasticity parameter for larger n (empirically) or even for $n \to \infty$ (theoretically)?

Even the empirical study of these questions (for few millions of primes) would be interesting: it might provide some conjectures on the randomness of the distribution of the prime numbers, even in the case, where these conjectures' proofs would wait some centuries (like it happened to the Legendre distribution, quoted above).

Among others randomly looking objects of number theory one might consider the quadratic residues x distributions (in $\mathbb{Z}_N$). I have no doubts that they are distributed uniformly (with density 1/2) – say, the set of pairs (x, N) is uniformly distributed in the angle $0 \leq x \leq N$ of the plane. But it would be interesting to see whether they are really random, computing the Kolmogorov's parameters values for these subsets of $\mathbb{Z}_N$. The conjecture is that the quadratic residues are genuinely randomly chosen residues in the sense of the Kolmogorov stochasticity parameter.

11 Problem 11. Algorithmic unsolvability of problems of higher dimensional continued fractions

The n-dimensional continued fraction is defined by any open simplicial cone K, bounded by n hyperplanes, containing the origin $0 \in \mathbb{R}^n$. It measures the interrelations of this cone K to the standard sublattice $\mathbb{Z}^n$ (of the points with integer coordinates).

The additive semigroup $P = K \cap \mathbb{Z}^n$ defines its convex hull $\widehat{P}$ (which is the minimal convex body, containing P). The boundary of this convex body

$$S = \partial \widehat{P},$$

is an (infinite) polyhedral hypersurface, called the *sail of* K (and being the higher dimensional version of the ordinary continued fractions).

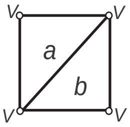
Consider now an integer elements matrix of order $n \times n$ $A \in \mathrm{SL(n, Z)}$, defining a linear operator $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ mapping the lattice of the integer points $\mathbb{Z}^n$ onto itself isomorphically.

Suppose that A has n different positive eigenvalues $(\lambda_1, \ldots, \lambda_n)$. The n corresponding invariant hyperplanes (generated each by $(n-1)$ eigenvectors) subdivide $\mathbb{R}^n$ into 2^n simplicial invariant cones. Choose one of these cones, K.

Operator A sends K, $\mathbb{Z}^n$, P, $\widehat{P}$ and the sail $S = \partial P$ onto themselves isomorphically. The group of such isomorphisms, containing those matrices $B \in \mathrm{SL(n, \mathbb{Z})}$ which commute with A and preserve K, having the same eigenvectors as A, form the commutative symmetry group G of the sail S, isomorphic to $\mathbb{Z}^{n-1}$ (as Dirichlet proved).

Consider the orbits space S/G. Topologically the action of group $G \approx \mathbb{Z}^{n-1}$ on $S \approx \mathbb{R}^{n-1}$ is homemorphic to the action of the translations of space $\mathbb{R}^{n-1}$ by its integer vectors, and therefore the quotient space $S/G \approx T^{n-1}$ is the $(n-1)$-torus.

This torus inherits from the sail S its polyhedral structure: T^{n-1} is decomposed into "convex faces", intersecting along convex boundaries of the faces and so on, so it will be called "*the triangulation* (of the periodic n-dimensional continued fraction of operator A)".

Example. The "3-dimensional golden ratio" matrix $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ provides the triangulation  of 2-torus into two triangles a and b, three segments and one vertex V.

Which other triangulations are possible? Some triangulations are not realizable by the continued fractions of any operator A.

No algorithm to decide whether a given triangulation of T^2 (of T^n) is isomorphic to the triangulation generated by the continued fraction of any operator $A \in \mathrm{SL}(2,\mathbb{Z})$ (of $\mathrm{SL}(\mathrm{n}+1,\mathbb{Z})$) is known.

The problem 11 is to determine whether there exists any of such algorithm. The conjecture is that it does not exist.

This problem for the triangulations of the tori T^n might be easier than for the particular case of T^2, since it might be possible that, say, starting from T^{17} there exists no such algorithm, while for T^2 the problem is algorithmically solvable (by a very difficult algorithm).

12 Problem 12. Periods of continued fractions of roots of quadratic equations

Consider the quadratic equation with integer coefficients

$$x^2 + px + q = 0. \tag{6}$$

According to a theorem of Lagrange, the continued fractions developments of the (real) irrational roots of such equations are periodical.

Example. For the equation

$$x^2 + x = 28$$

the period of the continued fraction of x is of length $T = 7$,

$$\frac{\sqrt{113}-1}{2} = 4 + \frac{1}{1+\frac{1}{4+\ldots}} = [4 + [1,4,2,2,4,1,9]]$$

($a_{k+7} = a_k$ for $k \geq 1$).

The problem is to describe *which sequences* $[a_{k+1}, \ldots a_{k+T}]$ *of* T *natural numbers are the periods of the continued fractions of roots of equation* (6) *with integer coefficients* $(p,q) \in \mathbb{Z}^2$.

One wishes to find some restrictions, whose fulfillment might be easily verified.

Example. One peculiar property of all such sequences is the *palindromicity* property of the period.

The *palindrome* is a sequence which can be read back (from right to left): it remains the same. The infinite periodic sequence of the preceding example

$$(\dots,1,4,2,2,4,1,9,1,4,2,2,4,1,9,1,4,2,2,1,4,1,9,\dots$$

coincides with the same sequence read back, which is

$$(\dots,9,1,4,2,2,4,1,9,1,4,2,2,4,1,9,1,4,2,2,1,4,1,\dots$$

The palindromicity property is verified by the continued fractions of the roots of all quadratic equations (6) with integer coefficients. It is also verified by some other quadratic irrational numbers' continued fractions, including all the (irrational) square roots of rational numbers.

Example.

$$\sqrt{11/8} = [1 + [5,1,3,1,5,2]]$$

is a periodic continued fraction of period, consisting of $T = 6$ elements. This period is palindromic:

$$\dots,5,1,3,1,5,2,5,1,3,1,5,2,5,1,3,1,3,1,5,2,\dots$$

is the same (infinite) periodic sequence as the inverse one,

$$\dots,2,5,1,3,1,5,2,5,1,3,1,5,2,5,1,3,1,3,1,5,\dots.$$

The palindromicity property of continued fractions of roots of equations (6) and

$$rx^2 + q = 0. \tag{7}$$

is a non evident fact. The periods of the continued fractions of the roots of general quadratic equations with integer coefficients $(p,q,r) \in \mathbb{Z}^3$

$$rx^2 + px + q = 0. \tag{8}$$

are just arbitrary finite sequences of integers.

It is known that the sequences forming the periods of the continued fractions of the roots of equation (6) or of the roots of equation (6) (or of the roots of equation (7)) have a lot of peculiar properties, distinguishing them from the periods of the continued fractions of other equations (9), but these properties remain unknown (except their palindromicity).

The statistics of the sequences, forming the periods of the continued fractions of roots of equations (6) and (7) also might differ from the statistics of the same sequences as forming subsets of these periods (which coincides with the Gauss-Kuz'min statistics of Problem 4): the frequency of the period $[1, 2, 3]$ among the periods might be quite different from the part (1,2,3) of longer periods.

Remark. I think that the quantity of the triples $(r, p, q) \in \mathbb{Z}^3$ for which the periodic continued fraction of the root of equation

$$rx^2 + px + q = 0$$

is palindromic is small, in the sense that the number of such triples in the ball $r^2+p^2+q^2 \leq R^2$ form a small part of the whole set of integer points in this ball for $R \to \infty$. But this conjecture is neither proved nor confirmed by the empirical data.

13 Problem 13. Statistics of lengths of periods of continued fractions of quadratic irrational numbers

The problem is either to confirm or to reject the empirically discovered conjecture on the growth rate of the lengths $T(p, q)$ of the periods of the periodic continued fractions, for the root(s) of the quadratic equations with integer coefficients

$$x^2 + px + q = 0 \qquad (p, q) \in \mathbb{Z}^2. \tag{9}$$

To evaluate the growth rate one calculates the values of the period's lengths $T(p, q)$ for those $N(R)$ integer points of the ball $p^2 + q^2 \leq R^2$ of radius R, where the roots are real: $\Delta > 0$, where Δ is the discriminant value

$$\Delta(p, q) = p^2 - 4q \tag{10}$$

The mean values

$$\widehat{T}(R) = \frac{\left(\sum_{p^2+q^2\leq R^2} T(p, q)\right)}{N}$$

behave more regularly than the "chaotically oscillating" function $T(p, q)$.

The empirical conclusion is the *linear growth rate conjecture*

$$\widehat{T}(R) \sim cR \qquad (\text{with } c \approx 0,15).$$

Examples.

R	20	40	60	80	100
$\widehat{T}$	$3,8$	$6,9$	$9,8$	$12,5$	$15,1$

Of course, the numerical continuation of these empirical calculations would be a useful thing, while even at present I have no doubts that the asymptotics ought be correct.

Remark. The proof might be difficult, and I am not including its requirement into the problems formulation.

However, some parts of the statement might be easier. For instance, one of such parts is the conjectural estimation from above

$$\widehat{T}(R) < \mathrm{const} R$$

(or even $T(p,q) < \mathrm{const} R$ for $p^2+q^2 \leq R^2$).

The value $T(p,q)$ equals 0 for the equations with rational roots, where $\Delta(p,q)$ is a square of an integer. For instance, $T(p,q)$ vanishes at those points (p,q), which belong to the double roots parabola, $\Delta = 0$ (where $q = -4p^2$).

This fact suggests that $T(p,q)$ might behave like some function of $\Delta(p,q)$, at least asymptotically. The constancy of the value $T(p,q)$ along the parabolas $\Delta =$const follows from the action of the shift $x \mapsto x+1$ on the coefficients p and q of equation (9).

The empirical study suggests that *this function might behave like* $T \approx \sqrt{\Delta}$.

Therefore the problem 13 includes the study (at least at the empirical level, say, for $R \leq 1000$) of the ratio $T(p,q)/\sqrt{\Delta(p,q)}$.

This study might provide the information on the mean values of such ratios (along the discs of radius R), on the distribution of other values and on the behaviour of the " level lines", where

$$T(p,q)/\sqrt{\Delta(p,q)} = \mathrm{const},$$

for different constants.

14 Problem 14. Random matrices' characteristic polynomials distributions

Consider some natural set of matrices with integer elements, like the set of all $n \times n$ matrices with integer elements, $\mathrm{End}(\mathbb{Z}^{\mathrm{n}})$, or the group $\mathrm{SL(n,\mathbb{Z})}$ of the unimodular $n \times n$ matrices A with integer elements (where $\det A = 1$).

In the radius $\sqrt{M}$ " ball", consisting of the matrices of our set, whose sum of the squares of the elements does not exceed M, there is a finite set of matrices. Each of these matrices has its characteristic equation with integer coefficients

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0.$$

The problem is to study the behaviour (for $M \to \infty$) of the numbers $N(a_1, \dots, a_n; M)$ of those matrices A, belonging to the ball of radius $\sqrt{M}$, whose characteristic equations have fixed coefficients.

Example. For the case $A \in \mathrm{SL}(2, \mathbb{Z})$ the characteristic equation has the form

$$x^2 + px + 1 = 0, \qquad p = -\mathrm{tr} A,$$

and the problem is to study the distribution of the traces $a + d$ of the matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, for which $ad - bc = 1$ and $a^2 + b^2 + c^2 + d^2 \leq M$.

Is it true, in this case, that the growth rates of the numbers of matrices of different traces p ($p \leq 2\sqrt{M}$) for $M \to \infty$ are similar?

The total number of matrices of $\mathrm{SL}(2, \mathbb{Z})$ in the ball of radius $\sqrt{M}$ grows like LM for some constant L, and therefore the equipartition of the traces $-p$ (in the interval $|p| \leq 2\sqrt{M}$) would provide the growth rate of the number $N_p(M)$ of matrices of fixed trace $-p$ of the form $N_p(M) \sim \sqrt{M}$.

One can prove, overcoming some difficulties, the inequalities $N_p(M) \leq L_p M^{5/6}$, and also, for $p = 2$, $N_2(M) \leq K\sqrt{M} \ln M$.

But what are the genuine asymptotics of $N_p(M)$ is not clear even for the case $n = 2$ of 2×2 matrices in $\mathrm{SL}(2, \mathbb{Z})$ of different traces $-p$.

For the arbitrary determinants second order matrices case, $A \in \mathrm{End}\mathbb{Z}^2$, the total number of matrices in the radius $\sqrt{M}$ ball grows like its volume, M^2, and the question is whether $N_p(M)$ behave like the value, suggested by the equipartition of the traces, $M^{3/2}$.

For the higher order matrices A ($n > 2$) the characteristic polynomials have more coefficients, and their distribution statistics for $M \to \infty$ is more complicated, but one might suppose it to be similar to the Euclidean geometry distribution of the masses for the preimages of different points $a \in \mathbb{R}^n$ in the image space of the mapping, sending each matrix A to its characteristic polynomial, $(a_1, \dots, a_n) \in \mathbb{R}^n$.