

Avdiyev Marat Alexandrovich

chief

‘Siberian Mediation Center’ Union

Surgut, KMAD – Ugra

FERMAT'S LAST THEOREM AND ABC-CONJECTURE IN THE SCHOOL OF THE XXI CENTURY

Abstract: *in 1637, Pierre de Fermat wrote at the margins of Diophantus' Arithmetica that he had found a truly wonderful proof of the insolvability of the Diophantine equation $a^n + b^n = c^n$, where $n > 2$, but the narrow margins of the books did not allow him to give the full proof. Is there a short and easy way to prove Fermat's Last Theorem? The following ABC conjecture states that for three co-prime numbers A , B , and C which satisfy $A + B = C$, the product of the prime factors of ABC is usually not much less than C . Both theorems are formulated very simply, but are extremely difficult to prove. Hundreds of pages have been spent by eminent mathematicians of Western world searching for proofs, and the search for proofs continues. The author found new methods of proof that are generally understandable, even to schoolchildren on the basis of a synthesis of several sciences, including physics. Number theory plays an interesting role in pedagogy.*

Keywords: *Fermat's Last theorem, ABC-conjecture, entropy, physics, symmetry, pedagogy, universe.*

Авдыев Марат Александрович

аспирант, директор

Союз «Сибирский Центр медиации»

г. Сургут, ХМАО – Югра

DOI 10.21661/r-561630

ВЕЛИКАЯ ТЕОРЕМА ФЕРМА И АВС-ГИПОТЕЗА В ШКОЛЕ XXI ВЕКА

Аннотация: *в 1637 году Пьер де Ферма написал на полях «Диофантовой арифметики», что он нашел воистину чудесное доказательство неразрешимости уравнения Диофанта $a^n + b^n = c^n$, где натуральный*

показатель степени $n > 2$, но узкие поля книг не позволили ему привести полное доказательство. Существует ли короткий и простой способ доказать последнюю теорему Ферма? Следующая ABC гипотеза утверждает, что для трёх взаимно-простых чисел A , B и C , удовлетворяющих соотношению $A + B = C$, произведение простых делителей A , B и C обычно ненамного меньше C . Обе теоремы формулируются очень просто, но чрезвычайно сложно доказываются. Сотни страниц были потрачены выдающимися математиками Западного мира на поиск доказательств, и процесс поиска доказательств продолжается. В этой работе автор приводит методы доказательства, понятные школьникам и студентам на основе синтеза ряда наук, включая физику. Теория чисел играет интересную роль в педагогике.

Ключевые слова: Великая теорема Ферма, ABC-гипотеза, энтропия, физика, симметрия, педагогика, Вселенная.

Chapter I. Fermat's Last Theorem on the School Globe

Introduction. Prejudice

There is a widespread prejudice in Western world that there is no concise proof of Fermat's Last Theorem. The author is convinced that Pierre de Fermat's statement about the ‘truly wonderful proof’ he found is not a figure of speech, but his words should be taken literally by a schoolboy of the XXI century who is familiar with the basics of logic, geometry, dimension theory, set theory. This knowledge is taught at school, without mathematical formalism, mainly in physics, computer science and other school subjects. From the point of view of the engineering approach, the listed scientific disciplines can be replaced by 3D modeling, design and generalization of results to a multidimensional space using the axioms of Euclid and topology. Students study the layered structure of the Earth's crust, the Pythagorean theorem, the scalar product of vectors, the law of conservation of matter in chemical reactions, congruence of shapes, symmetry. A number of basic concepts from higher mathematics, such as the concept of *uniformity*, *isotropy of space*, are taught in science, work and physics classes. The principles of *symmetry* can be used to solve problems that generally require knowledge

of the basics of integration and differentiation, and of solving systems of partial differential equations. Finally, lessons in drawing and astronomy, the study of World Heritage monuments, the basics of logic and philosophy will all help in the search for proofs of the theorem from the point of view of school knowledge. An interdisciplinary approach has its advantages. From these points of view, the extraordinary beauty and conciseness of the theorem's formulation helps to focus attention and stimulate interest in creative learning among schoolchildren and students.

Fermat's Last Theorem was formulated by Pierre de Fermat in 1637, it states that the Diophantine equation is true:

$$a^n + b^n = c^n \quad (1.1)$$

has no solutions in whole numbers \mathbb{Z} , except for zero values, for $n > 2$. The case degree of two is known in the school course under the name Theorem Pythagoras. Euler in 1770 proved Theorem (1.1) for $n = 3$, Dirichlet and Legendre in 1825 – for $n = 5$, Lame – for $n = 7$. In 1994, Prof. of Princeton University, Andrew Wiles, [1], [2] proved, for all n , but this proof, contains over one hundred and forty pages, understandable only to high qualified specialists in the field of number theory, cylindrical functions.

Such a proof cannot be recounted in a school lesson. At the same time, there are Olympian methods of proving the theorem, one of which was engraved by the author with a laser on a wooden cube for children's development. Six facets proved to be sufficient where American science needed over a hundred. The tendency to seek such extensive proofs can also be seen in the five-hundred-page (!) work of the Japanese mathematician Shinichi Mochizuki [4], a specialist in number theory and arithmetic geometry at Kyoto University. His 'mega-work' is devoted to the search for a proof of the *ABC conjecture* in number theory [4] – see chapter II.

The verification of such controversial proofs is done with the help of artificial intelligence, AI, and a human is left to take the proof on faith, which undermines the principle of positivism and contradicts the philosophical concept of the *knowability of the world*. In fact, there is no need to spend a hundred or more pages where there is enough creative imagination.

First of all, we note, that, without changing the generality, it is possible to order the natural numbers in formula by the relations: $a < b < c$. (The situation of equality of the edges $a = b$ is unacceptable due to irrationality of $\sqrt[3]{2}$). The case of negative numbers can be considered by transferring a term to another part of the equation and replacing variables – it is enough to prove the Theorem (1.1) for the case $a, b, c \in \mathbb{N}_1$ (index 1 denotes the set of natural numbers other than 0, as opposed to \mathbb{N}_0 in the era of digits, where the same arrays are numbered from zero in a number of programming languages) and generalize the result to integers.

Engineering Approach

Let's try to create a construction based on the Theorem. The project consists of three n-cubes *a-Small*, *b-Medium*, *c-Large*, whose common centre coincides with the origin of coordinates with edges (radii) $a < b < c \in \mathbb{N}_1$, provided that $V(B_a^n) = V(B_c^n) - V(B_b^n)$. V is the volume. Alternatively, three concentric balls with similar radii can be used, with radii a, b, c . Note that any cube can be transformed into a sphere by continuous and reversible deformations (*homeomorphism*) and vice versa. In topology such figures are called homeomorphic [Chap. 2, 10.4]. Therefore, for good comprehension the presentation below goes from a n-cube to a multidimensional balls denoted as B_a^n where a is the radius of the sphere, n is the dimension of the space. Remember that for $n \geq 3$, instead of volume, we should use an axiomatically defined concept of *measure*, an important property of which is the positivity of the values and *additivity*, i.e. when the numbers are combined, the measures add up, which is not only convenient for physical measurements of mass, length, area, volume, but also expresses the property of the matter of our world (for example, the law of conservation of the quantity of matter in motion and in chemical reactions).

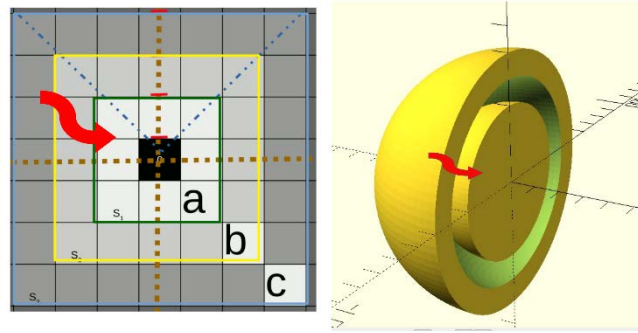


Fig. 1. On the left, the piercing of concentric n -cubes by a two-dimensional plane, on the right – balls with integer edges. The arrows symbolize the mapping of the sets

It is not difficult to see that this construction is not feasible. In fact, if the construction in Fig. 1, denoted by the quantifier \exists , exists, then by virtue of the postulated *homogeneity* of Euclidean space it is possible, as symbolized by the arrows in Fig. 1, to swap places, to match \forall – any point/unit cube of a small ball to another point/unit cube of the studied subset of space between the medium and large spheres, so as not to destroy the symmetry of the structure, to avoid inhomogeneities or voids. But these subsets (in other words, parts of the construction) are not equivalent, i.e. the condition is not feasible, where the sign \setminus means subtraction of sets and the sign \cong is a homeomorphism, which will be discussed later.

Figure 1 on the right uses the cross-platform application OpenSCAD, which, in addition to pure engineering applications, is an excellent tool for studying set theory, operations on sets, and associative law. The most common metal hammer with a wooden handle can be viewed as the result of subtraction operations on the sets of difference, union – union, the intersection of the handle and the actual metal part of the hammer. There are many educational examples of program code for visual presentation of the concepts of set, subset, universal set U , empty \emptyset set, complement of set, etc. Isn't it about time in the XXI century to study the basics of set theory in school through 3D modeling?

Why is a project illustrating Fermat's Last Theorem not feasible?

The obstacle lies in the fact that the figure modeled by formula (1.1) must have the property of central symmetry and cannot contain inhomogeneities. These restrictions mean that each layer in Figure 1 is not comparable to any other layer in space with dimension $n \geq 3$. The conditions of *additivity* and the *axiom of measure* do not work here – the very

formulation of the question of layering is illegitimate, since it requires symmetry of construction, addition/subtraction/reduction of volumes. The mathematicians of ancient Greece introduced the concept of the *incommensurability* of linear segments such as $\sqrt{2}$ and 1. Here we encounter a similar, but at the same time new phenomenon. Let's make sure that the pursuit of the condition of equality of volumes of subsets of the studied construction $V(B_a^n) = V(B_c^n) - V(B_b^n)$ and the central symmetry exclude each other.

Understanding of Symmetry, Isotropy and Uniformity of Space

We will distinguish between an open ball and a sphere that encloses it. The sphere is the so-called *layer* for the ball or n-cube enclosed by it. A sphere has one dimension less than a ball. This is well known from school mathematical lessons – let's remember the formulae for the circumference $2\pi R$ and area of a circle πR^2 , the area of a sphere $4\pi R^2$ and the volume of a sphere $\frac{4}{3}\pi R^3$. It is interesting to imagine that, in a one-dimensional world, an open sphere would take the form of a segment without its end points, or of zero-dimensional spheres located at a distance r from the origin. (In metric

space the formula of ball is $\sum_{i=1}^n x_i^2 < r^2$ and for n-cube $\max(x_1, \dots, x_n) < r$. The set of all B_a^n is the *base of topology*. The structures defined by these bases are the same.)

Imagine the projection of an n-dimensional sphere onto a hyperplane. Let's start with the simple case of 4D space (x, y, z, w) . What will happen as a result? – A three-dimensional sphere and a sphere covering it with the centre at the origin, as a result of the intersection of a four-dimensional sphere with a three-dimensional subspace – $(x, y, z, 0)$. The open Northern and Southern hemispheres differ, while the equator is excluded. (The equator becomes the meridian when the figure is rotated at right angles.) For the 3D case, we get an ordinary ball, similar to a football, enclosed by a 2D sphere. Again, we exclude any meridian from the sphere. Furthermore, in the case of a two-dimensional plane $(x, y, 0, 0)$, we get the usual circle and circle, excluding diametrically opposite points. Finally, for a one-dimensional sphere $(x, 0, 0, 0)$, the circle is the diameter, and the sphere is only 2 points separated by a radius from the origin. As a result, a sequence of non-intersecting geometric elements appeared on a 3D sphere with

a number of dimensions from 1 to $n-1$. These are the so-called *hypermeridians*, which will still be useful in the search for proofs.

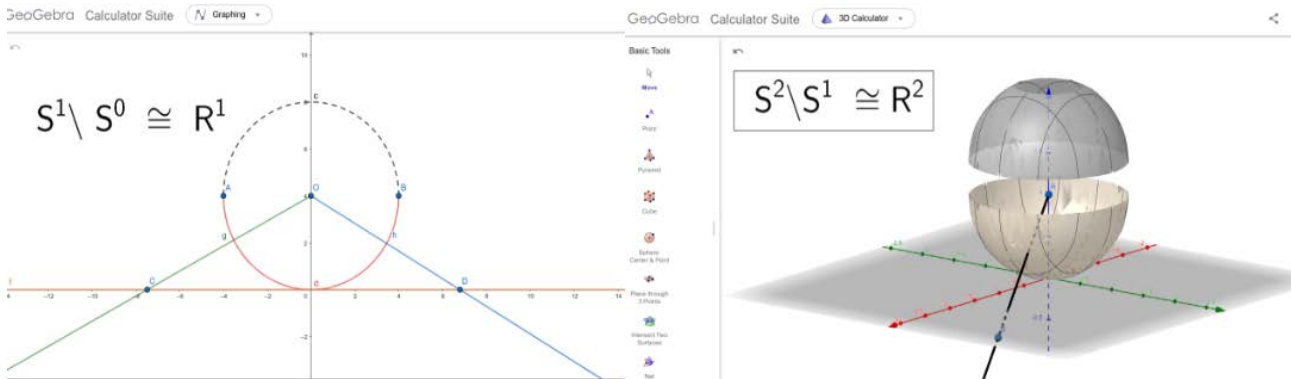


Fig. 2. One-to-one correspondence between a (hyper) plane and an open hemisphere using a beam from a point light source}

The dimension of the figure is checked by a one-to-one comparison of a (hyper)plane and an open hemisphere, for example using a beam from a point light source as in Fig. 2. Now imagine a cascade of successive spheres and meridians, each covering a sphere of the same radius from this set. If the triple of natural numbers is $a, b, c \in \mathbb{N}$, then in this case \forall point – arbitrary point in space, as well as the sphere in the interval between the b *Medium* and c *Large* spheres, should be mapped to a set of spheres

belonging to the a -Small ball: $dr = k \int_0^a r^{n-1} dr$, where k is some coefficient depending on the dimension of the space n , $|S^m| = kr^m$. Due to the symmetry of the construction, any set of these spheres can be mapped into a set of other spheres of smaller radius, the number of small spheres obviously having to be greater due to the equality of measures condition in (1.1) or the equivalence of the mapping (compare: the principle of *incompressibility* in physics, the volume of each particle remains constant during deformation. For example, liquids are conventionally considered to be incompressible. The principle of *conservation of mass*: during plastic deformation of a material, its mass is preserved). When explaining the concept of an integral or differential, teachers work with the concepts of dividing a figure, a sphere, into separate parts, like cutting a lemon, which are then put back together to form a whole. But the use of higher mathematics is unnecessary here: just write: $S_j \Delta R = S_i \Delta r$ for the radii *preimage* and *image* and equate $\Delta R = 1$. It is sufficient to write down the equation given here in terms of set theory and equivalence relations:

$$\{S_j\} \rightarrow \{\dots S_i\} \quad (1.2)$$

which means that a set of a single element containing a layer is mapped to a set of layers containing several elements using a continuous equivalence function (remember the ray from a point source in Fig. 2), in this case \forall point is mapped to a point, an elementary cube – to an elementary cube: $1^m \rightarrow 1^m$. In other words, \exists an equivalence function $F(X) \rightarrow Y$ that maps an open set of *X-preimage* to a set of *Y-images*. In topology, a mapping F is said to be continuous if the preimage \forall of an open subset of space Y is an open subset of space X . This is a kind of definition of continuity. (Unfortunately, in school the continuity of a function is determined in the only way – by *Cauchy*, omitting the more obvious ones by *Heine*, etc.) Let there also be an inverse continuous function $F^{-1}(Y) \rightarrow X$. This gives the homeomorphism $X \cong Y$. Speaking of homeomorphisms, we note an interesting transformation of a mug into a torus and vice versa – a popular video in topology.

In our construction, a set of consecutive spheres with integer radii form open spheres. In primary school, students study the texture of a sawn tree, the layered structure of the Earth's crust in a section, traces of layers on a probe taken from an ultra-deep well – all are examples of induced topology [4], when an image similar to the original is repeated in subspace. Concentric spheres, concentric layers – see Fig. 1 – are repeated in all *subspaces*, also called *hyperplanes* in this article.

It is surprising that the theorem excludes the existence of the symmetrical construction of three nested spheres under the conditions of equivalence of volumes/measures and continuous sequence of layers!

Table 1

**Any spheres from preimage can be mapped into image
by each element separately**

Sphere	Element ^m -hyper- meridian	...	Element ² - sphere	Element ¹ -circumference	Radius
S_j^m	$S_j^m \setminus S_j^{m-1}$...	$S_j^2 \setminus S_j^1$	$S_j^1 \setminus S_j^0$	j
\downarrow	\downarrow	...	\downarrow	\downarrow	...
S_i^m	$S_i^m \setminus S_i^{m-1}$...	$S_i^2 \setminus S_i^1$	$S_i^1 \setminus S_i^0$	i
S_{i-1}^m	$S_{i-1}^m \setminus S_{i-1}^{m-1}$...	$S_{i-1}^2 \setminus S_{i-1}^1$	$S_{i-1}^1 \setminus S_{i-1}^0$	$i-1$
...

(In simple terms, the construction is filled without gaps, layer by layer.) To understand this phenomenon, it is necessary to analyze equation (1.2) by elements:

(Note that $S^m \setminus S^{m-1} \cong \mathbb{R}^m$, $\dim(S^m) = m$ where \mathbb{R}^m is the space dimension of m or in short $\dim(m)$, see fig. 2 above.)

Each layer contains elements of dimension from 1 to $m = n-1$. Here the structure under study was filled with layers from the periphery of $j \dots i$, $(i-1)$ to the center, so the indices are listed in descending order. The equivalence function maps each element of the layer S_j separately \downarrow with a set of elements of the corresponding dimension. It is impossible to ensure simultaneous correspondence of layer elements in more than one dimension due to the unsolvability for $n \geq 3$ of the system of $n-1$ equations formulated below, where i and j are the natural radii of concentric spheres:

$$\begin{cases} j^{n-1} = i^{n-1} + (i-1)^{n-1} + \dots \\ \dots \\ j^2 = i^2 + (i-1)^2 + \dots \\ j = i + (i-1) + \dots \end{cases} \quad (1.3)$$

Each equation contains two or more terms to the right. This series of equations extends from $n-1$ to the power of 1 and only for the two-dimensional case contains a single equation which is always solvable (see below). Remember (see figure 2) that an element of any dimension can only be mapped by an equivalence function onto an element of the same dimension. The structure of the layers is determined solely by the dimension of the space and does not depend on the geometric size of the layer. The condition of equivalence or equality of the power volumes of the set of points for any n -dimensional space) of the mapped layers means that any arbitrary layer of a large sphere can be mapped onto a set of sequentially following small spheres (1.2) according to the columns of Table 1.

(The cardinality of a set is the number of elements in it, e.g. $|S^m| = kr^m$. For a discrete space, the analogue of a point is 1^n the number of such elementary cubes and is the power). The number of layers in the image to the right of each equation is ≥ 2 as the following natural number after the unit. The system of equations (1.3) is unsolvable even in real numbers \mathbb{R} . It is easy to check this using the *triangle inequality* $\|x + y\| \leq \|x\| + \|y\|$ for Euclidean space, where the sum of the lengths of the cathetus

of a right triangle is greater than the length of the hypotenuse. The case of equality occurs only with a zero cathetus and is excluded under the conditions of Theorem. It is important to remember that Euclidean space in the modern sense means either a finite-dimensional real vector space with a scalar product defined on it, or a metric space corresponding to such a vector space. If the number of terms in the system of equations (1.3) is more than two, we can concentrate on the last pair of equations, assuming x^2 is the square of the module \forall of the vector, and x is the projection of this vector onto any fixed axis, and then, by applying the associative law to the terms, reduce the problem to the one already considered.

So, an engineering project in the form of a construction of three nested concentric spheres such that, where $a, b, c \in \mathbb{Z}$ is not feasible. Formula (1.1) is valid for triples of integers; the units of measurement are irrelevant, be they metres, centimetres, millimetres, micrometres, nanometres, etc. This implies not only equality, but also identity of equation (1.1), or in other words, independence of the result from the partition of space (\implies The theorem holds for $a, b, c \in \mathbb{Q}$). The conclusion is that at least one of the values in the triple a, b, c must be an irrational number. But such a construction will necessarily have a *central_symmetry_defect*, since a, b, c will become *incommensurable*, i.e. project will become impossible. A partially filled layer is a symmetry defect. The theorem is proved from the opposite.

Why do Pythagorean triples only exist on the 2D plane?

Pythagorean triples were used by the ancient Sumerian civilisation millennia before Pythagoras himself was born to solve trigonometric problems, including the addition/subtraction of angles. But why do these triples \exists only on the two-dimensional plane? To give the answer, and at the same time to get an idea of the above-mentioned integration, let's use as an 'integrator' a 'good old' tape recorder, where the tape is rewound from the left to the right reel, or vice versa. In this case, the sum of the areas of the left and right reels of the tape remains constant and is equal to the thickness of the tape multiplied by its length. Therefore, $R\Delta R = r\Delta r$ for the radii *preimage* and *image* of the left and right reels respectively. Without loss of generality, we assume that the left spool/coil has a radius j greater than the right spool/coil i , in other words

$j \geq i + 1$, otherwise you can simply rewind the tape or swap the spools. When we talk about the need to preserve the symmetry of the figure when the layer is continuously displayed or moved, this means that one turn of the tape on the left must fit into an integer number of turns of the reel on the right: two, three, four, etc., but not a fraction, otherwise the condition for preserving the symmetry when moving or displaying layers (1.2) will be violated. In fact, a partially filled layer has a symmetry defect, it is asymmetric. The postulated homogeneity of Euclidean space is violated.

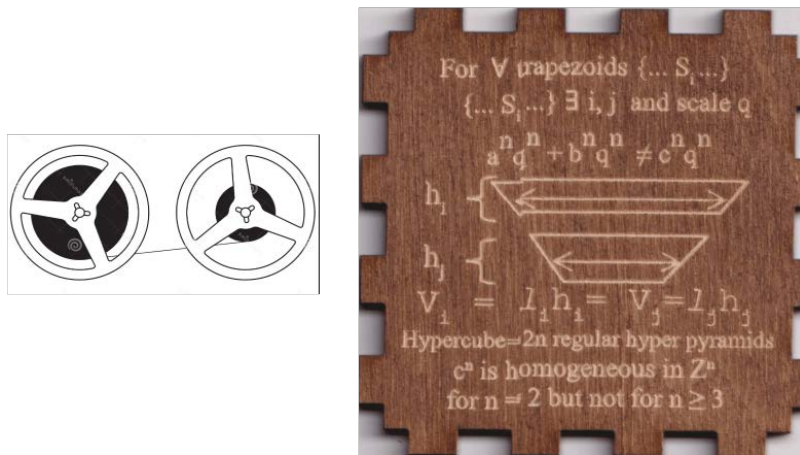


Fig. 3. Equivalence of layers on a 2D plane for circles and squares.

Photo of the face of the author's 3D cube according to the patent application for the Prom. Example No. 2021501435/49 dated 20.03.2021}

Assuming that it is possible to produce a tape of any thickness, then for any given ratio of the radii of the circles of the coils, it is easy to achieve symmetry in this model, namely: one layer or turn on a large coil is equivalent to a set of full turns or layers on a small coil. (In general, we are not talking about a single layer, but an ordered set of layers, where in the preimage S_j is the first element and in picture S_i is the last element. Due to symmetry, it is sufficient to study the *restriction* to one element $\{ S_j \}$ ($S_j \Delta R = S_i \Delta r$ and assume $\Delta R = 1$).

Note that the picture is similar in the case of an n -cube of dimension 2 or, more simply, a square on a plane. Because of the symmetry, it is convenient to focus on the study of a single segment formed by rays emanating from two adjacent vertices and to apply the formula for the area of the trapezoid by the product of the length of the median line by its height or the number of elements in the set of layers Fig. 3, right is analog of $S_j \Delta R = S_i \Delta r$, where median line is S and height is $\Delta R, \Delta r$. As a result, the

volume/measurement of a subset of layers mapped into a set of successive and continuous layers closer to the centre and the symmetrical shape of the structure under study are preserved. What prevents us from carrying out a similar experiment in the case of three-or-more dimensional space? – One can answer: Tab. 2. an unsolvable system of equations (1.3) which expresses the equality of measures for each dimension separately, when edges are mapped on edges, faces on faces, hypermeridian on hypermeridian of the corresponding dimension. But strictly speaking, for $n \geq 2$ there are *pairwise disjoint equivalent classes* [5] whose number is $n - 1$, and as mentioned above, the equivalence function works for each class separately. Look at the photo of the Taj Mahal, a monument of Indian architecture! It is symmetric about the vertical axis passing through its centre. Symmetry is an excellent demonstration of the equivalence function, along with other examples: the motion of a rigid body, its rotation, etc. The monument is made up of different elements: columns are shown in columns, domes in domes, and there is no confusion about this division into pairwise disjoint equivalent classes. Another example is chess (the board itself is a visual explanation of the Cartesian product of sets \times). While the game continues on a parity basis, pawns are exchanged for pawns, light pieces: a bishop, a knight for light pieces, a rook for a rook, and a queen for a queen. These are also pairwise disjoint equivalent classes. The equivalence function works with each class separately, it is easy to prove strictly mathematically [5]. In a concise form, Tab. 1. can be expressed by mapping the layer into pairwise disjoint equivalent classes via partitioning:

$$S \rightarrow S / \sim = \{S_{n-1} \setminus S_{n-2}, \dots, S_2 \setminus S_1, S_1 \setminus S_0\} \quad (1.4)$$

(Note that $S^m \setminus S^{m-1} \cong \mathbb{R}^m$, $\dim(S^m) = m$; where \mathbb{R}^m is the space dimension of m or in short $\dim(m)$, see fig. 2 above.)

This process is called *factorization*, which is convenient for analysis. In the formula above, each layer is considered as an indivisible unit – these are pairwise disjoint equivalent classes, which are convenient for research. Go back to the paragraphs explaining Fig. 2.

Cube and layer

Since the n-cube, denoted I_n , fits into the balls and vice versa, these figures are homeomorphic, which is easy to establish using a point source in their common centre (see Fig. 2). Like the hypermeridians of a sphere, the plane of the n-cube is composed of elements of different dimensions. The formula that expresses its structure is particularly clear:

$$S_i = (i+1)^n - i^n = I_{i+1}^n - I_i^n = I_{i+1}^n \setminus I_i^n = \bigcup_{k=1}^{k=n-1} C_n^k i^k \times 1^{n-k} \quad (1.5)$$

Here I_i^n is a n-cube with an edge i , and should be understood as a \times Cartesian product of segments resulting in a cuboid, and 1^{n-k} as a dimensionless multiplier. This ensures that the axioms of topology [5, Chapters 1, 2, 4] and Euclid's postulates [5] are satisfied for elements of arbitrary dimensions. The Cartesian product of a 1D-segment $i \times 1^{n-1}$ and a 2D-square forms $i^2 \times 1^{n-2}$ a 3D-cube $i^3 \times 1^{n-3}$ and so on. (The Cartesian product is closely related to the *projection* on the set $proj_x : X \times Y \rightarrow X : (x, y) \rightarrow x$.)

This is an illustration of the principle of dimensionality, well known from school physics lessons. The above formulae are derived from Newton's binomial and describe the plane as the difference of hypercubes with edges differing by one and a common vertex coinciding with the origin. An alternative way of representing the structure under consideration is to have the origin of the coordinates at the common centre of the hypercubes. This can be obtained by n reflections of hyperplanes, i.e. by multiplying the last formula by 2^n , which does not fundamentally change the nature of the last formula. Both geometric constructions are transformed into each other by reflections of hyperplanes perpendicular to each of the n coordinate axes, or by dissection and scaling of the figure.

A *multiset* (1.5) is a modification of the concept of a set that allows the same element to be included in the set several times. If we exclude repeating elements that do not play a fundamental role for us, i.e. reduce the binomial coefficients in (1.3), we obtain a linearly ordered set:

$$I_0^n \prec S_1 \prec S_2 \cdots \prec S_a \cdots \prec S_b \cdots \prec S_c \in U \quad (1.6)$$

where the first n -cube I^n can be a point in the case of spheres, 1^n or 2^n , depending on the parity, but with the caveats two paragraphs above, this is not essential. U is the name of a *universal set*, or *c-Large* ball or n -cube.

The last formula defines *chain of sets* (also called a strictly ordered set \prec , a linearly ordered set, denoting order relations) in the form of successive layers – spheres or layers of an n -cube. This is another clear proof of the homeomorphism of the cube layer to the sphere, which helps to establish the correctness of the representation of the piercing (Fig. 1.) of the hypercube by a two-dimensional plane.

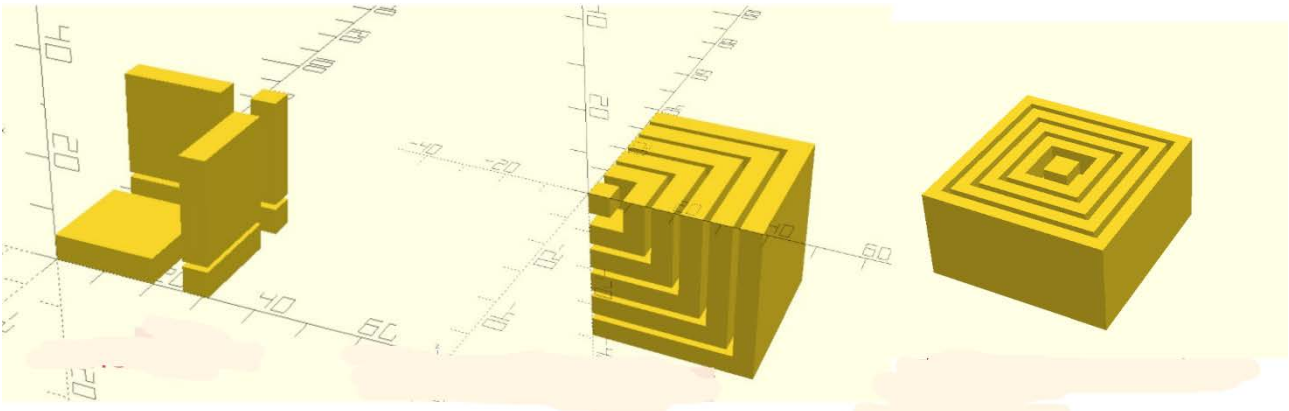


Fig. 4. On the left is a layer containing elements of dimension $1 \dots (n - 1)$.

A set of layers in octant (+, +, -) is at the centre.

On the right is a set of layers that form an n -cube 3D cube in section

Look at this corner (figure above, left). And this is a cascade of corners (centre) that together form a cube (right). The Cartesian product of a linear segment per square forms a cube. This shows that the principle of dimension is respected. The above formulae are derived from the Newtonian binomial and describe the plane as the difference of successive hypercubes with a common vertex coinciding with the origin.

Starting from the tesseract, a two-dimensional plane passing through the centre of an n -cube and parallel to two arbitrary coordinate axes can be closed, like the eye of a needle covering a thread, by an arbitrarily small continuous contour lying in a plane orthogonal to the passing plane ($n - 2 \geq 2$ when $n \geq 4$ is more consistent and easier to prove by the basics of linear algebra). As a result, the tesseract is not divided into two unrelated pieces as in the case of the 3D cube, but is *pierced*. The image is like a baked apple on a skewer. From these positions, the 2D plane in Figure 1. – This is a probe for

studying the layered structure of an \forall n-cube $n \geq 3$ or the Cartesian product of two orthogonal probes.

So, Pythagorean triples are only on a two-dimensional plane. Euclidean geometry is also called plane geometry. In the digital age, it is possible to supplement Euclid's postulates with elements derived from elementary 1n cubes lined up in a row, in a rectangle, in a 3D cube, and so on. (This implies working in n-dimensional space). The dimensions of the shapes created will be the same as if we were working with points.

For example, $\frac{dx^n}{dx} = nx^{n-1}$ it is consistent with the formula describing the structure of the layer (1.3), which again confirms the consistency of the axiom system of formal theory. It is implicitly implied that it is complemented by well-known axioms of topology, which make it possible to construct elements of different dimensions through the operations on sets mentioned above [4, p. 23]. However, leaving formal definitions aside for the sake of illustration, let us note that when working with decorative material, such as siding for a house, it is common to measure it in square meters $i^2 \times 1$. Baguette or baseboards, are measured in metres $i \times 1^2$. In reality, however, all the materials listed have specific dimensions in length, height and width. However, smaller sizes are neglected in the axiom system (the unit scale can be arbitrarily small). Therefore, 1^n has either zero dimension or the dimension of space n itself, depending on the partition, the scale. Furthermore, $1^n \cong 1^m$ if and only if $n = m$, which logically follows from the principle of uniformity of space.

The test of millennia

Is it legitimate to choose a figure from a sphere, its element – a hyper-meridian, an equatorial circle? – The answer is yes, because this technique is known from the ‘Elements’ of Euclid, formulated in the III century BC, and from school geometry courses on planes, stereometrics and drawing. To illustrate, let's imagine a correct three-dimensional pyramid and its image, for example, by holography. The principle of equivalence is expressed in the fact that the 3D pyramid preimage as a whole is mapped onto a 3D pyramid image, an arbitrarily chosen 2D surface onto a 2D surface, a 1D edge onto a 1D edge. Drawing and architecture are based on this principle, which

has stood the test of time for at least six millennia. Why, in this situation, is it impossible to isolate a hyper-meridian on a sphere and expect each of these elements to map onto a set of elements of the corresponding dimension? – The question is rhetorical.

Some eminent scientists argue that piercing the n -cube with a two-dimensional plane in Fig. 1. is too simple and distorts the extraordinarily complex structure of a multidimensional cube, bearing in mind the well-known beautiful projections with the parallax effect. This misunderstanding is caused by a misconception of what an n -cube is. In order to study the structure of the n -cube, the author used the method of virtual reality experiments [7, pp. 198–199].

Journey to the vertex of the n -cube in mental experiment

Astronaut Fyodor went into multidimensional space to make a trip from the origin of coordinates and simultaneously the center of the 10-dimensional cube, whose faces are orthogonal to the coordinate axes, to an arbitrary vertex of the hypercube. The movements made by the astronaut-imposed restrictions: allowed movements only in direction or against direction of any of the n coordinate axes. Fyodor made n jumps: the first one from the origin of coordinates to the center of the face of the highest dimensionality $n-1$, and all the rest – across the first jump. At each step or jump, Fyodor changed the direction of motion, turning at right angles.

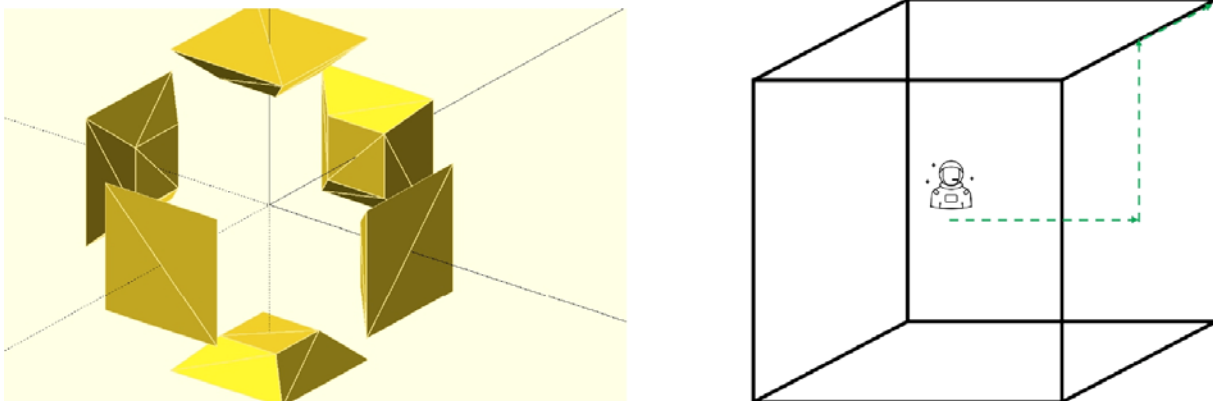


Fig. 5. Fyodor made n jumps: the 1st one from the origin of coordinates to the center of the face

Analysis of the experiment

Fyodor's report is as follow: *Start from Dekeract. Moving to Enneract. Moving to Octeract. Moving to Hepteract. Moving to Hexeract. Moving to Penteract. Moving to Tesseract. Moving to Cube. Moving to Square. Moving to Segment. Moving to vertex. Finish!*

In the last three steps we, as three-dimensional beings, could see the sudden appearance of Fyodor in the center of a three-dimensional cube from the invisible to us four-dimensional space, followed by a jump to the center of the wall, ceiling or floor – at the discretion of Fyodor, after which from a 2-dimensional square – a jump to the center of the edge and finally, the last step – arrival at the top.

Analyzing the two-dimensional and three-dimensional cases and generalising the result to the n-dimensional case, it is easy to calculate the length by formula $\frac{i}{2} \sqrt{n}$ of the distance from origin of coordinate to vertex. The edge of an n-dimensional cube lies in the hyperplane perpendicular to the height just constructed and passing through the base of this height – the point of intersection of the straight line with this edge. Note, that from Fyodor's point of view, all faces of the n-dimensional cube are perceived not as n-dimensional, but as *flat* figures.

Without changing the generality, we can arrange the indexes of the coordinate axis in ascending / descending order and get the following result in the form of the ship's Log of cosmonaut Fyodor. At the end of the flight, we turn to the data of the autonavigator, where Fyodor's observations and directions of movement along each axis were recorded. From the basics of combinatorics, it is easy to see that the number of elements – the power of the set, consisting only of binary elements zero and one, is 2^n – exactly the same number of vertices in the n – dimensional cube. In this experiment, Fyodor started from the center of the ten-dimensional cube moved only in the direction (denotes as one) or against the direction (denoted as zero) successively explored all of its faces of dimension from nine to one.

It is important to remember that Fyodor observed multidimensional cubes with the effect of parallax on his way. There are a number of examples from astronomy,

everyday life, and works of visual art to explain this effect. Images of an n-dimensional cube are mesmerizing in their beauty. Let's imagine a projection of an n-dimensional cube onto a two-dimensional plane, for example, passing through the X, Y axes, without the parallax effect [7, pp. 198–199]. What figure will it be? The answer is shown on Figure 1. So, when we talked about piercing an n-dimensional cube with a probe, the Cartesian product of two probes perpendicular to different faces, we worked with induced topology.

In fact, Fyodor investigated nested topological subspaces.

$\emptyset \subset \Omega^1 \subset \Omega^2 \subset \Omega^3 \dots \subset \Omega$ in the original n-dimensional space Ω they all pass through the origin, the index from above means the dimension of the subspace. For example, in 4D Ω :

$$\emptyset \subset (x, 0, 0, 0) \subset (x, y, 0, 0) \subset (x, y, z, 0) \subset \Omega.$$

For simplicity, we can imagine the last three subsets in this chain: in 3D – 2D- 1D subspaces, respectively: ball B^3 – circle B^2 -a linear segment B^1 . Due to the equivalence condition of measure (1.1) for each subspace, this led us to the unsolvable system of equations described above.

A paradox

From the algebraic decomposition it follows that figures in the form of a cube with a defect in the form of a missing point, corresponding to the expression $i^n - 1^n$ a set of layers or $2^n - 1^n$ layer (in \mathbb{Z}^n provided central symmetry), indifferent to elementary transformations, have a dimensionality one unit less than the space in which these figures are created. Given the above, such expressions and figures correspond to $B_r^n \setminus \text{point}$ or S_r^{n-1} [5, 10.20]. However, only in the plane is it possible to simultaneously fulfill the conditions of symmetry preservation and volume equality inherent in the Pythagorean Theorem.

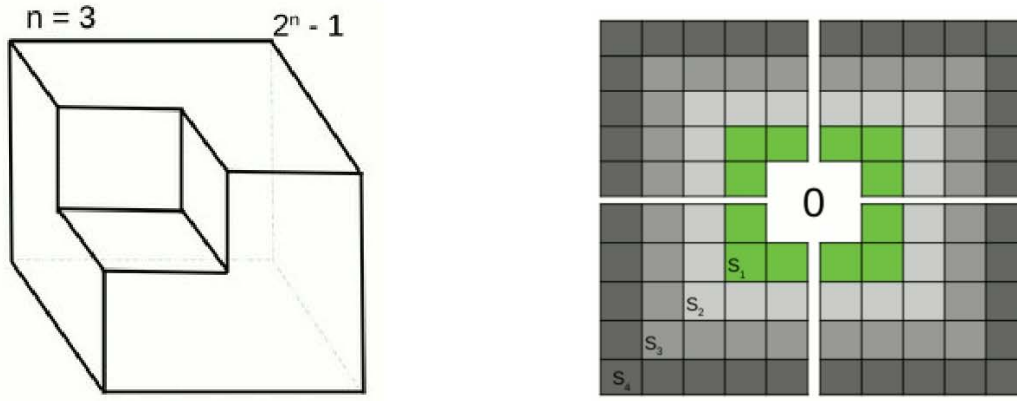


Fig. 6. Dimension of defect cube $2^n - 1$ is $n-1$ in central symmetry space of \mathbb{Z}^n

Note that $2^2 - 1$ is also a 1D figure in the central symmetry space of \mathbb{Z}^n , which is mapped to the 1D perimeter of the square I^2 layer by layer. This situation is normal, because of the condition of one class equivalence for $n = 2$.

Dimension is not a trivial question in mathematics. There is a great visualisation of *Malderbrot's fractal set*, which is defined by a recurrence formula, and such fractal sets differ from our usual sets in that as you increase the scale of the set, the same features are repeated infinitely. Imagine you're looking at a map from a bird's eye view, then you go down, you get closer to this object, you see more and more details, so these details show a more complex structure, it repeats all the time. There are no arcs, no smooth curves or straight lines. So, the dimensionality of this fractal set is *fractional*. It's amazingly beautiful!

Chapter 2. The ABC Conjecture from the Perspective of Physics

Formulations of ABC Conjecture

The ABC conjecture was formulated independently in number theory by the mathematicians David Masser in 1985 [8] and Joseph Esterle [9] in 1988. Further in the text: *Oesterlé-Masser conjecture*. It is stated in terms of three positive whole A , B C that are co-prime and satisfy $A + B = C$. The conjecture essentially states that the product of the distinct prime factors of ABC is usually not much smaller than C or $\max(A, B, C) < K_\varepsilon \text{Rad}(ABC)^{1+\varepsilon}$, where K_ε depend only on some positive real number ε . Rad function is the radical of the numbers A, B and C equal to the product of the

primes forming these numbers, but raised to the first power, e.g. $\text{Rad}(8) = \text{Rad}(2^3) = 2$, $\text{Rad}(1000000) = \text{Rad}(2^6 \cdot 5^6) = 30$.

Equivalent formulation of the Oesterlé-Masser conjecture involves the *quality* q_{ABC} of the triple which is defined as: $q_{ABC} = \frac{\log(C)}{\log \text{Rad}(ABC)}$. For every positive real number ε , there exist only finitely many triples A, B, C , of co-prime positive integers with $A + B = C$ such that $q_{ABC} > 1 + \varepsilon$.

Many hundreds of pages have been spent by eminent mathematicians in Western world searching for proofs, and the process of searching for proofs continues. But there is another approach. Minhyong Kim, a mathematician at the University of Oxford said: *It should be possible to use ideas from physicists to solve problems in number theory, but we haven't thought carefully enough about how to set up such a framework* [10].

Construction

Developing the approach adopted earlier, we compare this expression with a set of manifolds (for simplicity's sake we will also use the engineering term construction) in a finite-dimensional Euclidean space with a given metric and measure, then simply an n -dimensional space in the form of a multidimensional parallelepiped (hereafter referred to as a *cuboid*). Imagine that $A + B = C$ by decomposition into prime numbers, respecting the rules of the dimension of the elements in the space of integers $Z_n = Z * Z * \dots * Z$ (n -times *Cartesian product* of the axis of integers).

$$a_1^{\alpha_1} a_2^{\alpha_2} \dots a_k^{\alpha_k} * 1^{n-k} + b_1^{\beta_1} b_2^{\beta_2} \dots b_l^{\beta_l} * 1^{n-l} = c_1^{\gamma_1} c_2^{\gamma_2} \dots c_m^{\gamma_m} * 1^{n-m} \quad (2.1)$$

Here above is the sum of the prime numbers of the degrees of the multipliers and the degree of an elementary unit cube of the form 1^{n-k} complement each other up to the dimension of the space n . In other words, the sums of the exponents of the degrees of all factors are equal to k, l, m , respectively, i.e. the dimensions of the A, B, C manifolds:

$$\alpha_1 + \dots + \alpha_k = k, \beta_1 + \dots + \beta_l = l, \gamma_1 + \dots + \gamma_m = m \quad (2.2)$$

It is easy to understand the dimension of each term (2.1) and the minimum required dimension $n = \max(k, l, m)$ of the space in which the structure under study can be located. In this representation, the above formula can also be written as a condition

for the homeomorphism of the figures $A \cup B \cong C$, and the structure under study is a set of manifolds in the form of three cuboids with different edges, whose lengths are prime numbers.

Due to the central symmetry, it is easy to determine the center of mass of each cuboid, assuming that the cuboid is filled with a homogeneous material. Let's set these centres to the origin and orient the hyper faces perpendicular to the axes. In the following we will refer to such a construction as concentric parallelepipeds. In each of the cuboids, it is easy to see hyperplanes of symmetry orthogonal to the coordinate axes. Their number is equal to the sum of the exponents of the degrees of the factors forming the cuboid.

\exists a continuous invertible equivalence function f that maps each element as an elementary cube 1^n of the sets A and B in Z^n into C , i.e. $f(A, B) \rightarrow C$, which can also be expressed from a topological point of view by the homeomorphism condition of the following figures:

$$A \cong C \setminus B \quad (2.3)$$

All the prime factors in formula (2.1) are different, otherwise it would be possible to reduce the numbers A, B, C by the greatest common divisor GCD, but under the condition that A, B, C are co-prime. Each of the three terms in formula (2.1) is the product of a set of prime numbers which can be matched with one-to-one integer polynomials of similar degree, as will be shown below. In other words, formula (2.1) models a particular system consisting of subsystems that are not connected but share a common resource (energy, phase space, entropy) in a particular way.

What physical phenomenon can this correspond to? – The *phase space* and *entropy* of a mixture of ideal gases whose molecules do not interact with each other. For example, according to *Dalton's Law*, which states that the pressure of a mixture of chemically non-reacting gases is equal to the sum of the partial pressures of the individual gases – and this is clear because of the absence of intermolecular interaction.

The key point in the author's proof of Fermat's Last Theorem is the concept of layers comprising pairwise disjoint equivalence classes, the possibility of homeomorphism of figures of equal dimensions. The n -ball is homeomorphic to an n -

cuboid, and the *sphere* of dimension $n-1$ (to avoid further confusion of the notation with entropy, we will refer to the already familiar layers S as *Sphere*) is homeomorphic to an n -cuboid layer. From the symmetry conditions of the structure under consideration, there is at least one term of degree not greater than two in equation (see 1.3). Otherwise, there would be three hyper-cylinders with a base in the form of a hypercube of degree $n \geq 3$ and such coefficients would be found in the form of real numbers $k_1 = 1/x^3$, $k_2 = 1/y^3$, $k_3 = 1/z^3$ with the condition: $(k_1 A)x^3 + (k_2 B)y^3 = (k_3 C)z^3$ and layers – spheres belonging respectively to the cuboids A , B and C of dimension $n-1$, for which:

$$k_1 \text{Sphere}_A + k_2 \text{Sphere}_B = k_3 \text{Sphere}_C \quad (2.4)$$

where for the case $n \geq 3$ two or more pairwise disjoint equivalence classes are compared simultaneously, which is impossible due to the presence in each layer of elements of different dimensions from 1 to $n-1$ and the unsolvability of the system of equations due to the triangle inequality (substitution $k_1 = k_1 A \dots B$ and C correspondingly). See *Why do Pythagorean triples only exist on the 2D plane* in Chapter 1.

It is easy to see that high powers above two must cancel each other out, so that $A \cong C \setminus B$ and the result would be comparable to a summand of degree not higher than two, in other words, the reverse equivalence function $f^1(C, B) \rightarrow A$ works so that layers of higher dimensions than two are reduced and also the number of symmetry planes has been reduced to two for \forall prime number in (2.1). Excluding the trivial case of the first degrees of the Oesterlé-Masser conjecture, we further assume, without changing the generality, that $\dim(A) \leq 2$ and $\dim(A) \leq \dim(B)$. But this reasoning is not powerful enough for proof.

Physical Methods Applied to Number Theory

The difficulty of finding a proof of Oesterlé-Masser conjecture from purely mathematical approaches forces us to choose an alternative path of research using the phenomena of physics, based on reliably confirmed experimental data. It is appropriate to ask what is a rather large class of known physical phenomena described by a formula similar to (2.1)?

First the author proposes to consider Oesterlé-Masser conjecture (2.1) from the point of view of the phase space of a thermodynamic system with $2s$ dimensions, along the coordinate axes of which the values of *s-generalised coordinates* q and *s-impulses* p of this system (s is the number of degrees of freedom) are stored in the *Hamiltonian*, known to a student of mechanics under the name of the kinetic energy or work-energy theorem (The operator form of the Hamiltonian in canonical form is quite beautiful, universal, described in the literature, but requires little experience with partial derivatives [10]. More precisely, the application of the statistical approach in physics requires knowledge of the basics of quantum mechanics [11, 12], the concept of the *Heisenberg uncertainty principle*, the *Schrödinger equation*, the density matrix, the Hamiltonian operator, the remarkable properties of *hermitian operators*, the rules for working with commutators and Poisson brackets, the large *canonical Gibbs distribution* [13, 14], the basic formulas of thermodynamics, but the presentation below will be based mainly on knowledge of the materials of the school curriculum).

Phase Space

Each point in the phase space corresponds to a certain state of the system. During the evolution of the system, the phase point representing it describes a certain phase trajectory in the phase space. The product of differentials of coordinate and impulses:

$$dqdp = dq_1 dq_2 \dots dq_s dp_1 dp_2 \dots dp_s \quad (2.5)$$

is considered as an 'element of volume' of the phase space Γ . In relation to the formula under study, taking into account the shape of the cuboid and the equality of the increment of the linear function to its differential, the sign Δ will be used further along with the designation of the differential. Let's introduce the probability

$$dw = \rho(q_1, q_2, \dots, q_s, p_1, p_2, \dots, p_s) dpdq \quad (2.5.a)$$

have values lying in the specified infinitesimal intervals between q_i and $q_i + dq_i$, p_i and $p_i + dp_i$. The volume of the phase space is expressed in terms of energy according to the formula:

$$\Delta\Gamma = \frac{d\Gamma(\bar{E})}{dE} \Delta E \dots; \dots \rho(\bar{E}) \Delta q \Delta p = 1 \quad (2.6)$$

Here above, in the left formula before semicolon, the number of quantum microstates with energies less than or equal to E is denoted by $\Gamma(E)$. The above formula

helps to find the number of states of interest with energy between E and $E + dE$. It is assumed that an inertial frame of reference is chosen in which the total momentum and angular momentum of the whole subsystem under consideration is zero. The dash above the energy E indicates the average value of the energy of the subsystem for all quantum states.

Probability Density Function

Statistical averaging is also used for a macrostate. The second part of the formula (2.6) after semicolon introduces the concept of the classical probability density distribution p , which, when multiplied by the volume of the phase space, forms the total probability unit (here and below the notation Δq and Δp are products of the form (2.5) over all degrees of freedom). From the perspective of the quantum mechanical approach, the integral over all probabilities of the subsystem states also forms the total probability $\int W(E)dE = 1$ – this is the *normalization condition*.

Due to the nature of quantum mechanics and the statistical physics based on it, we can only talk about finding the probability distribution for coordinates or impulses separately, not both together, since the coordinates and momentum of a particle cannot have certain values at the same time. The desired probability distributions must take into account both the statistical uncertainty and the uncertainty directly inherent in the quantum mechanical description.

Conservation of Phase Volume. Entropy

According to *Liouville's theorem*, the volume of phase space has the property of invariance with respect to canonical transformations that preserve integrals of motion, in particular energy, momentum and angular momentum [17, pp. 192–193]. Similarly, in statistical physics, the time derivative of the statistical matrix w_{km} , the operator \hat{w} , must be commutative with the Hamiltonian of the system. For the physical phenomena under consideration, this condition is fulfilled. And the result is a quantum mechanical analogue of Liouville's theorem: the commutativity of an operator on any quantity with a Hamiltonian is precisely the quantum mechanical expression of the persistence of that quantity.

For the quasi-classical case the laws of quantum mechanics can be replaced by Hamilton's equations, which are a reformulation of the well-known Newton equations from the school. At the same time, the volume of the phase space

$$\Delta\Gamma = \frac{\Delta q \Delta p}{(2\pi)^s} \quad (2.7)$$

The value of $\Delta\Gamma$ is called the statistical weight of the macroscopic state of the subsystem, and its logarithm is called entropy. In turn, entropy is determined through the average value, denoted by parentheses $\langle \rangle$, of the logarithm Γ both for the case of describing a subsystem from the position of its quantum state from the probability distribution function in energy E_n , and in the quasi-classical approach from the probability density, according to the formulas:

$$S = -\langle \lg w(E_n) \rangle = -\sum_n w_n \ln(w_n) \quad (2.8)$$

$$S = -\langle \ln [(2\pi\hbar)^s \rho] \rangle = -\int \rho \cdot \ln [(2\pi\hbar)^s \rho] dpdq \quad (2.9)$$

The entropy defined in this way is, like the statistical weight itself, a dimensionless quantity. The minus sign before the two formulas above takes into account the *normalization condition* for probability. (In school physics and in some theoretical physics textbooks, the Boltzmann constant $k_B = 1.38 \cdot 10^{-23}$ Joules per Kelvin J/K is placed in front of the logarithmic sign in the last formula, which makes it possible to measure the absolute temperature in Kelvin rather than in energy units, but for simplicity it will be more convenient to omit the Boltzmann constant without necessity).

Let's consider a closed system as a whole (i.e. isolated from other systems, e.g. in a thermostat), where $\Gamma_1, \Gamma_2, \Gamma_3 \dots$ -are the statistical weights of its different subsystems. If each of the subsystems can be in one of the Γ_α states, then the phase volumes of the subsystems are multiplied and the entropies of the subsystems add up, which is illustrated by the expressions

$$\Delta\Gamma = \prod_{\alpha} \Delta\Gamma_{\alpha}; \dots S = \sum_{\alpha} S_{\alpha} \quad (2.10)$$

In other words, phase space allows the separation of variables. In practice, it is often necessary to deal with cases where not all of the microscopic particle motion is quasi-classical, but only the motion corresponding to a part of the degrees of freedom denoted as s , while for the rest of the degrees of freedom (2.5, 2.5.a) of the subsystem due to the motion is quantum (for example, the translational motion of molecules with the quantum nature of the intramolecular motion of atoms can be quasi-classical).

Since GCD – Greatest Common Divisor (GCD) – is simply the product of the terms in formula (2.1), it can be expressed as:

$$LCM \left(e^{\alpha_1 x_1} e^{\alpha_2 x_2} \dots e^{\alpha_k x_k} + e^{\beta_1 y_1} e^{\beta_2 y_2} \dots e^{\beta_l y_l} = e^{\gamma_1 z_1} e^{\gamma_2 z_2} \dots e^{\gamma_k z_l} \right) \quad (2.11)$$

where each real number x, y, z are chosen such that only one prime number is selected from those listed in formula (2.1). This is easily achieved by taking the natural logarithm by replacing the variable $\ln(\lambda_i)/\lambda_i = \ln(a_i)$ (See below *Eigenvalues and eigenvectors*. Looking ahead, we note that in classical mechanics, the variable x cannot be chosen unambiguously because entropy is determined up to an arbitrary term). The powers α, β, γ also taken directly from the original formula (2.1). Suppose generalised impulse Δp_i for the role of the multiplier taken into account in formula (2.1) and let Δq_i equal to 1 unit, provided *uncertainty principle* $\Delta p_i \Delta q_i > \frac{1}{2}\hbar$, where \hbar is the reduced Planck constant equal to $1.055 \cdot 10^{-34}$ Joules Kelvin *sec. One can control the entropy by choosing the thermodynamic parameters: pressure, volume, temperature, quantity of moles of ideal gas in the mixture. Quantum states in which all the impulses have certain values correspond to the free movement of all the particles. It is possible to construct a coordinate density matrix. As a result, the Oesterlé-Masser conjecture is a mathematical interpretation of a complex system made up of simple subsystems that do not interact but share common resources. where the principle of additivity of entropy (2.7, 2.8) and multiplication of volumes of phase spaces of subsystems is fulfilled.

Formula (2.1) is suitable for expressing the law of conservation of entropy / energy of a subsystem in the separation of variables and *adiabatic processes*, which are quite slow compared to the relaxation time / establishment of local equilibrium. In practice, this time passes enough quickly and is comparable to the propagation time of

a sound wave in a container filled with gas (for air the velocity of waves 330 meters per second at normal atmospheric pressure and a temperature of 20°C). Regarding of the exponent in formula (2.1), we note that the well-known formulae for the *distribution of the probability density* of finding particles are expressed in terms of an

exponent: for example, the *Maxwell distribution* $\left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T}}$, the *Boltzmann distribution* $n_0 e^{-\frac{mgh}{T}}$ studied in physics and mathematics schools.

The Canonical Gibbs Distribution

The conclusion of the *canonical Gibbs distribution* is quite simple, based on an equally probability distribution over all quantum states of the system, which can be attributed to both mathematics (foundations of probability theory) and physics [14]:

$$w_n = \exp\left(\frac{F - E_n}{T}\right), \dots \rho = (2\pi h)^{-s} \exp\frac{F - E(p, q)}{T} \quad (2.12)$$

The probability distribution function for the energy of the subsystem in the E_n state is given before the decimal point in the quantum mechanical representation, and after it in the quasi-classical representation. Here $w_n(p, q)$ is the probability of being in a state corresponding to the wave function ψ_p with eigenvalue E_n . F is the free energy of the gas (In thermodynamics, work is done due to the difference between the free energies of the final and initial states of the gas, which explains the convenience of working with this variable), ρ is the probability density of being in a given phase difference, $E(p, q)$ is the energy as a function of generalised coordinates, s is the number of degrees of freedom. In thermodynamics, free energy F , entropy S , energy E , volume V , pressure P , temperature T and a number of other variables are derived and controlled. The first three are additive functions. For example, e_k is the energy of a gas $E(p, q)$ divided by the number of gas molecules.

When not all the motion of microscopic particles is quasi-classical, but only the motion corresponding to a part of the degrees of freedom, while for the rest of the degrees of freedom the motion is quantum (for example, the translational motion of molecules can be quasi-classical with the quantum nature of the intramolecular

dynamics phenomena). In this case, the energy levels of the subsystem can be written as functions of the quasi-classical coordinates and momentums $E(p,q)$.

In physics and mathematics schools, the *Boltzmann distribution*, a *barometric formula* for a rarefied gas in a gravitational field, and the *Maxwell distribution* are studied in terms of velocities (impulses). These are special cases of the *canonical Gibbs distribution*. In the first case the potential energy is replaced by a negative sign in the numerator of the exponent fraction, in the second case by kinetic energy. In both cases the denominator remains T (or $k_B T$).

For example, due to the identity of the gas molecules in the Maxwell distribution, it is sufficient to consider the phase space of only one molecule, expressing the energy in the impulse representation in the coordinates $E = \frac{1}{2} m(p_x^2 + p_y^2 + p_z^2) + \varepsilon_k$, where ε_k is the energy of a gas divided by the number of gas molecules, the k -th energy level of a molecule, due to its rotational, vibrational degrees of freedom, the intrinsic angular momentum of the elementary particles, the spin, etc. Next, the integration over the physical volume of the system is replaced by a simple multiplication of the integral over the coordinates of the generalised impulses by the volume of the vessel. This greatly simplifies the calculations. As a result, to find the free energy F of an ideal gas the following formula is used [13]:

$$F = -NT \ln \left[\frac{eV}{N} \left(\frac{mT}{2\pi\hbar^2} \right)^{3/2} Z \right] = -NT \ln \frac{eV}{N} + Nf(T) \quad (2.13)$$

where N is the number of molecules, T is the absolute temperature in energy units, e is the Euler number 2.718.... V is the volume of the vessel, m is the mass of the molecule, \hbar is the reduced Planck constant equal to $1.055 \cdot 10^{-34}$ Joules *sec. Z is a *statistical sum* depending on the energy levels of the molecule. The definition is given below. The initial temperature, volume and number of particles are considered as externally set parameters of the subsystem under consideration. An alternative representation of just mentioned formula, written after the decimal point, is to put all the parameters into a separate function that depends only on the temperature $f(T)$.

The entropy S is defined from thermodynamic relations as a partial derivative with a negative sign of the free energy:

$$S = -\frac{\partial F}{\partial T} = N \ln \frac{eV}{N} - N \frac{df}{dT}(T) \quad (2.14)$$

Based on the basic equations of thermodynamics, it is possible to find other thermodynamic variables of the subsystem under study, than to completely determine or set its macroscopic motion.

Degenerate Energy Levels

When different wave functions ψ_p have the same eigenvalue of energy E_n , this means degenerate energy levels phenomena. The exponent in formula (2.1) corresponds to the multiplicity of degeneracy of the energy level, denoted as the *statistical weight* g_k . In this case, part of the formula (2.13) contains a statistical sum determined by the formula [14]

$$Z = \sum_k g_k e^{-\epsilon_k/kT} \quad (2.15)$$

The normalisation condition for formula (2.8) allows us to calculate the free energy F via the probabilities of the quantum states:

$$1 = \sum_n w_n = e^{F/T} \sum_n e^{-E_n/T} \Rightarrow F = -T \ln \sum_n e^{-E_n/T} \quad (2.16)$$

Let us use the general formula (2.13) to calculate the free energy F of an ideal gas obeying Boltzmann statistics. By writing the energy of E_p as the sum of the energies of the molecules. For example, we can reduce the summation over all the states of a gas to the summation over all the states of a single molecule, as in the case of the Maxwell distribution. Each state of the gas is determined by a set of N (N is the total number of molecules in the gas) values of ϵ_k , which in the Boltzmann case can all be considered to be different from each other (in each microstate – no more than one unit, due to the restrictions imposed by the barometric formula). Writing $e^{-E_n/T}$ as the product of the multipliers $e^{-\epsilon_k/T}$ for each of the molecules and summing independently over all the states of each molecule, we obtain an expression corresponding to N

$$\left(\sum_k e^{-\epsilon_k/T} \right)^N = \sum_n e^{-E_n/T} = \frac{1}{N!} \left(\sum_k e^{-\epsilon_k/T} \right)^N \quad (2.17)$$

At the same time, all sets of N different values, differing only in the distribution of identical gas molecules over the levels of ϵ_k , correspond to the same quantum state of the gas. In the statistical sum in the above formula, each of the states should be taken into account only once. Therefore, we must also divide the expression obtained by the number of possible permutations of N molecules with each other, i.e. by $N!$ and evaluate the value using the *Stirling formula* $n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$. As a result of this estimation, we obtain a formula for the free energy of a *Boltzmann ideal gas*:

$$F = -NT \ln \left[\frac{e}{N} \sum_k e^{-\epsilon_k/T} \right] \quad (2.18)$$

In the case of degenerate energy levels, the number of repetitions of the corresponding value of ϵ_k is equal to the degree of degeneracy. What happens to formulae (2.14) – (2.15) for a mixture of gases?

$$\begin{aligned} w_{n,N} &= A \exp \frac{\Omega + \mu N - E_{n,N}}{T}; \dots \\ w_{n,N_1,N_2,\dots} &= A \exp \frac{\Omega + \sum_i \mu_i N_i - E_{n,N_1,N_2,\dots}}{T} \end{aligned} \quad (2.19)$$

Here above, before the semicolon, is the distribution function of the subsystem of an ideal gas over two variables – the energy value E_n and the number of particles N . Ω is the *thermodynamic potential*, T is the absolute temperature in energy units, μ is the *chemical potential* of the molecule. After the semicolon, the distribution function of the subsystem of a mixture of ideal gases of a gas is determined by variables – the energy value E_n and the number of particles N_1, N_2, N_3, \dots . As in the case of a single gas, degenerate energy levels are possible in a mixture of gases. The energy $E_{n,N}$ can also be expressed as the sum of the energies of the molecules ϵ_k , reducing the summation over all states of the gas to the summation over all states of a single molecule.

For a mixture of ideal gases, the above formulae remain valid, but instead of N , the number of molecules of the corresponding gas is substituted, and the number of molecules, energy, entropy and partial pressure of each gas are added. According to the rules of thermodynamics, the volume and temperature (due to *equipartition theorem*) are the same for all molecules of each gas in the mixture. Since there is no interaction

between the molecules, the entropy values are summed and the volumes of the phase spaces are multiplied accordingly (2.10).

It is easy to trace the analogy of formula (2.15) with the logarithm of formula (2.1) and verify their similarity. Formulas modelling the free energy and entropy of a mixture of gases differ from the logarithm of formula (2.1), but this should not be discouraging. Formula (2.8, 2.9, 2.10, 2.19) modelling phase space and entropy is sufficient. Pairs of the product of $\mu_i N_i$ in the exponential sum enter symmetrically. In formula (2.19) it is possible to sum over all states of energy levels instead of molecules. The state of the energy levels is determined by the quantum theory of atom.

What is meant by the radicals of the numbers A, B, C?

The Oesterlé-Masser conjecture works with the concept of the quality q_{ABC} based on the decimal logarithm also for the natural logarithm to, since the q_{ABC} is defined as a fraction and the simultaneous multiplication of numerator and denominator by $1/\ln 10$ to change from the decimal to the natural logarithm does not change the quotient.

As an analogue to the radical of the product of the numbers A, B, C $\text{Rad}(ABC)$, a mixture of gases with non-degenerate energy levels can be chosen, which is achieved at a sufficiently low temperature when only the translational and rotational degrees of freedom for polyatomic molecules are activated. Suppose there are no degenerate energy levels. Taking into account the above, this means the first power of the prime numbers of formula (2.1).

By controlling the *Clapeyron relation* $PV = k_B NT$ it is possible to control the temperature T, the gas pressure P by adjusting the initial temperature, the number of molecules N in the mixture (concentration) and the volume of the subsystem. Thermodynamic relations are used to easily calculate the energy of the gas molecules and the entropy, and thus the volume of the phase space $\Gamma_i = \exp(S(E_i))$ (select $\Gamma_1, \Gamma_2, \Gamma_3, \dots$ – where the statistical weights of gases 1, 2, 3 in the mixture are each equal to the first degree prime numbers of formula (2.1)) – the radical $\Gamma_0 = \text{Rad}(ABC)$. Let's call such a value of the volume of the phase space Γ_0 and the corresponding entropy $S_0 = \ln(\text{Rad}(ABC))$. If each of the subsystems can be in one of the Γ_a quantum states, then

the phase volumes of the subsystems are multiplied and the entropies of the subsystems add up.

In the case of adiabatic compression (expansion), the fundamental thermodynamic equation takes the form that the internal energy of the gas decreases (increases) only by the amount of work done by the gas (over the gas). According to the formula of *the first law of thermodynamics*, $dQ = dE + PdV$, where dQ is the amount of heat, P is the gas pressure, dV is a small increase in volume, the multiplier pdV is equal to the R work done. In an adiabatic process, $dQ = 0$, there is no heat exchange, there are no *dissipative processes*, and therefore all changes in the subsystem remain reversible.

In thermodynamics, the relationship between temperature and volume in an adiabatic process is $PV^\gamma = \text{const}$ or $TV^{\gamma-1} = \text{const}$, where $\gamma > 1$ is the quotient of the heat capacity at constant pressure divided by the heat capacity at constant gas volume C_p/C_v . in the usual case. During adiabatic compression, the entropy of the system S_C remains constant, but degenerate energy levels of e_k are ‘turned on’ due to an increase in temperature. This is accompanied by the appearance of degrees above one in formula (2.1).

Next, the phase volume corresponding to some generalised coordinates/impulses, e.g. related to quantum effects, must be assigned to a separate subset of the studied phase space and labelled with the letter A and B i.e. $S_{AB} = S_A + S_B$ correspond to the same microstate. The remaining phase volume corresponding to the generalised coordinates of classical motion is labelled with the letter B. As stated above in formula (2.9), entropy is an additive function so $S_A + S_B = S_C$.

Because of (2.8), the adiabatic compression preserving entropy and the conditions of the experiment:

$$S_A + S_B = S_C = S_0 = \ln \text{Rad}(ABC) \quad (2.20)$$

Since entropy is preserved during the adiabatic process, it is possible to return the subsystem to its initial state $S_A + S_B = S_C \rightarrow S_0$. Therefore, the equal sign in the last formula can be replaced by the identity sign.

In other words, for formula (2.1) there is a whole class of physical subsystems: ideal gases, solutions, vibrations of the crystal lattice with quasi-particles – phonons, etc., where the principles of multiplication of volumes of phase spaces and addition of entropy are implemented. On the basis of the statistical matrix it is possible to calculate the superficiality of the numerical coincidence $S_A + S_B = S_C$ and the volume of the phase space in the studied subsystem with the formula (2.1). There is therefore an equivalence function f that maps each element as an elementary cube 1^n from the sets A and B into C , i.e. $f(A,B) \rightarrow C$.

What is the quality indicator of the triple q or of the deviations from the stated principles? To answer this question, let's look at the whole system, including the subsystem under consideration and the *fluctuations*.

Fluctuations

Taking into account the arguments above and formula (2.20), it is necessary to find a suitable interpretation for the quality indicator of the triple q_{ABC} . As will be shown below, asymptotically approaches to one $q \rightarrow 1$ as N increases (In thermodynamics, the number of molecules is operated on the order of the number/Avogadro constant – the number of particles contained in one mole of any substance $6.022 \cdot 10^{23}$). Most well-known examples of finding a triple of numbers q_{ABC} exceeding one are obviously less than Avogadro's constant.

Assuming $q = 1$ as the reference value of the three numbers A, B, C , for large N , it is more convenient to replace the word ‘quality’ with a more appropriate ‘defect’ according to the formula of fluctuation from one, namely $\Delta q = \Delta S/S_0$. For an isolated subsystem, the statement of fluctuations in energy, entropy, temperature, number of particles and other thermodynamic parameters is not appropriate. It is possible to talk about fluctuations only by considering the system as a whole. Let the equilibrium entropy of a body $S(E,V)$ be defined as a function of its (average) energy and volume.

We will understand by the fluctuation of entropy the change in the function $S(E,V)$, considered formally as a function of the exact (fluctuating) values of energy and volume. It is known that the probability of fluctuation in the system is proportional to the exponential deviation of entropy. This refers to the entropy of the system as a whole.

$$w = e^{\Delta S} \quad (2.21)$$

To continue let's find the average square of the fluctuation in the number of particles of an ordinary ideal gas located in a relatively small volume isolated in the gas. Based on the uniform distribution of gas molecules throughout the entire volume V_0 and the conditions for the smallness of the subsystem under study in comparison with the entire V/V_0 system, the fluctuation of the number of particles in the subsystem is determined by the formulas:

$$\langle (\Delta N)^2 \rangle = N; \dots \frac{(\langle (\Delta N)^2 \rangle)^{1/2}}{N} = \frac{1}{\sqrt{N}} \quad (2.22)$$

The relative fluctuation of the number of particles is equal to the inverse square root of the average number of particles. All these formulae are well known from textbooks of mathematical statistics.

Based on the average square of the fraction in a given volume of gas, a Gaussian probability distribution of the fluctuation of the number of particles is found:

$$w(N)dN = \frac{1}{\sqrt{2\pi\bar{N}}} \exp\left(-\frac{(N - \bar{N})^2}{2\bar{N}}\right) dN \quad (2.23)$$

This formula is valid for small deviations. For large deviations, it is more appropriate to use the *Poisson formula*, known from mathematical statistics.

$$W_n = \frac{\bar{N} e^{-\bar{N}}}{N!} \quad (2.24)$$

In the case we are interested in (marked volume) $V \ll V_0$ (initial volume), the number of particles in the assigned volume, although it may differ significantly from its average value, is assumed to be small compared to the total number of particles in the gas. Taking into account the Stirling formula, formula (2.24) becomes formula (2.23) with small deviations.

Regarding the question of entropy fluctuations, we note the following. Let S_{total} be the total entropy of the body (system) together with the medium (supersystem); if the body is in equilibrium with the environment, then S_{total} is a function of their total energy. If there is no equilibrium with the medium, the total entropy is different from the value of S_{total} . A minimum input of R_{min} from an external source is required to move the system from a state of equilibrium with the environment to a new state of

equilibrium with the environment. The entropy deviation can be expressed as a smooth increasing function of the minimum work with a negative sign. It is known from the basic equations of thermodynamics that the derivative $dS_{\text{total}}/dE_{\text{total}} = T_0$ is the equilibrium temperature of the system. As a result, we have:

$$\Delta S_{\text{total}} = -\frac{R_{\text{min}}}{T_0} = -\frac{\Delta E - T_0 \Delta S + P_0 \Delta V}{T_0} \quad (2.25)$$

Omitting the intermediate calculations with thermodynamic variables from (2.18) and (2.22), we eventually come to the expression of the probability of transition from one state to another due to fluctuations [13]:

$$w = \text{Const} \cdot \exp \left[\frac{1}{2T} \left(\frac{\partial V}{\partial P} \right)_S \Delta P^2 - \frac{1}{2C_p} (\Delta S)^2 \right] \quad (2.26)$$

From which, in particular, the statistical independence of fluctuations in pressure P and entropy S follows. The value of the entropy fluctuation that interests us is determined by the formula:

$$\langle (\Delta S)^2 \rangle = C_p \approx k_B N \quad (2.27)$$

where C_p is the heat capacity of the gas (mixture) at constant pressure, k_B is the Boltzmann constant, N – number of molecules. This value is constant for an ideal gas/mixture and is proportional to the number of particles in the subsystem under study. (For example, for a monatomic gas $C_p = 5/2 k_B N$ for a diatomic gas $C_p = 7/2 k_B N$, where it is easy to find the heat capacity of a mixture of gases by averaging using weighted coefficients taking into account the composition of the mixture).

According to *Chebyshev's theorem* [15, chapter 9.3.], given a sufficiently large number of independent random variables with limited variances, an event can be considered almost reliable, which means that the deviation of the arithmetic mean of the random variables from the arithmetic mean of their mathematical expectations will be arbitrarily small in absolute terms.

As a result, it is possible to apply the Gaussian probability distribution of entropy fluctuations by analogy with the formula for the fluctuation of the number of particles (2.23) in the subsystem under study, to estimate $q_{\text{ABC}} = 1 + \Delta q = 1 + \Delta S/S_0$ using the

Laplace function $\Phi(x)$, assuming that the variance $\sigma = \frac{\sqrt{k_B N}}{\ln(Rad(ABC))}$ of the fluctuation is:

$$|\Delta q_{ABC} < \epsilon| \leq 2\Phi\left(\frac{\epsilon}{\sigma}\right) \dots \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz \quad (2.28)$$

and for probability density function $p(\Delta q) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\Delta q^2}{2\sigma^2}\right)$ (2.29)

Taking into account the rapid decrease of the Laplace function $\Phi(x)$, for any predetermined ϵ there are only a finite number of triples of numbers for which the defect index q_{ABC} is not included in the interval $1 - \epsilon \leq q_{ABC} \leq 1 + \epsilon$ (ϵ as an arbitrary parameter should not be confused with the energy of a single molecule!) It is easy to prove that this conclusion is equivalent to another representation of the *defect* q_{ABC} , where $K(\epsilon)$ is a certain constant depending only on ϵ .

To estimate the maximum deviation q_{ABC} , we can take into account the discrete nature of the phase space and the principle of equal probability of microstates. Taking into account the abscissa axis the ordinate axis of the histogram of probability distribution of q_{ABC} the scale division price should be $1/\Gamma$ and $\frac{1}{\Gamma \ln \Gamma}$ respectively.

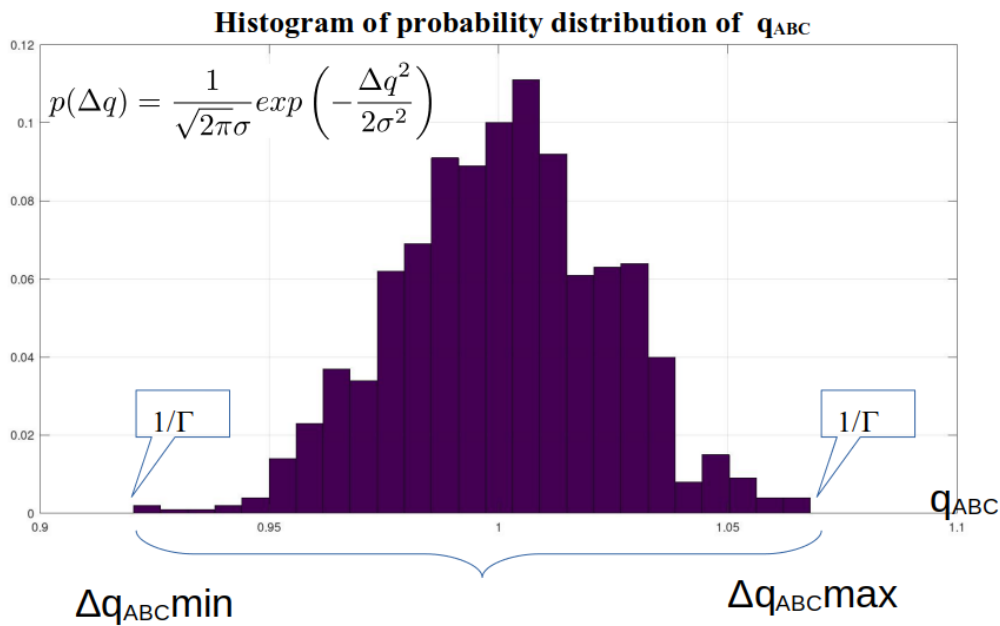


Fig. 2.1. This histogram illustrates the discrete nature of the phase space $q \rightarrow 1$. For $\Delta q_{ABC} > \Delta q_{ABCmax}$ there is no single triple that satisfies the Oesterlé-Masser conjecture

So, a lower limit other than zero, corresponding to the maximum deviation q according to the law of normal distribution (2/29) by formula:

$$\frac{1}{\Gamma} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\Delta q_{max}^2}{2\sigma^2}\right) \quad (2.30)$$

and resolving this equation:

$$\begin{aligned} \Delta q_{max} &= \sigma \sqrt{\left(2 \ln \frac{\Gamma}{\sqrt{2\pi}\sigma}\right)} = \\ &= \frac{\sqrt{k_B N}}{\ln(Rad(ABC))} \sqrt{2 Rad(ABC) + \ln \ln Rad(ABC) - \frac{1}{2} \ln(k_B N)} \end{aligned} \quad (2.31)$$

In the above, it is assumed that the variance σ of the fluctuation $\Delta q = \Delta S/S_0$ (2.27) is

defined as follows $\sigma = \frac{\sqrt{k_B N}}{\ln(Rad(ABC))}$. The larger N , the narrower the range of possible values of q (Note, Avogadro N_A constant – the number of particles contained in one mole of any substance $6.022 \cdot 10^{23}$).

where $\ln(k_B N_A) = 8.31 \text{ J/mol} \cdot \text{K}$. As a rule, the logarithm of this value can be neglected in comparison with other values below the radical. An arbitrary parameter here is N – the total number of particles in the gas mixture. It is much smaller than the number of phase states $Rad(ABC)$ and $\lim_{x \rightarrow \infty} \ln(x)/x = 0$ can also be neglected.

As a result, for 1 mole of a mixture of ideal gases, whose volume under normal conditions is 22.4 litres, we obtain Δq_{max} in according with formula (2.22):

$$\Delta q_{max} \sim \frac{4}{\sqrt{\ln Rad(ABC)}} \quad (2.32)$$

With the assumptions made above it follows from the above arguments that for $\Delta q_{ABC} > \Delta q_{ABCmax}$ there is no single triple that satisfies the Oesterlé-Masser conjecture.

Chapter 3. The ABC Conjecture from the Point of View of Pure Mathematics

Is it possible to find a proof of the Oesterlé-Masser conjecture from the point of view of pure mathematics? – It is possible if we use the wave function representations of quantum mechanics, the operator form of the Hamiltonian, estimates of entropy [14, 1.18] by a *statistical matrix* [13, pp. 28–29] in the quantum mechanical representation, and the rules

of linear algebra. The rapid development of quantum computing and cryptography is also helping to popularise knowledge of quantum mechanics. To broaden the horizons of a student of the XXI century, it would be logical to include the basics of quantum mechanics in the high school curriculum, as well as the chemical formula based on the same principle.

The school curriculum includes the basics of combinatorics and probability theory, pendulum motion, oscillator equations, basics of thermodynamics, internal energy, quantity of heat, first law of thermodynamics, wave interference, signal spectrum, resonance phenomena of wave-particle dualism and photons, phenomenon of photoelectric effect, atomic orbitals and spin, Heisenberg's uncertainty principle, basics of laboratory measurements and errors, and so on. – This extensive list of concepts allows for an analogy from the school curriculum in the course of the presentation below. *Hilbert spaces* were studied in the first decade of the 20th century by David Hilbert, Erhard Schmidt and Frigyes Rees in the theories of partial differential equations and quantum mechanics.

Based on the conditions of the Oesterlé-Masser conjecture let's choose *additive function*. The most suitable candidate is the phase space used in differential control (2.6–2.7) and entropy, defined as its logarithm (2.8). From the conservation of entropy in *adiabatic processes* it follows that the volume of the phase space is conserved. This simple statement is the key to the proof and requires almost no mathematical calculations.

Since formula (2.1) implies a particular system consisting of subsystems that are not connected but share a common resource (energy, phase space) in a particular way, this formula can be compared to an algebraic expression in the form of co-prime polynomials and then decomposed as characteristic polynomials of a system of differential equations.

The set of all expressions that can be constructed from the variable x using addition and multiplication operations is the algebra of polynomials of one variable with natural coefficients. It is possible to determine the homomorphism between the natural numbers of formula (2.1) and the algebra of polynomials. Homomorphism preserves operations. As will be shown below, for the system under study it is more convenient to consider the evolution of the phase space from the theory of ordinary differential equations.

$$(x - a_1)^{\alpha_1} (x - a_2)^{\alpha_2} \cdots (x - a_k)^{\alpha_k} + (x - b_1)^{\beta_1} (x - b_2)^{\beta_2} \cdots (x - b_l)^{\beta_l} + (x - c_1)^{\gamma_1} (x - c_2)^{\gamma_2} \cdots (x - c_m)^{\gamma_m} \quad (3.1)$$

This polynoms also are *co-prime* [16, p. 333]. Then replace these polynomials with the *characteristic polynomials*. Consider matrix equations for square matrices of block diagonal type such as:

$$(\mathcal{A}-\lambda\mathbf{I})^k \mathbf{x} = \mathbf{0}, \dots (\mathcal{B}-\lambda\mathbf{I})^l \mathbf{y} = \mathbf{0}, \dots (\mathcal{C}-\lambda\mathbf{m})^l \mathbf{z} = \mathbf{0} \quad (3.2)$$

where λ are the eigenvalues of the matrices $\mathcal{A}, \mathcal{B}, \mathcal{C}$. Our focus is on the matrix \mathcal{A} , each eigenvalue of λ corresponds to a root vector \mathbf{x} . The set of all root vectors corresponding to the eigenvalue λ forms an *invariant subspace* R_i .

Construction of Invariant Subspace

This subspace has a non-zero dimension if and only if λ are the eigenvalues of the mapping. Let $\lambda_1, \lambda_2, \dots, \lambda_i$, – denote different eigenvalues. Each eigenvalue may correspond to an eigenvector \mathbf{l}_i , or some eigenvalues may be repeated several times (you can choose a subspace basis from the root vectors, the dimension of which will correspond to the multiplicity of repetitions of the eigenvalue. The set of such eigenvalues and the repetition multiplicity are taken from formula (3.1). If $P(\lambda) = \det|\mathcal{A} - \lambda\mathbf{I}|$ is a characteristic polynomial, then the *Hamilton-Calley theorem* $P(\mathcal{A}) = 0$, in other words, every square matrix satisfies its *characteristic equation* [16, p. 333]

From the above follows the possibility to decompose the space R into a direct sum of subspaces of root vectors:

$$R = R_1 \oplus R_2 \oplus R_3 \oplus \dots R_i.. \quad (3.3)$$

The mapping \mathcal{A} has a single eigenvalue λ_i in each of the subspaces. In a basis consisting of proper and attached vectors, the matrix \mathcal{A} takes a *block diagonal* form composed of *Jordan cells*. The cell sizes k correspond to the dimension of the subspace of vectors found from the solution $(\mathcal{A}-\lambda\mathbf{I})^k \mathbf{x} = \mathbf{0}$ and are equal to $n - rang (\mathcal{A}-\lambda\mathbf{I})^k$. The order of the Jordan cells does not matter. All of the above also applies to matrices. $\mathcal{A}, \mathcal{B}, \mathcal{C}$ [16, p. 336].

Eigenvalues and Eigenvectors

Starting with the simple case of the first powers of $k = l = m = 1$ we place the eigenvalues on the diagonal of the matrices by solving the equation $\ln(\lambda_i)/\lambda_i = \ln(a_i)$ and, correspondingly for $\ln(b_i)$ and $\ln(c_i)$, which can be solved, for example, graphically or by the method of successive approximations. Then, by replacing the variables $w_i = 1/\lambda_i$, we

obtain, new diagonal matrix \mathcal{D} consisting of the combination of the three matrices mentioned, a matrix of the form: $\begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}$ the diagonals of which have eigenvalues for each invariant subspace. In this case, the *trace of the matrix* denoted as Sp is equal to:

$$Sp(D) = - \sum_{i=1}^{k+l+m} w_i \ln w_i = \prod_{i=1}^{k+l+m} a_i = \ln Rad(ABC) \quad (3.4)$$

By extending this matrix to infinity, both in rows and columns, we obtain the familiar statistical matrix for the system under investigation (2.8), (2.20). We present the matrix in more detail:

$$\mathbf{D} = \begin{pmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & a_{22} & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & b_{11} & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & b_{22} & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & b_{33} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & c_{11} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & c_{22} & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{33} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (3.5)$$

This matrix is block diagonal. The size of each cell is 1. A quantum mechanical system can be mapped onto this linear mapping. Let's move from Euclidean space to *Hilbert space*. Here, orthogonal vectors are an orthonormal system of functions, so that any operator \hat{Q} is expressed in terms of $\langle n|\hat{Q}|m \rangle = \int \psi_n^* \hat{Q} \psi_m dq$. If \hat{Q} is explicitly independent of time and commutes with the *Hamiltonian* $[\hat{H}, \hat{Q}] = 0$ then its matrix elements $\langle n|\hat{Q}|m \rangle$ do not change with time. This is the quantum form of *integrals of motion*. It is important to cancel out that for our case all eigenvalues – energy are real, this means that the Hermite operator \hat{H} – the corresponding matrix, being *transposed* and *conjugated*, will turn into itself again. This condition provides an important property of the commutativity of the Hamiltonian with a number of operators, including energy, entropy and the derivation of the statistical matrix by time $[\hat{H}, \hat{w}] = 0$.

The eigenvalue corresponds to the energy of the subsystem.

$$\hat{H}\psi = E\psi \quad (3.6)$$

The eigenvalues of the Hamiltonian of a system consisting of two subsystems take the form: $\hat{H} |n, m\rangle = (E_1 + E_2) |n, m\rangle$, this formula can be extended to the case of a number of subsystems and verify the analogy with invariant subspaces, see (3.3).

With sufficiently slow adiabatic processes (see the last paragraph *Conservation of phase volume. Entropy*) the volume of the phase space does not change, but states of degenerate energy levels arise in the subsystems of the system under consideration. From a mathematical point of view, this can be reduced to a block diagonal matrix $C = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, where the size of each Jordan cell corresponds to the multiplicity of the root of the eigenvalue in the characteristic polynomial:

$$C = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & \dots \\ a_{21} & a_{22} & 0 & 0 & 0 & \dots \\ 0 & 0 & b_{11} & b_{12} & b_{13} & \dots \\ 0 & 0 & b_{21} & b_{22} & b_{23} & \dots \\ 0 & 0 & b_{31} & b_{32} & b_{33} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (3.7)$$

The multiple repetition of the eigenvalue corresponds to the multiplicity of the degeneracy of the energy level and the degree of the simple multiplier from (3.1.).

In this case, the total values of the matrix are distributed values, where the multiplicity $\alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_l, \gamma_1, \gamma_2, \dots, \gamma_m$ and derive the ratio $\alpha_k \ln(\lambda_i)/\lambda_i = \alpha_k \ln(a_k)$ and the ratio $\beta_l \ln(\lambda_i)/\lambda_i = \beta_l \ln(b_\beta)$, $\ln(\lambda_i)/\lambda_i = \gamma_m \ln(c_\gamma)$, where the values α, β, γ they coincide with case of each power equals to one in formula (3.1). In the statistical matrix, the probability distribution density per eigenvalue decreases as the size of the Jordan cells increases.

As a result, formula (3.4), which expresses the conservation of entropy as well as the volume of the phase space in *reversible* adiabatic processes, remains valid. Thus, the conclusion can be formulated in matrix form in the form of the following equation:

$$Sp(\hat{w}(\ln \hat{w}))_A + Sp(\hat{w}(\ln \hat{w}))_B = Sp(\hat{w}(\ln \hat{w}))_C = S_0 = \ln Rad(ABC) \quad (3.8)$$

(Note that the formula above used statistical matrices generated from the original one A, B, C). All matrices are infinite and reduced to Jordan form, they contain cells with repeating eigenvalues. (See also the last sentence in Chapter 2 Construction above). Taking into account the *fluctuation* estimate above, the Oesterlé-Masser conjecture is proved. Since entropy is conserved during the adiabatic process, it is possible to return the subsystem to

its initial state $S_A + S_B = S_{AB} = S_C \equiv S_0 = \ln \text{Rad}(ABC)$ where $S_{AB} = S_A + S_B$ correspond to different blocks of the same matrix.

Summing up

Based on the theory of ordinary differential equations

It is interesting to make some analogies of the quantum mechanical approach outlined above with the theory of ordinary differential equations (ODE). Let the Hamiltonian H of the system under study be clearly independent of time for this case.

$$i\hbar\dot{\psi} = \hat{H}\psi \quad (3.9)$$

The *unitary operator* of the finite time shift is determined by the formula [12, 19.4.]:

$$\hat{U}(t) = e^{i\hat{H}t/\hbar}$$

The transformation operator $\hat{U}(t)$ translates a vector $\psi(0)$ into a vector $\psi(t)$ as follows:

$$\begin{aligned} \psi(t) &= \hat{U}(t)\psi(t_0) \\ \psi(t_0) &= \hat{U}^{-1}(t)\psi(t) \dots \psi(t_0) = \hat{U}^+(t)\psi(t) \end{aligned} \quad (3.10)$$

Let's compare the last two formulas with some system of ordinary differential equations expressed in matrix form:

$$\dot{\mathbf{Y}} = \mathbf{A}(t)\mathbf{Y} \quad (3.11)$$

where A is a square matrix of linear mapping (3.2), generally time-dependent t . We assume that the matrix A is non-degenerate, and all eigenvalues are different. For general case determinant of the such matrix is equal to the product of the eigenvalues: $\det|A| = \lambda_1, \lambda_2, \dots, \lambda_i$. It is easy to verify that the initial volume of the phase space $\Gamma_0 = \text{Rad}(ABC)$ because the phase volume can be easily calculated from the geometric shape of the manifolds (2.1)).

Floquet proved in a well-known theorem that for periodic systems with a period T (in the sense of time, not temperature) [18, chapter VIII (16)]:

$$\mathbf{Y}(t, 0) = \Phi(t)\exp(t\mathbf{M}) \quad (3.12)$$

where $\Phi(t)$ is a periodic function with period T and $\mathbf{Y}(T, 0) = \Phi(T)\exp(T\mathbf{M})$. Eigenvalues $\mathbf{Y}(T, 0)$ are called multipliers of the *T-periodic system*. Based on the above,

in our case, all these multipliers are modulo no more than one. The formula (3.12) corresponds to the unitary operator of the finite time shift (3.10).

By the Liouville-Ostrogradsky theorem [18, chapter VIII (18)]:

$$\det |\mathbf{Y}(t)| = \det |\mathbf{Y}(t_0)| \exp \int_{t_0}^t \text{Sp} A(\tau) d\tau \quad (3.13)$$

(a formula similar in meaning has been found for the determinant of the Polish philosopher and mathematician *Vronsky*, which is convenient for finding the derivative of a vector function in a system of linear differential equations of the form 3.11.) This means that along with the phase space used in the theory of ordinary differential equations, it is advisable to operate with the concept of the trace of the A matrix.

In analytical mechanics, the functions of the dynamic variables $f(p, q, t)$ that remain constant during the evolution of the system are called *integrals of motion*. The condition of constancy of some function f $df/dt = 0$ is expressed through *Poisson brackets*, similar to the above-mentioned switch in quantum mechanics:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{H, f\} \quad (3.14)$$

For the *integral of motion* f , the Poisson brackets with the Hamiltonian H must become equal to zero.

With an adiabatic change, the phase volume remains constant $\Gamma_0 = \text{Rad}(ABC)$. Therefore, the trace of the matrix in (3.11) must be zero. Since the integral enters under the sign of the exponent, this means that the trace matrix (3.13) corresponds to the concept of entropy and is associated with the phase space of the formula $S = \ln \Gamma_0 = \text{Rad}(ABC)$ (meaning, the operation of differentiation by the upper limit of the integral (3.13)). In other words, for adiabatic processes, entropy remains constant too $\Delta S = 0$.

The fluctuations of the phase space caused by the uncertainty principle and its maximum value Δq_{\max} (2.22) can be calculated from normal distribution function, additivity of entropy, discrete character of phase space and *law of large numbers* [19]. As a result, the the Oesterlé-Masser conjecture was proved based on the theory of ODE.

And one more his quotation of Minhyong Kim: *We're at a point where our understanding of physics is mature enough, and there are enough number theorists interested in it, to make a push*) [17]. So physical laws helped in the search for

evidence, like Ariadne's thread. Chapter I, II can be presented by an experienced teacher based on knowledge of the school curriculum in physics, chemistry and mathematics.

Look at a 3D cube!

For a creative student, it is enough to look at a 3D cube with its faces or a globe with its equator to understand the proof of Fermat's Last Theorem. Imagine a construction corresponding to the expression in n -dimensional space $a^n = c^n - b^n$. On the left is a symmetrical figure of dimension n . On the right is a set of faces with one dimension less. This set will be asymmetric or have an irreparable symmetry defect or a violation of the continuity of the layers for a space of dimension greater than two. Therefore, an object whose fundamental properties are described in a contradictory way there is no such thing in nature. Such a phenomenon was called an *aporia* in ancient Greece, or a logically consistent construction that corresponds to nothing. The proposed proof of (1.1) forces us to make broader generalisations about the properties of the universe itself, and therefore Fermat's Last Theorem should be included in the curriculum of general education schools of the XXI century.

The author's simple proof helps us understand the asymmetry of our universe. Look at the faces of beautiful women and notice that man, a child of the universe, must be slightly asymmetrical to be beautiful. Is this by chance?

Conclusion

Interestingly, in the concentric sphere theories of the ancient Greeks Eudoxus, Callippus, and Aristotle, the Earth was at the center of the universe and surrounded by symmetrical spheres, all of which were considered perfectly symmetrical. But it turns out that even when we worked with our symmetrical spheres, we were convinced that this design did not correspond to any physical object in nature. Such a phenomenon was called in ancient Greece the *aporia* or logically consistent constructions that do not correspond to anything.

It is surprising that we came to the same conclusion both for the case of working with n -cubes and balls. Absolutely symmetrical Universe is impossible. For the origin of matter, it is necessary to operate with the concept of volume/measure, and this is

possible only in a 'slightly' asymmetric, anisotropic space at n more than two, but not in Euclidean space. The deep nature of this conflict stems from the fundamental physical properties of our Universe. If it were perfectly symmetrical, like the studied construct of concentric spheres, then matter with its intrinsic properties of matter/measure conservation could not arise in it.

This way, by investigating number theory, we have approached questions of physics and worldview without even applying higher mathematics, general relativity theory, and complex differential equations to illustrate the principles of the origin of our Universe. The assumptions that were made within the framework of Euclidean geometry in relation to cosmic scales become inaccurate here non-Euclidean geometry is already at work.

In anisotropic space, the conflict between the form and content of the structures we study is constructively resolved, but to do so, we must abandon Euclidean geometry! Look at portrait of the Russian scientist Nikolai Ivanovich Lobachevsky, the rector of Kazan University, a mathematician who carefully studied the fifth postulate of Euclid, whose axioms we have been considering, and formulated his own so-called imaginary geometry. This geometry was later called Lobachevsky's geometry. It turned out that Lobachevsky's geometry describes surprisingly well the concept of *space-time interval*, the *Lorentz transformation*. The echoes of the Big Bang manifest themselves in the form of relic radiation, low-temperature photons with a temperature of 2.72 Kelvin, which can be studied with the latest equipment and get information about the anisotropy of our Universe.

And indeed, 1983–1984 four, the Russian relic experiment confirmed the anisotropy of relic radiation – the 'echo' of the Big Bang. The measurements were carried out with a radiometer developed at the Institute of Space Research of the USSR Academy of Sciences, in the group of Igor Arkadyevich Strukov (the general management of the program was carried out by Academician Nikolay Semenovich Kardashev) on the launched satellite 'Prognoz-9'. In 2006 the discovery of Russian scientists was 'evaluated on merit' by awarding the Nobel Prize in Physics to Americans George Smoot and John Mather for repeating the results of Russian

physicists, but with greater precision [20] The names of Russian scientists remained in the shadows.

Number theory is extremely useful in school physics classes. It also plays an interesting role in pedagogy.

References

1. Wiles, A.J. Fermat's Last Theorem and Modular Elliptic Curves // *Annals of Mathematics*. 1995. №141. P. 443–551.
2. Boston, N. The Proof of Fermat's Last Theorem. Washington DC: ASM Press, 1994.
3. Belova, L.Y., Belov, Y.A. Elements of the Theory of Sets and Mathematical Logic. Theory and Problems: a manual. Yaroslavl: Yaroslavl State University Press.
4. Mochizuki, S. Mochizuki's Proof of ABC Conjecture // *Publications of the Research Institute for Mathematical Sciences*. 2021. №1/2 (57). P. 12
5. Viro, O.Ya. Ivanov, O.A. Netsvetaev, N.Yu. Kharlamov, V.M. Elementary Topology, Elementary Topology Problem: textbook. 2020.
6. Euclid, M.L. Beginnings. GTTI. 1948.
7. Avdiyev, M.A. Education and Upbringing of Children and Adolescents: from theory to practice. Ulyanovsk: Zebra. 2020.
8. Masser, D.W. Open Problems // *Proceedings of the Symposium on Analytic Number Theory*. London: Imperial College. 1985.
9. Oesterlé, J. Nouvelles Approches du 'Théorème' de Fermat // *Séminaire N. Bourbaki*. – 1988. – Vol. 694. – P. 165–186. – ISSN 0303–1179.
10. Landau, L.D., Lifshitz E.M. Theoretical Physics In 10 vol. Vol. 1. Mechanics. Moscow: Fizmatlit. 2004.
11. Zelevinsky, A.V. Quantum Mechanics: NSU lectures, 6th semester. 2002
12. Fermi, Enrico. Lectures on Quantum Mechanics: 2nd edition. Izhevsk: Regular and Chaotic Dynamics.

13. Landau, L.D., Lifshitz E.M. Theoretical Physics In 10 vol. Vol. 5. Statistical Physics. Moscow: Fizmatlit. 2004. ISBN 9785922115100
14. Kubo, R., Ichimura H., Usui, T., Hashitsume, N. Statistical Mechanics: an advanced course with problems and solutions: 7th edition. 1988.
15. Gmurman, V.E. Fundamentals of Probability Theory and Mathematical Statistics. American Elsevier Pub. Co. 1968.
16. Szidarovszky, F. & S. Introduction to Matrix Theory with Applications to Business and Economics. 2002.
17. Hartnett, K. Secret Link Uncovered between Pure Math and Physics. New York: Quanta Magazine. 2017.
18. Myshkis, A.D. Introductory Mathematics for Engineers: lectures on Higher Mathematics. Mir. 1972. ISBN 9780828507271, 0828507279
19. Sedor, K. The Law of Large Numbers and its Applications. Mathematical Sciences in Conformity with the Requirements. Lakehead University Press. 2015.
20. Skulachev, D.P. Correlation of Data on the Anisotropy of Relic Radiation in the WMAP and Relict-1 Experiments // Uspekhi Fizicheskikh Nauk. 2020. №8 (10). P. 389–392.