Sequences with zero sums

I. Bogdanov, K. Kokhas, K. Kuyumzhiyan

Problems

1 *Zero-sequences*

For any nonnegative integer n let $\mathbb{Z}_n$ denote the set of residues modulo n equipped with operation "+" (addition mod n). We say that the sequence in $\mathbb{Z}_n$ is a zero-sequence if its sum equals $0 \in \mathbb{Z}_n$.

1.1. Let n be a nonnegative integer. Let k be the minimal number such that every sequence of length k in $\mathbb{Z}_n$ contains a zero-subsequence. Prove that $k = n$.

1.2. Describe all the sequences in $\mathbb{Z}_n$ of length $n-1$ that do not contain zero-subsequences.

1.3. Describe all the sequences in $\mathbb{Z}_n$ of length $n-2$ that do not contain zero-subsequences.

1.4. What is minimal m such that every sequence of length m containing at least 81 different elements of $\mathbb{Z}_{100}$ has a zero-subsequence of length 100?

1.5. Find a minimal m such that every sequence of length m in $\mathbb{Z}_{100}$, consisting of exactly 4 distinct elements, has a zero-subsequence of length 100.

1.6. Prove that every sequence in $\mathbb{Z}_{12}$ of length 23 contains a zero-subsequence of length 12.

1.7. Let S be a sequence of length 502 in $\mathbb{Z}_{541}$ that has exactly 10 different elements. Prove that S contains a zero-subsequence.

1.8. Let S be a sequence of length 10 in $\mathbb{Z}_{17}$ that does not contain zero-subsequences. Prove that some element of $\mathbb{Z}_{17}$ occurs in S at least 4 times.

1.9. Let S be a sequence of n integers coprime with n. Prove that every residue modulo n is a sum of some subsequence in S.

1.10. Let S be a sequence in $\mathbb{Z}_n$ of length $2n-1$. Assume that some element a occurs at least $[n/2]$ times in S. Prove that S contains a zero-sum subsequence of length n.

1.11. Let p be an odd prime number. Consider the following sequence in $\mathbb{Z}_p$: $0,\ 0,\ 1,\ 1,\ 2,\ 2,\ \ldots,\ p-1,\ p-1$. How many zero-subsequences of length p does this sequence have?

1.12. Let $n \geqslant 2$ be an integer. Let S be a sequence of length n in $\mathbb{Z}_n \setminus \{0\}$ with nonzero sum. Prove that there exist at least $n-1$ different zero-subsequences in S.

2 *Extremal problems*

For every nonnegative integer n by $\mathbb{Z}_n^d$ denote the set of all arrays of the form $(m_1, m_2, \ldots, m_d)$, where each m_i is a residue modulo n. We equip this set with operation "+" (addition modulo n in each coordinate). A zero-sequence in $\mathbb{Z}_n^d$ is a sequence that has zero sum (i.e. the sequence whose sum equals $(0,0,\ldots,0) \in \mathbb{Z}_n^d$). By $g(n,d)$ denote the minimal number M such that in every subset of $\mathbb{Z}_n^d$ consisting of M elements there exist n elements with zero sum. By $s(n,d)$ denote the minimal number M such that in every sequence $a_1, \ldots, a_M \in \mathbb{Z}_n^d$ there exists a zero-subsequence of length n. In other words there exist different indices $i_1, \ldots, i_n$, where $1 \leqslant i_1 < i_2 < \cdots < i_n \leqslant M$, such that $a_{i_1} + \ldots + a_{i_n} = 0$. Thus different sequences may coincide if considered simply as lists.

2.1. Prove that $s(2,d) = 2^d + 1$.

2.2. Prove that $(n-1)2^d + 1 \leqslant s(n,d) \leqslant (n-1)n^d + 1$.

2.3. Prove that $s(n_1 n_2, d) \leqslant s(n_1, d) + n_1\big(s(n_2,d) - 1\big)$.

2.4. Prove that что $g(3,3) \geqslant 10$, $s(3,3) \geqslant 19$. (Actually $g(3,3) = 10$, $s(3,3) = 19$.)

2.5. Prove that $g(n,2) \geqslant \begin{cases} 2n-1 & \text{for odd } n; \\ 2n+1 & \text{for even } n. \end{cases}$

2.6. Consider a square 3×3 on the grid paper. Let 9 points be marked in the nodes of the grid (including the boundary of the square). Prove that the center of mass of some 4 of these points is a node of the grid too. In other words prove that $g(4,2) = 9$.

2.7. Prove that $s(2048, d) = 2047 \cdot 2^d + 1$.

2.8. Prove that $s(432, 2) = 1725$.

3 *Erdős – Ginzburg – Ziv theorem and related questions*

3.1. [Cauchy – Davenport theorem] Let p be a prime number, A and B be two nonempty subsets in $\mathbb{Z}_p$. Prove that $|A+B| \geqslant \min\{p,\ |A|+|B|-1\}$.

3.2. [Erdős – Ginzburg – Ziv theorem] Prove that every sequence in $\mathbb{Z}_n$ of length $2n-1$ contains a zero-subsequence.

We can use these theorems in the subsequent problems.

Many of the following problems are valid in a more general context. A commutative finite group is a finite set equipped with the operation "+" that satisfies the usual axioms for this operation: commutativity $a+b=b+a$; associativity $a+(b+c)=(a+b)+c$; existence of zero element with the property $0+a=a$ for all a; and existence of inverse element: for any a there exists b such that $a+b=0$ (we write $b=-a$). We can define a difference of two elements using the last axiom: $a-b$ is defined to be equal to $a+(-b)$. A typical finite commutative group looks like $\mathbb{Z}_d^n$: fix the set of numbers $k_1, \ldots, k_n$ and a set of "vectors" $(x_1, x_2, \ldots, x_n)$, where $x_i \in \mathbb{Z}_{k_i}$, and the operation "+" is usual coordinatewise addition (modulo k_i for i-th coordinate).

We mark the problem by the sign † if its statement is valid for an arbitrary finite commutative group. Solutions that do not use specific properties of $\mathbb{Z}_n$ and remain valid in the case of an arbitrary finite commutative group are especially welcome.

3.3. Let p be a prime number; let $A_1, A_2, \ldots, A_k$ be nonempty subsets in $\mathbb{Z}_p$. Prove that

$$|A_1 + A_2 + \ldots + A_k| \geqslant \min\{p,\ \left(\sum_{i=1}^{k} |A_i|\right) - k + 1\}.$$

3.4. Let p be a prime number and let $S = (a_1, \ldots, a_{2p-1})$ be a sequence in $\mathbb{Z}_p$ such that its elements $a_1, \ldots, a_s$ are pairwise distinct ($s \geqslant 2$). Prove that S has a zero-subsequence of length p that contains at most one of the elements $a_1, \ldots, a_s$.

3.5. a) Take a regular 12-gon on the plane. We consider all symmetries and rotations of the plane which preserve the 12-gon. Prove that for any 47 transformations there exist 24 of them such that their composition (in some order) is an identity transformation.

b) Prove a similar statement for the symmetric group S_4.

3.6. Let p be a prime number, T be a sequence in $\mathbb{Z}_p \setminus \{0\}$ of length p, h be a maximal multiplicity of elements in T. Prove that every element of $\mathbb{Z}_p$ is a sum of at most h elements of T.

3.7. Assume $m \geqslant k \geqslant 2$, and let m be divisible by k. Prove that every sequence of integers of length $m+k-1$ contains a subsequence of length m whose sum is divisible by k.

3.8.† [Kemperman – Scherk theorem] Let n be a nonnegative integer. Let A and B be two subsets of $\mathbb{Z}_n^d$, such that $0 \in A$, $0 \in B$ and the equation $a+b=0$ has only the trivial solution $a=b=0$ if $a \in A$, $b \in B$. Prove that $|A+B| \geqslant \min\{n^d,\ |A|+|B|-1\}$.

3.9.† Let k and r be nonnegative integers and let $A = \{a_1, \ldots, a_{k+r}\}$ be a sequence in $\mathbb{Z}_k$, such that 0 is not a k-sum of this sequence. Prove that the sequence has at least $r+1$ different k-sums.

3.10.† Let S be a sequence of length n^d in $\mathbb{Z}_n^d$, let h be a maximal multiplicity of elements in this sequence. Prove that S contains a zero-subsequence of length at most h.

3.11. Let p be a prime number and $2 \leqslant k \leqslant p-1$. Consider a sequence of length $2p-k$ in $\mathbb{Z}_p$, such that every p elements of this sequence have a nonzero sum. Prove that some element occurs in the sequence at least $p-k+1$ times.

3.12. Let B_1, B_2, ..., B_h be a collection of subsets in $\mathbb{Z}_n^d$. Let $m_i = |B_i|$. Assume that $\sum_{i=1}^h m_i \geqslant n^d$. Prove that for each j we can choose at most one element $b_j \in B_j$ in such a way that the sum of the resulting nonempty collection of chosen elements is $0 \in \mathbb{Z}_n^d$.

3.13. Let $n > 4$ be an odd number. Let S be a sequence of length k in $\mathbb{Z}_n$, where $\frac{n+1}{2} \leqslant k \leqslant n$. Assume that S does not contain zero-subsequences. Prove that some element of $\mathbb{Z}_n$ is contained in S with multiplicity $2k-n+1$.

3.14.† Let $A \subset \mathbb{Z}_n^d$, $|A| \geqslant n^d/k$. Prove that there exists a number r, $1 \leqslant r \leqslant k$, and a sequence a_1, ..., a_r (of not necessarily different elements of A) such that $\sum a_i = 0$.

3.15. Prove that the sequence of length $2n-1$ in $\mathbb{Z}_n$ has a unique zero-subsequnce of length n only if it consists of n copies of some number a and $(n-1)$ copies of number b.

3.16.† Let S be a subset of $\mathbb{Z}_n^d$ containing k elements. Assume that S does not contain a subset with zero sum. Prove that there exist at least $2k-1$ distinct elements of $\mathbb{Z}_n^d$ that can be represented as a sum of several elements of S.

3.17. Let $M = \{(a_1, b_1), (a_2, b_2), \ldots, (a_{2p-1}, b_{2p-1})\}$ be a subset in $\mathbb{Z}_p^2$. Assume that every element of $\mathbb{Z}_p$ is contained in the sequence a_1, ..., a_{2p-1} at most twice. Prove that the set M contains a zero-subsequence of length p.

3.18. Prove that the minimal m such that every sequence of length m in $\mathbb{Z}_p^d$ contains a zero-subsequence is equal to $d(p-1)+1$.

4 Semifinal. Additional problems

2.9. Find $g(8,2)$ and, if possible, $g(2^n, 2)$.

2.10. Prove that there exists a function $f(d)$ such that $s(n,d) \leqslant f(d)n$ for all n, d.

3.19. Prove that if $3p$ elements of $\mathbb{Z}_p^2$ with zero sum are given, then one can choose p of them with zero sum.

3.20. Let S be a sequence in $\mathbb{Z}_n$ of length at least $(n+3)/2$, $n \geqslant 3$ without zero-subsequences. Prove that S contains an element coprime with n.

3.21. Let p be a prime number. Let $A \subset \mathbb{Z}_p$ such that $|A| > 2\sqrt{p}$.
a) Prove that A contains a subset with zero sum.
b) Prove that each residue mod n can be represented as a sum of some subset of A.

5 Open questions

4.1. $g(p,2) = 2p-1$ for odd prime p.

For $p = 3$, 5, 7 it has been proved by Kemnitz (Kemnitz A., Exremalprobleme für Gitterpunkte. Ph.D.Thesis. Technische Universität Braunschweig, 1982), and for $p \geqslant 67$ by Gao and Thangadurai. The first step in the latter is problem 3.17.

4.2. $g(n,2) = \begin{cases} 2n-1 & \text{for odd } n \\ 2n+1 & \text{for even } n \end{cases}$

4.3. $s(n,2) = 4n-3$.

This is Kemnitz's conjecture. It was has been proved last year by combination of algebraic and combinatorial methods. The sequence of $4n-4$ elements without zero-subsequences of length n is the following: $(0,0)^{n-1}$, $(0,1)^{n-1}$, $(1,0)^{n-1}$, $(1,1)^{n-1}$.

4.4. Let p be a prime number, A and B be two nonempty subsets in $\mathbb{Z}_p$. Denote by $A \dotplus B$ the set of all sums of the form $a+b$, where $a \in A$, $b \in B$ и $a \neq b$. Prove that $|A \dotplus B| \geqslant \min\{p,\ |A|+|B|-3\}$.

This statement is the conjecture of Erdős–Heilbronn. It is proved though no combinatorial proof is known.

4.5. Let S be a sequence of $4n-4$ elements of $\mathbb{Z}_n^2$. If S does not contain a zero subsequence of length n, then S consists of 4 distinct elements with multiplicity $n-1$ each.

This is Gao's conjecture. He has checked that this conjecture is multiplicative, i.e. if it is valid for $n = k$ and $n = \ell$, then it is valid for $n = k\ell$.

Solutions

1.1. The sequence 1, 1, ..., 1 (n units) does not contain zero-subsequences. Therefore $k \geqslant n$. Now let a_1, a_2, ..., a_n be an arbitrary sequence of length n. We prove that it contains a zero-subsequence. Consider n sums

$$\begin{aligned}
&a_1\,,\\
&a_1 + a_2\,,\\
&\ldots \qquad \ldots\\
&a_1 + \ldots + a_{n-1}\,,\\
&a_1 + \ldots + a_{n-1} + a_n\,.
\end{aligned}$$

If neither sum is equal to zero then two sums are equal (as elements of $\mathbb{Z}_n$). Their difference determines a zero-subsequence.

1.2. Answer: these are sequences containing an element a with multiplicity $n-1$, where a coprime with n.

Indeed, let a_1, a_2, ..., a_{n-1} be a sequence that does not contain zero susequences. Suppose that $a_1 \neq a_2$. Consider n sums

$$\begin{aligned}
&a_1\,,\\
&a_2\,,\\
&a_1 + a_2\,,\\
&a_1 + a_2 + a_3\,,\\
&\ldots \qquad \ldots\\
&a_1 + \ldots + a_{n-1}\,,\\
&a_1 + \ldots + a_{n-1} + a_n\,.
\end{aligned}$$

Since neither of the sums is equal to zero, there are two sums that coincide as elements of $\mathbb{Z}_n$. Then their difference determines a zero-subsequence. A contradiction. Therefore the sequence could not contain two different elements.

1.3. Answer: these are sequences consisting of $n-2$ copies of some element a, or sequences consisting of $n-3$ copies of some element a and one element $2a$.

We found this fact in [16].

The case $n \leqslant 5$ is left to the reader. Let the sequence contain two distinct elements $a_1 \neq a_2$. For any other element a_3 we conclude analogously to the previous solution that there are two identical elements among the elements a_1, a_2, $a_1 + a_3$, $a_2 + a_3$, $a_1 + a_2 + a_3$, ..., $a_1 + \cdots + a_{n-2}$. But in this set of sums only the sums $a_1 = a_2 + a_3$ or $a_2 = a_1 + a_3$ may coincide, whence

$$a_3 = \pm(a_1 - a_2)\,.$$

Thus for every element a_3 we have an equality $a_3 = \pm(a_1 - a_2)$. Since our sequence could not contain elements that are equal to both $a_1 - a_2$ and $-(a_1 - a_2)$ (because their sum is 0), we have the same sign for all possible elements a_3 in the last equality. That is all elelemnts of the sequence except a_1 and a_2 are equal to the same number a. But $a_1 \neq a_2$, therefore one of numbers a_1, a_2, say a_1, is not equal to a. Then analogously all the elements except a_1 and a are equal, i. e. for $n > 5$ we obtain that $a_2 = \cdots = a_{n-2} = a$. And in the case $a_1 \neq a, 2a$ a zero-subsequence can be easily constructed.

1.4. Answer: $m = 102$.

Let $A = \{a_1, a_2, \ldots, a_{81}\}$ be distinct elements of some sequence. If the sequence has length 101 and the sum of all elements does not belong to A, then the sum of all elemnts except one could not be equal to 0. Now let the sequence have length at least 102. Without loss of generality we may think that the length is exactly 102. Let its sum equal to s. It is not difficult to choose two elements a_i and a_j in the set A such that $a_i + a_j = s$. Now by removing these elements we obtain a zero-subsequence of length 100.

1.5. Answer: 197 coins.

1.6. From the given sequence choose 11 pairs of numbers with even sum of numbers in each pair. Divide all the sums by 2. We obtain numbers b_1, b_2, ..., b_{11}. From these numbers choose 5 pairs with even sum of elements in each pair. Again, divide the sums by 2. We obtain numbers $c_1, c_2, \ldots, c_5$. Now consider these numbers modulo 3. It is easy to see that either some resudue mod 3 has multiplicity 3 in this sequence or all 3 residues occur in it. In both cases these three residues have sum 0 mod 3. And recalling the way we have obtained the residues we obtain a desired zero-subsequence mod 12 of length 12.

1.7. This is a partial case of the theorem of Erdős and Eggleton [7]. Denote the number 502 by n, then $541 = n+39$. Let $a_1, a_2, \ldots, a_n$ be given sequence. Suppose that it does not contain a zero-subsequence. Fix for a moment a number m, $1 \leqslant m \leqslant n$. Observe that all the sums obtained as sums of subsequences of first m numbers of the sequence S could not be equal to sums of the form $\sum_{i=1}^{r} a_i$, where $m+1 \leqslant r \leqslant n$ (because the difference of such equal sums would determine a zero-subsequence) . We will show that it is possible to choose m in such a manner that the first m members of S will determine at least $m+39$ different sums.

Since S contains only 10 different elements, there is an element $a \neq 0$ that occurs in S at least 51 times. 541 is a prime number, hence we have the operation of division in $\mathbb{Z}_{541}$. So we can divide all elements of S by a. We will interpret the result of each division as a number from 0 to 540. So we may assume that $a_1 = a_2 = \ldots = a_{51} = 1$ and that all the other elements of S are numbers from 0 to 540. If all these numbers do not exceed 51 then every number from 1 to $\sum_{i=1}^{n} a_i$ can be represented as a sum of some a_i's. But $\sum_{i=1}^{n} a_i \geqslant 1+2+\ldots+10+(n-10) > n+39$. A contradiction.

If otherwise S contains a number greater then 51 we may assume that $a_{52} > 51$. (Observe that $a_{52} < 488$ because in the opposite case we can construct a zero-sequence containing a_{52} and several 1's). Let $m = 52$. It is obvious that the first m members of the sequence S determine at least $m+51$ sums, as we wish.

1.8. Partial case of 3.13.

1.9. Let $a_1, a_2, \ldots, a_n$ be a given sequence. It is trivially checked by induction that the set of sums of all subsequences $a_1, a_2, \ldots, a_i$ mod n contains at least i different residues mod n.

1.10. We take this proof from [16]. Let S be a given sequence of length $2n-1$. Assume $a \in \mathbb{Z}_n$ occurs $s \geqslant [n/2]$ times in S. We may also assume that $s \leqslant n-1$ because otherwise the statement is obvious.

Consider shifted sequence $S-a$, it contains zero s times. Let T_1 be a subsequence of S that consists of all nonzero elements of S, it contains at least $2n-1-s \geqslant n$ elements. By the result of the problem 1.1, we know that T_1 contains a zero-subsequence. Denote it by T_2. Assume T_2 contains t_2 elements, $2 \leqslant t_2 \leqslant n$. We may also assume that the sequence T_2 has maximal possible length (among the sequences of length at most n). If $s+t_2 \geqslant n$, then we can construct a zero-subsequence of length n by adding several 0's to T. Therefore we may assume that $s+t_2 < n$ whence $t_2 \leqslant [n/2]$ (because we still have a lot of zeroes).

Now we can easily obtain a contradiction. Since $s+t_2 < n$, it follows that $T_1 \setminus T_2$ contains at lest n elements and has a zero-subsequence T_3 of length t_3, $t_3 \leqslant t_2$ by the maximality of T_2. But $t_2 \leqslant [n/2]$, whence $t_3+t_2 \leqslant n$. This contradicts the maximality of T_2.

1.11. A n s w e r: $\frac{1}{p}(C_{2p}^{p}-2)+2$. This problem is from IMO1995. We take a solution in [1]. We may assume that the sequence is a union of two identical sets B and C, where $B = C = \{0, 1, 2, \ldots, p-1\}$.

Split all the p-element subsequences of our sequence except B and C into the groups of p subsequences. For every set $X = \{x_1, \ldots, x_k\} \subset B$ denote by $X+\ell$ the set $\{x_1+\ell, \ldots, x_k+\ell\} \subset B$, where addition is mod p (i.e. $X+\ell$ is a usual cyclic shift of the set X) For a sequence A we place into one group the following subsequences

$$A_0 = A;\ A_1 = \big((A\cap B)+1\big)\cup(A\cap C);\ A_2 = \big((A\cap B)+2\big)\cup(A\cap C);\ \ldots, A_{p-1} = \big((A\cap B)+p-1\big)\cup(A\cap C)\,.$$

If $|A\cap B| = q$, $0 < q < p$, then the sums of subsequences A_i and A_j differ by $(j-i)q$. Therefore all the sums of subsequences in one group are pairwise distinct. Hence the sum of exactly one subsequence is divisible by p.

Thus we have $\binom{2p}{p}$ p-element sequences, two of them (with zero sum) we throw off and between others $1/p$ part has a zero sum. So the answer is $\frac{1}{p}(C_{2p}^{p}-2)+2$.

1.12. Let $A = (a_1, \ldots, a_n)$ be a given sequence. In the solution of problem 1.1 we describe how to construct a zero-subsequence elements of which run successively in a given numbering. I.e. the sequence we have obtained in such a way has a form $a_i, a_{i+1}, \ldots, a_j$, where $1 \leqslant i < j \leqslant n$. The sequences of this type we will be referred to as intervals. After we have constructed several zero-sequences we can try to “mix” all numbers by changing enumeration in order to each of our sequences would not be an interval. In case of success of these attempts we can construct one more zero-subsequence by the method of the problem 1.1. This subsequence will be an interval and therefore it will not concide with those that we have constructed previously. Proceeding in this manner we will construct $n-1$ desired subsequences.

In order that this plan works, it remains to prove the following statement.

Lemma. Let $k \leqslant n-2$ subsequences (of the initial sequence A) be given such that they have length n and contain from 2 to $n-1$ elements each. Then we can rearrange the elemnts of A in such a manner that none of subsequences will be an interval in a new enumeration.

Proof. When we do not look after the order of elements we will use the word "set" instead of "sequence".

We use induction by n. Base of induction $n \leqslant 4$ can be easily checked by consideration of all possibilities. Suppose that the statement is proven for some $n \geqslant 4$. Then we will show that it is valid for $n+1$ by proving the following statement.

Let the sequence $B = (b_1, b_2, \ldots, b_{n+1})$ be given. Let $D_1, \ldots, D_k$ be its subsequences containing from 2 to n elements each, $k \leqslant n-1$. Then there exists a new enumeration of the sequence B such that all sequences D_i are not intervals.

Since we have at most $n-1$ sets D_i, then there exists an element b_k that belongs to at most one two-element set D_i; without loss of generality we may assume that this is b_{n+1}. Remove the element b_{n+1} from the initial sequence and from all subsets containing it. If a pair D_i containing b_{n+1} exists then drop it. If a set $D_j = \{b_1, \ldots, b_n\}$ exists then drop it too. If neither first nor second set exists then drop an arbitrary set D_m.

We can apply induction hypothesis to the remaining elements and sets because now all the sets contain from 2 to $n-1$ elements each. We try to find a place in the obtained sequence where we can insert an element b_{n+1} in order to satisfy all conditions. If an arbitrary D_m has been dropped it is sufficient to insert b_{n+1} in such a manner that elements of D_m would not run successively. Clearly it is possible. Otherwise if we have had a pair D_i then we can not insert b_{n+1} on the two positions adjacent to the second element of the pair. And finally if we have had the set D_j then now the "prohibited" positions are at the beginning and at the end of the sequence. Since total number of positions for b_{n+1} is $n+1 \geqslant 5$, we have at least one "allowed" position. Insertion of b_{n+1} in this position provides us a desired sequence.

2.1. The elements of $\mathbb{Z}_2^d$ are all possible collections of d zeros and ones equipped by the operation of addition modulo 2 in each coordinate. There are 2^d such collections, and the sum of any two different collections is not a zero collection. If we take 2^d+1 collections, then two of them are identical. Their sum will be equal to zero.

2.2. The upper estimate follows from the fact that if a sequence consists of $(n-1)n^d+1$ elements of $\mathbb{Z}_n^d$ then some element occurs at least n times. For the lower bound let us give an example. We consider 2^n collections of the form $(m_1, \ldots, m_d)$, where each m_i takes value 0 or 1. We construct a sequence containing each of these collections exactly $n-1$ times. It is clear that the sum of any n elements of this sequence is not equal to 0.

2.3. Suppose a sequence containing $s(n_1,d) + n_1(s(n_2,d)-1)$ elements is given. We will consequtively choose groups, containing n_1 elements, with the sum in each group divisible by n_1, using the definition of $s(n_1,d)$. We can choose $s(n_2,d)$ groups in this way. Now divide the sum of elements in each group by n_1 thus obtaining a sequence of $s(n_2,d)$ elements. After this we can choose n_2 elements of this sequence with the sum divisible by n_2.

2.4. Here we give an example of nine elements of $\mathbb{Z}_3^3$, such that no three of them have zero sum. It was given in [10]:

$$\begin{pmatrix}2\\1\\2\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}1\\2\\2\end{pmatrix}.$$

Furthermore, if we take each vector two times, we get an example of the sequence of 18 elements, such that no three of them have zero sum.

2.5. See [8]. For odd n we can take the following subset of $\mathbb{Z}_n^2$:

$$(0,0), (0,1), \ldots, (0, n-2),\ (1,1), (1,2), \ldots, (1, n-1).$$

This set conyains $2n-2$ elements. If its subset contains n elements then it has nonzero sum. For even n we can similarly consider the following subset of $\mathbb{Z}_n^2$:

$$(0,0), (0,1), \ldots, (0, n-1),\ (1,0), (1,1), \ldots, (1, n-1).$$

This set conyains $2n$ elements. If its subset contains n elements then it has nonzero sum.

2.6. This fact is taken in [8] but the proof there is too technical. If we consider integer vectors a_1, a_2, a_3, a_4 then $\frac{a_1+a_2+a_3+a_4}{4} = \frac{\frac{a_1+a_2}{2}+\frac{a_3+a_4}{2}}{2}$. Split 9 marked points into 4 groups according to the parity of both coordinates. In each group $x_1, x_2, \ldots, x_k$ compute $\frac{x_1+x_2}{2}, \frac{x_1+x_3}{2}, \ldots, \frac{x_1+x_k}{2}$. It yields $k-1$ integer vectors which are pairwise distinct as elements of $\mathbb{Z}_2^2$. Having considered all 4 groups, we get at least $9-4=5$ vectors with integer coordinates.

We can choose two vectors with the coinciding parities of both coordinates of them. The pairs of initial vectors corresponding to them were received from different groups, so they provide the required 4-tuple.

2.7.

2.8. This statement has been proven in AlonDubiner2 by non elementary methods.

3.1. The simple case $|A|+|B|>p$ is left to the reader. We may therefore consider subsets satisfying the condition $|A|+|B|\leqslant p$.

Assume that, in contradiction with our assumptions, the set A and B satsfy the inequality

$$|A+B|<|A|+|B|-1\,. \tag{$*$}$$

For definiteness, assume that $|B|\geqslant|A|$. Among all pairs satisfying the above inequality $(*)$ we choose pairs such that $|A+B|$ achieves its minimum possible value. Among those pairs take an arbitrary pair A, B such that the cardinality of the smaller A achieves its minimum.

If $|A|=1$ then the inequality is obvious.

Consider the case $|A|\geqslant 2$. If instead of the set A we consider for an arbitrary $c\in\mathbb{Z}_p$ the shifted set $A+c$, then the set $A+B$ is also shifted by c, and in particular its number of elements remains the same.

Assume that we may shift the set A in such a way that the sets A and B have a nonempty intersection, but $A\not\subset B$. Consider the sets $A_1=A\cap B$, $B_1=A\cup B$. By definition $|A_1|+|B_1|=|A|+|B|$. But $A_1+B_1\subset A+B$, whence $|A_1+B_1|\leqslant|A+B|$ and the pair A_1, B_1 satisfies the inequalty $(*)$. Since $|A_1|<|A|$, the set A is not minimal, which contradicts our assumptions.

Now we prove that we may shift the set A in such a way that the sets A and B have a nonempty intersection but $A\not\subset B$. Indeed, assume that such a shift is impossible, i.e. any shift of A is either contained in B or does not intersect B. Let a, $b\in A\cap B$, $A\subset B$. Shift the set A by $b-a$ and denote $A'=A+(b-a)$ (all arithmetical operations in our formulas should be understood mod p). The point a is taken by our shift to the point b, and therefore A' and B have a nonempty intersection, whence $A'\subset B$ and b, $b+(b-a)\in A'\subset B$. We now shift A' by $b-a$ and obtain the set A''. Observe that A'' and B have a nonempty intersection. Proceeding inductively we obtain that all points

$$a,\ b,\ b+(b-a),\ b+2(b-a),\ b+3(b-a),\ \ldots$$

belong to the set B. But the set of all such points exhausts the whole set $\mathbb{Z}_p$, and we arrive at a contradiction.

3.2. First assume that we have proved our statement for prime values of n and derive it for all n by induction on the number of prime divisors of n. Let $n=pm$, where p is prime. Let a_1, a_2, $\ldots$, a_{2n-1} be our sequence.

By our statement for prime p we may choose an array I_1 of p elements with zero sum mod p. From the remaining $2(m-1)p-1$ elements we may choose another such array I_2. Proceeding in this way we obtain $2m-1$ distinct arrays I_1, I_2, $\ldots$, I_{2m-1}. (Note that after $2m-2$ arrays have been chosen the remaining number of elemnts is at least $2pm-1-(2m-2)p=2p-1$.) For each i, $1\leqslant i\leqslant 2m-1$, denote $a'_i=\frac{1}{p}\sum_{j\in I_i}a_j$. By the inductive hypothesys the sequence a'_i contains a zero sum sequence of m elements. We denote this elements by a'_{i_1}, a'_{i_2}, $\ldots$, a'_{i_m}. The union of arrays $I_{i_1}, \ldots I_{i_m}$ is the desired set of pm elements with sum divisible by n.

We now prove our statement for prime p. Arrange our sequence in increasing order: $0\leqslant a_1\leqslant a_2\leqslant\ldots\leqslant a_{2p-1}$. If $a_i=a_{i+p-1}$ for some i, then $a_i+a_{i+1}+\ldots+a_{i+p-1}=pa_i=0$ (in $\mathbb{Z}_p$) and the statement is proved. In the opposite case denote $A_i=\{a_i,a_{i+p-1}\}$ for $1\leqslant i\leqslant p-1$. Using the Cauchy−Davenport theorem several times we obtain $|A_1+A_2+\ldots+A_{p-1}|=p$ (one may also use problem 3.3 here). Therefore, every element of $\mathbb{Z}_p$ may be represented as a sum of $p-1$ elements of the sequence a_1, a_2, $\ldots$, a_{2p-2}. But then the element $-a_{2p-1}$ also admits such a representation, and we obtain the desired zero subsequence.

3.3. An obvious induction.

3.4. This problem is taken from [8].

The statement is obvious if some element occurs p times in S. Assume that each element occurs at most $p-1$ times. Denote $A_1=\{a_1,\ldots,a_s\}$, and split other elements of S into $p-1$ sets A_2, $\ldots$, A_{p-1}. By the Cauchy–Davenport theorem,

$$|A_1+\ldots+A_p|\geqslant\min\{p,\sum|A_i|-p+1\}=\min\{p,2p-1-p+1\}=p\,.$$

Hence $A_1+\ldots+A_p=\mathbb{Z}_p$, which implies the desired statement.

3.5. A composition of two axial symmetries is a rotation. We can interpret a rotation of a 12-gon as an element of the group $\mathbb{Z}_{12}$. Let us decompose our collection of transformations into pairs of two symmetries or two rotations (there will be one transformation without a pair; we will not use it). Fix an arbitrary order of elements in each pair. There are 23 pairs, and the composition of elements in each pair is a rotation; these compositions form a collection of 23 elements of $\mathbb{Z}_{12}$. Applying the Erdős – Ginzburg – Ziv theorem to this collection, we obtain the required result.

3.6. Let us split all the elements of T into h sets $A_1, A_2, \ldots, A_h$ in the following *canonical* way: we denote by A_i the set of all elements appearing in T at least i times. Then any element has equal multiplicities in T and in the sets $A_1, \ldots, A_h$; moreover, A_1 contains all different elements of T. All the sums of not more than h elements of T form the set $A_1 + \sum\limits_{i=2}^{h}\big(A_i \cup \{0\}\big)$. By the Cauchy–Davenport theorem, we obtain

$$\Big|A_1 + \sum_{i=2}^{h}\big(A_i \cup \{0\}\big)\Big| \geqslant \min\big\{p,\ |A_1| + |A_2 \cup \{0\}| + \ldots + |A_h \cup \{0\}| - h + 1\big\} = p$$

as required.

3.7. Using the Erdős – Ginzburg – Ziv theorem one can find a zero subsequence of k numbers. Delete them from the sequence and repeat the procedure. This process can be repeated until there are less than $2k-1$ numbers left in the sequence. When this is the case, we obtain m/k groups with zero sums as required.

3.8. This statement is proved in [13].

Assume the contrary. As in the proof of the Cauchy–Davenport theorem, we may restrict ourselves to the case $|A| + |B| - 1 \leqslant n^d$. From all pairs (A, B) contradicting the statement of the problem, choose a pair with minimal $|A|$. Note that there exists an element $b^* \in B$ such that $b^* + A \not\subset B$. Actually, assuming the converse we have the equality $a_1 + B = B$ for any $0 \neq a_1 \in A$. In particular, $a_1 + b_i = 0 \in B$ for some $b_i \in B$ which is impossible.

We will use this element to redistribute our pair of sets. Denote by A^* the set of all $a \in A$ for which $a + b^* \notin B$. Let $A' = A \setminus A^*$, $B' = B \cup (b^* + A^*)$. Obviously, $0 \notin A^*$, hence $0 \in A'$. Moreover $0 \in B' \supset B$.

The equation $a' + b' = 0$ has only a trivial solution on the sets A', B'. Actually, if $a_1 + (b^* + a_2) = 0$, where $a_1 \in A'$, $a_2 \in A^*$, then $(a_1 + b^*) + a_2 = 0$, which is impossible since $a_1 + b^* \in B$, $a_2 \in A$, $a_2 \neq 0$. On the other side, if $a_1 + b_1 = 0$ for $0 \neq a_1 \in A'$, $b_1 \in B$, then we obtain a nontrivial solution of the initial equation again.

Additionally, recall that $A' + B' \subset A + B$.

Finally, we obtain a new pair (A', B') satisfying the same conditions though with $|A'| < |A|$. Contradiction with minimality of $|A|$.

The solution is valid for any finite abelian group.

3.9. This statement is taken from [11]. The Erdős – Ginzburg – Ziv theorem implies that $r \leqslant k - 2$.

If we add some number to all elements (this operation will be referred to as *a shift* of a sequence), then all k-sums remain the same, so one can consider only the case when 0 has the maximal multiplicity in A. Denote by L a subsequence of A, consisting of all zeroes. Obviously, $\ell \leqslant k - 1$. Consider all subsequences in $A \setminus L$ of length not exceeding $k - 1$. Let S be the subsequence of maximal length in this set. Denote $s = |S|$, $\ell = |L|$ (possibly $S = \varnothing$ and $s = 0$). Then $\ell + s \leqslant k - 1$ (conversely, we obtain a zero subsequence by adding to S some zeroes from L).

Then, $|A \setminus (L \cup S)| \geqslant r + 1$. Let T be an arbitrary subsequence of length r in $A \setminus (L \cup S)$. Let h be the maximal multiplicity of an element in T. Then $h \leqslant \ell$ by the definition of ℓ. Split T into h sets $X_1, \ldots, X_h$ in a canonical way and denote $X'_i = X_i \cup \{0\}$ $(i = 1, \ldots, h)$. Note that $0 \notin T$ and for any $1 < j \leqslant h$ no j elements of T have zero sum. On the contrary, we can add these elements to S obtaining a new zero subsequence of length $j + s \leqslant h + s \leqslant \ell + s \leqslant k - 1$, which contradicts to the choice of S. Hence, we have verified the conditions of the Kemperman–Scherk theorem. Applying this theorem, we obtain

$$|X'_1 + \ldots + X'_h| \geqslant |X_1| + \ldots + |X_h| + 1 = r + 1\,.$$

In other words, if we add h zeroes from L into T, then the obtained sequence has at least $r + 1$ h-sums. Adding the remaining $(k + r) - (r + h)$ elements of A to all these sums, we obtain $r + 1$ different k-sums as required.

The Erdős – Ginzburg – Ziv theorem is in fact a corollary of this statement. The arguments are valid for any finite abelian group.

3.10. This result is a consequence of the Kemperman–Scherk theorem. Actually, there is nothing to prove if $0 \in S$. Suppose $0 \notin S$. Distribute elements of S into h subsets $B_1, \ldots, B_h$ in a canonical way and let $A_i = \{0\} \cup B_i$ for $1 \leqslant i \leqslant h - 1$. If an equation $a_1 + a_2 + \ldots + a_{h-1} = 0$ has a nontrivial solution with $a_i \in A_i$, then we obtain the required result. In the opposite case we have

$$|A_1 + A_2 + \ldots + A_{h-1}| \geqslant \Big(\sum_{i=1}^{h-1} |A_i|\Big) - (h-1) + 1 = (n^d + h - 1 - |B_h|) - (h-1) + 1 = n^d + 1 - |B_h|.$$

This implies that there exists an element of $\mathbb{Z}_n^d$ of the form $(-b)$, where $b \in B_h$, which can be represented as a sum of the form $a_1 + a_2 + \ldots + a_{h-1}$, $a_i \in A_i$. Then $a_1, a_2, \ldots, a_{h-1}, b$ is a required zero subsequence.

3.11. This statement is found in [8].

The p-sums remain the same under a shift. Hence, having made a shift we may assume that 0 has the maximal multiplicity h. Let T be a subsequence consisting of all nonzero elements, and s be the sum of its elements. Assume that the statement is not valid and $h \leqslant p-k$. Then $|T| \geqslant p$. Since $0 \notin T$, s can be represented as the sum of not more than h elements of T by problem 3.6. Let Q be a subsequence formed by these elements, and set $T_1 = T \setminus Q$. Clearly, T_1 is the zero sequence, and $p-h \leqslant |T_1| \leqslant |T|-1$.

Suppose $|T_1| < p$. Then one can easily obtain a zero sequence by adding some zeroes to T_1. Hence the only case remaining is $|T_1| > p$. Again, applying problem 3.6, we can eliminate not more than h elements from T_1, obtaining the subsequence T_2 with zero sum, and so on. Finally, for some T_i we will obtain $p-h \leqslant |T_i| \leqslant p$, QED.

3.12. This statement is equivalent to the problem 3.10.

3.13. Found in [5].

Assume the contrary: each element has multiplicity not more than $2k-n$ in S. Note that for $a, b \in S$, $a \neq b$, elements a, b and $a+b$ are pairwise distinct since S does not contain zero subsequences.

Lemma. Suppose $a, b, c \in S$ are pairwise distinct elements. Then there are at least 6 different sums modulo n among all the sums of the set $\{a, b, c\}$.

Proof. Consider the following equations in $\mathbb{Z}_n$:

$$a+b=c, \quad a+c=b, \quad b+c=a\,.$$

Suppose that two of them are satisfied (wlog 1st and 2nd). Adding we obtain $2a=0$ which is impossible since n is odd. If, on the contrary, two of these equalities (again 1st and 2nd) are NOT satisfied, then one can easily check that the sums a, b, c, $a+b$, $a+c$, $a+b+c$ are pairwise distinct. Actually, if, for instance, $a+b+c=a$ then $b+c$ is a zero sum. The lemma is proved.

Let us choose one by one nonintersecting triples of pairwise distinct elements (and eliminate them from the sequence). Assume that we can choose j triples $A_1, A_2, \ldots, A_j$, and the rest of the sequence does not contain such a triple. Then the rest contains only 2 different elements a and b with multiplicities λ and μ, $\lambda \geqslant \mu \geqslant 0$. Our assumption implies that $\lambda \leqslant 2k-n$. We can make μ pairs $\{a,b\} = A_{j+1} = \cdots = A_{j+\mu}$ and $\lambda-\mu$ singles $\{a\} = A_{j+\mu+1} = \cdots = A_{j+\lambda}$ from these elements.

We have split our sequence into $j+\lambda$ sets A_i. So,

$$k = 3j + \lambda + \mu\,.$$

For any A_i we introduce the set ΣA_i of all sums of subsets of A_i. By Lemma we have $|\Sigma A_i| \geqslant 6$ for triples; moreover, $|\Sigma A_i| = 3$ for all pairs and $|\Sigma A_i| = 1$ for singles. It follows that

$$\sum_{i=1}^{j+\lambda} |\Sigma A_i| = 6j + 3\mu + (\lambda-\mu) = 2(3j+\mu+\lambda) - \lambda = 2k - \lambda \geqslant n\,.$$

It remains to apply the statement of the problem 3.12.

Remark. The statement of the problem is also valid for even values of n. For this case, one should add some technical corrections into the solution; these corrections will deal with the case when $n/2$ is present in S.

3.14. A straightforward consequence of problem 3.12 where all sets are supposed to be equal.

3.15. Suppose a sequence of $2n-1$ elements of $\mathbb{Z}_n$ is given in such a way that it has a unique zero subsequence of length n. Let us remove any element appearing in this zero subsequence. Then the remaining sequence S does not have a zero subsequence of length n at all. Let a be an element with maximal multiplicity in S, and s be its multiplicity. Again, we will consider the shifted sequence $T = S - a$ instead of S, 0 being its member with multiplicity s. Let T_1 be the subsequence consisting of all nonzero elements of T, and S_0 be the subsequence consisting of all zeroes.

Clearly, $s < n$. If $s = n-1$, then $|T_1| = n-1$ and T_1 does not contain nonempty zero subsequences at all. By problem 1.2, T_1 consists of $n-1$ equal elements. Hence S has only two different elements, each being with multiplicities $n-1$.

Suppose $s \leqslant n-2$. Then $|T_1| \geqslant n$, and the multiplicities of its elements do not exceed s. By problem 3.10, we may choose a zero subsequence S_1 in T_1 of length $s_1 \leqslant s$. If $s + s_1 \geqslant n$, then we may construct a zero sequence of length n, adding some zeroes to S_1. If $s+s_1 \leqslant n-2$, then we may find another zero subsequence $S_2 \subset T_2 = T_1 \setminus S_1$ of length $s_2 \leqslant s$, and so on. Finally, we will obtain either $s+s_1+\cdots+s_k \geqslant n$ (hence we may find a zero subsequence of length n by adding some zeroes to $S_1 \cup \cdots \cup S_k$) or $s+s_1+\cdots+s_k = n-1$.

If $T_{k+1} = T_1 \setminus (S_1 \cup \cdots \cup S_k)$ does not contain zero subsequences, then by problem 1.2 all its elements are equal, which is impossible since the maximal multiplicity is $s < n-1$. Otherwise there exists a zero subsequence U of length u, and $u+s+s_1+\cdots+s_k \geqslant n$. Removing from the sequence $S_0 \cup S_1 \cup \ldots S_k \cup U$ a suitable number of S_i and then a suitable number of zeroes, we will obtain a zero subsequence of length n. Contradiction.

So, we have shown that S has only two different elements, each being with multiplicities $n-1$. Now let us return the removed element and throw another one appearing in the zero subsequence. Again, we will obtain a sequence with exactly two values; this means that the first deleted element was equal to some other. It follows that the sequence has the required form.

3.16. This is a particular case of another theorem of Erdős–Eggleton [7].

Let $f(k)$ be the least possible number of elements of $\mathbb{Z}_n^d$ which may be represented as a sum of some subsequence of S, where S is an arbitrary k-sequence without zero subsequences. Let us prove that $f(k) \geqslant 2k-1$ by induction on k. Obviously, $f(1) = 1$.

Suppose that $f(k) \geqslant 2k-1$. Consider a sequence $S = (a_1, a_2, \ldots, a_{k+1})$ without zero subsequences. We should check that there are at least $2k+1$ elements which can be represented as a sum of some its elements.

Case 1. Some element of S (say a_{k+1}) cannot be represented as a sum of some other elements of S. By the induction hypothesis, the set of all sums of $S \setminus \{a_{k+1}\}$ consists of at least $2k-1$ elements, and this set contains neither a_{k+1} nor $\sum_{i=1}^{k+1} a_i$ (otherwise the difference of this sum and its representation yields a zero subsequence). Moreover, these elements are obviously different. Hence we have found $2k+1$ different sums.

Case 2. Each element of S can be represented as a sum of some other elements. Recall that S does not contain zero subsequences, hence we can apply the Kemperman–Sherk theorem to the sets $A = B = \{0, a_1, \ldots, a_{k+1}\}$ obtaining $|A+B| \geqslant 2k+3$. Moreover, each element of the form $2a_i$ can be represented as the sum of different elements of S since one can change a_i by its representation as a sum of some other elements. Hence, each nonzero element of $A+B$ can be represented as the sum of some elements of S, hence we have in fact found $2k+2$ of such elements.

3.17. A result from [8]. Since we deal with p-sums, we may shift our sequence again. Then we can assume that there is exactly one 0 and exactly two entries of any other element of $\mathbb{Z}_p$ in our sequence. Then

$$(a_n) = ((0,z),\ (1,x_1),(1,y_1),\ (2,x_2),(2,y_2),\ \ldots,\ (p-1,x_{p-1}),(p-1,y_{p-1}))\,,$$

where $x_i \neq y_i$ for all i (since (i,x_i) and (i,y_i) are different elements of M). Note that $0+1+2+\ldots+(p-1) = 0$ in $\mathbb{Z}_p$. Denote $C_i = \{x_i, y_i\}$. By the Cauchy–Davenport theorem 3.3, we have

$$|C_1 + C_2 + \ldots + C_{p-1}| \geqslant 2(p-1) - (p-1) + 1 = p\,.$$

This means that zero may be represented as $0 = \sum_{i=0}^{p-1} z_i$, where $z_0 = z$ and for each i either $z_i = x_i$ or $z_i = y_i$. The corresponding elements of our sequence form a required subsequence.

3.18. This is a result of Olson [14].

Let us show that the number in this statement cannot be made smaller. Take d “basic” vectors

$$e_i = (0,0,\ldots,0,1,0,\ldots,0)$$

(the unit is on the ith place in the vector e_i), each having multiplicity $p-1$.

The proof of the remaining part of the statement is much more complicated.

3.19. This fact is proved algebraically in [4].

3.20.

3.21.

References

[1] Школьные олимпиады. Международные математические олимпиады / Сост. А. А. Фомин, Г. М. Кузнецова. М: Дрофа, 1998.

[2] Alon N. Subset sums // J. Number Theory. 1987. Vol. 27. P. 196–205.

[3] Alon N., Dubiner M. A lattice point problem and additive number theory // Combinatorica. 1995. Vol. 15. P. 301–309.

[4] Alon N., Dubiner M. Zero-sum sets of prescribed size. In: Combinatorcs, Paul Erdős is Eighty. Bolyai Society, Math. Studies, Keszthely. Hungary, 1993. P. 33–50.

[5] Bovey J. D., Erdős P., Niven I. Conditions for a zero sum modulo n // Canad. Math. Bull. 1975. Vol. 18. № 1. P. 27–29.

[6] Davenport H. On the addition of residue classes // J. London Math. Soc. 1935. Vol. 10. P. 30–32.

[7] Eggleton R. B., Erdős P. Two combinatorial problems in group theory // Acta Arithmetica. 1972. Vol 19. P. 111-116.

[8] Gao W.D., Thangadurai R. A variant of Kemnitz conjecture // J. of Comb. Theory. Ser. A. 2004. Vol. 107. P. 69–86.

[9] Halberstam H., Roth K.F., Sequences. New York, Heidelberg, Berlin: Springer-Verlag, 1966.

[10] Harborth H. Ein Extremalproblem für Gitterpunkte // J. Reine Angew. Math. 1973. Vol. 262/263. P. 356-360.

[11] Hong Bing Yu. A simple proof of a theorem of Bollobás and Leader // Proc. of the AMS. 2002. Vol. 131. № 9. P. 2639–2640.

[12] Kemnitz A. Exremalprobleme für Gitterpunkte. Ph.D.Thesis. Technische Universität Braunschweig, 1982.

[13] Moser L., Scherk P. Distinct elements in a set of sums // Amer. Math. Monthly. 1955. Vol. 62. P. 46–47.

[14] Olson J.E., A combinatorial problem on finite Abelian groups, II // J. Number Theory. 1969. Vol. 1. P. 195–199.

[15] Olson J.E., An addition theorem modulo p // J. Comb. Theory. 1968. Vol. 5. P. 42–52.

[16] Thangadurai R. Non-canonical extensions of Erdős–Ginzburg–Ziv theorem // Integers: Electronic J. of Comb. Number Theory. 2002. V. 2.

Алгебраические группы и проблема Бернсайда

А. Я. Канель-Белов   И. А. Иванов-Погодаев   А. С. Малистов

Данный проект посвящен комбинаторно-геометрическому методу, позволившему решить несколько сложных проблем в теории групп, в том числе построить бесконечную конечно-порожденную группу с тождеством $\boldsymbol{x^n = 1}$ (ограниченная проблема Бернсайда). Эта конструкция является главной целью настоящего цикла. В основном, мы будем следовать построениям А. Ю. Ольшанского. Суть метода состоит в существовании наглядной геометрической интерпретации вывода следствий при изучении абстрактных алгебраических объектов.

После введения некоторых полезных понятий, мы перейдем к построению диаграмм на плоскости, обычно представляющих собой карту из многоугольников. В этом смысле полезны следующие вводные задачи.

♦ **A1.** *Выпуклый* 1993*-угольник разрезан на выпуклые семиугольники. Докажите, что найдутся четыре соседние вершины* 1993*-угольника, принадлежащие одному семиугольнику.* (*Вершина семиугольника не может лежать внутри стороны* 1993*-угольника.*)

♦ **A2.** *Можно ли разрезать плоскость:* a) *на выпуклые семиугольники;* b) *на одинаковые выпуклые семиугольники*?

♦ **A3.** *Можно ли разрезать плоскость на выпуклые семиугольники так, чтобы любой единичный круг пересекал не более миллиона из них*?

♦ **A4.** *Плоскость разбита на выпуклые семиугольники, диаметры которых не превышают* 1*. Пусть $n(R)$ — количество семиугольников, попавших в круг радиуса R с центром в начале координат. Доказать, что существуют числа R_0 и $\lambda > 1$ такие, что для всех $R > R_0$ верно $n(R) \geqslant \lambda^R$.*

Рассмотрим конечный алфавит $\boldsymbol{L}$, состоящий из букв

$$\boldsymbol{a,\ a^{-1},\ b,\ b^{-1},\ c,\ c^{-1},\ \ldots}$$

В наш алфавит буквы входят парами: $\boldsymbol{a}$ и $\boldsymbol{a^{-1}}$, $\boldsymbol{b}$ и $\boldsymbol{b^{-1}}$ и так далее. Такие буквы будем называть *обратными*. Из букв можно составлять слова —- произвольные конечные последовательности букв — например: $\boldsymbol{aba}$, $\boldsymbol{aba^{-1}ab^{-1}c}$, $\boldsymbol{a^{-1}a^{-1}bbca^{-1}}$. Слова можно преобразовывать, убирая или вставляя пары рядом стоящих обратных букв. Так, слово $\boldsymbol{aba^{-1}ab^{-1}c}$ приводится к виду $\boldsymbol{ac}$: $\boldsymbol{ab\underline{a^{-1}a}b^{-1}c \equiv a\underline{bb^{-1}}c \equiv ac}$. Некоторые слова (например, $\boldsymbol{b^{-1}aa^{-1}b}$) можно привести к пустому слову, которое будем обозначать как $\mathbf{1}$.

Зададим несколько слов, равных $\mathbf{1}$ по определению, например: $\boldsymbol{\{aba^{-1}b^{-1} = 1;\ ca = 1\}}$. Такие равенства будем называть *определяющими соотношениями.* Из них можно извлекать следствия. Например, возьмем соотношение $\boldsymbol{aba^{-1}b^{-1} = 1}$ и допишем справа в обоих частях этого равенства слово $\boldsymbol{ba}$ (это называется “умножим справа на $\boldsymbol{ba}$”). Получим равенство $\boldsymbol{aba^{-1}b^{-1}ba = ba}$. Сокращаем в левой части обратные буквы и получаем $\boldsymbol{ab = ba}$. Теперь допишем в начале обоих частей равенства $\boldsymbol{c}$ (“умножим слева на $\boldsymbol{c}$”). Получим $\boldsymbol{cab = cba}$. Пользуясь $\boldsymbol{ca = 1}$, получаем $\boldsymbol{b = cba}$.

Итак, задание определяющих соотношений приводит к тому, что некоторые слова становятся *эквивалентными* или *равными*: одно можно привести к другому пользуясь определяющими соотношениями и сокращениями обратных букв. Допустим, нам известны определяющие соотношения. Можно ли узнать по заданной паре слов, равны они или нет? Оказывается, общего алгоритма для этого не существует, и это очень серьезный результат в высшей алгебре. Однако, если определяющие соотношения удовлетворяют некоторым ограничениям, такой алгоритм может существовать.

Основной вопрос. Пусть для любой пары определяющих соотношений $\boldsymbol{A} = \mathbf{1}$ и $\boldsymbol{B} = \mathbf{1}$ их общее начало либо пусто, либо по длине составляет менее $^1/_6$ от длины $\boldsymbol{A}$ и $\boldsymbol{B}$. Тогда существует алгоритм, позволяющий для любой пары слов выяснить, равны они или нет.

К основному вопросу мы вернемся, сначала попрактиковавшись на более простых примерах. Везде далее мы предполагаем, что алфавит состоит только из букв, упоминаемых в определяющих соотношениях.

Рассмотрим еще один вывод следствия из определяющих соотношений. Пусть $\boldsymbol{aba} = \mathbf{1}$, $\boldsymbol{bab} = \mathbf{1}$. Докажем, что $\boldsymbol{a} = \boldsymbol{b}$. Действительно, $\boldsymbol{a} \equiv \boldsymbol{abaa^{-1}b^{-1}} \equiv \boldsymbol{a^{-1}b^{-1}} \equiv \boldsymbol{a^{-1}b^{-1}bab} \equiv \boldsymbol{b}$. Таким образом, при данных определяющих соотношениях существует только три неэквивалентных слова: $\mathbf{1}$, $\boldsymbol{a}$, $\boldsymbol{a^2}$.

В дальнейшем, если в слове встречается повторяющаяся буква, будем использовать в записи степень: например, писать $\boldsymbol{a^3}$ вместо $\boldsymbol{aaa}$. Для отрицательных степеней по определению будем считать $\boldsymbol{x^k} = (\boldsymbol{x^{-1}})^{\boldsymbol{-k}}$. К примеру, $\boldsymbol{a^{-3}} = (\boldsymbol{a^{-1}})^{\mathbf{3}}$.

◆ **В1.** *Пусть имеется одно определяющее соотношение* $\boldsymbol{ba} = \boldsymbol{ab}$. *Докажите, что тогда любое слово можно привести к виду* $\boldsymbol{a^m b^n}$, *где* $\boldsymbol{m}$, $\boldsymbol{n}$ *– некоторые целые числа.*

◆ **В2.** *Сколько существует различных слов, если заданы определяющие соотношения:* $\boldsymbol{a^4} = \boldsymbol{b^3} = (\boldsymbol{ab})^{\mathbf{2}} = \mathbf{1}$?

◆ **В3.** *Сколько существует различных слов, если заданы определяющие соотношения:* $\boldsymbol{aba^{-2}ba} = \boldsymbol{b^3} = \mathbf{1}$?

◆ **В4.** *Сколько существует различных слов, если заданы определяющие соотношения:* $\boldsymbol{a^2} = \boldsymbol{b^2} = (\boldsymbol{ab})^{\boldsymbol{n}} = \mathbf{1}$?

Существует способ, позволяющий графически изобразить вывод следствия из определяющих соотношений. Представление о таком способе можно получить из рисунков, где показан вывод следствия $\boldsymbol{a^2b^3} = \boldsymbol{b^3a^2}$ из соотношения $\boldsymbol{ab} = \boldsymbol{ba}$ и вывод следствия $\boldsymbol{b^6} = \mathbf{1}$ из соотношений $\boldsymbol{b^2} = \boldsymbol{a}$ и $\boldsymbol{a^3} = \mathbf{1}$: при обходе любой области карты читается одно из заданных определяющих слов, а при обходе границы всей карты читается следствие, если при движении против стрелки буква считается как обратная.

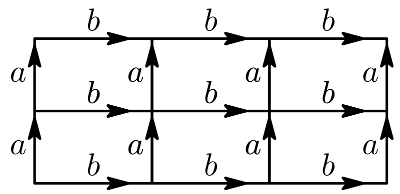


рис. 1

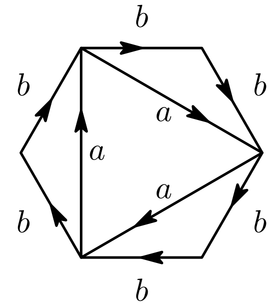


рис. 2

Итак, плоскость разбивается на несколько многоугольников. Границы многоугольников будем называть *ребрами*. Область карты, лежащую внутри какого-либо многоугольника, будем называть *клеткой*. Рядом с каждым ребром на карте пишем букву алфавита так, что слова, отвечающие обходу любой клетки отвечают какому-либо определяющему соотношению.

♦ **В5.** *Нарисуйте карту, иллюстрирующую вывод следствия* $\boldsymbol{a^2b^2c^2 = 1}$ *из соотношений* $\boldsymbol{a^3 = 1}$, $\boldsymbol{b^3 = 1}$, $\boldsymbol{c^3 = 1}$, $\boldsymbol{cba = 1}$.

♦ **В6.** *Нарисуйте карту, иллюстрирующую вывод следствия* $\boldsymbol{ab^{-1}aba^{-1}b = 1}$ *из соотношений* $\boldsymbol{a^3 = 1}$, $\boldsymbol{b^3 = 1}$, $\boldsymbol{abab = 1}$.

На самом деле, обсуждаемые нами структуры слов с соотношениями имеют свое название. Пусть имеется множество G. Пусть на этом множестве определен некоторый закон, по которому каждой упорядоченной паре элементов (x, y) сопоставлен некоторый элемент z. Будем обозначать это соответствие с помощью значка $*$: $z = x*y$. При этом говорят, что на множестве определена операция $*$. Заметим, что в общем случае результат операции зависит от порядка двух элементов: $x * y$ и $y * x$ различны.

Обычно рассматриваются такие операции, что $(x * y) * z = x * (y * z)$ для любых элементов x, y, z. В этом случае говорят, что выполнена *ассоциативность* (*или операция* $*$ *ассоциативна*).

На множестве слов в конечном алфавите можно ввести операцию произведения: произведением двух слов $\boldsymbol{a_1a_2\dots a_k}$ и $\boldsymbol{b_1b_2\dots b_n}$ считается слово $\boldsymbol{a_1a_2\dots a_kb_1b_2\dots b_n}$ — результат приписывания второго слова к первому. Очевидно, что эта операция ассоциативна. Смысл ассоциативности состоит в том, что результат произведения $\boldsymbol{x_1 * x_2 * \dots * x_n}$ не зависит от расстановки скобок, то есть произведения $\boldsymbol{(x_1 * x_2) * (x_3 * x_4)}$ и $\boldsymbol{x_1 * ((x_2 * x_3) * x_4)}$ дают один и тот же результат.

Определение. Множество $\boldsymbol{G}$ с определенной на нем операцией $*$ называется *группой*, если выполнены три условия:

(i) Операция $*$ ассоциативна, то есть для любых $\boldsymbol{x, y, z} \in \boldsymbol{G}$ выполнено равенство $\boldsymbol{(x * y) * z = x * (y * z)}$;

(ii) В $\boldsymbol{G}$ существует элемент, обозначаемый $\boldsymbol{1}$, такой, что $\boldsymbol{x * 1 = 1 * x = x}$ для любого элемента $\boldsymbol{x} \in \boldsymbol{G}$; элемент $\boldsymbol{1}$ называется единицей;

(iii) Для любого элемента $\boldsymbol{x} \in \boldsymbol{G}$ существует обратный элемент $\boldsymbol{x^{-1}} \in \boldsymbol{G}$, такой что $\boldsymbol{x * x^{-1} = x^{-1} * x = 1}$.

Для удобства записи, если понятно о какой операции идет речь, знак $*$ опускают, и пишут, например, $ab = c$ или $xy = yx$. Мы будем этим пользоваться.

Слова в алфавите, которые мы обсуждали, образуют группу: единицей в ней является пустое слово (которое мы обозначаем значком 1), а обратное слово получается из данного заменой всех букв на обратные и выписыванием слова в обратном порядке. Например, обратным для $abcdxyz$ будет слово $z^{-1}y^{-1}x^{-1}d^{-1}c^{-1}b^{-1}a^{-1}$, а для $ab^{-1}cab^{-1}a^{1}$ будет слово $aba^{-1}c^{-1}ba^{-1}$. Если приписать обратные слова друг к другу, получившиеся произведение можно преобразовать до пустого слова (единицы).

После введения определяющих соотношений некоторые слова становятся эквивалентными: можно считать, что они представляют один элемент группы. При выполнении операций с каким-то элементом группы можно выбирать любое из представляющих этот элемент эквивалентных слов.

Группа G называется *конечной*, если конечно множество ее элементов. Число элементов конечной группы называют ее *порядком* и обозначают $|G|$.

Группа называется *абелевой* (или *коммутативной*) если $xy = yx$ выполнено для всех $x, y \in G$.

♦ **B7.** *Докажите, что если для любого элемента x выполнено $x^2 = 1$, то группа абелева.*

Определение. *Подгруппой* группы G называются непустое подмножество $H \subset G$ такое, что:

(i) если $a, b \in H$, то $ab \in H$;

(ii) если $a \in H$, то $a^{-1} \in H$.

Из определения сразу вытекает, что единица группы G содержится в H (возьмем $a \in H$, тогда $a^{-1} \in H$, и $aa^{-1} \in H$). Следовательно, H является группой относительно операции, определенной в G.

Заметим, что в любой группе есть две тривиальные подгруппы: вся группа и группа из одной единицы. Эти подгруппы называются *несобственными*. Остальные называются *собственными*. Как находить собственные подгруппы? Возьмем некоторый элемент $a \in G$ и будем возводить его в разные степени. Подмножество $\{a^k\}$, где k — целое число является абелевой (коммутативной) подгруппой в G. Такие подгруппы называются *циклическими* подгруппами группы G, а элемент a – *порождающим* элементом. Подгруппа, *порожденная* элементом a обозначается $\langle a \rangle$. Если $\langle a \rangle = G$, то есть a порождает всю группу, G называется *циклической группой*.

Если $a^k = 1$ при каком-либо натуральном k, то существует такое минимальное n, что $a^n = 1$. Такое n называется *порядком* элемента a. Если же никакая степень элемента a не равна единице, то говорят, что элемент имеет бесконечный порядок.

Циклическая группа порождается одним своим элементом. Что получится, если в качестве порождающего множества использовать несколько элементов?

Пусть S — подмножество в G. Обозначим как $\langle S \rangle$ подмножество в G, состоящее из всевозможных конечных произведений

$$g_1{}^{\alpha_1} g_2{}^{\alpha_2} \cdots g_k{}^{\alpha_k},$$

где $g_i \in S$, $\alpha_i = \pm 1$, $i = 1, \dots, k$. Ясно, что $\langle S \rangle$ является подгруппой в G, причем обратным элементом к выписанному произведению будет

$$g_k{}^{\beta_k} g_{k-1}{}^{\beta_{k-1}} \dots g_1{}^{\beta_1},$$

где $\beta_i = -\alpha_i$. Подгруппа $\langle S \rangle$ называется подгруппой, порожденной в G подмножеством S, а сами элементы S – порождающими элементами для этой подгруппы.

До сих пор мы рассматривали слова как конечные последовательности букв. Однако, когда мы изображаем какое-нибудь определяющее отношение $a_1 a_2 \dots a_n = 1$ графически в виде цикла стрелок, фактически теряется информация, какая буква является первой в слове, так как для циклических сдвигов нашего слова ($a_2 a_3 \dots a_n a_1$, $a_3 a_4 \dots a_n a_1 a_2$ и так далее) цикл стрелок будет точно такой же. Таким образом, имеет смысл вместо обычного слова рассматривать *циклическое слово*: совокупность всех его циклических сдвигов. Подслово (нециклическое!) циклического слова X это подслово одного из циклических сдвигов обычного слова X. Например, в список подслов длины 3 циклического слова a^2ba входят a^2b, aba, ba^2, a^3. Будем называть слово *циклически несократимым*, если все его циклические сдвиги несократимы в обычном смысле.

В группах циклические сдвиги имеют свой аналог. Два элемента a и b группы G будем называть *сопряженными*, если существует такой элемент $x \in G$, что $a = xbx^{-1}$. Ясно, что все циклические сдвиги одного слова попарно сопряжены.

♦ **B8.** *Докажите, что в группе, заданной некоторыми определяющими соотношениями, каждое слово сопряжено с циклически несократимым.*

♦ **B9.** *Пусть группа задана определяющими соотношениями $U_1 = 1$, ..., $U_k = 1$. Докажите, что если $W \equiv 1$ (слово W приводится к пустому), то существуют такие слова X_1, ..., X_n, что слово*

$$X_1 U_{i_1}^{\pm 1} X_1{}^{-1} X_2 U_{i_2}^{\pm 1} X_2^{-1} \dots X_n U_{i_n}^{\pm 1} X_n^{-1},$$

где $1 \leqslant i_j \leqslant k$, приводится к W только сокращениями рядом стоящих взаимно-обратных элементов.

Если в некоторой группе выполнены соотношения $a^3 = 1$, $bab^{-1} = c$, то очевидно, выполняется и соотношение $c^3 = 1$. Этот вывод можно отразить с помощью рисунка 3. Обход внутренней треугольной клетки против часовой стрелки дает слово a^3, обход любой из четырехугольных клеток — $cba^{-1}b^{-1}$, а проводя обход границы всего рисунка получаем слово c^3, левую часть следствия соотношений $a^3 = 1$ и $cba^{-1}b^{-1} = 1$.

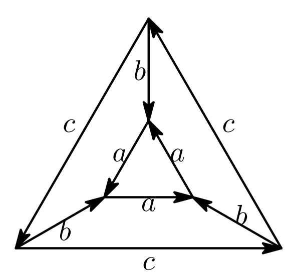


рис. 3

Опишем, как строятся подобные примеры. Для этого подробнее проведем вывод следствия $a^3b^{-1}a^2b^3 = 1$ из соотношений $a^3 = 1$ и $b^2 = a$ (то есть $b^2a^{-1} = 1$). Сначала запишем слово $a^3b^{-1}a^2b^3$ в виде $(a^3)(b^{-1}a^3b)(b^{-1}a^{-1}b^2b)$. Тем самым мы записываем наше слово в виде произведения определяющих слов и их сопряженных. Изобразим сомножители в виде трех последовательных лепестков (см рис. 4.)

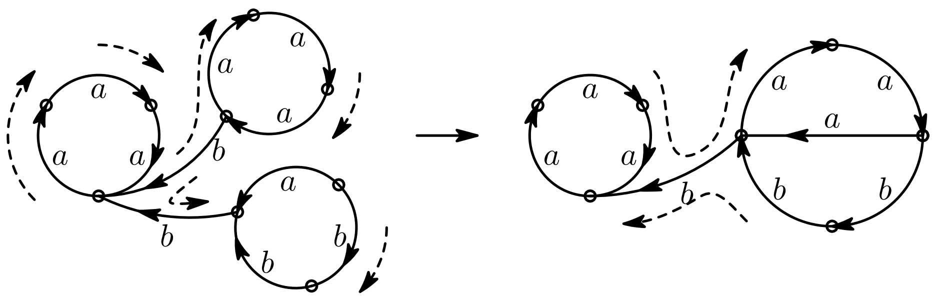


рис. 4

Каждый из них нарисован в виде круга с ножкой (возможно, пустой). Окружности размечаются так, чтобы прочитать соответствующее определяющее слово (в данном примере, a^3 или $a^{-1}b^2$), а на ножке написано сопрягающее слово (в данном примере, b^{-1} или пустое слово). Тогда, обходя последовательно лепестки, можно прочитать правую часть равенства

$$a^3b^{-1}a^2b^3 = (a^3)(b^{-1}a^3b)(b^{-1}a^{-1}b^2b).$$

Чтобы получить слово, графически совпадающее с левой частью этого равества, нужно дополнительно провести сокращения. Эти сокращения в слове, записанном на контуре рисунка, можно осуществить путем последовательных склеиваний соседних ребер контура с одинаковыми метками (и согласованно направленными стрелками). В данном примере, сначала склеиваются ножки, потом по одному ребру второй и третьей окружностей. В результате получается диаграмма, на контуре которой написано в точности слово $a^3b^{-1}a^2b^3$.

Таким же способом можно построить диаграмму вывода любого следствия определяющих соотношений. Причем, даже для одного следствия $W = 1$ диаграммы могут быть клеточно неэквивалентны.

♦ **C1. Лемма ван Кампена.** *Пусть W – произвольное непустое слово в алфавите $\bar{L} = L \cup L^{-1} \cup 1$. Слово $W = 1$ в группе с определяющими соотношениями U тогда и только тогда, когда существует диаграмма над U, метка контура которой графически равна W.*

Пусть группа задана определяющими соотношениями $U = \{U_1 = 1, \ldots, U_k = 1\}.$ Условимся, что каждое слово U_i выбирается не только несократимым, но и циклически несократимым. Пусть, кроме того, вместе с каждым словом R в систему определяющих соотношений входит и графически обратное слово R^{-1} и если XY — некоторое определяющее слово, то YX — тоже некоторое определяющее слово. Систему отношений, в которой выполняются все эти условия, назовем *симметризованной.* Ясно, что добавление обратных слов и циклических сдвигов не меняет множества всех следствий, а значит и группы G.

Если XY_1 и XY_2 — различные слова из U с общим началом X, то X называется *куском* относительно U. Говорят, что группа G, заданная определяющими соотношениями U является *группой малых сокращений*, если любой кусок X имеет длину, меньшую чем $^1/_6$ длины любого из слов в которое он входит. Это условие означает, что в произведении любых двух определяющих слов U_iU_j сокращается малая часть. Этим оправдывается название «условие малых сокращений».

Роль условия малых сокращений состоит в том, что в следствиях остается много «следов» определяющих соотношений.

♦ **C2.** *Пусть* $\boldsymbol{G}$ *— группа с условием малых сокращений. Рассмотрим диаграмму над* $\boldsymbol{G}$*. Пусть метка* $\boldsymbol{\phi(q)}$ *контура циклически несократима и в циклическом слове* $\boldsymbol{\phi(q)}$ *нет собственных подслов, равных в* $\boldsymbol{G}$ *единице. Тогда существует клетка* $\boldsymbol{P}$*, внешняя дуга которой имеет длину более половины периметра клетки.*

♦ **C3.** *Пусть* $\boldsymbol{G}$ *— группа с условием малых сокращений, заданная набором определяющих соотношений. Докажите, что существует алгоритм, позволяющий по заданным двум элементам определить, равны они или нет, то есть эквивалентны ли слова, представляющие эти элементы.*

Добавления

♦ **A5.** *Можно ли разрезать плоскость на выпуклые семиугольники с диаметром не более 1 так, чтобы любой единичный круг пересекал не более миллиона из них?*

♦ **B7.5a.** *Рассмотрим группу с алфавитом* $\boldsymbol{\{a, b\}}$*. Докажите, что если для любого элемента* $\boldsymbol{x}$ *из группы выполнено* $\boldsymbol{x^3 = 1}$*, то группа конечна.*

♦ **B7.5b.** *Рассмотрим группу с алфавитом* $\boldsymbol{\{a, b, c\}}$*. Докажите, что если для любого элемента* $\boldsymbol{x}$ *из группы выполнено* $\boldsymbol{x^3 = 1}$*, то группа конечна.*

Дополнительные соображения

Введем формальные определения. Произвольное разбиение плоскости на многоугольники-клетки $\boldsymbol{U}$ в дальнейшем будем называть *картой*. Ориентированные стороны разбиения $\boldsymbol{U}$ называем *ребрами карты*. Таким образом, вместе с каждым ребром $\boldsymbol{e}$ в $\boldsymbol{U}$ появляется и ребро $\boldsymbol{e^{-1}}$ с противоположной ориентацией (состоящее из тех же точек поверхности, что и сторона в $\boldsymbol{U}$). Примем соглашение, по которому контуры всех клеток обходятся по часовой стрелке. Если компонента $\boldsymbol{Y}$ края состоит из $\boldsymbol{n}$ сторон, то, в соответствии с определенной на $\boldsymbol{Y}$ ориентацией, эти стороны можно задать ребрами $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\ldots$ $\boldsymbol{e_n}$, такими что $\boldsymbol{e_1} \ldots \boldsymbol{e_n}$ – петля, которую назовем *контуром* карты $\boldsymbol{U}$. Аналогично можно определить контур клетки. Контур карты мы будем часто рассматривать с точностью до циклического сдвига. Если ребро $\boldsymbol{e}$ входит в контур $\boldsymbol{e_1} \ldots \boldsymbol{e_n}$ (клетки или карты), то будем говорить, что $\boldsymbol{e}$ *принадлежит контуру* (соответственно, клетки или карты). Цепочку ребер $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\ldots$, $\boldsymbol{e_n}$, такую, что конец $\boldsymbol{e_i}$ совпадает с началом $\boldsymbol{e_{i+1}}$ для всех $i = 1, \ldots, n-1$, будем называть *путем*.

Понятие подпути аналогично понятию подслова: путь $\boldsymbol{p}$ – *подпуть* $\boldsymbol{q}$, если $\boldsymbol{q = p_1 p_2}$ для некоторых путей $\boldsymbol{p_1}$, $\boldsymbol{p_2}$. Пусть есть некоторый конечный алфавит $\boldsymbol{L}$. Обозначим $\boldsymbol{\bar{L} = L \cup L^{-1} \cup 1}$, где $\boldsymbol{L^{-1}}$ – алфавит обратных букв.

Пусть далее каждому ребру $\boldsymbol{e}$ карты $\boldsymbol{U}$ сопоставляется некоторая буква $\boldsymbol{\phi(e)}$ из $\boldsymbol{\bar{L}}$. Если при этом $\boldsymbol{\phi(e^{-1}) = \phi(e)^{-1}}$, то карту $\boldsymbol{U}$ назовем *диаграммой над* $\boldsymbol{U}$. Для пути

$p = e_1 \dots e_n$ в диаграмме U над L *меткой* $\phi(p)$ называется слово $\phi(e_1)\dots\phi(e_n)$ в алфавите $\bar{L}$. Если $n = |p| = 0$, то $\phi(p) = 1$ по определению. Метка контура клетки или диаграммы определена с точностью до циклического сдвига, то есть это циклическое слово.

Назовем клетку K R-*клеткой*, если метка $\phi(p)$ ее контура графически равна (с точностью до циклического сдвига) некоторому слову из числа определяющих соотношений или обратному к ним с точностью до вставки нескольких символов $\mathbf{1}$. (Ясно, что выбирая начало и направление обхода и игнорируя символ $\mathbf{1}$ всегда можно прочитать слово из определяющего соотношения.)

Клетку K назовем 0-*клеткой*, если метка W ее контура $e_1 \dots e_n$ графически равна $\phi(e_1)\dots\phi(e_n)$, где все $\phi(e_i) = 1$ (графическое равенство) либо для некоторых $i \neq j$ $\phi(e_i) = a$, $\phi(e_j) = a^{-1}$, где a — буква из алфавита, а для остальных $k \neq i,\ j$ $\phi(e_k) = 1$. Ребра с меткой $\mathbf{1}$ назовем $\mathbf{0}$-*ребрами*, а ребра, метка которых нетривиальное (не равное $\mathbf{1}$) слово из алфавита, назовем U-*ребрами. Длина* $|p|$ произвольного пути определяется как число его U-ребер. Периметр клетки или диаграммы – это длина ее контура.

В рассмотренных выше примерах не встречается 0-клеток. Но иногда их удобно вводить по следующей причине. Примеры на рисунках 1–3 являются полноценными дисковыми диаграммами: при удалении контура они не распадаются на две части. Изображение на рисунке 4 не является диском, при удалении контура распадается на две компоненты связности. Это приводит к некоторым техническим трудностям, например, при вырезании поддиаграмм для проведения индуктивных рассуждений. С помощью 0-клеток диаграмму на рисунке 4 можно сделать дисковой (рис. 5). В дальнейшем можно представлять себе 0-клетки в виде "очень тонких" клеток (или в виде "толстых ребер"), а 0-ребра как "очень короткие" ребра по сравнению с ребрами, представляющими букву алфавита.

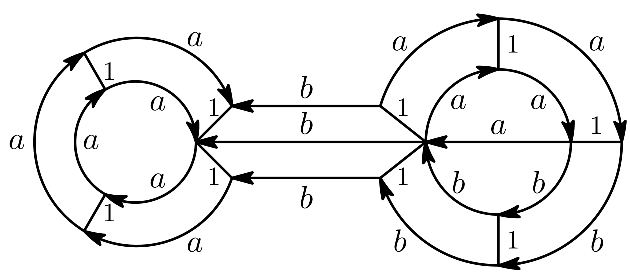


рис. 5

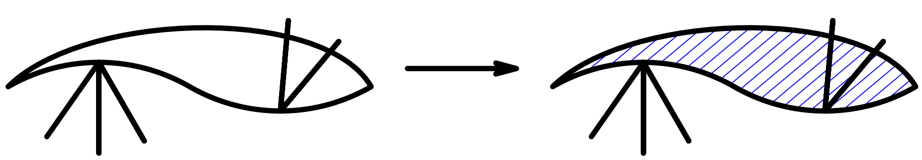


рис. 6

Итак, *диаграммой над группой* G заданной соотношениями $R_1 \dots R_n$ назовем всякую диаграмму над нашим алфавитом U, каждая клетка которой является R-клеткой или 0-клеткой.

Иногда бывает полезно проводить 0-*измельчение* диаграммы. Пусть клетка представляет собой многоугольник. Расположим внутри многоугольника меньший по размеру подобный многоугольник и соединим их соответствующие вершины. Внутренний многоугольник разметим буквами аналогично внешнему. Добавленные ребра отмечаем $\mathbf{1}$ как

0-ребра. Получаем 0-измельчение клетки. Аналогично можно провести *раздвоение* пути. Каждое ребро некоторого пути разделяем на два ребра, с аналогичной разметкой. Получается два пути, с совпадающими началом и концом.

Задачи цикла $\boldsymbol{D}$ являются подготовительными к основным задачам цикла $\boldsymbol{E}$.

♦ **D1.** *Дана бесконечная периодическая последовательность с наименьшим периодом $\boldsymbol{n}$ и два её одинаковых подслова длины $\boldsymbol{n-1}$.*

1. Докажите, что их начальные буквы находятся на расстояниях, кратных $\boldsymbol{n}$.

2. Верно ли аналогичное утверждение для двух одинаковых подслов длины $\boldsymbol{n-2}$?

♦ **D2.** *Рассматриваются слова над конечным алфавитом. Имеется конечный словарь нехороших слов. Известно, что имеется бесконечное слово без нехороших подслов. Докажите, что имеется бесконечное периодическое слово без нехороших подслов.*

♦ **D3.** *Докажите, что в алфавите из двух букв существуют слова сколь угодно большой длины, не содержащие трёх одинаковых подслов, идущих подряд (бескубные слова).* **Указание:** *рассмотреть слова $\boldsymbol{a}$, $\boldsymbol{ab}$, $\boldsymbol{abba}$, $\boldsymbol{abbabaab}$, $\boldsymbol{\dots}$ Чтобы получить следующее слово, в предыдущем производятся замены $\boldsymbol{a \to ab}$, $\boldsymbol{b \to ba}$.*

♦ **D4.** *Даны две различные периодические последовательности с минимальными периодами $\boldsymbol{n}$ и $\boldsymbol{m}$ соответственно. Докажите, что если они имеют общий кусок длины $\boldsymbol{m+n-1}$, то они имеют сколь угодно большие общие куски.*

♦ **D5.** *Укажите точную оценку в* D4, *рассмотрев случаи взаимно простых и не взаимно простых $\boldsymbol{m}$ и $\boldsymbol{n}$.*

♦ **D6.** *Треугольник разбит на выпуклые четырёхугольники. Докажите, что среди них найдётся четырёхугольник с углом не менее $\boldsymbol{120^\circ}$.*

♦ **D7.** *Существует ли многогранник, у которого каждая грань имеет не менее* **6** *сторон?*

♦ **D8.** *У многогранника не менее $\boldsymbol{7n}$ граней. Докажите, что у него найдётся $\boldsymbol{n}$ граней с одинаковым числом сторон.*

♦ **D9.** *Плоскость разбита на выпуклые $\boldsymbol{k}$-угольники диаметра, не превосходящего 1. Зафиксирована точка $\boldsymbol{O}$. Пусть $\boldsymbol{S_k(R)}$ – количество $\boldsymbol{k}$-угольников, попавших в круг радиуса $\boldsymbol{R}$* с центром в точке $\boldsymbol{O}$. *Докажите, что существует $\boldsymbol{R_0}$ такое, что для всех $\boldsymbol{R > R_0}$ верно $\boldsymbol{S_k(R) \geqslant \lambda^R}$, где вместо $\boldsymbol{\lambda}$ можно подставить $\boldsymbol{k/10}$. Постарайтесь получить лучшую оценку для $\boldsymbol{\lambda}$.*

♦ **D10.** *Полуплоскость разбита на выпуклые $\boldsymbol{k}$-угольники диаметра, не превосходящего 1. Среди граничных $\boldsymbol{k}$-угольников выделено $\boldsymbol{L}$ соседних. Первым слоем назовем такие $\boldsymbol{k}$-угольники, которые граничат с выделенными. Граничащие с первым слоем $\boldsymbol{k}$-угольники назовем вторым слоем. Докажите, что во втором слое k-угольников не менее $\boldsymbol{L(k/10)^2}$.*

♦ **Проблема Бернсайда.** *Пусть задано натуральное число $\boldsymbol{n}$. Существует ли конечнопорожденная (то есть в конечном алфавите) бесконечная группа, в которой любой элемент $\boldsymbol{x}$ удовлетворяет соотношению $\boldsymbol{x^n = 1}$.*

А. Ю. Ольшанский построил такую группу для нечетного $\boldsymbol{n}$, большего, чем $\boldsymbol{10^{10}}$. Мы будем следовать основным идеям этого построения.

Итак, наша задача разбивается на две части. Первая часть посвящена самому процессу введения определяющих соотношений. Этот процесс состоит из счетного числа шагов. На каждом шаге вы вводим несколько определяющих соотношений вида $\boldsymbol{A^n = 1}$, при этом следим за тем, чтобы любое слово на каком-либо этапе процесса стало периодичным (то есть для любого слова $\boldsymbol{X}$ соотношение $\boldsymbol{X^n = 1}$ было выводимо из определяющих соотношений, введенных за конечное число шагов). Два слова в полученной группе считаются *равными*, если одно может быть приведено к другому с использованием конечного числа определяющих соотношений, введенных на каких-либо шагах. В результате процесса получается группа, каждый элемент которой будет периодичным.

Вторая часть касается **доказательства бесконечности** группы. Для этого необходимо доказать, что после каждого очередного шага найдется слово, не равное $\mathbf{1}$. Для этого нужно доказать, что не может существовать карты, все клетки которой являются периодическими словами с достаточно большим периодом, а по периметру написано слово, не содержащее периодических подслов с большим периодом (например, бескубное или бесквадратное слово). Для доказательства этого факта проводится исследование различных случаев примыкания клеток друг к другу. Клетки отвечают введенным определяющим соотношениям на разных этапах, на этапе с большим номером слово длиннее. Таким образом, клетки могут быть разных размеров.

Сначала доказываем, что "большие" клетки не могут сильно "примыкать" друг к другу. Затем рассматривается случай, когда к большой клетке примыкает несколько "слоев" маленьких клеток. Здесь используется экспоненциальный рост количества многоугольников в задачах D9 и A4.

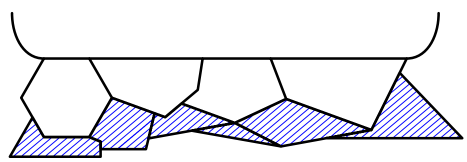


рис. 7

Таким образом, согласно задаче D9, можно за счет выбора $\boldsymbol{k}$ добиться того, чтобы уже во втором слое было очень много клеток. Выбор $\boldsymbol{k}$ (количества сторон у многоугольников) мы можем производить за счет выбора $\boldsymbol{n}$ – основной степени периодичности группы. То есть, на первом этапе длина введенных соотношений равна $\boldsymbol{n}$, на $\boldsymbol{i}$-ом этапе – $\boldsymbol{ni}$. Таким образом, многоугольники будут содержать достаточно много углов, и при экспоненциальном росте уже во втором слое будет много клеток.

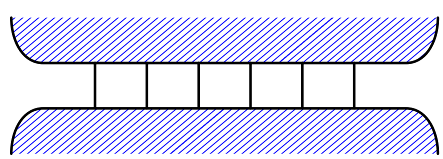


рис. 8

В оставшемся случае, между двумя большими клетками есть один слой малых клеток. Введем полезное понятие.

Клетки $\boldsymbol{A}$ и $\boldsymbol{B}$ называются *сократимыми*, если:

1. $\boldsymbol{A}$ и $\boldsymbol{B}$ имеют друг с другом общую границу, (либо соединяются через цепочку 0-клеток) ;

2. $\boldsymbol{A}$ и $\boldsymbol{B}$ имеют одинаковые метки.

В случае, если в диаграмме встречаются сократимые клетки, можно проделать следующую операцию: вырезать из диаграммы диск – объединение этих двух клеток – и вставить вместо него несколько $\mathbf{0}$-клеток. Тем самым в диаграмме уменьшено число нетривиальных клеток. Диаграмма, не содержащая сократимых клеток, называется *приведённой.*

♦ **E1. Пример узкой полосы, если показатель — четный.** *Докажите, что существует приведённая диаграмма* $\boldsymbol{D}$*, все клетки которой суть соотношения* $\boldsymbol{X^k = 1}$*, имеющая следующую структуру:*

1. $\boldsymbol{D}$ содержит две клетки $\boldsymbol{A}$ и $\boldsymbol{B}$, такие, что все остальные клетки соседствуют с ними;

2. Периметры клеток $\boldsymbol{A}$ и $\boldsymbol{B}$ в любое наперед заданное число раз больше периметров остальных клеток.

♦ **E2.** *Пусть в приведённой диаграмме имеются клетки двух видов: «большие» из* $\boldsymbol{m}$ *букв и «малые» из* $\boldsymbol{n}$*. Метка каждой клетки – периодическое слово вида* $\boldsymbol{A^n}$*. Какой наибольший общий участок могут иметь две большие клетки?*

♦ **E3.** *Пусть все клетки приведенной диаграммы имеют периметр либо* $\boldsymbol{m}$*, либо* $\boldsymbol{n}$*, причем* $\boldsymbol{n \gg m}$ *и метки всех клеток периодичны. Рассмотрим клетку* $\boldsymbol{A}$ *с периметром* $\boldsymbol{n}$*. Пусть к ней примыкают только клетки с периметром* $\boldsymbol{m}$*. Назовем* первым слоем *непосредственно примыкающие к* $\boldsymbol{A}$ *клетки. Назовем* $\boldsymbol{k}$*-тым слоем* $(\boldsymbol{k > 1})$ *клетки, примыкающие к* $\boldsymbol{k-1}$ *слою. Пусть* $\boldsymbol{A_k}$ *количество клеток в* $\boldsymbol{k}$*-том слое. Докажите, что* $\boldsymbol{A_k \geqslant (m/100)^k}$

♦ **E4.** *Пусть все клетки приведенной диаграммы имеют периметр либо* $\boldsymbol{m}$ *либо* $\boldsymbol{n}$*, причем* $\boldsymbol{n \gg m}$ *и метки всех клеток периодичны с нечетным периодом. Пусть слой из малых клеток зажат между двумя большими клетками (см рис. 8). Верно ли, что метка контура диаграммы периодична?*

Группы и мозаики

Решения

♦ **A1.** Эта задача предлегалась на осеннем туре турнира городов в 1993 году.

Подсчитаем двумя способами средний угол. С одной стороны, средний угол семиугольника равен $\frac{5}{7}\pi$

С другой стороны, если несколько углов сходятся во внутренней вершине, то их по крайней мере 3, и их средний угол не превосходит $\frac{2}{3}\pi$. Если же несколько углов примыкают к стороне, то ихкрайней мере 2 и соответствующий среднийугол не превосходит $\frac{\pi}{2}$

Оценим средний угол, приходящийся на вершину 1993-угольника. Сумма таких углов равна 1991π, а если никакие 4 подряд идущих вершины не принадлежат одному семиугольнику, то количество примыкающих углов не менее $1993\frac{3}{2}$, потому срднийугол оказываетсяне более $\pi\frac{1991}{1993}\frac{3}{2} < \frac{3}{2}\pi$. В итоге получается, что средний угол у семиугольников меньше $\frac{2}{3}\pi$, чего не может быть.

♦ **A2.** а) Ответ: да, соответствующая конструкция легко строится.

б) Ответ: нет. Проходят те же рассуждения со средним углом, что и в задаче $\boldsymbol{A1}$. Пусть D – диаметр семиугольника.

Рассмотрим круг с центром O и радиусом R. Тогда количество семиугольников, пересекающихся с границей этого круга, не превосходит $4\pi DR$, а количество семиугольников, попавших внутрь не менее чем $\pi D(D-R)^2/S$, где S – площадь семиугольника.

Средний угол в семиугольнике, с одной стороны равен $\frac{5\pi}{7}$, с другой стороны, средний угол в круге не превосходит $\frac{2}{3}\pi$.

Легко видеть, что граничными эффектами можно пренебречь. Подробно об этом методе см. в книге А. Я. Канель-Белов, А. К. Ковальджи "Как решают нестандартные задачи".

♦ **A3.** Ответ: нельзя. Проходит решение $\boldsymbol{A2}$, поскольку количество семиугольников, пересечённых кругом радиуса R, растёт в этом случае не быстрее, чем CR, где C – некоторая константа.

♦ **A4.** Будем раасуждать, как и в предыдущих пунктах. Пусть N – количество углов, попавших внутрь круга радиуса R. Их среднее арифметическое не превосходит $\frac{2}{3}\pi$. С другой стороны, рассмотрим семиугольники с этими углами. Среднее арифметическое их углов равно $\frac{5}{7}\pi = \frac{2}{3}\pi + \frac{1}{14}\pi$.

Легко видеть, что нужно добавить дополнительно более $N\frac{3}{14}$ дополнительных углов (если они все равны π, то ровно столько).

Таким образом, количество углов, попавших в полосу $R \leq r \leq R+1$ не менее $N\frac{3}{14}$.

Пусть $N(R)$ – количествоуглов в круге радиуса R. Тогда мы имеем, что $N(R+1) > N(R)(1+\frac{3}{14})$

Таким образом, количество углов, попавших в круг радиуса R, а вместе с ним и число семиугольников, растёт экспоненциально.

♦ **B1.** есть легкое упражнение.

♦ **B2.** Ответ: 24.

Докажем, что их не более $\mathbf{24}$. Будем рассматривать только положительные степени букв, потому что мы моеж сделать их такими, прибавляя к степени $\boldsymbol{a}$ по $\mathbf{4}$, к степени $\boldsymbol{b}$ по $\mathbf{3}$. Рассмотрим следующие преобразования:

1. $\boldsymbol{a^4 \to 1}$
2. $\boldsymbol{b^3 \to 1}$
3. $\boldsymbol{bab \to aaa}$, так как $\boldsymbol{abab = aaaa = 1}$
4. $\boldsymbol{aba \to bb}$, так как $\boldsymbol{abab = bbb = 1}$
5. $\boldsymbol{aaabb \to ba}$, так как $\boldsymbol{aaabb = aaaaba = ba}$
6. $\boldsymbol{aabbaa \to baab}$, так как $\boldsymbol{aabbaa = aaabaaa = babbbab = baab}$

Для каждого слова рассмотрим равное ему с наименьшим количеством букв $\boldsymbol{a}$, среди них с наименьшим количеством букв $\boldsymbol{b}$. Заметим, что каждая операция либо уменьшает первый параметр, либо уменьшает второй, не изменяя первый. Теперь посчитаем все возможные слова

1. $\boldsymbol{e}$
2. (слова с буквы $\boldsymbol{a}$): $\boldsymbol{a}$ – одно слово
3. (сочетания $\boldsymbol{ab}$.) $\boldsymbol{ab \to^4 abb \to^2 abba \to^3 abbaa \to^7 abbaab \to^4 abbaabb \to^2 abbaabba}$. По $\mathbf{(3)}$ слеующая буква не $\boldsymbol{b}$, по $\mathbf{(6)}$ – не $\boldsymbol{a}$. Каждый шаг однозначно определяет следующую букву, поэтому всего не более 7 слов, сочетаний $\boldsymbol{ab}$.
4. (сочетание $\boldsymbol{aab}$.) $\boldsymbol{aab \to^4 aabb \to^3 aabba}$. Не $\boldsymbol{a}$ по $\mathbf{(6)}$, не $\boldsymbol{b}$ по $\mathbf{(4)}$. Получаем 3 слова.
5. (сочетание $\boldsymbol{aaab}$.) Далее не $\boldsymbol{a}$ по $\mathbf{(3)}$, не $\boldsymbol{b}$ по $\mathbf{(5)}$. Всего одно слово.

   Слова с буквы $\boldsymbol{b}$:
6. $\boldsymbol{b}$ – одно слово
7. (Сочетание $\boldsymbol{ba}$.) $\boldsymbol{ba \to^3 baa}$; $\boldsymbol{baa \to baaa \to^1 baaab}$, не $\boldsymbol{a}$ по $\mathbf{(4)}$, не $\boldsymbol{b}$ по $\mathbf{(5)}$; $\boldsymbol{baa \to baab \to^4 baabb \to^7 baabba}$, не $\boldsymbol{a}$ по $\mathbf{(6)}$, не $\boldsymbol{b}$ по $\mathbf{(3)}$. 7 слов.
8. (Сочетание $\boldsymbol{bba}$.) $\boldsymbol{bba \to^3 bbaa \to^7 bbaab}$, не $\boldsymbol{a}$ по $\mathbf{(4)}$, не $\boldsymbol{b}$ по $\mathbf{(8)}$. Всего три слова.

Мы получили такие $\mathbf{24}$ слова, что любое другое равно одному из них.

Теперь покажем, что их не менее чем $\mathbf{24}$. Рассмотрим в группе $\boldsymbol{S_4}$ перестановки $\mathbf{(1,2,3,4)}$ и $\mathbf{(1,3,4)}$. Легко проверить, что этими двумя перестановками порождается вся группа, и соотношения для них выполнены. Это значит, что и сначала различных слов было не менее чем $\mathbf{24}$.

♦ **В3.** Ясно, что при указанных преобразованиях количество букв $\boldsymbol{a}$ в слове не меняется, поэтому слова с разным числом букв $\boldsymbol{a}$ оказываются разными. Таким образом, количество элементов в группе бесконечно (счётно).

♦ **В4.** Ответ: $\boldsymbol{2n}$.

Сначала покажем, что количество элементов не превосходит $2n$. Ясно, что нетсмысла рассматривать слова, содержащие подряд две буквы a или две буквы b. Таким образом, достаточно рассмотреть множество слов $(ab)^k, (ba)^k, a(ba)^k, b(ab)^k, k < n$

Далее легко заметить, что $(ab)^k(ba)^k = e$, отсюда $(ab)^k = (ba)^{n-k}$ и $b(ab)^k = a(ba)^{n-k-1}$. Таким образом, множества $\{(ab)^k\}_{k=1}^n$ и $\{(ba)^k\}_{k=1}^n$, а также $\{b(ab)^k\}_{k=1}^n$ и $\{a(ba)^k\}_{k=1}^n$, совпадают и достаточно рассмотреть $2n$ элементов $(ab)^k$ и $b(ab)^k$.

Остаётся проверить, что группа содержит по крайней мере $2n$ элементов. Для этого рассмотрим правильный n-угольник. Элементом a будет симметрия относительно прямой, проходящей через центр и вершину, b –симметрия относительно прямой, проходящей через центр и соседнюю вершину. Легко провеить, что полученная группа удовлетворяет всем соотношениям и содержит $2n$ элементов.

♦ **B9.** Рассмотрим элементарные преобразования четырех типов:

1 Вставка взаимнообратных букв aa^{-1};

2 Удаление взаимнообратных букв aa^{-1};

3 Вставка определяющего слова U_i;

4 Удаление определяющего слова U_i.

Пусть слово W выводится из пустого слова 1 с помощью нескольких операций указанного вида. Если вывод состоит из одной операции, то существование требуемого вида у W очевидно. Пусть требуемый вид существует у слова W, выводимого с помощью N операций:

$$W = X_1 U_{i_1}{}^{\pm 1} X_1^{-1} X_2 U_{i_2}{}^{\pm 1} X_2^{-1} \dots X_k U_{i_k}{}^{\pm 1} X_k^{-1}.$$

Рассмотрим четыре случая выполнения еще одной операции–перехода к слову W'. Предыдущее слово W представляется в виде произведения блоков вида $XU_i^{\pm 1}X^{-1}$ для различных определяющих слов U_i и сопрягающих слов X. Слово W получается из такого вида сокращением некоторых подряд идущих обратных букв. Покажем как получить аналогичный вид для слова W'. Последняя операция – это вставка или удаление некоторого определяющего слова U или двух обратных букв aa^{-1} в наше слово. Рассмотрим вставки. Вставка приходится либо между блоками $XU_i^{\pm 1}X^{-1}$ (очевидный случай), либо внутрь некоторого X (X^{-1}) либо внутрь некоторого U_i. Если это вставка взаимнообратных букв aa^{-1} в некоторый блок X, аналогичную вставку можно сделать в блок X^{-1}. Получим требуемый вида у W'. Если же aa^{-1} оказывается внутри некоторого U (то есть $U = Yaa^{-1}Z$), то этот блок меняем на блок $XYaa^{-1}Y^{-1}YZYaa^{-1}Y^{-1}X^{-1}$. То есть получается сопряжение определяющего слова YZ с помощью $XYaa^{-1}Y^{-1}$. Пусть теперь последняя операция это вставка определяющего слова U. Пусть вставка производится в какое-то слово $U_i{}^{\pm 1}$ (то есть $U_i{}^{\pm 1} = AB$ и $U_i{}^{\pm 1} \to AUB$). Тогда блок $XU_i{}^{\pm 1}X_j^{-1}$ можно заменить на два подряд идущих блока $XAUA^{-1}X^{-1}XABX^{-1}$ (мы вставили слово $A^{-1}X^{-1}XA$). Если же вставка производится в какое-то слово X (то есть $X = YZ$, $X \to YUZ$) то блок XUX^{-1} меняем на два подряд идущих блока $YUY^{-1}YZUZ^{-1}Y^{-1}$ (мы вставили слово YY^{-1}).

Подобным образом разбираются случаи удаления определяющего слова или подряд идущих обратных букв (упражнение).

♦ **C1.** **1**. Пусть D диаграмма с контуром p. Если клетка в диаграмме только одна, то все очевидно. Если клеток более одной, то D разрезается некоторым путем x на две диаграммы с меньшим числом клеток. Пусть p_1x и p_2x^{-1} контуры для этих двух диаграмм, и $p = p_1p_2$. По предположению индукции пути p_1x и p_2x^{-1} отвечают словам равным единице. Тогда $\boldsymbol{p} = \boldsymbol{p_1 p_2} = \boldsymbol{p_1 x x^{-1} p_2}$ тоже отвечает единичному слову. **2**. Пусть $\boldsymbol{W = 1}$ в $\boldsymbol{G}$. Используя $\boldsymbol{B9}$, представляем $\boldsymbol{W}$ в виде

$$\boldsymbol{X_1 U_{i_1}{}^{\pm 1} X_1^{-1} X_2 U_{i_2}{}^{\pm 1} X_2^{-1} \ldots X_k U_{i_k}{}^{\pm 1} X_k^{-1}}.$$

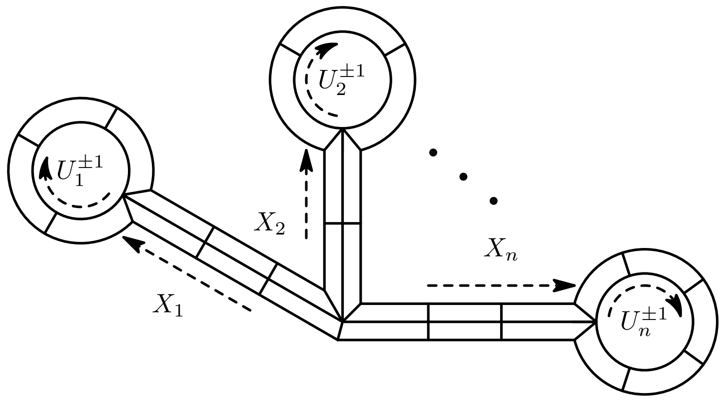


рис. 1

Построим на плоскости ломаную $\boldsymbol{p_1}$, разделим ее на отрезки и напишем на них буквы, так чтобы вдоль $\boldsymbol{p_1}$ было написано слово $\boldsymbol{X_1}$. Окружность $\boldsymbol{c_1}$, прикрепленную к концу ломаной разметим так, чтобы при ее обходе по часовой стрелке читалось определяющее слово $\boldsymbol{U_{i_1}}$. Чтобы получить именно дисковую диаграмму, пририсуем 0-клетки к $\boldsymbol{p_1}$, $\boldsymbol{c_1}$, $\boldsymbol{c_1}^{\boldsymbol{-1}}$ (см. рисунок 1). Получим диаграмму, по контуру которой написано слово $\boldsymbol{X_1 U_{i_1}{}^{\pm 1} X_1^{-1}}$. По последнему ребру приклеим к ней аналогичную диаграмму для блока $\boldsymbol{X_2 U_{i_2}{}^{\pm 1} X_2^{-1}}$. Аналогичные операции произведем со всеми блоками $\boldsymbol{X_j U_{i_j}{}^{\pm 1} X_j^{-1}}$. В итоге получим диаграмму, отвечающую слову

$$\boldsymbol{X_1 U_{i_1}{}^{\pm 1} X_1^{-1} X_2 U_{i_2}{}^{\pm 1} X_2^{-1} \ldots X_k U_{i_k}{}^{\pm 1} X_k^{-1}}.$$

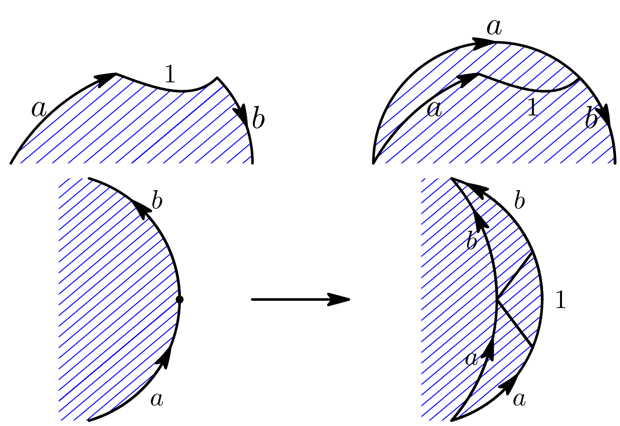


рис. 2

От данного слова можно перейти к слову $\boldsymbol{W}$ с помощью следующих преобразований: а) вычеркивание **1** в некотором месте; б) вставка **1**; в) сокращение взаимнообратных букв; г) вставка пары взаимно обратных букв. Каждое из этих преобразований может быть проведено с помощью приклеивания нескольких 0-клеток, как показано на рисунке 2. В итоге получаем требуемую диаграмму.

♦ **C2.** Построим вспомогательный граф $\boldsymbol{H}$. Вершины $\boldsymbol{H}$ выберем внутри значимых клеток (не 0-клетки) диаграммы плюс одну вершину расположим вне диаграммы. Вершины соединим ребром, если они примыкают друг к другу либо если могут быть соединены через цепочку 0-клеток. Вершина вне диаграммы соединяется с вершиной внутри какой-либо клетки, если эта клетка имеет участок на контуре, либо соединяется с контуром через цепочку 0-клеток. В полученном графе нет петель, так как все клетки отвечают циклически несократимым словам, и нет кратных ребер, так как любые две клетки мы соединяли не более чем одним ребром. Пусть $\boldsymbol{n}$ – число вершин в графе, $\boldsymbol{r}$ – число ребер. Из формулы Эйлера следует, что $\boldsymbol{r \leqslant 3n-6}$. Теперь, в предположении, что в диаграмме нет граничных клеток, у которых более половины периметра примыкает к контуру, докажем неравенство, противоречащее $\boldsymbol{r \leqslant 3n-6}$. Каждое ребро, соединяющее две вершины графа, находящиеся внутри клеток, разделим пополам, и будем считать, что к двум клеткам, соответствующим этому ребру относится по половине ребра. Если ребро соединяет вершину в клетке с вершиной вне диаграммы, то полностью отнесем это ребро к данной клетке. Таким образом, все ребра распределены.

Не граничащая с периметром клетка диаграммы соседствует не менее чем с семью другими клетками, так как каждый участок примыкания у нее менее $\boldsymbol{1/6}$ ее периметра. Значит, к такой клетке относится не менее 3,5 ребер. Пусть клетка граничит с периметром по одной дуге. Тогда эта дуга не более половины периметра, и соседей у такой клетки не менее $\boldsymbol{4}$. Значит, к ней относится не менее $\boldsymbol{4/2+1=3}$ ребер (внешнее ребро отходит к этой клетке полностью). Если у клетки два внешних ребра, то на ее периметре чередуются участки, являющиеся контуром диаграммы и не являющиеся таковыми. Тогда у нее, как минимум, двое соседей среди клеток и к ней относится не менее $\boldsymbol{2/2+2=3}$ ребер. Если у клетки более трех контурных участков, то к ней относится не менее 3 ребер. Итак, можно сделать вывод, что $\boldsymbol{r \geqslant 3(n-1)}$ (к вершине вне диаграммы мы ребер не относили, а к остальным отнесли, как минимум, по 3). Получаем противоречие с $\boldsymbol{r \leqslant 3n-6}$.

♦ **C3.** Слова $\boldsymbol{U}$ и $\boldsymbol{W}$ равны тогда и только тогда, когда $\boldsymbol{UW^{-1}=1}$. Опишем как определить, равно ли слово $\boldsymbol{1}$. Ясно, что можно считать слово циклически несократимым. По предположению индукции можно проверить равны ли $\boldsymbol{1}$ слова, являющиеся подсловами данного. Если такие слова есть, то можно уменьшить длину исходного слова. Пусть таких слов нет. Используя C2 заключаем, что Если в слове есть подслово $\boldsymbol{X}$, такое что $\boldsymbol{XY=1}$ – определяющее соотношение и $\boldsymbol{|X| > \frac{1}{2}|XY|}$, то $\boldsymbol{X}$ можно заменить на $\boldsymbol{Y^{-1}}$ с понижением длины. Если же такого подслова не найдется, то используя C2 заключаем, что для $\boldsymbol{W}$ нет диаграммы, значит $\boldsymbol{W \not\equiv 1}$.

♦ **D1.** Для решения задачи достаточно заметить, что количества букв любого сорта в периоде не зависит о того места, откуда период начинается. Поэтому, если взять отрезок периода длиной $\boldsymbol{n-1}$, то из соображений баланса однозначно определена следующая буква. Поэтому, если совпадают два отрезка длиной $\boldsymbol{n-1}$, то совпадают все последующие буквы. Поэтому сдвиг, переводящий одно вхождение слова $\boldsymbol{V}$ в сверхслово $\boldsymbol{W}$ в любое другое, кратен минимальному периоду. Поскольку $\boldsymbol{n}$ – минимальный период, это сдвиг на величину, кратную $\boldsymbol{n}$.

б) Пример: рассмотрим последовательность, период которой состоит из одной единицы и двух нулей. В ней можно найти два участка $\boldsymbol{n-2}$ нулей, находящихся на единичном расстоянии.

♦ **D2.** Пусть $n-1$ – максимальная длина плохого слова. Ясно, что некоторое подслово длины n в сверхслове W повторится более чем n раз. Тогда в W найдётся два не пересекающихся вхождения U, то есть подслово UVU. Тогда все подслова длины $\leq n$ периодической последовательности $UVUVUV\dots$ будут подсловами UVU, то есть подсловами W, то есть все они будут приличными.

♦ **D3.** Данная задача носит технический характер. Она давалась на основном туре Турнира Городов а также как один из технических пунктов на 4 летней конференции (задача "слова и хаос").

♦ **D4.** См. статью Самовола, Журавлева, Аннельбаума в журнале "Математическое просвещение" за 2006 г. Там эта задача разбирается. Решение задачи выводится с помощью очевидной индукции из следующей леммы.

Лемма. Дан участок длины $n+m+1$. Если символы, находящиеся на расстоянии m, равны, и символы, находящиеся на расстоянии n, равны, то равны и символы, находящиеся на расстоянии $m-n$.

♦ **D6.** Достаточно показать наличие тройной развилки. Иначе подсчитаем двумя способамисредний угол наших четырёхугольников. С одной стороны, он должен быть равен 90^o, с другой в четверной развилке он 90^o, в развилке большего порядка - меньше. Аналогичным образом Средние углы примыкающих к стороне 90^o, а средние углов, примыкающих к вершинам треугольника не больше 60^o. В итоге общее среднее оказывается сторого меньше 90^o, что и даёт нужное противоречие.

♦ **D7.** Делается аналогично. Если такой многогранник существует, то средний угол грани не менее 120^o. С другой стороны, в вершине сходятся не менее трёх граней и сумма соответствующих углов строго менее 360^o. Поэтому из оценки двумя способами среднего арифметического получаем нужное противоречие.

♦ **D8.** Идея решения: средний угол у грани строго меньше 120^o. Далее заметим, что средние углы у k-угольника при $k<6$ меньше 120^o, при $k=6$ ровно 120^o, а при $k>6$ строго больше 120^o, причём с ростом числа сторон вклад многоугольника в средний угол растёт. В качестве оптимального достаточно взять многоугольник с $n-1$ трёх-, четырёх-, пяти-, шести-, семи-, восьми-, девятиугольниками. У любого другого средний угол будет больше. Из рассмотрения этого многогранника получаем нужное противоречие.

♦ **D9.** Решение команды Омска (Кудык Никита, Матвеев Константин)

Пусть многоугольник с S сторонами разбит на n k-угольников. Внутри имеются вершины двух типов: те, которые вершины для всех содержащих их k-угольников (пусть их m), и те которые строго внутри стороны какого-нибудь k-угольника (пусть их l).

$$n(k-2)\pi = 2\pi m + \pi l + (s-2)\pi$$

(счёт углов)

$$n(k-2) = 2m + l + s - 2 \quad (1)$$

Кроме того,

$$nk \leq 3m + 2l + s \quad (2)$$

(количество точек, внутренних и на краю). Вычтем и получим:

$$2n \leq m + l + 2 > m + l \quad (3)$$

Из (1) и (3) выводим $\frac{n(k-2)-s+2}{2} < 2n$ (так как $m + l > m + \frac{l}{2}$, то есть $n(k-6) < s-2 < s$).

Пусть следующий слой состоит из t k-угольников, s' — количество строк у нового (рассмотренного многоугольника. $s' > (k-6)(n+t)$, так как $kt > s + s'$ (очевидно, так как kt — число сторон), то есть $kt > (k-6)(n+t) + (k-6)n \to t > \frac{k-6}{3}n$, то есть в каждом дополнительном слое мнооугольников в $\frac{k}{3}$ раз больше, чем было в s-угольнике, значит, внутри круга радиуса R, начиная с некоторого R_0, будет $> \left(\frac{k}{10}\right)^R$ многоугольников.

♦ **D10.** Отразим полуплоскотсь на другую половину, получим $2k$ отмеченных многоугольников. Используя $D9$, получаем, что число примыкающих к ним k-угольников (первый слой), больше $\frac{k-6}{3}2l$, во втором слое $\left(\frac{k-6}{3}\right)^2 2l$ и так далее, $\dots$, в i-ом слое более $\left(\frac{k-6}{3}\right)^i 2l$, учитывая, что для больших k $\frac{k-6}{3} > \frac{k}{10}$, получаем, что требуется.

Groups And Burnside Problem

A. Ya. Kanel-Belov   I. A. Ivanov-Pogodaev   A. S. Malistov

This project is focused on the combinatorial-geometrical method that has enabled to solve several complicated problems within the group theory like the construction of a infinite finitely generated group with identity $\boldsymbol{x^n = 1}$ to name one (The restricted Burnside problem). The goal of this project is the construction of this group. The following results were obtained by A. Yu. Olshanskii. The fact that group relations' consequences could be presented geometrically is the essence of the method.

We should introduce some useful ideas and definitions. Then we begin constructing diagrams on a plane. Usually the diagram is a polygon map. So it is interesting to handle the following preliminary problems.

◆ **A1.** *A convex* 1993*-gon was cut to convex* 7*-gons. Prove that exist four neighboring vertices of the* 1993*-gon which belong to same* 7*-gon.* (*A vertex of a* 7*-gon cannot belong to an edge of the* 1993*-gon.*)

◆ **A2.** *Can one cut a plane: a) to convex* 7*-gons? b) to equal convex* 7*-gons?*

◆ **A3.** *Can one cut a plane to convex* 7*-gons such that any unite circle intersects with less then million of them?*

◆ **A4.** *Let us consider a plane divided into* 7*-gons which diameters are less or equal to* 1*. Fix a point O. Let $N(R)$ be the number of* 7*-gons falling into the circle with diameter R and center O. Prove that there exists $\lambda > 1$ such that $N(R) > \lambda^R$.*

Let us consider an alphabet $\boldsymbol{L}$. Let $\boldsymbol{L}$ contain the following letters:

$$\boldsymbol{a,\ a^{-1},\ b,\ b^{-1},\ c,\ c^{-1},\ \ldots}$$

For any letter in our alphabet there exists a pair: $\boldsymbol{a}$ and $\boldsymbol{a^{-1}}$, $\boldsymbol{b}$ and $\boldsymbol{b^{-1}}$ and so on. These letters are called *inverse* letters. These letters can be used to construct words (some finite sequences of letters). Examples: $\boldsymbol{aba}$, $\boldsymbol{aba^{-1}ab^{-1}c}$, $\boldsymbol{a^{-1}a^{-1}bbca^{-1}}$. Words can be transformed by either cancelling or inserting two neighboring inverse letters. Thus $\boldsymbol{aba^{-1}ab^{-1}c}$ transforms into $\boldsymbol{ac}$: $\boldsymbol{ab\underline{a^{-1}a}b^{-1}c \equiv a\underline{bb^{-1}}c \equiv ac}$. Some words (for example, $\boldsymbol{b^{-1}aa^{-1}b}$) can be transformed into an empty word. Let this word be labeled zero.

Suppose that some words equal to $\boldsymbol{1}$. For example, $\boldsymbol{\{aba^{-1}b^{-1} = 1;\ ca = 1\}}$. Such relations are said to be *defining relations*. We can use these equalities to obtain new ones. For instance, suppose that $\boldsymbol{aba^{-1}b^{-1} = 1}$. Let us add a word $\boldsymbol{ba}$ to both sides on the right. ("multiply by $\boldsymbol{ba}$ on the right"), We obtain $\boldsymbol{aba^{-1}b^{-1}ba = ba}$. Having cancelled inverse letters on the left side, we obtain $\boldsymbol{ab = ba}$. Having added $\boldsymbol{c}$ to both sides of the equality (on the left) we get $\boldsymbol{cab = cba}$. Since $\boldsymbol{ca = 1}$, we get $\boldsymbol{b = cba}$.

After considering the defining relations we see that some words are *equivalent* or *equal*. We can multiply an equality by the same factor or cancel inverse letters. Thus we obtain new relations (equal words). Suppose we have some defining relations. Let us consider two words.

Can we find out whether these words are equal? The thing is that there is no general algorithm. This is a very important result in higher algebra. Nevertheless if the defining relations satisfy certain terms, such an algorithm exists in fact.

The main issue. Let the common begining of any pair of defining relations $\boldsymbol{A} = \mathbf{1}$ and $\boldsymbol{B} = \mathbf{1}$ be either empty or its length constitutes less then $1/6$ of the $\boldsymbol{A}$ and $\boldsymbol{B}$ lenght. Thus there exists an algorithm enabling for any pair of words to find out if they are equal or not.

The main issue will be discussed below. First we should practice with simple examples. Herewith, we assume that the alphabet consists of the letters from the defining relations only.

Let us consider an example. Suppose $\boldsymbol{aba} = \mathbf{1}$, $\boldsymbol{bab} = \mathbf{1}$. Let us prove that $\boldsymbol{a} = \boldsymbol{b}$. Indeed, $\boldsymbol{a} \equiv \boldsymbol{abaa^{-1}b^{-1}} \equiv \boldsymbol{a^{-1}b^{-1}} \equiv \boldsymbol{a^{-1}b^{-1}bab} \equiv \boldsymbol{b}$. So there exist three nonequivalent words: ($\mathbf{1}$, $\boldsymbol{a}$, $\boldsymbol{a^2}$).

Consequently, if a letter occurs in a word twice, we shall use powers in our notation. So we shall write $\boldsymbol{a^k}$ instead of $\underbrace{\boldsymbol{aaa\ldots aa}}_{\boldsymbol{k}}$. If a power is negative, we put $\boldsymbol{x^k} = \boldsymbol{(x^{-1})^{-k}}$. In particular, $\boldsymbol{a^{-3}} = \boldsymbol{(a^{-1})^3}$.

♦ **B1.** *Let* $\boldsymbol{ba} = \boldsymbol{ab^k}$ *be a defining relation such that* $\boldsymbol{k}$ *is a nonzero integer. Prove that any word can be transformed to the form* $\boldsymbol{a^m b^n}$*; here* $\boldsymbol{m}$*,* $\boldsymbol{n}$ *are integers.*

♦ **B2.** *Consider defining relations* $\boldsymbol{a^4} = \boldsymbol{b^3} = \boldsymbol{(ab)^2} = \mathbf{1}$*. How many different words there are?*

♦ **B3.** *Consider defining relations* $\boldsymbol{aba^{-2}ba} = \boldsymbol{b^3} = \mathbf{1}$*. How many different words there are?*

♦ **B4.** *Consider defining relations* $\boldsymbol{a^2} = \boldsymbol{b^2} = \boldsymbol{(ab)^n} = \mathbf{1}$*. How many different words there are?*

There exists a way which lets illustrate consequences from defining relations. Look at the pictures illustrating the two following consequences: $\boldsymbol{a^2b^3} = \boldsymbol{b^3a^2}$ with the relation $\boldsymbol{ab} = \boldsymbol{ba}$; and $\boldsymbol{b^6} = \mathbf{1}$ with the relations $\boldsymbol{b^2} = \boldsymbol{a}$ and $\boldsymbol{a^3} = \mathbf{1}$. If we move around any cell then we read one of the defining words, and if we move around the whole map then we read the consequence. If we move in the opposite direction then we read letters as inverse ones.

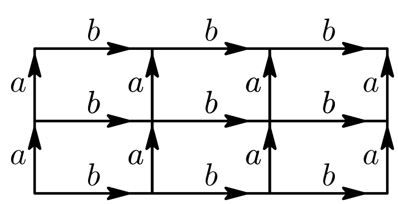


рис. 1

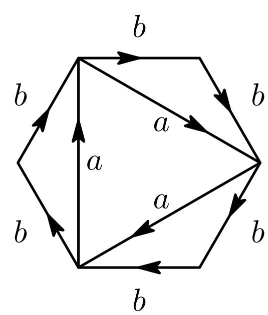


рис. 2

Hence the plane can be divided into some polygons. Now we shall give some definitions. The area inside any polygon is said to be *a cell.* The e dges of the polygons are called *the edges.* Let us write a letter near each edge on the map such that the words written arround any cell correspond to the defining relations words.

♦ **B5.** *Draw a map for the consequence* $\boldsymbol{a^2b^2c^2 = 1}$ *using the defining relations* $\boldsymbol{a^3 = 1}$, $\boldsymbol{b^3 = 1}$, $\boldsymbol{c^3 = 1}$, $\boldsymbol{abc = 1}$.

♦ **B6.** *Draw a map for the consequence* $\boldsymbol{ab^{-1}aba^{-1}b = 1}$ *using the defining relations* $\boldsymbol{a^3 = 1}$, $\boldsymbol{b^3 = 1}$, $\boldsymbol{abab = 1}$.

In fact, the discussed above structures of words with relations have their own name.

Let us consider finite or infinite set of elements G. We assume the law, according to which a third element z is uniquely obtained from any two equal or different elements x and y of such a set, is known. This operation is called *composition* or *symbolic multiplication* of elements and it's result is called the *product* of the elements x and y. We denote the product of x and y by $x * y$. We note that the law of composition may be such that the result depends on the order of the elements multiplied. So $x * y$ is not in general equal to $y * x$.

We will discuss the operations such that an equality $(x * y) * z = x * (y * z)$ holds for any x, y, z from G. In these cases we say that the operation $*$ is *associative*.

Let $\boldsymbol{G}$ be a set of all words in our finite alphabet. We can assign a product operation on $\boldsymbol{G}$: a word $\boldsymbol{W}$ is called the product of two words $\boldsymbol{A}$ and $\boldsymbol{B}$ if $\boldsymbol{W}$ is a result of attaching of $\boldsymbol{B}$ to $\boldsymbol{A}$. Specifically, If $\boldsymbol{A = a_1a_2\dots a_k}$, $\boldsymbol{B = b_1b_2\dots b_n}$ then $\boldsymbol{W = a_1a_2\dots a_kb_1b_2\dots b_n}$. Obviously, this operation is associative. It is easy to understand the meaning of the associativity. The overall result of a product $\boldsymbol{x_1 * x_2 * \dots * x_n}$ is not depends on brackets placement. So the products $\boldsymbol{(x_1 * x_2) * (x_3 * x_4)}$ and $\boldsymbol{x_1 * ((x_2 * x_3) * x_4)}$ are equal.

Definition. A set $\boldsymbol{G}$ with an operation $*$ is called a *group* if the following three conditions hold:

(i) For any $\boldsymbol{x, y, z \in G}$ the following condition holds: $\boldsymbol{(x * y) * z = x * (y * z)}$. So the operation is assotiative;

(ii) There exists an element $\boldsymbol{1}$ in $\boldsymbol{G}$ such that $\boldsymbol{x*1 = 1*x = x}$ for any $\boldsymbol{x \in G}$. This element is called an *identity*;

(iii) For any $\boldsymbol{x \in G}$ there exists a $\boldsymbol{x^{-1} \in G}$ such that $\boldsymbol{x * x^{-1} = x^{-1} * x = 1}$.

We shall often ommit the $*$ sign and write (for example) $\boldsymbol{ab = c}$ or $\boldsymbol{xy = yx}$.

The set of all words in a finite alphabet is a group: an empty word is an identity. If we replace any letter in a word with inverse letter and rewrite the word in the reverse order then we obtain an inverse word. For example, the word $\boldsymbol{z^{-1}y^{-1}x^{-1}d^{-1}c^{-1}b^{-1}a^{-1}}$ is inverse for the word $\boldsymbol{abcdxyz}$. And $\boldsymbol{aba^{-1}c^{-1}ba^{-1}}$ is inverse for $\boldsymbol{ab^{-1}cab^{-1}a^1}$. If we write two inverse words one after another then we can obtain an empty word (identity) by several transformations.

Now assume defining relations. It is easy to see that some words become equivalent. Let us consider these words as the same element of the group. So if we want to make some operations with this element we can choose any of these words.

The number of different elements of a group may be finite, in which case the group is called *finite* and the number of its elements is called its *order*. The order of a group is denoted by $|\boldsymbol{G}|$.

A group is abelian (or commutative) if $\boldsymbol{xy = yx}$ for all $\boldsymbol{x, y \in G}$.

♦ **B7.** *Prove that if for any x from a group $x^2 = 1$, then the group is abelian.*

Definition. A nonempty subset $H \subset G$ is called a *subgroup* of a group G if the following conditions hold:

(i) if $a, b \in H$ then $ab \in H$;

(ii) if $a \in H$ then $a^{-1} \in H$.

It is easy to see that the identity of G belongs to H (suppose $a \in H$, then $a^{-1} \in H$ and $aa^{-1} \in H$). Hence H is a group with respect to the mother-group G operation.

Note that there are two trivial subgroups in any given group: the whole group and the group with one element $\{1\}$. These two groups are called *improper*. Any other group is called *proper*. How can one find any proper group? Suppose $a \in G$. Let us find some powers of a. The set $\{a^k\}$ is an abelian (commutative) subgroup in G. This group is called a *cyclic subgroup* of G and is denoted by $\langle a \rangle$. The element a is said to be a *generator* of $\langle a \rangle$. If $\langle a \rangle = G$, then we say that G is a *cyclic group*.

Suppose n is the minimal integer such that $a^n = 1$. The integer n is called an *order* of element a. If there are no such n, then we say that a has an infinite order.

A cyclic group is generated by its unique element. Let us try to generate a group by several elements.

Suppose that the subset $\langle S \rangle$ in G possesses all finite products

$$g_1{}^{\alpha_1} g_2{}^{\alpha_2} \dots g_k{}^{\alpha_k},$$

$g_i \in S$, $\alpha_i = \pm 1$, $i = 1, \dots, k$. It is easy to see that $\langle S \rangle$ is a subgroup in G. The inverse element for the product above is

$$g_k{}^{\beta_k} g_{k-1}{}^{\beta_{k-1}} \dots g_1{}^{\beta_1},$$

$\beta_i = -\alpha_i$. If every element of a group G is the product of a finite number of elements and inverses of elements from S, then we call the subset S a *set of generators* of G and call the elements of S *generating elements*.

Up to this point we considered every word as a finite letter sequence. Nevertheless, let us have a diagram for the defining relation $a_1 a_2 \dots a_n = 1$. If we look at this diagram, than we cannot find out which letter is the first in the word. Hence it is good to consider a set of all cyclic shifts of a word. This set is called a *cyclic word*. Let a *subset of a cyclic word* (that is not cyclic!) be a subset of one of a word's cyclic shifts. For example, a^2b, aba, ba^2, a^3 are subwords of lenght 3 of the cyclic word a^2ba. A word is called *cyclic irredundant* if every of it's cyclic shifts is irredundant.

Cyclic shifts appear in groups too. Two elements a and b are called *conjuagate* if there exists $x \in G$ such that $a = xbx^{-1}$. It is easy to see that all cyclic shifts are pairwise conjuagate.

♦ **B8.** *Consider a group with some defining relations. Let us take a word. Prove that there exists a conjuagate cyclic irredundant one.*

♦ **B9.** *Пусть группа задана определяющими соотношениями $U_1 = 1, \ldots, U_k = 1$. Докажите, что если $W \equiv 1$ (слово W приводится к пустому), то существуют такие слова $X_1, \ldots X_k$, что слово*

$$X_1U_1X_1{}^{-1}X_2U_2X_2^{-1}\ldots X_kU_kX_k^{-1}$$

приводится к W только сокращениями рядом стоящих взаимно-обратных элементов.

♦ **B9.** *Consider the group with the following defining relations: $U_1 = 1, U_2 = 1, \ldots U_k = 1$. Suppose that $W \equiv 1$. Prove that there exist words $X_1, X_2, \ldots X_k$, such that one can transform the word*

$$X_1U_1X_1{}^{-1}X_2U_2X_2^{-1}\ldots X_kU_kX_k^{-1}$$

into an empty word using only the cancellations of neighboring inverse letters.

Suppose a group satisfies the following equalities: $a^3 = 1$, $bab^{-1} = c$. Hence we obtain the relation $c^3 = 1$. This conclusion can be illustrated by the picture 3. Indeed, if we move around any triangle cell, then we read the word a^3. If we move around any quadrangle cell, then we read the word $cba^{-1}b^{-1}$. If we move around the edges of the whole map, we read the word c^3. This word is a consequence of the equalities $a^3 = 1$ and $cba^{-1}b^{-1} = 1$

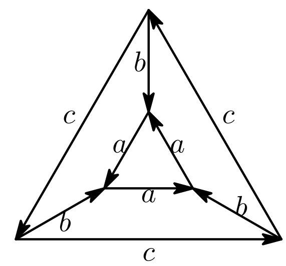


рис. 3

Let us describe how to construct such examples. In order to do this, we consider the conclusion $a^3 = 1$, $b^2 = a \to a^3b^{-1}a^2b^3 = 1$ in detail. First we convert $a^3b^{-1}a^2b^3 = 1$ into the form $(a^3)(b^{-1}a^3b)(b^{-1}a^{-1}b^2b)$. (Here we present our word as a product of some defining relation words and its' inverses.) Let us draw a petal for each product efficient. (see pict. 4.)

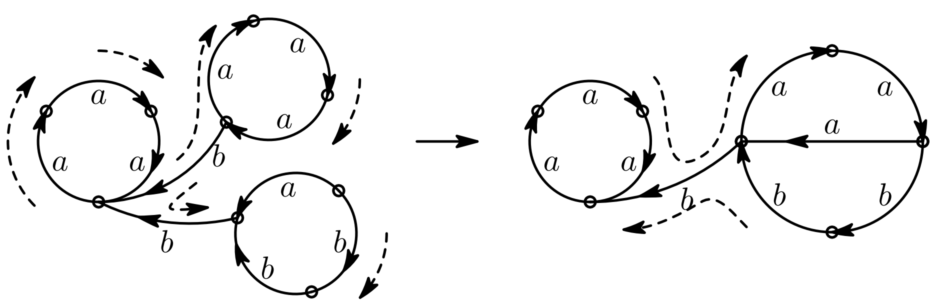


рис. 4

Every petal is a circle with a stem. Then we mark every circle with a letters. Finally we obtain defining relation words (in this example a^3 или $a^{-1}b^2$) on the circles and conjugating words (in this example b^{-1} or an empty word) on the stems. Now walk around all the petals. It is easy to see that we can read the right part of the equality

$$a^3b^{-1}a^2b^3 = (a^3)(b^{-1}a^3b)(b^{-1}a^{-1}b^2b).$$

We want to obtain the word graphically equal to the left part of this equality. So we need to do some cancellations. We do these cancelations by a conglutination of the stems and two edges of the second and the third circles. Finally we obtain the diagram with the word $a^3b^{-1}a^2b^3$ on the outline.

The construction of any other conclusion is performed in a similar way. However there are many different diagrams for one conclusion $W = 1$.

♦ **C1. The van Kampen lemma.** *Suppose that $\boldsymbol{W}$ is a nonempty word in the alphabet $\bar{\boldsymbol{L}} = \boldsymbol{L} \cup \boldsymbol{L}^{-1} \cup \mathbf{1}$; then $\boldsymbol{W} = \mathbf{1}$ iff there exists a diagram which label graphically equals $\boldsymbol{W}$.*

Consider the group with the following defining relations: $\boldsymbol{U} = \{\boldsymbol{U_1} = \mathbf{1}, \dots, \boldsymbol{U_k} = \mathbf{1}\}.$ Let us agree that every word $\boldsymbol{U_i}$ is a cyclic irredundant one and if a word $\boldsymbol{R}$ belongs to the system of the defining relations then $\boldsymbol{R}^{-1}$ also does. Also we may assume that if $\boldsymbol{XY} = \mathbf{1}$ is a defining relation, then $\boldsymbol{YX} = \mathbf{1}$ is a defining relation too. A system of defining relations is *symmetric* if it satisfies all these conditions. It is easy to see that an addition of inverse letters and cycle shifts does not change the set of all consequences. Hence it does not change the group $\boldsymbol{G}$.

If $\boldsymbol{X}$ is a common beginning of two different words $\boldsymbol{XY_1}$ и $\boldsymbol{XY_2}$ from $\boldsymbol{U}$, then we say that $\boldsymbol{X}$ is a *piece* with respect to $\boldsymbol{U}$. Consider the group $\boldsymbol{G}$ with the system of defining relations $\boldsymbol{U}$. Suppose that the lenght of every piece $\boldsymbol{X}$ is less than $^1/_6$ of the length of any word this piece belongs to. Then we say that the group $\boldsymbol{G}$ is a *group with small cancellations*. This condition means that only small part could be cancelled in the product $\boldsymbol{U_i U_j}$ of defining relations.

If we transform words in a group with small cancellations then we «leave many traces» of the defining relations.

♦ **C2.** *Suppose that $\boldsymbol{G}$ is a group with small cancellations. Let us consider a diagram over $\boldsymbol{G}$. Suppose that the label of the outline $\boldsymbol{\phi(q)}$ is a cyclic irredundant word and there are no proper subwords in the cyclic word $\boldsymbol{\phi(q)}$ which are equal to $\mathbf{1}$. Then there exists a cell $\boldsymbol{P}$ with an external arc which length is more than half of the cell perimeter.*

♦ **C3.** *Suppose that $\boldsymbol{G}$ is a group with small cancellations defined by a set of relations. Prove that there exists an algorithm enabling for any pair of words to find out if they are equal or not.*

Groups and Mosaics

Mistakes

It is reasonable to fix some mistakes in the problems text.

♦ **A4.** *Let us consider a plane divided into* `convex` *7-gons which diameters are less or equal to* 1. *Fix a point* O. *Let* $N(R)$ *be the number of 7-gons falling into the circle with diameter* R *and center* O. *Prove that there exists* $\lambda > 1$ *such that* $N(R) > \lambda^R$.

♦ **B1.** *Let* $ba = ab$ *be a defining relation. Prove that any word can be transformed to the form* $a^m b^n$*; here* m, n *are integers.*

Additional problems

♦ **A5.** *Can one cut a plane to convex 7-gons such that diameters of the 7-gons are less or equal to 1 and any unite circle intersects with less then million of them?*

♦ **B7.5a.** *Consider a group over an alphabet* $\{a, b\}$. *Prove the following statement: if for any* x *from a group* $x^3 = 1$, *then the group is finite.*

♦ **B7.5b.** *Consider a group over an alphabet* $\{a, b, c\}$. *Prove the following statement: if for any* x *from a group* $x^3 = 1$, *then the group is finite.*

Some additional technics

Now we shall give the following notation. Consider a cutting of a plane to some polygons (cells). This cutting is called a *map* and is denoted by U. An oriented edge of a cutting U is called an *edge of a map*. So if there exist an edge e, then there exist an opposite oriented edge e^{-1}. This edge e^{-1} belongs the same points of the plane as the e one. Let us make an agreement that outlines of all cells should be read по часовой стрелке. Consider the outline of all map. This outline belongs n edges. Let us denote these edges e_1, e_2, $\dots e_n$ using the orientation of the map. A lap $e_1 \dots e_n$ is called an *outline of a map* U. We can define the outline of a cell by the same way. We shall consider an outline of a map regardless of cyclic shift. Suppose an outline (of a cell, or of a map) $e_1 \dots e_n$ belongs an edge e. A chain of the edges e_1, e_2, $\dots e_n$ is called a *way* if the end of e_i is congruent to the begin of e_{i+1} for any $i = 1, \dots, n-1$.

A subpath is similar to a subword: p is a *subway* of g if $q = p_1 p_2$ holds for some ways p_1 and p_2. Consider a finite alphabet L. We denote $\bar{L} = L \cup L^{-1} \cup 1$, here L^{-1} is an alphabet of inverse letters.

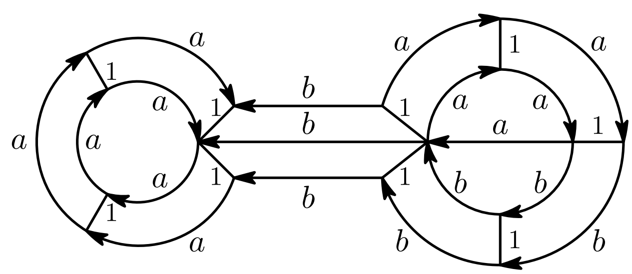
Let us assign a letter $\phi(e)$ from $\bar{L}$ to each edge e of map U. The map U is called a *diagram over* U if $\phi(e^{-1}) = \phi(e)^{-1}$. Consider a path $p = e_1 \dots e_n$ in the diagram U over L. A *label* $\phi(p)$ is a word $\phi(e_1) \dots \phi(e_n)$ in the alphabet $\bar{L}$. By definition, put $\phi(p) = 1$ if

$n = |p| = 0$. It is easy to see that a label of cell (or diagram) outline is defined up to cyclical shift. So it is a cyclical word.

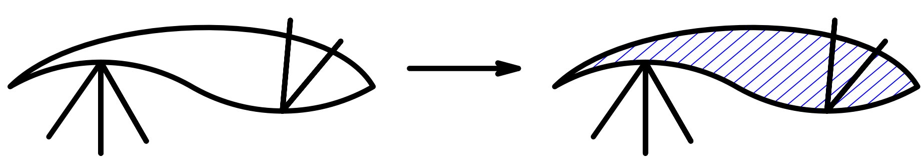
A cell K is called a R-*cell* if it's outline label $\phi(p)$ is graphically equal (up to cyclical shift) to some word from defining relations or its inverse (up to pasting some amount of symbols $\mathbf{1}$). It is clear that if we choose the beginning and direction of "reading" and ignore the symbol $\mathbf{1}$, then we can read a defining relation word.

A cell K is called a *0-cell* if it's outline label $e_1 \ldots e_n$ is graphically equal $\phi(e_1)\ldots\phi(e_n)$, where all $\phi(e_i) = \mathbf{1}$ (= means a graphical equality) OR if for some $i \neq j$ $\phi(e_i) = a$, $\phi(e_j) = a^{-1}$ a belongs to alphabet and for all other $k \neq i, j$ $\phi(e_k) = \mathbf{1}$. An edge is called a $\mathbf{0}$-*edge* if it's label is equal to $\mathbf{1}$. An edge is called a U-*edge* if it's label is non trivial word. *A length* $|p|$ of an arbitrary path is a number of it's U-edges. *Perimeter* of a cell or a diagram is just a length of it's outline.

There are no 0-cells in the examples presented above. However, it is reasonable to introduce them by following reason. The examples shown on the figures 1–3 are truly disk diagrams: if we delete the outline, then the diagram doesn't brake onto several parts. Diagram on the picture 4 is not a disk: if we delete the outline, then it divides into several parts. This problem cause some technical difficulties. For example, we need to cut the subdiagram for induction reason. If we use 0-cells, then we can make a disk diagram with the diagram on the picture 4. It is reasonable to think of 0-cells as "thin" cells (or "fat" edges) and 0-edges as extremely "short" edges.



pic. 5



pic. 6

Thus, a diagram over an alphabet U is called a *diagram over group* G defined by the relations $R_1 \ldots R_n$ if it's every cell is R-cell or $\mathbf{0}$-cell.

It is reasonable to do a $\mathbf{0}$-*cutting* of a diagram. Let a cell be a polygon. We draw a second polygon inside the first one. The second polygon is similar to the first one. Let us connect corresponding vertexes of the polygons with additional edges. Then we label the second polygon with letters as the first one. The additional edges are $\mathbf{0}$-edges. Let us label them with $\mathbf{1}$'s. So we obtain the $\mathbf{0}$-cutting of a cell. Similarly, we can obtain a *doubling* of a path. Draw an additional edge for every edge of a path. Then we label the additional edges by the same way as the corresponding original edges. So we obtain two paths with the same beginning and ending.

G Groups and mosaics

Corrections

♦ **B9.** *A group is presented by defining relations $U_1 = 1, \dots, U_k = 1$. Assume $W \equiv 1$ (i.e., the word W may be reducedto the empty word). Then words $X_1, \dots, X_n$ may be chosen in such a way that the word*

$$X_1 U_{i_1}^{\pm 1} {X_1}^{-1} X_2 U_{i_2}^{\pm 1} X_2^{-1} \dots X_n U_{i_n}^{\pm 1} X_n^{-1},$$

(here $1 \leqslant i_j \leqslant k$), may be reduced to W by using cancellations of elements adjacent to their inverses only.

Additional Problems

Problems of the cycle D prepare the problems of the main cycle E.

♦ **D1.** *We are given an infinite periodic sequence of minimal period n. Assume this sequence has two identical subwords of length $n-1$.*

1. Show that the distance between initial letters of these subwords is a multiple of n.
2. Does a similar statement hold for subwords of length $n-2$?

♦ **D2.** *We consider finite words over a finite alphabet. There is a finite dictionary of forbidden words. Assume that there exists an infinite sequence which does not contain forbidden subwords. Prove that there exists an infinite periodic sequence which does not contain forbidden subwords.*

♦ **D3.** *A finite word is called* cube-free *if it never contains three consecutive entries of the same subword. Show that there exist arbitrarily long cube-free words over the alphabet of two letters.*

Hint: Consider the words a, ab, $abba$, $abbabaab$, $\dots$ Here the next word is obtained from the previous one by using the substitution $a \to ab$, $b \to ba$.

♦ **D4.** *We are given two distinct periodic sequences, whose smallest periods are n and m respectively. Assume that these sequences have a common subword of length $m+n-1$. Show that they have arbitrarily long common subwords.*

♦ **D5.** *Give a precise estimate in* D4. *Consider separately the cases when m and n are coprime and not coprime.*

♦ **D6.** *A triangle is partitioned into convex quadrilaterals. Show that one of the quadrilaterals must have an angle of at least $120°$.*

♦ **D7.** *Does there exist a convex polyhedron whose every face has at least 6 edges?*

♦ **D8.** *A convex polyhedron has at least $7n$ faces. Show that at least n faces have the same number of edges.*

♦ **D9.** *The plane is partitioned into convex k-gons. The diameter of each k-gon is at most 1. For a given point O, let $S_k(R)$ be the number of k-gons contained in the circle of radius R centred at O. Show that there exists R_0 such that for all $R > R_0$ we have $S_k(R) \geqslant \lambda^R$. Show that we may take $\lambda = k/10$. Try to obtain a better estimate for λ.*

♦ **D10.** *The half-plane is partitioned into convex k-gons. The diameter of each k-gon is at most 1.*

We mark L adjacent k-gons at the boundary of the half-plane. We define *the first layer* to be the set of k-gons adjacent to the marked ones. The set of k-gons adjacent to those of the first layer is defined to be *the second layer*. Show that the number of k-gons in the second layer is at least $L(k/10)^2$.

♦ **The Burnside Problem.** *We are given a positive integer n. Does there exist a finitely generated infinite group such that its every element x satisfies the relation $x^n = 1$ ("finitely generated" here means that the group is defined over a finite alphabet).*

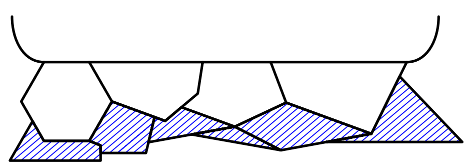
A. Yu. Olshansky has constructed such a group for odd $n > 10^{10}$. We shall follow the main ideas of his construction.

The solution to the Burnside problem is carried out in two steps. The first step is an inductive construction of defining relations. At each step we add several new relations of the form $A^n = 1$. We must verify that every word eventually becomes periodic (i.e., for any word X the relation $X^n = 1$ is deducible from defining relations introduced at some stage). Two words are declared *equal* if one may be reduced to the other using finitely many defining relations introduced.

We thus obtain a group whose every element is periodic.

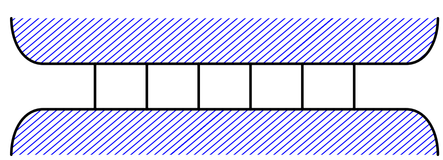
The second part of the argument is the proof that the group we have obtained is **infinite**. To this end we must show that after every step there exists a word not equal to **1**. For this we must show that there does not exist a chart all whose cells are periodic words with a sufficiently large period and along the perimetre we have a word without periodic subwords with a larger period (e.g., square-free or cube-free). To establish this fact one investigates different cases of the adjacency of cells. Cells correspond to defining relations introduced at various stages of the inductive process; at each consecutive stage the word is longer. Cells may thus be of different sizes.

First we show that "large" cells cannot have "large adjacency". Then we consider the case when a large cell is adjacent to several layers of small cells. Here we use exponential growth of the number of polygons in problems D9 and A4.



Pic. 7

Thus according to problem D9, by carefully choosing k one may achieve the effect that already the second layer already has hugely many cells. The choice of k (i.e., the number of sides of our polygons) is controlled by the choice of n (the degree of periodicity of the group). I.e., at the first stage the length of defining relations is n, at the i-th stage it is ni. Polygons will thus have sufficiently many angles, and the exponential growth will force a huge number of cells already in the second layer.



Pic. 8

It remains to consider the case when there is a layer of small cells between two large cells. We shall need the following definition.

Cells $\boldsymbol{A}$ and $\boldsymbol{B}$ are called *reducible* if the following holds:

1. $\boldsymbol{A}$ and $\boldsymbol{B}$ either have common boundary or are joined by a chain of 0-cells ;
2. $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same labels.

If our diagram has two reducbile cells, we may carry out the following operation. We cut out from our diagram the disk formed by the un ion of two cells and insert several $\mathbf{0}$-cells. This operation reduces the number of nontrivial cells. The diagram which does not contain reducible cells is said to be *reduced.*

♦ **E1. Example of a narrow strip for the even exponent.** *Show that there exists a reduced diagram $\boldsymbol{D}$ all whose cells have the form $\boldsymbol{X^k} = \mathbf{1}$, and, which, additionally, has the following structure:*

1. $\boldsymbol{D}$ contains two cells $\boldsymbol{A}$ and $\boldsymbol{B}$ such that all cells are adjacent to them;
2. The perimetres of cells $\boldsymbol{A}$ and $\boldsymbol{B}$ may be arbitrarily larger than perimetres of all other cells (i.e., the ratio of these perimetres may be arbitrarily large.

♦ **E2.** *Assume that a reduced diagram has cells of two types: «large» cells of $\boldsymbol{m}$ letters and «small» cells of $\boldsymbol{n}$ letters. Each cell is labelled by a periodic word of the form $\boldsymbol{A^n}$. What largest common segment may the two large cells have?*

♦ **E3.** *Assume that perimetres of all cells of a reduced diagram are equal either to $\boldsymbol{m}$ or to $\boldsymbol{n}$, where $\boldsymbol{n} \gg \boldsymbol{m}$ and labels of all cells are periodic. Consider the cell $\boldsymbol{A}$ with perimetre $\boldsymbol{n}$. Assume that all adjacent cells have perimetre $\boldsymbol{m}$. As before, the first layer is the set of cells adjacent to $\boldsymbol{A}$. The $\boldsymbol{k}$-th layer ($\boldsymbol{k} > \mathbf{1}$) is the set of cells adjacent to the $\boldsymbol{k} - \mathbf{1}$-th layer. Let $\boldsymbol{A_k}$ be the number of cells at the $\boldsymbol{k}$-th layer. Prove that $\boldsymbol{A_k} \geqslant (\boldsymbol{m}/\mathbf{100})^{\boldsymbol{k}}$.*

♦ **E4.** *Assume that perimetres of all cells of a reduced diagram are equal either to $\boldsymbol{m}$ or to $\boldsymbol{n}$, where $\boldsymbol{n} \gg \boldsymbol{m}$ and labels of all cells are periodic with an odd period. Assume the layer of small cells is squeezed between two large cells (see Pic. 8). Is it true that the label of this contour is periodic?*

ФУНКЦИОНАЛЬНЫЕ УРАВНЕНИЯ

Первый этап

A) **Функциональное уравнение** – это уравнение, которое содержит одну или несколько неизвестных функций (с заданными областями определения и значений). Решить функциональное уравнение – значит найти все функции, которые тождественно ему удовлетворяют. Функциональные уравнения возникают в самых различных областях математики, обычно в тех случаях, когда требуется описать все функции, обладающие заданными свойствами.

Вначале – некоторые типичные приёмы решения функциональных уравнений. Часто бывает полезен **метод подстановок**. Он состоит в том, что переменные заменяются некоторыми новыми функциями (возможно, константами), что позволяет привести уравнение к более удобному виду.

1. Решите следующие функциональные уравнения.

$$f(x) + 2f\left(\frac{1}{x}\right) = 3x. \tag{a}$$

$$f(x) + f\left(\frac{x-1}{x}\right) = 2x. \tag{b}$$

$$f(x+y) + f(x-y) = 2f(x)\cos\ y. \tag{c}$$

B) Рассмотрим теперь некоторые разновидности функциональных уравнений. Во многих функциональных уравнениях задана некоторая итерация искомой функции. Следующую задачу якобы любил давать Фейнман своим молодым сотрудникам.

2. Существует ли такая функция $f(x): R \to R$, что $f(f(x)) = x^2 - 2$ для всех вещественных x?

3. Найдите все функции $f: R^2 \to R$, удовлетворяющие функциональному уравнению

$$f(\ldots f(f(x_1, x_2), x_3)\ldots, x_{2006}) = x_1 + x_2 + x_3 + \ldots + x_{2006}.$$

C) При решении функционального уравнения результат часто зависит от того, требуется ли от искомых функций непрерывность. В наших задачах строгое определение непрерывности не потребуется. Нужно лишь знать, что любой многочлен, экспонента, логарифм (при положительных x), синус, косинус непрерывны и что непрерывная функция всегда обладает следующими свойствами.

(a) Если в точках a и b непрерывная функция принимает различные значения, то любое промежуточное между ними значение принимается хотя бы в одной точке отрезка $[a; b]$ (теорема о промежуточном значении).

(b) Если две непрерывные функции, заданные на вещественной оси, совпадают во всех рациональных точках, то они совпадают всюду.

(c) Если непрерывная функция взаимно однозначна, то она строго монотонна. (Разумеется, верно и обратное.)

(d) Непрерывная функция на отрезке ограниченна.

При решении функциональных уравнений часто бывает также важна монотонность функции.

4. Если функция $f(x)$ строго монотонна, что можно сказать о направлении изменения функции $f(f(x))$?

5. Существует ли такая непрерывная функция $f : R \to R$, что функция $f(f(x))$ строго убывает?

6. Непрерывная функция $f(x)$ такова, что $f(f(x)) = -x^2$ для всех вещественных x. Докажите, что $f(x) \leq 0$ для всех вещественных x.

Дополнительные вопросы к задаче 6.

(a) Существует ли вообще функция, удовлетворяющая условиям задачи?

(b) Существенно ли здесь условие непрерывности?

7. Существует ли такая непрерывная функция $f(x) : R \to R$, что $f(f(x)) = x^2 - 1/2$ для всех вещественных x? (Ср. с задачей 2.)

D) Теперь рассмотрим самое известное из функциональных уравнений - **аддитивное уравнение Коши**, которое часто называют просто **уравнением Коши**:

$$f(x+y) = f(x) + f(y) \quad (x, y \in R).$$

8. Найдите все непрерывные решения аддитивного уравнения Коши.

Какое из свойств (a)-(d) непрерывных функций здесь использовано? Способ решения функциональных уравнений в непрерывных функциях с использованием этого свойства называется **методом Коши**.

Как известно, $(xy)^n = x^n y^n$, $\exp(x+y) = \exp(x)\exp(y)$ $(x, y \in R)$ и $\ln(|xy|) = \ln(|x|) + \ln(|y|)$ $(x, y \in R \setminus \{0\})$. Пользуясь этими фактами и результатом задачи 8, решите следующую задачу.

9. Найдите все непрерывные решения **уравнений Коши**:

$$f(xy) = f(x) + f(y) \quad (x, y \in R \setminus \{0\}); \tag{a}$$

$$f(x+y) = f(xy) \quad (x, y \in R); \tag{b}$$

$$f(xy) = f(x)f(y) \quad (x, y \in R). \tag{c}$$

Базис Гамеля – это такое множество вещественных чисел, что любое вещественное число можно представить, причём единственным образом, в виде $r_1\alpha_1 + \ldots r_n\alpha_n$, где n – натуральное число, $r_1, \ldots, r_n$ – рациональные числа, а $\alpha_1, \ldots, \alpha_n$ принадлежат данному базису Гамеля. Существование базиса Гамеля доказывается с помощью аксиомы выбора, мы примем его как факт.

10. Существуют ли разрывные решения аддитивного уравнения Коши? Если да, то как описать всю совокупность решений этого уравнения? Как описать решения этого уравнения, неотрицательные при $x \geq 0$?

Теперь для сравнения решите следующую задачу.

11.

$$f(x+y) = f^n(x) + f^n(y) \ (x, y \in R;\ n \text{ – фиксированное натуральное число},\ n > 1).$$

E) **Уравнение Пексидера** получается из аддитивного уравнения Коши, если все вхождения функции f заменить на различные функции:

$$k(x+y) = g(x) + h(y).$$

12. (a) Решите уравнение Пексидера.

(b) Найдите все его непрерывные решения.

ФУНКЦИОНАЛЬНЫЕ УРАВНЕНИЯ

Второй этап

Поправка к первому этапу:
9b. Найдите все непрерывные решения уравнения Коши

$$f(x+y)=f(x)f(y) \quad (x,y\in R).$$

F) Пусть функция g удовлетворяет уравнению

$$g(x+y)=g(x)+g(y),$$

где точка (x,y) принадлежит некоторому подмножеству Z плоскости R^2. Если её можно продолжить до функции f, удовлетворяющей аналогичному уравнению при всех $x,y\in R^2$, то мы говорим, что f – *аддитивное продолжение* функции g.

13. Покажите, что если Z – единичный квадрат, то любая функция g, удовлетворяющая уравнению Коши на Z, допускает единственное аддитивное продолжение на всю плоскость.

14. Приведите пример бесконечного множества Z и функции g, которая удовлетворяет уравнению Коши на этом множестве, но не допускает аддитивного продолжения с Z на R^2.

Полученные результаты нашли применение, например, в следующей экономико-математической модели ([1], с. 95-96).

15. Пусть нужно разделить сумму денег, равную S, между $m>2$ конкурирующими проектами. Каждый из n экспертов даёт свои рекомендации (j-й эксперт предлагает дать i-му проекту сумму x_{ij}), и в итоге определяется согласованное распределение

$$\phi_i(x_{i1},\dots,x_{in}) \quad (i=1,\dots,m)$$

(для каждого проекта это распределение определяется суммами, которые выделялись экспертами именно на этот проект, причём вид зависимости может быть различным для различных проектов). Пусть выполнены следующие два естественных условия.

(i) Если никто из экспертов не выделяет денег на данный проект, то проект ничего не получает и в согласованном распределении:

$$\phi_i(0,\dots,0)=0 \quad (i=1,\dots,m).$$

(ii) Если каждое из предлагаемых распределений исчерпывает всю сумму S, то и согласованное распределение исчерпывает всю сумму: из $\sum\limits_{x_{ij}} = S$ $(j=1,\dots,n)$ следует $\sum\limits_{i=1}^{m}\phi_i(x_{i1},\dots,x_{in})=S$.

Покажите, что при указанных условиях все функции ϕ_i совпадают с некоторой функцией $\sum\omega_j\xi_j$, где $\omega_j\geq 0,\ \sum\omega_j=1$.

G) **16.** Найдите все непрерывные вещественные функции положительной вещественной переменной, удовлетворяющие уравнению

$$f(xy)=a(x)+b(x)c(y).$$

Многие школьники знают, что интеграл от степенной функции – степенная функция с коэффициентом (с точностью до аддитивной константы), за единственным исключением, когда вместо степенной функции получается логарифм. В стандартном курсе анализа этот факт доказывается с помощью дифференцирования, причём отдельно для показателя -1 и отдельно для остальных показателей.

17. Опираясь на результаты задач 16, 9a и 9c, найдите интеграл от x^a, где x – положительная вещественная переменная, a – произвольная константа. Дифференцировать нельзя!

H) **18.** Найдите все непрерывные решения **уравнения Даламбера**

$$f(\phi+\psi)+f(\phi-\psi)=2f(\phi)f(\psi)$$

при условии $f(\pi/4)=\sqrt{2}/2$.

19. Можете ли вы указать теперь функциональное уравнение:

(a) для синуса?

(b) для тангенса?

$*$**20.** Используя результаты задач 8 и 18, покажите, что сложение векторов в трёхмерном евклидовом пространстве – единственная операция над парами таких векторов, которая удовлетворяет следующим условиям:

(i) если оба вектора подвергаются одинаковому вращению, то и результат операции подвергается такому же вращению;

(ii) операция коммутативна и ассоциативна;

(iii) для векторов одинакового направления операция сводится к сложению длин;

(iv) сумма двух векторов равной длины непрерывно зависит от угла между ними.

Библиография

1. Я.Ацел, Ж.Домбр. Функциональные уравнения с несколькими переменными. М.: Физматлит, 2003.

2. Е. Пенцак, А.Юрчишин. Функцiйнi рiвняння. Львiв: ЛДУ, 1998.

3. Л.М.Лихтарников. Элементарное введение в функциональные уравнения. СПб: Лань, 1997.

4. Б.Р.Френкин. Интеграл от степени: неочевидное в очевидном. // Математическое просвещение, N 1. М.: МЦНМО, 1997.

ФУНКЦИОНАЛЬНЫЕ УРАВНЕНИЯ

(Решения)

Первый этап

А) **Функциональное уравнение** – это уравнение, которое содержит одну или несколько неизвестных функций (с заданными областями определения и значений). Решить функциональное уравнение – значит найти все функции, которые тождественно ему удовлетворяют. Функциональные уравнения возникают в самых различных областях математики, обычно в тех случаях, когда требуется описать все функции, обладающие заданными свойствами.

Вначале — некоторые типичные приёмы решения функциональных уравнений. Часто бывает полезен **метод подстановок**. Он состоит в том, что переменные заменяются некоторыми новыми функциями (возможно, константами), что позволяет привести уравнение к более удобному виду.

1. Решите следующие функциональные уравнения.

(a) $$f(x)+2f\left(\frac{1}{x}\right)=3x\ (x\neq 0).$$

(b) $$f(x)+f\left(\frac{x-1}{x}\right)=2x\ (x\neq 0).$$

(c) $$f(x+y)+f(x-y)=2f(x)\cos y.$$

Решение. (a) Положим $y=1/x$. Тогда $f\left(\frac{1}{y}\right)+2f(y)=\frac{3}{y}$ и $f(y)+2f(\frac{1}{y})=3y$. Отсюда $f(y)=\frac{2}{y}-y$.

(b) Положим $y=\dfrac{x-1}{x}$, затем $z=\dfrac{y-1}{y}$. Получим систему трёх линейных уравнений относительно $f(x), f(y), f(z)$, из которой находим

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

(c) Положив $y=\pi/2$, получаем $f(x+\frac{\pi}{2})+f(x-\frac{\pi}{2})=0$ для любого x, откуда $f(x+\pi)=-f(x)$. Заменив y на $y+\frac{\pi}{2}$, получаем

$$f(x+y+\frac{\pi}{2})+f(x-y-\frac{\pi}{2})=-2f(x)\sin y.$$

Заменив теперь $x-\frac{\pi}{2}$ на x, имеем

$$f(x+y+\pi)+f(x-y)=-2f(x+\frac{\pi}{2})\sin y$$

и с учётом предыдущего

$$f(x+y)-f(x-y)=2f(x+\frac{\pi}{2})\sin y.$$

Положив $x = 0$, получаем отсюда и из исходного уравнения

$$f(y) = f(0)\cos y + f(\frac{\pi}{2})\sin y.$$

Таким образом, искомая функция должна иметь вид $a\cos y + b\sin y$, где a, b — константы. Легко проверить, что любая такая функция удовлетворяет исходному уравнению.

B) Рассмотрим теперь некоторые разновидности функциональных уравнений. Во многих функциональных уравнениях задана некоторая итерация искомой функции. Следующую задачу якобы любил давать Фейнман своим молодым сотрудникам.

2. Существует ли такая функция $f(x) : R \to R$, что $f(f(x)) = x^2 - 2$ для всех вещественных x?

Решение. Нет. Пусть $g(x) := f(f(x)) = x^2 - 2$. У уравнения $g(x) = x$ два корня: $x = 2, -1$. Очевидно, корни $g(x)-x$ являются корнями $g(g(x))-x$. Разделив $g(g(x))-x$ на $g(x)-x$ и найдя корни полученного квадратного трёхчлена, видим, что $g(g(x))-x$ имеет корни

$$x_1 = 2,\ x_2 = -1,\ x_3 = \frac{-1+\sqrt{5}}{2},\ x_4 = \frac{-1-\sqrt{5}}{2}.$$

Пусть $i \in \{1, 2, 3, 4\}$. Рассмотрим последовательность $x_i, f(x_i), f(f(x_i)), \ldots$. Число 4 является её периодом (возможно, не наименьшим), поэтому все её члены являются корнями $g(g(x)) - x$. Так как $x_{1,2}$ – корни $g(x) - x$, то при $i = 1$ или $i = 2$ число 2 также является периодом этой последовательности, и потому в неё могут входить лишь x_1 и x_2. При $i = 3, 4$ число 4 является наименьшим периодом, так как $x_{3,4}$ не являются корнями $g(x) - x$. Значит, в последовательности $x_3, f(x_3), f(f(x_3)$ все три числа различны и потому какое-то из них равно x_1 или x_2. Но тогда по доказанному все последующие члены равны x_1 или x_2, тогда как пятый член последовательности должен быть равен x_3.

3. Найдите все функции $f : R^2 \to R$, удовлетворяющие функциональному уравнению

$$f(\ldots f(f(x_1, x_2), x_3)\ldots, x_{2006}) = x_1 + x_2 + x_3 + \ldots + x_{2006}.$$

Решение. $f(x, y) = f(0 + 0 + \ldots + 0 + x, y)$ (с 2005 нулями). Отсюда

$$f(x, y) = f(f(\ldots f(f(0, 0))\ldots, x)y) = f(0, 0) + 0 + \ldots + 0 + x + y = x + y + \text{const}.$$

Подставив в исходное уравнение, получаем $\text{const} = 0$, т.е. $f(x, y) = x + y$.

C) При решении функционального уравнения результат часто зависит от того, требуется ли от искомых функций непрерывность. В наших задачах строгое определение непрерывности не потребуется. Нужно лишь знать, что любой многочлен, экспонента, логарифм (при положительных x), синус, косинус непрерывны и что непрерывная функция всегда обладает следующими свойствами.

(a) Если в точках a и b непрерывная функция принимает различные значения, то любое промежуточное между ними значение принимается хотя бы в одной точке отрезка $[a; b]$ (теорема о промежуточном значении).

(b) Если две непрерывные функции, заданные на вещественной оси, совпадают во всех рациональных точках, то они совпадают всюду.

(c) Если непрерывная функция взаимно однозначна, то она строго монотонна. (Разумеется, верно и обратное.)

(d) Непрерывная функция на отрезке ограниченна.

(a) Если в точках a и b непрерывная функция принимает различные значения, то любое промежуточное между ними значение принимается хотя бы в одной точке отрезка $[a; b]$ (теорема о промежуточном значении).

(b) Если две непрерывные функции, заданные на вещественной оси, совпадают во всех рациональных точках, то они совпадают всюду.

(c) Если непрерывная функция взаимно однозначна, то она строго монотонна. (Разумеется, верно и обратное.)

(d) Непрерывная функция на отрезке ограниченна.

При решении функциональных уравнений часто бывает также важна монотонность функции.

4. Если функция $f(x)$ строго монотонна, что можно сказать о направлении изменения функции $f(f(x))$?

Ответ. Функция $f(f(x))$ строго возрастает.

5. Существует ли такая непрерывная функция $f : R \to R$, что функция $f(f(x))$ строго убывает?

Решение. Нет, не существует. Если функция $f(f(x))$ строго убывает, то она взаимно однозначна. Тогда $f(x)$ взаимно однозначна и по свойству (c) строго монотонна. С учётом результата задачи 4 $f(f(x))$ строго возрастает – противоречие.

6. Непрерывная функция $f(x)$ такова, что $f(f(x)) = -x^2$ для всех вещественных x. Докажите, что $f(x) \le 0$ для всех вещественных x.

Решение. В каждой из областей $x \ge 0, x \le 0$ функция $f(x)$ взаимно однозначна и по свойству (c) строго монотонна. Ввиду результата задачи 4 знак монотонности различен на двух полуосях. Так как $f(f(x))$ не ограничена снизу, это верно и для $f(x)$. Значит, $f(x)$ возрастает при $x \le 0$ и убывает при $x \ge 0$, а потому имеет максимум при $x = 0$. Остаётся заметить, что $f(0) = f(-0^2) = f(f(f(0))) = -f(0)^2 \le 0$.

Дополнительные вопросы к задаче 6.

(a) Существует ли вообще функция, удовлетворяющая условиям задачи?

Ответ. Да: $f(x) = -|x|^{\sqrt{2}}$.

(b) Существенно ли здесь условие непрерывности?

Ответ. Да: изменим функцию из п. (a), положив, например, $f(2) = 2^{\sqrt{2}}$.

7. Существует ли такая непрерывная функция $f(x) : R \to R$, что $f(f(x)) = x^2 - 1/2$ для всех вещественных x? (Ср. с задачей 2.)

Решение. Нет. Решение задачи 2 здесь неприменимо (почему?), зато можно использовать приёмы из решения задачи 6. Получаем, что функция $f(x)$ должна убывать при $x \le 0$ и возрастать при $x \ge 0$. Тогда в силу свойства (d) она при $x > 0$ возрастает неограниченно. Пусть $f(0) = a$. Тогда a — минимальное значение $f(x)$, $a \ne 0$ и потому $-1/2 = f(a) > f(0) = a$. Значит, существует такое $b > 0$, что $f(b) = -1/2$. Но $b = f(c)$ для некоторого $c \ne 0$, откуда $c^2 - \frac{1}{2} = f(f(c)) = -\frac{1}{2}$, т.е. $c = 0$ — противоречие.

D) Теперь рассмотрим самое известное из функциональных уравнений — **аддитивное уравнение Коши**, которое часто называют просто **уравнением Коши**:

$$f(x + y) = f(x) + f(y) \quad (x, y \in R).$$

8. Найдите все непрерывные решения аддитивного уравнения Коши.

Решение. Положим $c = f(1)$ и последовательно найдём $f(m)$, $f(m/n)$, $f(0)$ и $f(-m/n)$ для натуральных m, n, а затем применим свойство (b). Получим, что $f(x) = cx$.

Какое из свойств (a)-(d) непрерывных функций здесь использовано? Способ решения функциональных уравнений в непрерывных функциях с использованием этого свойства называется **методом Коши**.

Ответ: использовано свойство (b).

Как известно,

$$(xy)^n = x^n y^n, \ \exp(x+y) = \exp(x)\exp(y) \qquad (x, y \in R)$$

и

$$\ln(|xy|) = \ln(|x|) + \ln(|y|) \qquad (x, y \in R \setminus \{0\}).$$

Пользуясь этими фактами и результатом задачи 8, решите следующую задачу.

9. Найдите все непрерывные решения **уравнений Коши**:

(a) $$f(xy) = f(x) + f(y) \quad (x, y \in R \setminus \{0\});$$

(b) $$f(x+y) = f(xy) \quad (x, y \in R);$$

(b′) $$f(x+y) = f(x)f(y) \quad (x, y \in R).$$

(c) $$f(xy) = f(x)f(y) \quad (x, y \in R).$$

Решение. (a) Пусть вначале $x > 0$. Положим $g(x) = f(e^x)$. Тогда

$$g(x+y) = f(e^{x+y}) = f(e^x e^y) = f(e^x) + f(e^y) = g(x) + g(y),$$

т. е. $g(x)$ удовлетворяет аддитивному уравнению Коши. Так как e^x и $f(x)$ непрерывны, то и $g(x)$ непрерывна и согласно результату задачи 8 имеет вид cx, где c – константа. Тогда $f(x)$ имеет вид $c \ln x$.

В частности, $f(1) = 0$. Положив $x = y = -1$, получаем $f(1) = 2f(-1)$, откуда $f(-1) = 0$. Для произвольного $x < 0$ получаем $f(x) = f(-x) + f(-1) = f(-x)$. Отсюда $f(x) = c \ln |x|$ для произвольного $x \neq 0$.

(b) Положив $y = 0$, получаем $f(x) = f(0)$, т.е. $f(x) \equiv \mathrm{const}$. Очевидно, что любая константа подходит.

(b’) Если $f(x) = 0$ для некоторого x, то $f(z) = f(x)f(z-x) = 0$ для любого z. В противном случае функция, будучи непрерывной, всюду имеет один и тот же знак. Так как $f(2x) = (f(x))^2$, то этот знак положителен и можно рассмотреть непрерывную функцию $g(x) := \ln f(x)$. Имеем $g(x+y) = \ln(f(x)f(y)) = \ln f(x) + \ln f(y) = g(x) + g(y)$, т.е. выполнено аддитивное уравнение Коши. Отсюда $g(x) = cx$ для некоторого c, и $f(x) = e^{cx}$. Таким образом, либо $f(x) \equiv 0$, либо $f(x) \equiv e^{cx}$.

(c) Если $f(x)=0$ для некоторого $x\neq 0$, то $f(z)=f(x\cdot x^{-1}z)=f(x)f(x^{-1}z)=0$ для любого z. В противном случае пусть $x>0$. Тогда $x=z^2$ для некоторого z, и $f(x)=f(z^2)=(f(z))^2>0$. Рассмотрим функцию $\ln f(x)$. Она непрерывна, удовлетворяет аддитивному уравнению Коши и согласно результату задачи 9a имеет вид $c\ln x$ для некоторой константы c. Значит, $f(x)$ при $x>0$ имеет вид x^c.

Так как $f(x)$ должна быть непрерывна в нуле, то $c\geq 0$. При $x<0$ имеем:

$$f(x)=f((-1)\cdot(-x))=f(-1)f(-x)=f(-1)(-x)^c.$$

Кроме того,

$$1=f(1)=f((-1)\cdot(-1))=(f(-1))^2,$$

откуда $f(-1)=\pm 1$. Оба значения подходят. Таким образом, либо $f(x)\equiv 0$, либо $f(x)\equiv |x|^c$, либо $f(x)\equiv x^c$ при $x\geq 0$ и $f(x)\equiv -|x|^c$ при $x<0$ (c – неотрицательная константа).

Базис Гамеля – это такое множество вещественных чисел, что любое вещественное число можно представить, причём единственным образом, в виде $r_1\alpha_1+\dots r_n\alpha_n$, где n — натуральное число, $r_1,\dots,r_n$ — рациональные числа, а $\alpha_1,\dots,\alpha_n$ принадлежат данному базису Гамеля. Существование базиса Гамеля доказывается с помощью аксиомы выбора, мы примем его как факт.

10. Существуют ли разрывные решения аддитивного уравнения Коши? Если да, то как описать всю совокупность решений этого уравнения? Как описать решения этого уравнения, неотрицательные при $x\geq 0$?

Решение. Произвольно зададим значения функции на базисе Гамеля и продолжим по аддитивности. Неотрицательные решения имеют вид $f(x)=cx$, где $c\geq 0$. Действительно, функции такого вида подходят. Обратно, если решение $f(x)$ не имеет такого вида, то для каких-то двух чисел x,y из базиса Гамеля имеем

$$\frac{y}{x}=\alpha,\quad \frac{f(y)}{f(x)}=\beta<\alpha.$$

Пусть m/n — рациональное число, меньшее α и большее β. Тогда $\frac{mn}{x}<\alpha x=y$ и

$$f(\frac{m}{n}x)=\frac{m}{n}f(x)>\beta f(x)=f(y),$$

откуда

$$f(y-\frac{m}{n}x)=(\beta-\frac{m}{n})f(x)<0.$$

Теперь для сравнения решите следующую задачу.

11.

$$f(x+y)=f^n(x)+f^n(y)\ (x,y\in R;\ n - \text{фиксированное натуральное число},\ n>1).$$

Решение. Положив $x=y=0$, находим, что $f(0)=2f^n(0)$, откуда $f(0)$ равно 0, $2^{-1/(n-1)}$ или (при нечётном $n>1$) $-2^{-1/(n-1)}$. Постоянные функции с такими значениями удовлетворяют исходному уравнению. Положив теперь $y=0$ при произвольном x, получаем $f(x)=f^n(x)+f^n(0)$. Значит, $f(x)$ принимает не более n различных значений. Теперь положим $y=x$ при произвольном x. С учётом предыдущего $f(2x)=2f^n(x)=2f(x)-2f^n(0)=2f(x)-f(0)$. Поэтому если $f(x)>f(0)$, то $f(x) < f(2x)$. Значит, $f(2x) > f(0)$, $f(4x) > f(2x)$ и т.д., тогда как f принимает лишь конечное множество значений. Случай $f(x) < f(0)$ аналогичен.

E) **Уравнение Пексидера** получается из аддитивного уравнения Коши, если все вхождения функции f заменить на различные функции:

$$k(x+y) = g(x) + h(y).$$

12. (a) Решите уравнение Пексидера.

(b) Найдите все его непрерывные решения.

Решение. (a) Положим

$$a = g(0), \quad b = h(0), \quad f(z) = -a - b + k(z).$$

Тогда $f(x+y) = f(x) + f(y)$, $k(x) = f(x) + a + b$, $g(x) = f(x) + a$, $h(x) = f(x) + b$. Зададим $f(x)$ на некотором базисе Гамеля и продолжим по аддитивности.

(b) С учётом результата задачи 8 и п. (a) $k(z) = cx+a+b$, $g(x) = cx+a$, $h(x) = cx+b$, где a, b, c – произвольные константы.

ФУНКЦИОНАЛЬНЫЕ УРАВНЕНИЯ
(Решения)

Второй этап

F) Пусть функция g удовлетворяет уравнению

$$g(x+y) = g(x) + g(y),$$

где точка (x, y) принадлежит некоторому подмножеству Z плоскости R^2. Если её можно продолжить до функции f, удовлетворяющей аналогичному уравнению при всех $x, y \in R^2$, то мы говорим, что f – *аддитивное продолжение* функции g.

13. Покажите, что если Z – единичный квадрат, то любая функция g, удовлетворяющая аддитивному уравнению Коши на Z, допускает единственное аддитивное продолжение на всю плоскость.

Решение. Всякое число x представляется в виде ny, где $n \in N,\ y \in Z$. Положим $g(x) = ng(y)$. Корректность, аддитивность и единственность проверяются.

14. Приведите пример бесконечного множества Z и функции g, которая удовлетворяет уравнению Коши на этом множестве, но не допускает аддитивного продолжения с Z на R^2.

Пример. $Z = \{(x, \{x\})\} \mid x \in R$; $g(x) = \{x\}$, где $\{x\}$ – целая часть числа x.

Полученные результаты нашли применение, например, в следующей экономико-математической модели ([1], с. 95–96).

15. Пусть нужно разделить сумму денег, равную S, между $m > 2$ конкурирующими проектами. Каждый из n экспертов даёт свои рекомендации (j-й эксперт предлагает дать i-му проекту сумму x_{ij}), и в итоге определяется согласованное распределение

$$\phi_i(x_{i1}, \ldots, x_{in}) \ (i = 1, \ldots, m)$$

(для каждого проекта это распределение определяется суммами, которые выделялись экспертами именно на этот проект, причём вид зависимости может быть различным для различных проектов). Пусть выполнены следующие два естественных условия.

(i) Если никто из экспертов не выделяет денег на данный проект, то он ничего не получает в согласованном распределении:

$$\phi_i(0, \ldots, 0) = 0 \quad (i = 1, \ldots, m).$$

(ii) Если каждое из предлагаемых распределений исчерпывает всю сумму S, то и согласованное распределение исчерпывает всю сумму: из $\sum\limits_{i=1}^{m} x_{ij} = S\ (j = 1, \ldots, n)$ следует $\sum\limits_{i=1}^{m} \phi(x_{i1}, \ldots, x_{in}) = S$.

Покажите, что при указанных условиях все функции ϕ_i совпадают с некоторой функцией $\sum \omega_j x_i j$, где $\omega_j \geq 0,\ \sum \omega_j = 1$.

Решение. Для краткости будем обозначать $(S, \ldots, S)$ через $\mathbf{S}$, а $(x_{i1}, \ldots, x_{in})$ через $\mathbf{x_i}$. Заметим, что условие (ii) равносильно

$$\sum_{i=2}^{m} \phi_i(\mathbf{x_i}) + \phi_1(\mathbf{S} - \sum_{i=2}^{m} \mathbf{x_i}) = S.$$

При $\mathbf{x_2} = \ldots = \mathbf{x_n} = \mathbf{0}$ получаем отсюда $\phi_1(\mathbf{S}) = S$ $(i = 1, \ldots, m)$. Если же $\mathbf{x_3} = \ldots = \mathbf{x_m} = \mathbf{0}$, то получаем $\phi_1(\mathbf{S} - \mathbf{x}) = S - \phi_2(\mathbf{x})$ для произвольного $\mathbf{x}$. Положив $\mathbf{x_4} = \ldots = \mathbf{x_m} = \mathbf{0}$, $\mathbf{x_2} = \mathbf{x}$, $\mathbf{x_3} = \mathbf{y}$, получаем уравнение Пексидера (для каждой координаты):

$$\phi_2(\mathbf{x}+\mathbf{y}) = \phi_2(\mathbf{x}) + \phi_3(\mathbf{y}) \ (\mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y} \in [0; S]^n).$$

Положим здесь $\mathbf{x} = \mathbf{0}$. Тогда $\phi_2(\mathbf{y}) = \phi_3(\mathbf{y})$. Аналогично получаем, что все ϕ_i равны одной и той же функции ϕ. Приходим к уравнению $\phi(\mathbf{x}+\mathbf{y}) = \phi(\mathbf{x}) + \phi(\mathbf{y})$ для $\mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y} \in [0; S]^n$. Очевидно, $\phi(\xi_1, \ldots, \xi_n) \equiv \phi(\xi_1, 0, \ldots, 0) + \phi(0, \xi_2, 0, \ldots) + \ldots + \phi(0, \ldots, 0, \xi_n) := f_1(\xi_1) + f_2(\xi_2) + \ldots + f_n(\xi_n)$ (здесь $\xi_1, \ldots, \xi_n$ – вещественные переменные). Каждая из функций $f_1, \ldots, f_n$ удовлетворяет аддитивному уравнению Коши на $[0; S]$. Значения всех рассматриваемых функций по смыслу задачи неотрицательны, и с учётом результатов задач 13 и 10 получаем, что $\phi(\mathbf{x})$ имеет вид $\phi(\mathbf{x}) = \sum\limits_{j=1}^{n} \omega_j \xi_j$, где $\omega_j \geq 0$. Поскольку $\phi(\mathbf{S}) = S$, то $\sum\limits_{j=1}^{n} \omega_j = 1$. Обратно, функции такого вида удовлетворяют условиям задачи.

G) **16.** Найдите все непрерывные вещественные функции положительной вещественной переменной, удовлетворяющие уравнению

$$f(xy) = a(x) + b(x)c(y).$$

Ответ. 1) $f(x) \equiv a(x) \equiv K \quad (K$ – произвольная константа, $b(x) \equiv 0, c(y)$ – произвольная непрерывная функция.

2) $f(x) \equiv K_1, a(x) \equiv K_1 - b(x)K_2 \quad (K_1, K_2$ – произвольные константы), $b(x)$ – произвольная непрерывная функция, $c(y) \equiv K_2$.

3)

$$f(x) = K_1 \ln x + K_2, \quad a(x) = K_1 \ln x + K_2 - K_3 K_4,$$

$$b(x) \equiv K_3, \quad c(y) = \frac{K_1 \ln y}{K_3} + K_4$$

$(K_1, K_2, K_3, K_4$ – произвольные константы, $K_3 \neq 0)$.

4)

$$f(x) = K_1(x^\alpha - 1) + K_2, \quad a(x) = K_1(x^\alpha - 1) + K_2 - K_3 K_4 x^\alpha,$$

$$b(x) = K_3 x^\alpha, \quad c(y) = \frac{K_1(x^\alpha - 1)}{K_3} + K_4$$

$(K_1, K_2, K_3, K_4, \alpha$ — произвольные константы, $K_3 \neq 0$, $\alpha \neq 0)$.

Решение. Положим

$$f_1(x) := f(x) - f(1),\ c_1(y) := c(y) - c(1),\ a_1(x) := a(x) - f(1) + b(x)c(1).$$

Тогда

$$(') \qquad f_1(xy) = a_1(x) + b(x)c_1(y),$$

$$('') \qquad f_1(1) = c_1(1) = 0.$$

Положив $y = 1$, получаем $f_1(x) = a_1(x)$. Если $a_1 \equiv 0$, то b или c_1 – тождественный нуль, и мы получаем классы функций 1 и 2 из ответа к задаче (см. выше). В противном случае положим $x = 1$. Тогда

$$f_1(y) \equiv b(1)c_1(y),$$

откуда $b(1) \neq 0$. Положим

$$b_1(x) := b(x)/b(1).$$

Тогда $f_1(xy) = f_1(x) + b_1(x)f_1(y)$, откуда

$$f_1(xyz) = f_1(xy) + b_1(xy)f_1(z) = f_1(x) + b_1(x)f_1(y) + b_1(xyz)f_1(z),$$

$$f_1(xyz) = f_1(x) + b_1(x)f_1(yz) = f_1(x) + b_1(x)f_1(y) + b_1(x)b_1(y)f_1(z).$$

Сравнив эти равенства и взяв значение z, при котором $f_1(z) \neq 0$, получаем

$$b_1(xy) \equiv b_1(x)b_1(y).$$

Так как $b_1(0) \neq 0$, то согласно результату задачи 9с имеем $b_1(x) = x^\alpha$, где α – произвольная константа. Отсюда $f_1(xy) = f_1(x) + x^\alpha f_1(y)$. Если $a = 0$, то с учётом результата задачи 9а получаем класс функций 3. Пусть теперь $\alpha \neq 0$. Возьмём произвольные $x, y \neq 1$. Так как $f_1(xy) = f_1(x) + x^\alpha f_1(y)$, $f_1(xy) = f_1(y) + y^\alpha f_1(x)$, то

$$\frac{f_1(x)}{x^\alpha - 1} := \frac{f_1(y)}{y^\alpha - 1}.$$

Так как левая часть не зависит от y, а правая часть от x, то они являются константами, и мы получаем класс функций 4.

Многие школьники знают, что интеграл от степенной функции – степенная функция с коэффициентом (с точностью до аддитивной константы), за единственным исключением, когда вместо степенной функции получается логарифм. В стандартном курсе анализа этот факт доказывается с помощью дифференцирования, причём отдельно для показателя -1 и отдельно для остальных показателей.

17. Опираясь на результаты задач 16, 9а и 9с, найдите интеграл от x^a, где x – положительная вещественная переменная, a – произвольная константа. Дифференцировать нельзя, пока не получите функциональное уравнение!

Решение. Пусть $g(x) = x^a$, $f(x)$ – первообразная от $g(x)$. Можно считать, что $f(1) = 0$. Так как $g(xy) \equiv g(x)g(y)$, то

$$f(xy) = \int_1^{xy} g(t)dt = \int_1^x g(t) + \int_x^{xy} g(t)dt =$$

$$= f(x) + \int_1^y g(tx)d(tx) = f(x) + g(x)x\int_1^y g(t)dt =$$

$$= f(x) + x^{a+1}f(y).$$

Выполнены условия задачи 16, причём $f(x) = a(x) \neq \mathrm{const}$, $b(x) = f(x)$, $c(y) = x^{a+1}$. Если $a = -1$, то мы имеем случай 3 при $K_2 = K_4 = 0$, $K_3 = 1$, т.е. $f(x) = K_1 \ln x$. Если же $a \neq -1$, то получаем случай 4 при $\alpha = a + 1$, $K_2 = K_4 = 0$, $K_3 = 1$, т.е. $f(x) = K_1(x^{a+1} - 1)$.

Н) **18.** Найдите все непрерывные решения **уравнения Даламбера**

$$f(\phi+\psi)+f(\phi-\psi)=2f(\phi)f(\psi)$$

при условии

$$(**) \qquad f(\pi/4)=\sqrt{2}/2.$$

Решение. Положим $\psi=0$: получаем $f(0)=1$. Тогда для некоторого $C>0$ имеем $f(x)>0$ при $x\in[0,C]$. Положим $\psi=\phi=x/2$: получаем

$$(***) \qquad f(x)+1=2f(x/2)^2.$$

Пусть $f(C)<1$. Тогда $f(C)=\cos\alpha$ для некоторого $\alpha\in[0;\pi/2)$. Ввиду $(***)$ имеем $f(C/2)=\cos\alpha/2$ и по индукции $f(C/2^n)=\cos\alpha/2^n$ для всех натуральных n. Из исходного уравнения получаем:

$$f\left(\frac{k+1}{2^n}C\right)=2f\left(\frac{k}{2^n}C\right)f\left(\frac{1}{2^n}C\right)-f\left(\frac{k-1}{2^n}C\right)=$$

$$=2\cos\left(\frac{k}{2^n}\alpha\right)\cos\left(\frac{\alpha}{2^n}\right)-\cos\left(\frac{k-1}{2^n}\alpha\right)=\cos\left(\frac{k+1}{2^n}\alpha\right).$$

Тогда по непрерывности (свойство (b)) $f(Cx)=\cos\alpha x$ для любого x. Положив $c=\alpha/C$, $Cx=\phi$, имеем $f(x)=\cos c\phi$. Из условия $(**)$ следует, что $c=8k\pm1$.

Пусть теперь $f(C)>1$. Тогда аналогичные рассуждения показывают, что $f(\phi)>1$ при любом ϕ, поэтому условие $(**)$ не выполнено.

Комментарий. При $C>1$ и отсутствии условия $(**)$ уравнение имеет решение $f(x)=\operatorname{ch}(cx)$, где c – произвольная константа, $\operatorname{ch}(x):=\dfrac{e^x+e^{-x}}{2}$.

19. Можете ли вы указать теперь функциональное уравнение:

(a) для синуса?

(b) для тангенса?

Решение. (a)

$$f(\phi+\psi+\frac{\pi}{2})+f(\phi-\psi+\frac{\pi}{2})=2f(\phi+\frac{\pi}{2})f(\psi+\frac{\pi}{2})$$

при условии $(**)$.

(b) $f(x+y)=\dfrac{f(x)+f(y)}{1-f(x)f(y)}$ при условии, например, $f(\pi/4)=1$.

$*$**20.** Используя результаты задач 8 и 18, покажите, что сложение векторов в трёхмерном евклидовом пространстве — единственная операция над парами таких векторов, которая удовлетворяет следующим условиям:

(i) если оба вектора подвергаются одинаковому вращению, то и результат операции подвергается такому же вращению;

(ii) операция коммутативна и ассоциативна;

(iii) для векторов одинакового направления операция сводится к сложению длин;

(iv) сумма двух векторов равной длины непрерывно зависит от угла между ними.

Рассмотрите два случая:

(а) оба вектора имеют равную длину (скажем, единичную);

*(б) векторы имеют произвольную длину.

Схема решения (см. [1], с. 13–18). Обозначим рассматриваемую операцию через $\circ$, а её результат будем называть *суммой* векторов. Из (iii) следует, что $p \circ 0 = p$ для любого вектора p (0 – нулевой вектор). Применив (i), имеем $-p \circ p = 0$. С учётом (ii) получаем:

(v) относительно операции $\circ$ векторы образуют абелеву группу, в которой 0 является нулём, а $-p$ – противоположный элемент для p.

Из (i) и коммутативности операции $\circ$ следует, что сумма двух векторов одинаковой длины лежит на биссектрисе одного из двух углов между ними. Если направление векторов одинаково, то в силу (iii) это меньший (нулевой) угол. Если для каких-то векторов равной длины это больший угол, то по непрерывности (свойство (iv)) для каких-то векторов не противоположного направления сумма равна 0, но это противоречит (v). Поэтому сумма всегда лежит на биссектрисе меньшего угла между векторами.

Фиксируем теперь угол φ между двумя векторами. Если дана их длина x, то в силу (i) определена и длина их суммы $g(x)$. Длину вектора v обозначим $|v|$. Пусть векторы p_1 и p_2 имеют одинаковое направление, так же как и векторы q_1 и q_2, причём угол между p_1 и q_1 равен φ. Если $|p_1| = |q_1| = x$ и $|p_2| = |q_2| = y$, то согласно (iii) $|p_1 \circ p_2| = |q_1 \circ q_2| = x + y$. При этом $|p_1 \circ q_1| = g(x)$, $|p_2 \circ q_2| = g(y)$. Тогда

$$g(x+y) = |(p_1 \circ p_2) \circ (q_1 \circ q_2)| = |p_1 \circ (p_2 \circ q_1) \circ q_2| =$$
$$= |p_1 \circ (q_1 \circ p_2) \circ q_2| = |(p_1 \circ q_1) \circ (p_2 \circ q_2)| = g(x) + g(y).$$

Мы получили уравнение Коши для неотрицательных x, y, причём его решение неотрицательно. Согласно результату задачи 10 $g(x) \equiv cx$ для некоторого $c \geq 0$. На самом деле $c > 0$, т.к. сумма ненулевых векторов непротивоположного направления отлична от 0 (см. выше).

Если угол между векторами единичной длины равен 2ϕ, то длину их суммы обозначим $2f(\phi)$. Пусть даны векторы $p_1 . p_2, q_1, q_2$ единичной длины, причём угол между p_1, p_2 и между q_1, q_2 равен 2ψ, угол между p_1, q_1 равен $2(\phi + \psi)$, а угол между p_2, q_2 равен $2(\phi - \psi)$. Тогда угол между $p_1 \circ p_2, q_1 \circ q_2$ равен 2ϕ. Отсюда можно получить, что

$$f(\phi + \psi) + f(\phi - \psi) = 2f(\phi)f(\psi)$$

при $0 \leq \psi \leq \phi \leq \frac{\pi}{4}$. Функция $f(\phi)$ непрерывна в силу (iv), равна 0 при $\phi = \frac{\pi}{2}$ и отлична от 0 при $0 < \phi < \frac{\pi}{2}$. Из решения задачи 18 следует, что $f(\phi) \equiv \cos\phi$, откуда следует утверждение задачи для векторов равной длины. На случай неравной длины оно обобщается посредством геометрических рассуждений, не использующих функциональные уравнения; см. [1], с. 17–18.

Библиография

1. Я.Ацел, Ж.Домбр. Функциональные уравнения с несколькими переменными. М.: Физматлит, 2003.

2. Е. Пенцак, А.Юрчишин. Функцiйнi рiвняння. Львiв: ЛДУ, 1998.

3. Л.М.Лихтарников. Элементарное введение в функциональные уравнения. СПб: Лань, 1997.

4. Б.Р.Френкин. Интеграл от степени: неочевидное в очевидном. // Математическое просвещение, N 1. М.: МЦНМО, 1997.

FUNCTIONAL EQUATIONS

First stage

A) **A functional equation** is an equation including one or more unknown functions (having prescribed domain and range). To solve a functional equation means to find all functions which satisfy it identically. Functional equations arise in various areas of mathematics, usually when we have to describe all functions having given properties.

We start with some typical methods for solving functional equations. **The method of substitution** is often useful. It consists in replacing variables by some new functions (maybe constants) in order to reduce the equation to some more handy form.

1. Solve the following functional equations.

$$f(x) + 2f\left(\frac{1}{x}\right) = 3x. \tag{a}$$

$$f(x) + f\left(\frac{x-1}{x}\right) = 2x. \tag{b}$$

$$f(x+y) + f(x-y) = 2f(x)\cos\ y. \tag{c}$$

B) Now we shall consider some specific kinds of functional equations. Many functional equations include some iteration of the unknown function. The following Problem 2 is said to be often given by Feynman to his young colleagues.

2. Does there exist a function $f(x) : R \to R$ such that $f(f(x)) = x^2 - 2$ for all real x?

3. Find all functions $f : R^2 \to R$ which satisfy the functional equation

$$f(\ldots f(f(x_1, x_2), x_3) \ldots, x_{2006}) = x_1 + x_2 + x_3 + \ldots + x_{2006}.$$

C) The result of solving a functional equation often depends on whether we impose the requirement of continuity on the functions sought. We shall not need the strict definition of continuity here. It suffices to know that any polynomial, exponent, logarithm (for $x > 0$), sine, cosine are continuous, and that any continuous function has the following properties.

(a) If the values of a continuous function at the points a and b are different then any intermediate value is achieved in some point of the interval $[a; b]$ (the Intermediate Value Theorem).

(b) If two continuous functions defined on the real axis coincide at all rational points then they coincide everywhere.

(c) If a continuous function is 1-1 then it is strictly monotonic. (The inverse is obvious.)

(d) A function continuous on a closed interval $[a; b]$ is bounded on this interval.

Monotonicity of a function also is sometimes important for solving functional equations.

4. If a function $f(x)$ is strictly monotonic, what can you say about growth and decrease of the function $f(f(x))$?

5. Does there exist a continuous function $f : R \to R$ such that $f(f(x))$ strictly decreases?

6. A continuous function $f(x)$ is such that $f(f(x)) = -x^2$ for all real x. Prove that $f(x) \leq 0$ for all real x.

Supplementary questions for problem 6.
(a) Does there exist any function satisfying the conditions of the problem?
(b) Is the continuity condition essential here?
7. Does there exist a continuous function $f(x) : R \to R$ such that $f(f(x)) = x^2 - 1/2$ for all real x? (Compare with Problem 2.)

D) Now we shall consider the most famous functional equation, namely the **additive Cauchy equation**, often simply called the **Cauchy equation**:

$$f(x+y) = f(x) + f(y) \quad (x, y \in R).$$

8. Find all continuous solutions for the additive Cauchy equation.
Which of the properties (a)-(d) of continuous functions has been used here? The method of finding continuous solutions of functional equations using this property is called the **Cauchy method**.
As is well known, $(xy)^n = x^n y^n$, $\exp(x+y) = \exp(x)\exp(y)$ $(x, y \in R)$ and $\ln(|xy|) = \ln(|x|) + \ln(|y|)$ $(x, y \in R \setminus \{0\})$. Using this facts and the result of Problem 8, solve the following problem.
9. Find all continuous solutions for the **Cauchy equations**:

$$f(xy) = f(x) + f(y) \quad (x, y \in R \setminus \{0\}); \tag{a}$$

$$f(x+y) = f(xy) \quad (x, y \in R); \tag{b}$$

$$f(xy) = f(x)f(y) \quad (x, y \in R). \tag{c}$$

A Hamel basis is a set of real numbers such that any real number has a unique representation of the form $r_1\alpha_1 + \ldots r_n\alpha_n$ where n is a positive integer, $r_1, \ldots, r_n$ are rationals, and $\alpha_1, \ldots, \alpha_n$ belong to the given Hamel basis. The existence of a Hamel basis is proved using the axiom of choice, and here we assume it as a given fact.
10. Do there exist any discontinuous solutions of the additive Cauchy equation? If yes, then how can the set of all its solutions be described? How can one describe the solutions of this equation that are non-negative for $x \geq 0$?
For comparison, now solve the following problem.
11.

$$f(x+y) = f^n(x) + f^n(y) \ (x, y \in R;\ n \text{ is a fixed positive integer, } n > 1).$$

E) The **Pexider equation** is obtained from the additive Cauchy equation by replacing all occurences of f with different functions:

$$k(x+y) = g(x) + h(y).$$

12. (a) Solve the Pexider equation.
(b) Find all its continuous solutions.

FUNCTIONAL EQUATIONS

Second stage

The improvement for First Stage:
9b. Find all continuous solutions for the Cauchy equation

$$f(x+y) = f(x)f(y) \quad (x, y \in R).$$

F) Suppose a function g satisfies the equation

$$g(x+y) = g(x) + g(y)$$

where the tuple (x, y) belongs to some subset Z of the plane R^2. If g can be extended to a function f satisfying the same equation for all $x, y \in R^2$ then we say that f is an *additive extension* of g.

13. Show that if Z is the unit square then any function g satisfying the Cauchy equation on Z has a unique additive extension to the whole plane.

14. Give an example of an infinite set Z and a function g which satisfies the Cauchy equation on Z but has no additive extension from Z to R^2.

The above results have an application, for instance, in the following economic-mathematical model described in [1], pp. 95-96. Suppose we have to divide an amount S of money between $m > 2$ competing projects. Each of n experts makes a recommendation (expert j suggests to grant the project i with the sum ξ_{ij}), and finally the 'consensus' allocation is given by some function

$$\phi_i(\xi_{i1}, \ldots, \xi_{in}).$$

Observe that for each project the consensus allocation is determined by the sums recommended by the experts for this project only, but the form of this dependence may vary for different projects. We impose two natural requirements.

(i) If all experts allocate zero sum to some project then this project obtains 0 in the consensus allocation:

$$\phi_i(0, \ldots, 0) = 0 \quad (i = 1, \ldots, n).$$

b) If all the allocations recommended by the experts exhaust the sum S then this is true for the consensus allocation as well: $\sum\limits_{i=1}^{m} x_{ij} = S$ $(j = 1, \ldots, n)$ implies $\sum\limits_{i=1}^{m} \phi_i(x_{i1}, \ldots, x_{in}) = S$.

Show that under the above conditions, all the functions ϕ_i have the same (not depending on i) form $\sum \omega_j \xi_j$ where $\omega_j \geq 0$, $\sum \omega_j = 1$.

G) **16.** Find all continuous real functions of a positive real variable which satisfy the equation

$$f(xy) = a(x) + b(x)c(y).$$

Many of you know that the integral of a power function is again a power function (with a coefficient) with the only exception: the integral of $1/x$ is the logarithm (all integrals are, of course, defined up to an additive constant). In a standard course of calculus, this fact is proved with the help of differentiation, and the cases of degree -1 and of all other degrees are treated separately.

17. Find the integral of x^a where x is a positive real variable, a an arbitrary constant, using the results of Problems 16, 9a and 9c as the base for your argument. It is not allowed to differentiate!

H) **18.** Find all continuous solutions of the **d'Alembert equation**

$$f(\phi+\psi)+f(\phi-\psi)=2f(\phi)f(\psi)$$

under the condition $f(\pi/4)=\sqrt{2}/2$.

19. Now can you present a functional equation defining

(a) the sine function $\sin\, x$?

(b) the tangent function $\tan\, x$?

$*$**20.** Using results of Problems 8 and 18, show that the vector addition in 3-dimensional Eucledean space is the only operation on pairs of such vectors which satisfies the following conditions:

(i) if both vectors are subject to the same rotation then the result of the operation also is subject to the same rotation;

(ii) the operation is commutative and associative;

(iii) two vectors pointing in the same direction yield a vector of the same direction whose length is the sum of the lengths of our initial vectors;

(iv) the sum of two vectors of equal length depends continuously on their angle.

Bibliography

1. J.Aczél, J.Dhombres. Functional equations in several variables. Cambridge University Press 1989.

2. Е. Пенцак, А.Юрчишин. Функційні рівняння. Львів: ЛДУ, 1998.

3. Л.М.Лихтарников. Элементарное введение в функциональные уравнения. СПб: Лань, 1997.

4. Б.Р.Френкин. Интеграл от степени: неочевидное в очевидном. // Математическое просвещение, N 1. М.: МЦНМО, 1997.

FUNCTIONAL EQUATIONS

First stage

A) **A functional equation** is an equation including one or more unknown functions (having prescribed domain and range). To solve a functional equation means to find all functions which satisfy it identically. Functional equations arise in various areas of mathematics, usually when we have to describe all functions having given properties.

We start with some typical methods for solving functional equations. **The method of substitution** is often useful. It consists in replacing variables by some new functions (maybe constants) in order to reduce the equation to some more handy form.

1. Solve the following functional equations.

(a)
$$f(x) + 2f\left(\frac{1}{x}\right) = 3x.$$

(b)
$$f(x) + f\left(\frac{x-1}{x}\right) = 2x.$$

(c)
$$f(x+y) + f(x-y) = 2f(x)\cos y.$$

Solution. (a) Denote $y = 1/x$. Then $f\left(\frac{1}{y}\right) + 2f(y) = \frac{3}{y}$ and $f(y) + 2f(\frac{1}{y}) = 3y$. Hence $f(y) = \frac{2}{y} - y$.

(b) Denote $y = \dfrac{x-1}{x}$, then $z = \dfrac{y-1}{y}$. We obtain a system of three linear equations for $f(x), f(y), f(z)$, from which we find

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

(c) Put $y = \pi/2$, then $f(x + \frac{\pi}{2}) + f(x - \frac{\pi}{2}) = 0$ for any x, hence $f(x + \pi) = -f(x)$. Replace y with $y + \frac{\pi}{2}$, then

$$f(x + y + \frac{\pi}{2}) + f(x - y - \frac{\pi}{2}) = -2f(x)\sin y.$$

Now replace $x - \frac{\pi}{2}$ with x, then

$$f(x + y + \pi) + f(x - y) = -2f(x + \frac{\pi}{2})\sin y,$$

and according to the above

$$f(x+y) - f(x-y) = 2f(x + \frac{\pi}{2})\sin y.$$

For $x = 0$, in view of the original equation we have

$$f(y) = f(0)\cos y + f(\frac{\pi}{2})\sin y.$$

So the desired function must be of the form $f(y) = a\cos y + b\sin y$ where a, b are constants. We easily check that any such function satisfies the original equation.

B) Now we shall consider some specific kinds of functional equations. Many functional equations include some iteration of the unknown function. The following Problem 2 is said to be often given by Feynman to his young colleagues.

2. Does there exist a function $f(x) : R \to R$ such that $f(f(x)) = x^2 - 2$ for all real x?

Solution. No. Suppose $g(x) := f(f(x)) = x^2 - 2$. The equation $g(x) = x$ has two roots: $x = 2, -1$. The roots of $g(x) - x$ obviously are roots of $g(g(x)) - x$. Dividing $g(g(x)) - x$ by $g(x) - x$ and determining the roots of the resulting quadratic polynomial, we see that $g(g(x)) - x$ has roots

$$x_1 = 2, \quad x_2 = -1, \quad x_3 = \frac{-1 + \sqrt{5}}{2}, \quad x_4 = \frac{-1 - \sqrt{5}}{2}.$$

Suppose $i \in \{1, 2, 3, 4\}$. Consider the sequence $x_i, f(x_i), f(f(x_i)), \ldots$. It has a period 4 (maybe not the least one), thus all its terms are roots of $g(g(x)) - x$. Since $x_{1,2}$ are roots of $g(x) - x$, for $i = 1$ or for $i = 2$ the sequence has the period 2 as well, hence the sequence contains x_1 and x_2 only. For $i = 3, 4$ the number 4 is the least period because $x_{3,4}$ are not roots of $g(x) - x$. Thus in the sequence $x_3, f(x_3), f(f(x_3)$ all three numbers are distinct, so one of them equals x_1 or x_2. But then the above implies that all subsequent terms equal x_1 or x_2 while the fifth term must equal x_3.

3. Find all functions $f : R^2 \to R$ which satisfy the functional equation

$$f(\ldots f(f(x_1, x_2), x_3) \ldots, x_{2006}) = x_1 + x_2 + x_3 + \ldots + x_{2006}.$$

Solution. $f(x, y) = f(0 + 0 + \ldots + 0 + x, y)$ (with 2005 zeroes). Hence

$$f(x, y) = f(f(\ldots f(f(0, 0)) \ldots, x)y) = f(0, 0) + 0 + \ldots + 0 + x + y = x + y + \text{const}.$$

Substituting this formula in the original equation, we get const $= 0$, that is, $f(x, y) = x + y$.

C) The result of solving a functional equation often depends on whether we impose the requirement of continuity on the functions sought. We shall not need the strict definition of continuity here. It suffices to know that any polynomial, exponent, logarithm (for $x > 0$), sine, cosine are continuous, and that any continuous function has the following properties.

(a) If the values of a continuous function at the points a and b are different then any intermediate value is achieved in some point of the interval $[a; b]$ (the Intermediate Value Theorem).

(b) If two continuous functions defined on the real axis coincide at all rational points then they coincide everywhere.

(c) If a continuous function is 1-1 then it is strictly monotonic. (The inverse is obvious.)

(d) A function continuous on a closed interval $[a; b]$ is bounded on this interval.

Monotonicity of a function also is sometimes important for solving functional equations.

4. If the function $f(x)$ is strictly monotonic, what can you say about the direction of growth of the function $f(f(x))$?

Answer. The function $f(f(x))$ is strictly increasing.

5. Does there exist a continuous function $f : R \to R$ such that $f(f(x))$ strictly decreases?

Solution. No. If $f(f(x))$ is strictly decreasing then it is 1-1. Then $f(x)$ is 1-1 as well and thus it is strictly monotonic by property (c). By the result of Problem 4, $f(f(x))$ strictly increases – a contradiction.

6. A continuous function $f(x)$ is such that $f(f(x)) = -x^2$ for all real x. Prove that $f(x) \le 0$ for all real x.

Solution. In each of the domains $x \ge 0, x \le 0$, the function $f(x)$ is 1-1 and hence strictly monotonic by property (c). According to the result of Problem 4, the sign of monotonicity is different for these two areas. Since $f(f(x))$ is not bounded from below, the same is true for $f(x)$. Hence $f(x)$ increases for $x \le 0$ and decreases for $x \ge 0$ and thus has maximum at $x = 0$. It remains to observe that $f(0) = f(-0^2) = f(f(f(0))) = -f(0)^2 \le 0$.

Supplementary questions for problem 6.

(a) Does there exist any function satisfying the conditions of the problem?

Answer. Yes: $f(x) = -|x|^{\sqrt{2}}$.

(b) Is the continuity condition essential here?

Answer. Yes: modify the function from (a), putting for example $f(2) = 2^{\sqrt{2}}$.

7. Does there exist a continuous function $f(x) : R \to R$ such that $f(f(x)) = x^2 - \frac{1}{2}$ for all real x? (Compare with Problem 2.)

Solution. No. Here we cannot apply the solution of Problem 2 (why?) but we may use method of solving Problem 6. We obtain that $f(x)$ must decrease for $x \le 0$ and must increase for $x \ge 0$. Then by property (d) it is not bounded from above for $x > 0$. Suppose $f(0) = a$. Then a is the minimum value of $f(x)$, $a \ne 0$ and thus $-1/2 = f(a) > f(0) = a$. Hence there exists $b > 0$ such that $f(b) = -1/2$. But $b = f(c)$ for some $c \ne 0$ and so $c^2 - \frac{1}{2} = f(f(c)) = -\frac{1}{2}$, that is, $c = 0$ — a contradiction.

D) **8.** Find all continuous solutions for the additive Cauchy equation.

Solution. Put $c = f(1)$, find step by step $f(m)$, $f(m/n)$, $f(0)$ and $f(-m/n)$ for positive integers m, n and then apply property (b). We get $f(x) = cx$.

Which of the properties (a)–(d) of continuous functions has been used here? The method of finding continuous solutions of functional equations using this property is called the **Cauchy method**.

Answer: we have used property (b).

As is well known,

$$(xy)^n = x^n y^n, \quad \exp(x+y) = \exp(x)\exp(y) \qquad (x, y \in R)$$

and

$$\ln(|xy|) = \ln(|x|) + \ln(|y|) \qquad (x, y \in R \setminus \{0\}).$$

Using this facts and the result of Problem 8, solve the following problem.

9. Find all continuous solutions for the **Cauchy equations**:

(a) $$f(xy) = f(x) + f(y) \quad (x, y \in R \setminus \{0\});$$

(b) $$f(x+y) = f(xy) \quad (x, y \in R);$$

(b′) $$f(x+y) = f(x)f(y) \quad (x, y \in R).$$

(c) $$f(xy) = f(x)f(y) \quad (x, y \in R).$$

Solution. (a) To start with, let $x > 0$. Set $g(x) = f(e^x)$. Then

$$g(x+y) = f(e^{x+y}) = f(e^x e^y) = f(e^x) + f(e^y) = g(x) + g(y),$$

and so $g(x)$ satisfies the additive Cauchy equation. Since e^x and $f(x)$ are continuous, $g(x)$ also is continuous, and according to the result of Problem 8, the function $g(x)$ has the form $g(x) = cx$ where c is a constant. Then $f(x)$ is of the form $c \ln x$.

In particular, $f(1) = 0$. Putting $x = y = -1$, we obtain $f(1) = 2f(-1)$ and hence $f(-1) = 0$. For arbitrary $x < 0$, we obtain $f(x) = f(-x) + f(-1) = f(-x)$. Hence $f(x) = c \ln |x|$ for arbitrary $x \neq 0$.

(b) Put $y = 0$ and obtain $f(x) = f(0)$, so $f(x) \equiv$ const. Any constant obviously fits.

(b') If $f(x) = 0$ for some x then $f(z) = f(x)f(z-x) = 0$ for any z. Otherwise the function, being continuous, does not change its sign. Since $f(2x) = (f(x))^2$, this sign is positive and we may consider the continuous function $g(x) := \ln f(x)$. We have $g(x+y) = \ln(f(x)f(y)) = \ln f(x) + \ln f(y) = g(x) + g(y)$ and so the additive Cauchy equation is satisfied. Hence $g(x) = cx$ for some c, and $f(x) = e^{cx}$. Thus either $f(x) \equiv 0$ or $f(x) \equiv e^{cx}$.

(c) If $f(x) = 0$ for $x \neq 0$ then $f(z) = f(x \cdot x^{-1}z) = f(x)f(x^{-1}z) = 0$ for any z. Otherwise let $x > 0$. Then $x = z^2$ for some z, and $f(x) = f(z^2) = (f(z))^2 > 0$. Consider the function $\ln f(x)$. It is continuous and satisfies the additive Cauchy equation. According to the result of Problem 9a, our function is of the form $c \ln x$ for some constant c. Thus $f(x)$ for $x > 0$ has the form x^c.

Since $f(x)$ has to be continuous at zero, we have $c \geq 0$. For $x < 0$ we have:

$$f(x) = f((-1) \cdot (-x)) = f(-1)f(-x) = f(-1)(-x)^c.$$

At the same time, $1 = f(1) = f((-1) \cdot (-1)) = (f(-1))^2$, hence $f(-1) = \pm 1$. Both values do fit. Thus either $f(x) \equiv 0$, or $f(x) \equiv |x|^c$, or $f(x) \equiv x^c$ for $x \geq 0$ and $f(x) \equiv -|x|^c$ for $x < 0$ (where c is a non-negative constant).

A Hamel basis is a set of real numbers such that any real number has a unique representation of the form $r_1\alpha_1 + \ldots r_n\alpha_n$ where n is a positive integer, $r_1, \ldots, r_n$ are rationals, and $\alpha_1, \ldots, \alpha_n$ belong to the given Hamel basis. The existence of a Hamel basis is proved using the axiom of choice, and here we assume it as a given fact.

10. Do there exist any discontinuous solutions of the additive Cauchy equation? If yes, then how can the set of all its solutions be described? How can one describe the solutions of this equation that are non-negative for $x \geq 0$?

Solution. Let the values of the function on some Hamel basis by arbitrary. Extend the function by additivity. Non-negative solutions are of the form $f(x) = cx$ where $c \geq 0$. Indeed, functions of such form do fit. Conversely, if some solution $f(x)$ is not of this form then for some two numbers x, y from a Hamel basis we have

$$\frac{y}{x} = \alpha, \quad \frac{f(y)}{f(x)} = \beta < \alpha.$$

Suppose m/n is a rational number which is less than α and greater than β. Then $\frac{mn}{x} < \alpha x = y$ and

$$f(\frac{m}{n}x) = \frac{m}{n}f(x) > \beta f(x) = f(y)$$

and hence

$$f(y - \frac{m}{n}x) = (\beta - \frac{m}{n})f(x) < 0.$$

For comparison, now solve the following problem.

11.

$$f(x+y) = f^n(x) + f^n(y) \ (x, y \in R; \ n \text{ is a fixed positive integer, } \ n > 1).$$

Solution. Putting $x = y = 0$ we get $f(0) = 2f^n(0)$, hence $f(0)$ equals 0, $2^{-1/(n-1)}$ or (for odd $n > 1$) $-2^{-1/(n-1)}$. Constant functions with such values satisfy the original equation. Put now $y = 0$ for arbitrary x. We get $f(x) = f^n(x) + f^n(0)$. Thus $f(x)$ has not more than n distinct values. Now put $y = x$, x arbitrary. Taking the above into account, we get $f(2x) = 2f^n(x) = 2f(x) - 2f^n(0) = 2f(x) - f(0)$. Thus if $f(x) > f(0)$ then $f(x) < f(2x)$. Hence $f(2x) > f(0)$, $f(4x) > f(2x)$ etc., but f has only a finite number of values. The case $f(x) < f(0)$ is similar.

E) The **Pexider equation** is obtained from the additive Cauchy equation by replacing all occurences of f with different functions:

$$k(x+y) = g(x) + h(y).$$

12. (a) Solve the Pexider equation.

(b) Find all its continuous solutions.

Solution. (a) Denote

$$a = g(0), \quad b = h(0), \quad f(z) = -a - b + k(z).$$

Then $f(x+y) = f(x) + f(y)$, $k(x) = f(x) + a + b$, $g(x) = f(x) + a$, $h(x) = f(x) + b$. Take arbitrary values of $f(x)$ on some Hamel basis and extend it by additivity.

(b) In view of the result of Problem 8 and of (a), we have $k(z) = cx + a + b$, $g(x) = cx + a$, $h(x) = cx + b$ where a, b, c are arbitrary constants.

FUNCTIONAL EQUATIONS

Second stage

F) Suppose a function g satisfies the equation

$$g(x+y) = g(x) + g(y)$$

where the tuple (x, y) belongs to some subset Z of the plane R^2. If g can be extended to a function f satisfying the same equation for all $x, y \in R^2$ then we say that f is an *additive extension* of g.

13. Show that if Z is the unit square then any function g satisfying the Cauchy equation on Z has a unique additive extension to the whole plane.

Solution. An arbitrary number x may be represented as ny where $n \in N,\ y \in Z$. Set $g(x) = ng(y)$. Then correctness of the definition, additivity and uniqueness are easily checked.

14. Give an example of an infinite set Z and a function g which satisfies the Cauchy equation on Z but has no additive extension from Z to R^2.

Example. $Z = \{(x, \{x\}) \mid x \in R;\ g(x) = \{x\}\}$ where $\{x\}$ is the integer part of x.

The above results have an application, for instance, in the following economic-mathematical model described in [1], pp. 95–96.

15. Suppose we have to divide an amount S of money between $m > 2$ competing projects. Each of n experts makes a recommendation (expert j suggests to grant the project i with the sum ξ_{ij}), and finally the 'consensus' allocation is given by some function

$$\phi_i(\xi_{i1}, \ldots, \xi_{in}).$$

Observe that for each project the consensus allocation is determined by the sums recommended by the experts for this project only, but the form of this dependence may vary for different projects. We impose two natural requirements.

(i) If all experts allocate zero sum to some project then this project obtains 0 in the consensus allocation:

$$\phi_i(0, \ldots, 0) = 0 \quad (i = 1, \ldots, n).$$

b) If all the allocations recommended by the experts exhaust the sum S then this is true for the consensus allocation as well: $\sum\limits_{i=1}^{m} x_{ij} = S$ $(j = 1, \ldots, n)$ implies $\sum\limits_{i=1}^{m} \phi_i(x_{i1}, \ldots, x_{in}) = S$.

Show that under the above conditions, all the functions ϕ_i have the same (not depending on i) form $\sum \omega_j \xi_j$ where $\omega_j \geq 0,\ \sum \omega_j = 1$.

Solution. For brevity, denote $(S, \ldots, S)$ by $\mathbf{S}$ and $(x_{i1}, \ldots, x_{in})$ by $\mathbf{x_i}$. Observe that condition (ii) is equivalent to

$$\sum_{i=2}^{m} \phi_i(\mathbf{x_i}) + \phi_1(\mathbf{S} - \sum_{i=2}^{m} \mathbf{x_i}) = S.$$

For $\mathbf{x_2} = \ldots = \mathbf{x_n} = \mathbf{0}$ we deduce $\phi_1(\mathbf{S}) = S$ $(i = 1, \ldots, m)$. For $\mathbf{x_3} = \ldots = \mathbf{x_m} = \mathbf{0}$ we get $\phi_1(\mathbf{S} - \mathbf{x}) = S - \phi_2(\mathbf{x})$ for arbitrary $\mathbf{x}$. Putting $\mathbf{x_4} = \ldots = \mathbf{x_m} = \mathbf{0},\ \mathbf{x_2} = \mathbf{x},\ \mathbf{x_3} = \mathbf{y}$, we obtain the Pexider equation (for each coordinate):

$$\phi_2(\mathbf{x} + \mathbf{y}) = \phi_2(\mathbf{x}) + \phi_3(\mathbf{y})\ (\mathbf{x}, \mathbf{y}, \mathbf{x} + \mathbf{y} \in [0; S]^n).$$

Put $\mathbf{x} = \mathbf{0}$. Then $\phi_2(\mathbf{y}) = \phi_3(\mathbf{y})$. Similarly we deduce that each ϕ_i equals the same function ϕ. We get the equation $\phi(\mathbf{x}+\mathbf{y}) = \phi(\mathbf{x})+\phi(\mathbf{y})$ for $\mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y} \in [0; S]^n$. Obviously $\phi(\xi_1, \ldots, \xi_n) \equiv \phi(\xi_1, 0, \ldots, 0)+\phi(0, \xi_2, 0, \ldots)+\ldots+\phi(0, \ldots, 0, \xi_n) := f_1(\xi_1)+f_2(\xi_2)+\ldots+f_n(\xi_n)$ (here $\xi_1, \ldots, \xi_n$ are real variables). Each of functions $f_1, \ldots, f_n$ satisfies additive Cauchy equation on $[0; S]$. The sense of the problem implies non-negativity of values of $\phi, f_1, \ldots, f_n$. In view of the results of Problems 13 and 10, we see that $\phi(\mathbf{x})$ has the form $\phi(\mathbf{x}) = \sum_{j=1}^{n} \omega_j \xi_j$ where $\omega_j \geq 0$. Since $\phi(\mathbf{S}) = S$, we have $\sum_{j=1}^{n} \omega_j = 1$. Conversely, functions of this form satisfy the conditions of the problem.

G) **16.** Find all continuous real functions of a positive real variable which satisfy the equation

$$f(xy) = a(x) + b(x)c(y).$$

Answer. 1) $f(x) \equiv a(x) \equiv K \quad$ (K is an arbitrary constant, $b(x) \equiv 0, c(y)$ an arbitrary continuous function.

2) $f(x) \equiv K_1, a(x) \equiv K_1 - b(x)K_2 \quad$ (K_1, K_2 are arbitrary constants), $b(x)$ an arbitrary continuous function, $c(y) \equiv K_2$.

3)

$$f(x) = K_1 \ln x + K_2, \quad a(x) = K_1 \ln x + K_2 - K_3K_4,$$

$$b(x) \equiv K_3, \quad c(y) = \frac{K_1 \ln y}{K_3} + K_4$$

(K_1, K_2, K_3, K_4 are arbitrary constants, $K_3 \neq 0$).

4)

$$f(x) = K_1(x^\alpha - 1) + K_2, \quad a(x) = K_1(x^\alpha - 1) + K_2 - K_3K_4x^\alpha,$$

$$b(x) = K_3x^\alpha, c(y) = \frac{K_1(x^\alpha - 1)}{K_3} + K_4$$

($K_1, K_2, K_3, K_4, \alpha$ are arbitrary constants, $K_3 \neq 0,\ \alpha \neq 0$).

Solution. Put

$$f_1(x) := f(x) - f(1),\ c_1(y) := c(y) - c(1),\ a_1(x) := a(x) - f(1) + b(x)c(1).$$

Then

$$(') \qquad f_1(xy) = a_1(x) + b(x)c_1(y),$$

$$('') \qquad f_1(1) = c_1(1) = 0.$$

Putting $y = 1$, we get $f_1(x) = a_1(x)$. If $a_1 \equiv 0$ then either b or c_1 is zero constant and we obtain classes 1 and 2 of functions in the answer (see above). Otherwise put $x = 1$. Then

$$f_1(y) \equiv b(1)c_1(y),$$

whence $b(1) \neq 0$. Put

$$b_1(x) := b(x)/b(1).$$

Then $f_1(xy) = f_1(x) + b_1(x)f_1(y)$, hence

$$f_1(xyz) = f_1(xy) + b_1(xy)f_1(z) = f_1(x) + b_1(x)f_1(y) + b_1(xyz)f_1(z),$$

$$f_1(xyz) = f_1(x) + b_1(x)f_1(yz) = f_1(x) + b_1(x)f_1(y) + b_1(x)b_1(y)f_1(z).$$

Compare these equations and take a value of z such that $f_1(z) \neq 0$. We get

$$b_1(xy) \equiv b_1(x)b_1(y).$$

Since $b_1(0) \neq 0$, we have, in view of the result of Problem 9(c), $b_1(x) = x^\alpha$ where α is an arbitrary constant. Thus $f_1(xy) = f_1(x) + x^\alpha f_1(y)$. If $a = 0$ then in view of the result of Problem 9(a) we obtain the class 3 of functions. Now suppose $\alpha \neq 0$. Take arbitrary $x, y \neq 1$. Since $f_1(xy) = f_1(x) + x^\alpha f_1(y)$, $f_1(xy) = f_1(y) + y^\alpha f_1(x)$, we have

$$\frac{f_1(x)}{x^\alpha - 1} := \frac{f_1(y)}{y^\alpha - 1}.$$

Since the left side does not depend on y and the right one on x, both are constants, and we obtain the class 4 of functions.

Many of you know that the integral of a power function is again a power function (with a coefficient) with the only exception: the integral of $1/x$ is the logarithm (all integrals are, of course, defined up to an additive constant). In a standard course of calculus, this fact is proved with the help of differentiation, and the cases of degree -1 and of all other degrees are treated separately.

17. Find the integral of x^a where x is a positive real variable, a an arbitrary constant, using the results of Problems 16, 9a and 9c as the base for your argument. It is not allowed to differentiate until you obtain the functional equation!

Solution. Suppose $g(x) = x^a$, $f(x)$ is the antiderivative for $g(x)$. We may assume $f(1) = 0$. Since $g(xy) \equiv g(x)g(y)$, we have

$$f(xy) = \int_1^{xy} g(t)dt = \int_1^x g(t) + \int_x^{xy} g(t)dt =$$
$$= f(x) + \int_1^y g(tx)d(tx) = f(x) + g(x)x \int_1^y g(t)dt =$$
$$= f(x) + x^{a+1}f(y).$$

We are in the conditions of Problem 16 with $f(x) = a(x) \neq \text{const}$, $b(x) = f(x)$, $c(y) = x^{a+1}$. If $a = -1$, we have case 3 with $K_2 = K_4 = 0$, $K_3 = 1$, so $f(x) = K_1 \ln x$. If $a \neq -1$, we get case 4 with $\alpha = a + 1$, $K_2 = K_4 = 0$, $K_3 = 1$, and so $f(x) = K_1(x^{a+1} - 1)$.

H) **18.** Find all continuous solutions of the **d'Alembert equation**

$$f(\phi + \psi) + f(\phi - \psi) = 2f(\phi)f(\psi)$$

under the condition $f(\pi/4) = \sqrt{2}/2$.

Solution. Putting $\psi = 0$, we get $f(0) = 1$. Then for some $C > 0$ we have $f(x) > 0$ for $x \in [0, C]$. Putting $\psi = \phi = x/2$, we get

$$(***) \qquad f(x) + 1 = 2f(x/2)^2.$$

Suppose $f(C) < 1$. Then $f(C) = \cos\alpha$ for some $\alpha \in [0; \pi/2)$. By $(***)$ we have $f(C/2) = \cos\alpha/2$, and using induction, we get $f(C/2^n) = \cos\alpha/2^n$ for all positive integers n. The original equation implies:

$$f\left(\frac{k+1}{2^n}C\right) = 2f\left(\frac{k}{2^n}C\right)f\left(\frac{1}{2^n}C\right) - f\left(\frac{k-1}{2^n}C\right) =$$

$$= 2\cos\left(\frac{k}{2^n}\alpha\right)\cos\left(\frac{\alpha}{2^n}\right) - \cos\left(\frac{k-1}{2^n}\alpha\right) = \cos\left(\frac{k+1}{2^n}\alpha\right).$$

By continuity (property (b)) $f(Cx) = \cos\alpha x$ for any x. Putting $c = \alpha/C$, $Cx = \phi$, we have $f(x) = \cos c\phi$. Condition $(**)$ implies $c = 8k \pm 1$.

Now suppose $f(C) > 1$. Then similar argument shows that $f(\phi) > 1$ for any ϕ, thus condition $(**)$ fails.

Comment. For $C > 1$, if we omit condition $(**)$, the equation has the solution $f(x) = \mathrm{ch}(cx)$ where c is an arbitrary constant, $\mathrm{ch}(x) := \dfrac{e^x + e^{-x}}{2}$.

19. Now can you present a functional equation defining

(a) the sine function $\sin x$?

(b) the tangent function $\tan x$?

Solution. (a) $f(\phi+\psi+\frac{\pi}{2}) + f(\phi-\psi+\frac{\pi}{2}) = 2f(\phi+\frac{\pi}{2})f(\psi+\frac{\pi}{2})$ under condition $(**)$.

(b) $f(x+y) = \dfrac{f(x)+f(y)}{1-f(x)f(y)}$ under condition, for instance, $f(\pi/4) = 1$.

$*$**20.** Using results of Problems 8 and 18, show that the vector addition in 3-dimensional Euclidean space is the only operation on pairs of such vectors which satisfies the following conditions:

(i) if both vectors are subject to the same rotation then the result of the operation also is subject to the same rotation;

(ii) the operation is commutative and associative;

(iii) two vectors pointing in the same direction yield a vector of the same direction whose length is the sum of the lengths of our initial vectors;

(iv) the sum of two vectors of equal length depends continuously on their angle.

The pattern of the solution (see [1], p. 13–18). Denote the operation under consideration by $\circ$ and call its result *the sum* of the vectors. Condition (iii) implies that $p \circ 0 = p$ for any vector p (0 is the zero vector). Applying (i), we have $-p \circ p = 0$. Taking (ii) into account, we obtain:

(v) for the operation $\circ$, vectors form an Abelian group whose neutral element is 0, and $-p$ is the inverse element for p.

Condition (i) and commutativity of $\circ$ imply that the sum of two vectors of equal length lies on the bissector of one of two angles between them. If the vectors have the same direction then this is the smaller angle by (iii). If this is the greater angle for some two vectors of equal length then by continuity (property (iv)) some two vectors of non-opposite direction have zero sum but this contradicts (v). Thus the sum always lies on the bissector of the smaller angle between vectors.

Fix now the angle φ between two vectors. If their length x is given then the length $g(x)$ of their sum is determined by (i). Denote the length of a vector v by $|v|$. Suppose two vectors p_1 and p_2 have the same direction as well as two vectors q_1 and q_2, and suppose the angle between p_1 and q_1 equals φ. If $|p_1| = |q_1| = x$ and $|p_2| = |q_2| = y$ then in view of (iii) we have $|p_1 \circ p_2| = |q_1 \circ q_2| = x + y$. We also have $|p_1 \circ q_1| = g(x)$, $|p_2 \circ q_2| = g(y)$. Then

$$g(x+y) = |(p_1 \circ p_2) \circ (q_1 \circ q_2)| = |p_1 \circ (p_2 \circ q_1) \circ q_2| =$$
$$= |p_1 \circ (q_1 \circ p_2) \circ q_2| = |(p_1 \circ q_1) \circ (p_2 \circ q_2)| = g(x) + g(y).$$

We have obtained the Cauchy equation for non-negative x, y which has non-negative solution. According to the result of Problem 10, $g(x) \equiv cx$ for some $c \geq 0$. In fact $c > 0$ since the sum of two nonzero vectors of non-opposite direction is not 0 (see above).

If the angle between two vectors of unit length equals 2ϕ then denote the length of their sum by $f(\phi)$. Suppose vectors $p_1.p_2, q_1, q_2$ of unit length are given, and the angles between p_1, p_2 and q_1, q_2 equal 2ψ, the angle between p_1, q_1 equals $2(\phi+\psi)$, and the angle between p_2, q_2 equals $2(\phi-\psi)$. Then the angle between $p_1 \circ p_2, q_1 \circ q_2$ equals 2ϕ. One can deduce that

$$f(\phi+\psi)+f(\phi-\psi)=2f(\phi)f(\psi)$$

for $0 \leq \psi \leq \phi \leq \frac{\pi}{4}$. The function $f(\phi)$ is continuous by (iv), it equals 0 for $\phi = \frac{\pi}{2}$ and does not equal 0 for $0 < \phi < \frac{\pi}{2}$. The solution of Problem 18 implies that $f(\phi) \equiv \cos\phi$, and this in turn implies the assertion of the problem for vectors of equal length. It can be extended to the case of unequal length by means of geometrical argument not using functional equations; see [1], pp. 17–18.

Bibliography

1. J.Aczél, J.Dhombres. Functional equations in several variables. Cambridge University Press 1989.

2. Е. Пенцак, А.Юрчишин. Функцiйнi рiвняння. Львiв: ЛДУ, 1998.

3. Л.М.Лихтарников. Элементарное введение в функциональные уравнения. СПб: Лань, 1997.

4. Б.Р.Френкин. Интеграл от степени: неочевидное в очевидном. // Математическое просвещение, N 1. М.: МЦНМО, 1997.

Теорема о высотах треугольника и тождество Якоби

С. Дориченко и М. Скопенков

Цель данного цикла задач — познакомить с красивой идеей академика В. И. Арнольда о связи теоремы о высотах треугольника с тождеством Якоби [1].

Расскажем прежде всего об одном интересном применении этой идеи — обобщении теоремы о высотах на случай трехмерного пространства:

Теорема о высотах 'трехсторонника' [1]**.** Пусть a, b и c — три попарно непараллельные прямые в пространстве. Пусть a', b' и c' — три общих перпендикуляра к парам прямых b и c, c и a, a и b. Наконец, пусть a'', b'' и c'' — три общих перпендикуляра к парам прямых a и a', b и b', c и c' (дано, что упомянутые пары попарно непараллельны). Тогда три прямые a'', b'' и c'' имеют один общий перпендикуляр (т.е. прямую, пересекающую их все и перпендикулярную им всем).

0. Проверьте, что если прямые a, b и c лежат в одной плоскости, то эта теорема превращается в теорему о высотах треугольника.

Мы начнем со знакомства со сферической геометрией, в которой идея Арнольда проявляется наиболее явно.

Сюжет Первый. Сферическая геометрия и векторное произведение.

Рассмотрим единичную сферу в трехмерном пространстве. Назовем *большой окружностью* (*сферической прямой*) сечение этой сферы произвольной плоскостью, проходящей через ее центр. Отождествим диаметрально противоположные точки на нашей сфере.

Каждой точке на сфере сопоставим вектор, идущий из центра сферы в данную точку. Будем считать, что все векторы, получающиеся из данного умножением на число (положительное или отрицательное), соответствуют той же самой точке.

Каждой сферической прямой сопоставим любой вектор, перпендикулярный плоскости, содержащей данную сферическую прямую. Будем считать, что все векторы, получающиеся из него умножением на число (положительное или отрицательное), соответствуют той же самой сферической прямой.

В дальнейшем под точкой (прямой) мы будем понимать точку на сфере (соответственно, сферическую прямую). Мы оставляем в качестве упражнения определения других естественных объектов сферической геометрии — отрезка, треугольника, перпендикуляра и т. д. В задачах, где спрашивается о геометрическом смысле некоторого тождества, не предполагается однозначного ответа — таких смыслов может быть несколько.

Будем обозначать точки большими буквами, прямые — маленькими, а вектор, соответствующий точке (прямой) — той же самой буквой, что и саму точку (прямую), со значком вектора.

Если $\vec{A}$ и $\vec{B}$ — два вектора, то через $[\vec{A}, \vec{B}]$ будем обозначать их векторное произведение. Напомним, что *векторное произведение* двух векторов $\vec{A}$ и $\vec{B}$ — это вектор, перпендикулярный данным векторам и равный по модулю площади натянутого на них параллелограмма. Направление этого вектора определяется правилом правой руки. Через (A, B) будем обозначать скалярное произведение векторов.

1. Пусть A и B — две точки на сфере. Докажите, что вектор $[\vec{A}, \vec{B}]$ соответствует сферической прямой, проходящей через точки A и B.

2. Пусть a и b — две сферические прямые. Докажите, что вектор $[\vec{a}, \vec{b}]$ соответствует их точке пересечения.

3. Пусть A — точка, b — прямая. Докажите, что вектор $[\vec{A}, \vec{b}]$ (если он ненулевой) соответствует перпендикуляру, опущенному из точки A на прямую b.

4. Пусть A, B и C — три точки. Что означает геометрически условие $\vec{A} + \vec{B} + \vec{C} = 0$?

5. Пусть a, b и c — три прямые. Что означает условие $\vec{a} + \vec{b} + \vec{c} = 0$?

6. Докажите тождество $[\vec{A}, [\vec{B}, \vec{C}]] = \vec{B}(\vec{A}, \vec{C}) - \vec{C}(\vec{A}, \vec{B})$ (*'бац минус цаб'*).

7. Докажите *тождество Якоби* для векторов $\vec{A}$, $\vec{B}$ и $\vec{C}$:

$$[\vec{A}, [\vec{B}, \vec{C}]] + [\vec{B}, [\vec{C}, \vec{A}]] + [\vec{C}, [\vec{A}, \vec{B}]] = 0.$$

8. Пусть A, B и C — вершины сферического треугольника. Что означает геометрически тождество Якоби для векторов $\vec{A}$, $\vec{B}$ и $\vec{C}$?

---

[1]Это утверждение предлагалось в качестве задачи на отборочном туре мех-мата МГУ на студенческую международную олимпиаду в этом году. Попробуйте его доказать!

9. Докажите, что в Теореме о высотах 'трехсторонника' прямые a'', b'' и c'' параллельны одной плоскости.
10. Пусть A и B — две точки. Пусть $\vec{A}$ и $\vec{B}$ — *единичные* векторы, идущие в эти точки из центра сферы. Какой точке соответствует вектор $\vec{A}+\vec{B}$?
11. Какой геометрический смысл имеет тождество $[\vec{A},\vec{B}+\vec{C}]+[\vec{B},\vec{C}+\vec{A}]+[\vec{C},\vec{A}+\vec{B}]=0$?
12. Пусть $\vec{A}$, $\vec{B}$ и $\vec{C}$ — единичные векторы, идущие в точки A, B и C из центра сферы. Какой точке соответствует вектор $[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]$?
13. Пусть a и b - две прямые. Пусть $\vec{a}$ и $\vec{b}$ — соответствующие им векторы, причем $|\vec{a}|=|\vec{b}|=1$. Какой геометрический смысл имеет скалярное произведение $(\vec{a},\vec{b})$ и модуль векторного произведения $|[\vec{a},\vec{b}]|$? Найдите ответ на тот же вопрос для двух точек A и B; для точки A и прямой b.
14. Какой геометрический смысл имеет смешанное произведение $([\vec{A},\vec{B}],\vec{C})$ и тождество $(\vec{A},[\vec{B},\vec{C}])=(\vec{B},[\vec{A},\vec{C}])=(\vec{C},[\vec{A},\vec{B}])$?
15. Докажите, что в сферической геометрии не бывает подобных треугольников — углы треугольника однозначно определяют его стороны.
16. Попытайтесь найти геометрический смысл как можно большего числа алгебраических объектов и тождеств. Тем самым получите доказательство соответствующих теорем сферической геометрии (или, наоборот, самих алгебраических тождеств, если соответствующие геометрические теоремы уже известны).
17*. Назовем *окружностью* сечение сферы произвольной плоскостью. Точка на сфере и сферическая прямая — частные случаи окружности. Изучите, как можно обобщить наше соответствие между векторами и точками (прямыми) на сфере на случай произвольной окружности, и получите новые теоремы сферической геометрии. (Подробнее мы рассмотрим этот вопрос в задачах после промежуточного финиша.)

Сферическая геометрия — пример *неевклидовой геометрии*. Она базируется на тех же аксиомах, что и геометрия Евклида, за исключением 'пятого постулата' — аксиомы о параллельных прямых. В ней остаются справедливыми многие теоремы элементарной геометрии, не использующие пятого постулата.

18*. Попробуйте доказать теоремы сферической геометрии, установленные Вами выше (например, в задачах 8 и 11) чисто геометрически, исходя из аксиом геометрии, кроме аксиомы о параллельных.

Следующий сюжет посвящен доказательству Теоремы о высотах 'трехсторонника'.

Сюжет Второй. Прямые в пространстве и бивекторы.

Зафиксируем в (трехмерном, евклидовом) пространстве точку O. Сопоставим каждой прямой в пространстве пару векторов $(u;v)$ следующим образом: возьмем две точки A и B на этой прямой и положим $u=AB$, $v=[OA,OB]$. Пары векторов будем в дальнейшем называть *бивекторами*.

19. Проверьте, с точностью до умножения обоих векторов u и v на одно и то же число построенный бивектор не зависит от выбора точек A и B на данной прямой.
20. Проверьте, что построенный бивектор однозначно определяет исходную прямую.
21. Проверьте, что таким образом можно получить любой бивектор $(u;v)$ с $u\neq 0$ и $v\perp u$.

Продолжим наше соответствие между прямыми и бивекторами на бивекторы $(u;v)$ с $v \not\perp u$. Обозначим через $\operatorname{pr} v$ проекцию вектора v на плоскость, перпендикулярную вектору u. Будем считать, что бивектор $(u;v)$ соответствует той же прямой, что и бивектор $(u;\operatorname{pr} v)$. Будем обозначать бивектор и соответствующую ему прямую одной и той же буквой, над обозначением бивектора будем ставить 'крышечку' — значок $\hat{\ }$.

Пусть $\hat{a}=(u;v)$ и $\hat{b}=(u';v')$ — два бивектора. Определим их *сумму* покомпонентно: $\hat{a}+\hat{b}=(u+u';v+v')$. Определим их *произведение* формулой

$$[\hat{a},\hat{b}]=([u,u'];[u,v']+[v,u']).$$

Сумма и произведение двух бивекторов — снова бивекторы.

Замечание*. Это определение происходит из формулы для коммутатора в алгебре Ли группы движений трехмерного пространства. Одной из геометрических интерпретаций наших бивекторов являются *скользящие векторы* [3].

22. Пусть прямые a и b не параллельны. Докажите, что бивектор $[\hat{a},\hat{b}]$ соответствует общему перпендикуляру к прямым a и b.

23. Пусть прямые a и b не параллельны. Докажите, что прямая, соответствующая бивектору $\hat{a}+\hat{b}$, пересекает общий перпендикуляр к прямым a и b.
24. Докажите, что произведение бивекторов удовлетворяет тождеству Якоби.
25. Докажите Теорему о высотах 'трехсторонника'
26. Попробуйте получить другие теоремы стереометрии таким способом.

Теорема о высотах треугольника и тождество Якоби

А. Заславский и М. Скопенков

В этом сюжете мы покажем, как с помощью идеи Арнольда получать теоремы, в которой участвуют не только точки и прямые, но и окружности. Вот один из примеров:

Теорема о биссектрисах криволинейного треугольника. Пусть a, b и c — три попарно пересекающиеся окружности на плоскости. Через две точки пересечения окружностей a и b проведем окружность c', образующую с этими окружностями равные углы, и имеющую общие точки с пересечением внутренностей кругов, ограниченных окружностями a и b. Аналогично определим окружности a' и b'. Тогда три окружности a', b' и c' имеют общую точку.

27. Докажите Теорему о биссектрисах криволинейного треугольника.
28*. Зафиксируем некоторую окружность I на плоскости. Назовем *общим перпендикуляром* к паре окружностей a и b окружность, перпендикулярную всем трем окружностям a, b и I. Докажите, что Теорема о высотах 'трехсторонника' останется справедливой, если в ней слово 'прямая' (в пространстве) всюду заменить на слово 'окружность' (на плоскости).

Мы начнем, как обычно, с рассмотрения сферической геометрии, в которой идея Арнольда проявляется наиболее ясно. Задачи этого сюжета можно решать независимо от остальных.

Сюжет Третий. Суммирование окружностей.

Назовем *окружностью* сечение сферы произвольной плоскостью, не обязательно проходящей через центр сферы. Каждой окружности, сопоставим вектор, перпендикулярный этой плоскости, направленный в сторону этой плоскости, и равный по модулю $1/h$, где h — расстояние от этой плоскости до центра сферы.

Всюду в дальнейшем заглавные буквы обозначают некоторые окружности. Вектор, соответствующий некоторой окружности, обозначается той же буквой, что и сама окружность, со значком вектора. Будем обозначать через d_A длину касательной, проведенной к нашей сфере из конца вектора $\vec{A}$.

29. Что означает геометрически условие $\vec{A}+\vec{B}+\vec{C}=0$?
30. Докажите формулу $(\vec{A},\vec{A})=1+d_A^2$.
31. Что означает геометрически равенство $(\vec{A},\vec{B})=1$?
32. Какой окружности соответствует вектор

$$\frac{[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]}{(\vec{A},\vec{B},\vec{C})}?$$

33. Что означает геометрически условие $\vec{A}-\vec{B}+\vec{C}-\vec{D}=0$?
34. Какой окружности соответствует вектор

$$\frac{d_B}{d_B-d_A}\vec{A}-\frac{d_A}{d_B-d_A}\vec{B}?$$

35. Используйте данные факты и алгебраические тождества, чтобы получить теоремы планиметрии.

Дополнение. Представление окружностей точками пространства

Пусть окружность на плоскости задается уравнением $x^2+y^2+ax+by+c=0$. Поставим в соответствие ей точку пространства с координатами (a,b,c).

36. При каких a, b, c существует окружность соответствующая точке (a,b,c)?
37. Пусть дана окружность, соответствующая точке P. Найдите геометрическое место точек, соответствующих окружностям, перпендикулярным данной.
38. Какие пары точек соответствуют двум
а) пересекающимся;
б) не пересекающимся;
в) касающимся окружностям?
39. Докажите следующую теорему (В.Ю.Протасов).
Пусть даны три окружности: первая внутри второй, вторая внутри третьей. Рассматриваются цепочки окружностей, касающихся первой и третьей из данных, такие, что одна из точек пересечения соседних окружностей цепочки принадлежит второй из данных окружностей. Если эта цепочка замыкается при некоторой начальной окружности, то она будет замыкаться и при любой начальной окружности.
40. Дана сфера, касающаяся ее плоскость и четыре точки A, B, C, D в этой плоскости. Пусть D' — точка пересечения плоскостей, проходящих через прямые AB, BC, CA и касающихся сферы. Аналогично определяются точки A', B', C'. Докажите, что точки A', B', C', D' лежат в одной плоскости, касающейся сферы.

Теорема о высотах треугольника и тождество Якоби

М. Скопенков

Решения задач, предложенных до промежуточного финиша.

1. Обозначим через O центр нашей сферы. Рассмотрим плоскость, проходящую через точки O, A и B. Пусть c — сферическая прямая, получающаяся в сечении сферы этой плоскостью. Тогда c — это в точности сферическая прямая, проходящая через точки A и B.
С другой стороны, векторы, соответствующие точкам A и B — это векторы $\vec{A} = \overrightarrow{OA}$ и $\vec{B} = \overrightarrow{OB}$. По определению, вектор $[\vec{A}, \vec{B}]$ перпендикулярен обоим векторам $\overrightarrow{OA}$ и $\overrightarrow{OB}$. Значит, он перпендикулярен и нашей плоскости OAB. По нашему определению, все векторы, перпендикулярные некоторой плоскости, соответствуют сферической прямой, получающейся в сечении сферы этой плоскостью. Поэтому вектор $[\vec{A}, \vec{B}]$ соответствует сферической прямой c.
Итак, вектор $[\vec{A}, \vec{B}]$ соответствует сферической прямой, проходящей через точки A и B.
2. Две сферические прямые пересекаются по паре диаметрально противоположных точек на сфере. Пусть точка C — одна из точек пересечения сферических прямых a и b.
Докажем, что вектор $\overrightarrow{OC}$ параллелен вектору $[\vec{a}, \vec{b}]$. Обозначим через α плоскость сферической прямой a. По определению соответствия между сферическими прямыми и векторами вектор $\vec{a}$ перпендикулярен плоскости α. Отрезок O лежит в этой плоскости, поэтому $\overrightarrow{OC} \perp \vec{a}$. мналогично, $\overrightarrow{OC} \perp \vec{b}$. Значит, векторы $\overrightarrow{OC}$ и $[\vec{a}, \vec{b}]$ параллельны, а следовательно — пропорциональны.
Отсюда получаем, что вектор $[\vec{a}, \vec{b}]$ соответствует точке C (поскольку мы считаем, что все векторы, получающиеся из вектора $\overrightarrow{OC}$ умножением на число, соответствуют точке C.)
3. Пусть c — перпендикуляр, опущенный из точки A на прямую b, то есть сферическая прямая, проходящая через точку A и перпендикулярная прямой b. Нам достаточно показать, что вектор $\vec{c}$, соответствующий прямой c, перпендикулярен обоим векторам $\vec{A}$ и $\vec{b}$. (Тогда вектор $\vec{c}$ пропорционален $[\vec{A}, \vec{b}]$, а пропорциональные вектора соответствуют одной и той же прямой.)
Докажем, что $\vec{c} \perp \vec{A}$. Рассмотрим плоскость сферической прямой c. Так как c проходит через точку A, то A лежит в этой плоскости. Поэтому $\vec{c} \perp \overrightarrow{OA} = \vec{A}$.
Докажем, что $\vec{c} \perp \vec{b}$. Так как сферические прямые b и c перпендикулярны, то содержащие их плоскости также перпендикулярны. Но тогда угол между любым вектором, перпендикулярным первой плоскости, и любым вектором, перпендикулярным второй плоскости, также равен 90°, что и требовалось.
4. *Ответ:* три точки A, B и C лежат на одной прямой.
Решение. Из условия $\vec{A} + \vec{B} + \vec{C} = 0$ следует, что векторы $\vec{A}$, $\vec{B}$ и $\vec{C}$ параллельны некоторой плоскости π. Проведем через центр сферы плоскость, параллельную π. Тогда точки A, B и C лежат на сферической прямой, получающейся в сечении сферы этой плоскостью.
5. *Ответ:* три прямые a, b и c пересекаются в одной точке.
Решение. Из условия $\vec{a}+\vec{b}+\vec{c} = 0$ следует, что векторы $\vec{a}$, $\vec{b}$ и $\vec{c}$ параллельны некоторой плоскости π. Рассмотрим точку P на сфере, такую что $OP \perp \pi$. Проверим, что все три сферические прямые a, b и c проходят через точку P. Поскольку отрезок OP перпендикулярен π, то он перпендикулярен вектору $\vec{a}$, который лежит в этой плоскости. м это значит (смотри решение задачи 3), что точка P лежит на прямой a. мналогично доказывается, что P лежит на прямых b и c.
6. Ни левая, ни правая часть тождества $[\vec{A}, [\vec{B}, \vec{C}]] = \vec{B}(\vec{A}, \vec{C}) - \vec{C}(\vec{A}, \vec{B})$ не изменится, если к вектору $\vec{A}$ добавить любой вектор, пропорциональный вектору $[\vec{B}, \vec{C}]$. Поэтому можно считать, что три вектора $\vec{A}$, $\vec{B}$ и $\vec{C}$ параллельны одной плоскости.
Далее, обе части нашего тождества не изменятся, если к вектору $\vec{B}$ добавить любой вектор, пропорциональный вектору $\vec{C}$. С помощью данной операции можно сделать вектор $\vec{B}$ параллельным вектору $\vec{A}$ (если только вектор $\vec{C}$ не параллелен вектору $\vec{A}$). Поэтому достаточно доказать тождество только для этих двух случаев — $\vec{B} \parallel \vec{A}$ или $\vec{C} \parallel \vec{A}$.
Пусть, для определенности, $\vec{B} \parallel \vec{A}$. Добавляя к вектору $\vec{C}$ вектор $\vec{B}$, умноженный на подходящее действительное число, мы можем добиться условия $\vec{C} \perp B$. Поэтому наше тождество достаточно доказать для случая $\vec{C} \perp B \parallel A$. В этом простейшем случае оно легко проверяется непосредственно: обе части равны $|\vec{A}| \cdot |\vec{B}| \cdot \vec{C}$.
Замечание. Данное тождество получается также из соображений линейности.
7. Тождество Якоби получается суммированием трех тождеств, получаемых из тождества задачи 6 циклической заменой переменных.

Теорема о высотах треугольника и тождество Якоби

А. Заславский и М. Скопенков

Решения задач.

Приведем вначале таблицу, содержащую ответы ко всем задачам, в которых спрашивается о геометрическом смысле алгебраических объектов:

Алгебраический объект	Геометрический смысл
$\vec{A}$	точка A
$\vec{a}$	сферическая прямая a
$[\vec{A},\vec{B}]$	прямая, проходящая через A и B
$[\vec{a},\vec{b}]$	точка пересечения прямых a и b
$[\vec{A},\vec{b}]$	перпендикуляр, опущенный из точки A на прямую b
$\vec{A}+\vec{B}$	середина отрезка AB
$(\vec{a},\vec{b})$	косинус угла между a и b
$\|[\vec{a},\vec{b}]\|$	синус угла между a и b
$(\vec{A},\vec{B})$	косинус длины отрезка AB (то есть дуги большой окружности между A и B)
$\|[\vec{A},\vec{B}]\|$	синус длины отрезка AB
$(\vec{A},\vec{b})$	синус расстояния от A до b
$\|[\vec{A},\vec{b}]\|$	косинус расстояния от A до b
$\vec{A}+\vec{B}+\vec{C}=0$	1) точки A, B и C лежат на одной прямой 2) центры окружностей A, B и C лежат на одной прямой
$\vec{a}+\vec{b}+\vec{c}=0$	прямые a, b и c пересекаются в одной точке
$[\vec{A},[\vec{B},\vec{C}]]+[\vec{B},[\vec{C},\vec{A}]]+[\vec{C},[\vec{A},\vec{B}]]=0$	высоты треугольника ABC пересекаются в одной точке
$[\vec{A},\vec{B}+\vec{C}]+[\vec{B},\vec{C}+\vec{A}]+[\vec{C},\vec{A}+\vec{B}]=0$	медианы треугольника ABC пересекаются в одной точке
$[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]$	центр описанной окружности треугольника ABC
$(\vec{A},\vec{B},\vec{C})$	синус длины отрезка AB · синус расстояния от C до AB
$(\vec{A},[\vec{B},\vec{C}])=(\vec{B},[\vec{C},\vec{A}])=(\vec{C},[\vec{A},\vec{B}])$	теорема синусов для треугольника ABC
$(\vec{A},\vec{B})=1$	окружности A и B перпендикулярны
$\frac{[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]}{(\vec{A},\vec{B},\vec{C})}$	окружность, перпендикулярная окружностям A, B и C
$\vec{A}-\vec{B}+\vec{C}-\vec{D}=0$	существует окружность, перпендикулярная окружностям A, B, C и D
$\frac{d_B}{d_B-d_A}\vec{A}-\frac{d_A}{d_B-d_A}\vec{B}$	биссектриса окружностей A и B

8. Пусть $\vec{A}$, $\vec{B}$ и $\vec{C}$ — векторы, соответствующие вершинам сферического треугольника ABC. Согласно задаче 1 вектор $[\vec{A},\vec{B}]$ соответствует сферической прямой, проходящей через A и B. Тогда по задаче 3 Вектор $[\vec{A},[\vec{B},\vec{C}]$ соответствует перпендикуляру, опущенному из точки A на прямую BC, то есть прямой, содержащей высоту h_A треугольника ABC. Аналогично векторы $[\vec{B},[\vec{C},\vec{A}]$ и $[\vec{C},[\vec{A},\vec{B}]$ соответствуют прямым, содержащим две другие высоты треугольника ABC. Поэтому тождество Якоби, ввиду задачи 5, означает, что построенные три прямые пересекаются в одной точке, то есть мы получаем теорему о высотах треугольника в сферической геометрии.

9. Обозначим через $\vec{a}$ *направляющий вектор* прямой a, то есть любой вектор, параллельный прямой a. Аналогично пусть $\vec{b}$ и $\vec{c}$ — направляющие векторы прямых b и c. Тогда, очевидно, вектор $[\vec{b},\vec{c}]$ параллелен общему перпендикуляру к прямым b и c, то есть прямой a'. Поэтому вектор $[\vec{a},[\vec{b},\vec{c}]]$ параллелен общему перпендикуляру к прямым a и a', то есть прямой a''. Аналогично векторы $[\vec{b},[\vec{c},\vec{a}]]$ и $[\vec{c},[\vec{a},\vec{b}]]$ параллельны прямым b'' и c'' соответственно. Из тождества Якоби следует, что три последних вектора параллельны одной плоскости. Значит, и три прямые a'', b'' и c'' параллельны одной плоскости.

10. Из соображений симметрии легко следует, что вектор $\vec{A}+\vec{B}$ соответствует середине дуги большой окружности, проходящей через точки A и B. Эту точку естественно считать *серединой* сферического отрезка с концами A и B.

Замечание. Поскольку сферическая прямая представляет собой окружность, и мы не различаем диаметрально противоположные точки на сфере, то пара точек A и B на прямой определяет не один, а два отрезка. Если *один* из единичных векторов $\vec{A}$ и $\vec{B}$ заменить на противоположный, то их сумма будет соответствовать середине другого отрезка с концами A и B.

11. Пусть $\vec{A}$, $\vec{B}$ и $\vec{C}$ — единичные векторы, идущие в точки A, B и C из центра сферы. Тогда по задаче 10 вектор $\vec{B}+\vec{C}$ соответствует середине сферического отрезка с концами B и C. Тогда по задаче 1 вектор $[\vec{A},\vec{B}+\vec{C}]$ соответствует

медиане сферического треугольника ABC. А тогда, по задаче 5, тождество $[\vec{A}, \vec{B}+\vec{C}]+[\vec{B}, \vec{C}+\vec{A}]+[\vec{C}, \vec{A}+\vec{B}]=0$ означает, что медианы сферического треугольника ABC пересекаются в одной точке.
Теорему о медианах сферического треугольника можно доказать также, рассматривая вектор $\vec{A}+\vec{B}+\vec{C}$.
Замечание. (А. Мафусалов) Оказывается, у теоремы о медианах сферического треугольника есть 'внешние' аналоги (как у теоремы о биссектрисах треугольника), отсутствующие в евклидовой геометрии. А именно, назовем *внешней медианой* m_A сферического треугольника ABC сферическую прямую, проходящую через точку A и середину дуги BC', где точка C' на сфере диаметрально противоположна точке C. Тогда, оказывается, что *две внешние медианы* m_A *и* m_B*, и одна внутренняя (то есть обычная) медиана* m_C *пересекаются в одной точке.* Эта теорема есть в точности утверждение о пересечении (обычных) медиан, но для треугольника ABC'.
12. Докажем сначала, что векторы $\vec{A}-\vec{B}$, $\vec{B}-\vec{C}$ и $\vec{C}-\vec{A}$ соответствуют *серединным перпендикулярам* к сторонам сферического треугольника ABC. Действительно, поскольку

$$(\vec{A}-\vec{B}, \vec{A}+\vec{B})=(\vec{A}, \vec{A})-(\vec{B}, \vec{B})=1-1=0,$$

то прямая, соответствующая вектору $\vec{A}-\vec{B}$, проходит через середину сферического отрезка AB. А поскольку

$$(\vec{A}-\vec{B},[\vec{A}, \vec{B}])=(\vec{A},[\vec{A}, \vec{B}])-(\vec{B},[\vec{A}, \vec{B}])=0,$$

то эта прямая перпендикулярна отрезку AB. Значит, вектор $\vec{A}-\vec{B}$ соответствует серединному перпендикуляру к отрезку AB.
Обозначим $\vec{V}=[\vec{A}, \vec{B}]+[\vec{B}, \vec{C}]+[\vec{C}, \vec{A}]$, и пусть V — точка на сфере, которая соответствует этому вектору. Проверим, что точка V лежит на всех трех построенных серединных перпендикулярах. Действительно, поскольку

$$(\vec{V}, \vec{A}-\vec{B})=([\vec{B}, \vec{C}], \vec{A})-([\vec{C}, \vec{A}], \vec{B})=0,$$

то точка V лежит на серединном прямой, соответствующей вектору $\vec{A}-\vec{B}$. Аналогично доказывается, что V лежит на двух других серединных перпендикулярах. Итак, V — центр описанной окружности треугольника ABC.
Замечание. То, что три серединных перпендикуляра к сторонам треугольника пересекаются в одной точке, следует также из тождества $(\vec{A}-\vec{B})+(\vec{B}-\vec{C})+(\vec{C}-\vec{A})=0$ и задачи 5.
13. Дадим вначале два естественных определения, необходимых, чтобы сформулировать ответ к этой задаче (приведенный выше). Назовем *углом* между двумя сферическими прямыми угол между касательными к ним в их точке пересечения. (Это то же самое, что угол между плоскостями, в которых лежат данные сферические прямые). Назовем *длиной* сферического отрезка AB длину дуги большой окружности, проходящей через точки A и B.
Решение этой задачи сразу получается из формул $(\vec{a}, \vec{b})=|\vec{a}| \cdot|\vec{b}| \cdot \cos \gamma$ и $|[\vec{a}, \vec{b}]|=|\vec{a}| \cdot|\vec{b}| \cdot \sin \gamma$, где γ — угол между векторами $\vec{a}$ и $\vec{b}$.
14. Пусть векторы $\vec{A}$, $\vec{B}$ и $\vec{C}$ равны по модулю 1. Тогда из задачи 13 получаем, что $|([\vec{A}, \vec{B}], \vec{C})|=\sin AB \sin h_C$, где h_C - длина высоты треугольника ABC, проведенной из вершины C. Тождество $(\vec{A},[\vec{B}, \vec{C}])=(\vec{B},[\vec{C}, \vec{A}])=(\vec{C},[\vec{A}, \vec{B}])$ означает, что $\sin AB \sin h_C=\sin BC \sin h_A=\sin CA \sin h_B$ (1). Записывая аналогичное тождество для единичных векторов, соответствующих прямым, содержащим стороны треугольника ABC, поучим равенство $\sin \angle C \sin h_C=\sin \angle A \sin h_A= \sin \angle B \sin h_B$ (2). Поделив равенство (1) на равенство (2), получим сферическую *теорему синусов*:

$$\frac{\sin AB}{\sin \angle C}=\frac{\sin BC}{\sin \angle A}=\frac{\sin CA}{\sin \angle B}.$$

Замечание. Смешанное произведение трех векторов имеет прозрачный геометрический смысл — объем параллелепипеда, построенного на этих трех векторах, со знаком. Нам не известно столь же естественной интерпретации этого числа в терминах сферической геометрии.
15. *Первое решение.* То, что *стороны* сферического треугольника однозначно определяют его углы, доказывается так же, как в геометрии Евклида. Теперь остается рассмотреть треугольник, 'двойственный' данному (то есть такой треугольник, что идущие в его вершины векторы перпендикулярны плоскостям, содержащим стороны исходного треугольника.)
Второе решение. Аналогично решению задачи 6 можно доказать тождество $([\vec{a}, \vec{c}],[\vec{b}, \vec{c}])=(\vec{a}, \vec{b})(\vec{c}, \vec{c})-(\vec{a}, \vec{c})(\vec{b}, \vec{c})$. Применяя результаты задачи 13, получаем отсюда формулу (*теорема косинуса для сферического треугольника*)

$$\sin \angle A \sin \angle B \cos AB=\cos \angle C-\cos \angle A \cos \angle B.$$

Отсюда уже непосредственно следует, что углы сферического треугольника однозначно определяют его стороны.
Замечание.* Полученные выше теоремы сферической геометрии сохраняют силу также для геометрии Лобачевского, поскольку она переходит в сферическую при умножении всех координат на мнимую единицу i. При этом слова 'прямые пересекаются в одной точке' нужно заменить на 'прямые принадлежат одному пучку', а '$\cos AB$' во всех формулах — на '$\operatorname{ch} AB$'. Если представить плоскость Лобачевского как единичную сферу в псевдоевклидовом пространстве, то формально проходят предыдущие рассуждения с векторным произведением $[\vec{A}, \vec{B}]:=*(\vec{A} \wedge \vec{B})$ [1].

19. Нетрудно убедиться, что направления обоих векторов u и v не зависят от выбора точек A и B: вектор u параллелен нашей прямой, а вектор v перпендикулярен плоскости, проходящей через нашу прямую и точку O. При этом модуль вектора u равен AB, а модуль вектора v равен удвоенной площади треугольника ABO, то есть $AB \cdot h$, где h — расстояние от точки O до нашей прямой. Поэтому при изменении точек A и B векторы u и v умножаются на одно и то же число.

20. Действительно, наша прямая обязана лежать в плоскости, проходящей через O и перпендикулярной вектору v. Наша прямая также должна быть параллельна вектору v, а расстояние от нее до точки O должно равняться $|v|/|u|$. При этом вектор $[u, v]$ указывает, в какой 'стороне' относительно точки O расположена данная прямая. Тем самым наша прямая определяется парой векторов u и v однозначно.
21. Построение требуемой прямой по векторам $\vec{u}$ и $\vec{v}$ фактически приведено в решении предыдущей задачи.
22. (эта одна из самых трудных задач из данного списка) Определим *скалярное произведение* бивекторов $\hat{a} = (u, v)$ и $\hat{b} = (u', v')$ формулой

$$(\hat{a}, \hat{b}) = ((u, v); (u, v') + (u', v)).$$

Скалярное произведение двух бивекторов — это *пара* чисел. Нетрудно проверить, что $(\hat{a}, [\hat{a}, \hat{b}]) = (\hat{b}, [\hat{a}, \hat{b}]) = (0; 0)$. Поэтому наша задача сразу следует из такой леммы:

Лемма. Если $(\hat{a}, \hat{b}) = (0; 0)$, то прямые a и b пересекаются, и угол между ними — прямой.

Доказательство леммы. То, что прямые a и b перпендикулярны, следует сразу из того, что обращается в нуль первая компонента скалярного произведения бивекторов $\hat{a}$ и $\hat{b}$: $(u, u') = 0$. Остается показать, что если $(\hat{a}, \hat{b}) = (0; 0)$, то прямые a и b пересекаются.
Рассмотрим вначале случай, когда $u \perp v$ и $u' \perp v'$.
Проверим, что скалярное произведение бивекторов не зависит от выбора точки O. Предположим, что в самом начале мы фиксировали другую точку O_1. Пусть $\hat{a}_1 = (u_1, v_1)$, $\hat{b}_1 = (u'_1, u'_1)$ — бивекторы, которые мы построили по прямым a и b, считая фиксированной точкой точку O_1. Тогда непосредственно проверяется, что $u_1 = u$, $v_1 = v + [u, \overrightarrow{OO_1}]$ и $u'_1 = u'$, $v'_1 = v' + [u', \overrightarrow{OO_1}]$. Поэтому

$$(\hat{a}_1, \hat{b}_1) = ((u, u'); (u, v' + [u', \overrightarrow{OO_1}]) + (v + [u, \overrightarrow{OO_1}], u')) = ((u, u'); (u, v') + (v, u')) = (\hat{a}, \hat{b}),$$

так как $(u, [u', \overrightarrow{OO_1}]) + ([u, \overrightarrow{OO_1}], u') = 0$. Итак, действительно, скалярное произведение бивекторов не зависит от выбора точки O.
Поэтому можно без ограничения общности считать, что точка O лежит на прямой a. Тогда $v = 0$. Поэтому условие $(\hat{a}, \hat{b}) = (0; 0)$ означает, что вектор u перпендикулярен вектору v'. Если $v' = 0$, то все доказано, так как тогда O — общая точка прямых a и b. Если $v' \neq 0$, то тогда обе прямые лежат в плоскости, проходящей через точку O перпендикулярно вектору v', а следовательно — пересекаются. Тем самым случай $u \perp v$ и $u' \perp v'$ полностью разобран.
Рассмотрим теперь случай, когда не обязательно $u \perp v$ и $u' \perp v'$. (Этот случай особенно важен для дальнейшего, поскольку даже если бивекторы $\hat{a}$ и $\hat{b}$ обладали указанным свойством, для бивектора $[\hat{a}, \hat{b}]$ оно может не выполняться.) В этом случае, по определению, бивектор $\hat{a} = (u, v)$ соответствует той же прямой, что и бивектор $(u, pr\, v)$. Обозначим последний бивектор через $\hat{a}_\perp$. Заметим, что $pr\, v = v + \alpha u$ для некоторого числа α. Аналогично определим бивектор $b_\perp$ и число β. Тогда, если $(\hat{a}, \hat{b}) = (0; 0)$, то $(\hat{a}_\perp, \hat{b}_\perp) = ((u, u'); (u, v') + \beta(u, u') + (v, u') + \alpha(u, u')) = (\hat{a}, \hat{b}) = (0; 0)$, поскольку $(u, u') = 0$. Тем самым второй случай в нашей лемме сводится к первому, и лемма доказана.
Замечание.* Определение произведения бивекторов происходит из формулы для коммутатора в алгебре Ли группы движений трехмерного пространства. Одной из геометрических интерпретаций наших бивекторов являются *скользящие векторы* [3].
23. Заметим, что $(\hat{a} + \hat{b}, [\hat{a}, \hat{b}]) = (\hat{a}, [\hat{a}, \hat{b}]) + (\hat{b}, [\hat{a}, \hat{b}]) = 0$. Поэтому по лемме из решения задачи 22 прямая, соответствующая вектору $\hat{a} + \hat{b}$, пересекает прямую, соответствующую вектору $[\hat{a}, \hat{b}]$. А по задаче 22 последняя прямая — это общий перпендикуляр к прямым a и b.
24. Тождество Якоби для произведения бивекторов получается непосредственно из тождества Якоби для произведения векторов.
25. Применяя несколько раз задачу 22, получаем, что бивекторы $[\hat{a}, [\hat{b}, \hat{c}]]$, $[\hat{b}, [\hat{c}, \hat{a}]]$ и $[\hat{c}, [\hat{b}, \hat{a}]]$ соответствуют прямым a'', b'' и c'' (в таком порядке). Из тождества Якоби следует, что $[\hat{c}, [\hat{b}, \hat{a}]] = [\hat{a}, [\hat{b}, \hat{c}]] + [\hat{b}, [\hat{c}, \hat{a}]]$. Тогда по задаче 23 прямая c'' пересекает общий перпендикуляр к прямым a'' и b'', который мы обозначим через h. По задаче 9 угол между c'' и h — прямой. Поэтому h — общий перпендикуляр к a'', b'' и c'', и Теорема о высотах 'трехсторонника' доказана.

27. В решении данной задачи мы будем опираться на некоторые факты, доказанные в последующих задачах.
Требуемое утверждение достаточно доказать для окружностей на сфере (так как плоскость можно отобразить на сферу с помощью стереографической проекции). Будем обозначать вектор, соответствующий данной окружности, той же буквой, что и саму окружность, со значком вектора. Тогда по задаче 34 $\vec{c}' = \frac{d_B}{d_B - d_A}\vec{a} - \frac{d_A}{d_B - d_A}\vec{b}$, аналогичные формулы можно написать для $\vec{a}'$ и $\vec{b}'$. Тогда

$$\frac{d_C - d_B}{d_C d_B}\vec{a}' + \frac{d_A - d_C}{d_A d_C}\vec{b}' + \frac{d_B - d_A}{d_B d_A}\vec{c}' = 0.$$

Применяя лемму из решения задачи 34, получаем, что три окружности a, b и c проходят через одну точку.
Замечание.* В зависимости от того, образуют ли окружности 'выпуклый' или 'вогнутый' треугольник, эта теорема означает теорему о биссектрисах в сферической геометрии или геометрии Лобачевского соответственно.
28*. (Решение А. Мафусалова) Приведем план решения данной задачи, опуская технические детали. Решение основано на следующей элегантной теореме:

Теорема о высотах криволинейного треугольника. Пусть A, B и C — тройка попарно пересекающихся окружностей на плоскости. Через две точки пересечения окружностей A и B проведем окружность C', перпендикулярную окружности C. Аналогично определим окружности A' и B'. Тогда три окружности A', B' и C' принадлежат одному пучку. (В частности, если две из них пересекаются, то все три окружности проходят через одну точку.)

Можно проверить, что утверждение задачи 28 — это в точности теорема о высотах для криволинейного треугольника $a'b'c'$.

Доказательство теоремы. Пусть D — радикальный центр трех окружностей A, B и C. Обозначим через d степень точки D относительно данных окружностей. Рассмотрим сферу радиуса $\sqrt{|d|}/2$, касающуюся плоскости в точке D. Произведем стереографическую проекцию плоскости на данную сферу. Возможны три случая:
Случай $d < 0$. Тогда можно проверить, что окружности A, B и C перейдут в три большие окружности на нашей сфере. В этом случае наша теорема следует из теоремы о высотах сферического треугольника (задача 8).
Случай $d = 0$. Сделаем инверсию относительно точки D. Тогда наша теорема перейдет в теорему о высотах евклидова треугольника.
Случай $d > 0$. В этом случае можно проверить, что окружности A, B и C перейдут в три окружности на сфере с центрами на одной сферической прямой. Обозначим эту сферическую прямую через p. Рассмотрим сечение сферы плоскостью, содержащей p. Сделаем ортогональную проекцию нашей сферы на эту плоскость. Тогда три построенных окружности на сфере перейдут в три хорды окружности p. Обозначим эти хорды через KL, MN и PQ. Окружность C' при композиции нашей стереографической проекции и ортогональной проекции переходит в некоторую хорду $P'Q'$. Можно проверить, что прямая XY должна проходить через точку пересечения прямых KL и MN, а также точку пересечения касательных к окружности p в точках P и Q. Аналогично строятся хорды $K'L'$ и $M'N'$. Нам нужно показать, что прямые $K'L'$, $M'N'$ и $P'Q'$ пересекаются в одной точке. Это утверждение можно доказать, например, применяя теорему Чевы в тригопометрической форме. (На самом деле полученное утверждение есть теорема о высотах треугольника в геометрии Лобачевского, сформулированная в модели Клейна.)
29. Пусть A_1, B_1 и C_1 — *'центры'* окружностей, соответствующих векторам $\vec{A}$, $\vec{B}$ и $\vec{C}$ (то есть такие точки на сфере, что касательная плоскость в каждой точке параллельна плоскости соответствующей окружности). Тогда векторы $\overrightarrow{OA_1}$, $\overrightarrow{OB_1}$ $\overrightarrow{OC_1}$ параллельны векторам $\vec{A}$, $\vec{B}$ и $\vec{C}$. Рассуждая аналогично решению задачи 4, получаем, что точки A_1, B_1, C_1 и O лежат в одной плоскости, то есть, что 'центры' окружностей A, B и C лежат на одной сферической прямой.
30. Пусть A_1 — такая точка, что $\overrightarrow{OA_1} = \vec{A}$. Обозначим через d_A длину касательной, проведенной из точки A_1 к нашей сфере. Возьмем произвольную точку C на окружности A. По теореме Пифагора для треугольника OA_1C получаем $OA_1^2 = OC^2 + A_1C^2$, то есть $(\vec{A}, \vec{A}) = 1 + d_A^2$, что и требовалось.
31. Решение данной задачи получается из следующей более общей формулы для угла γ между окружностями A и B:

$$\cos\gamma = \frac{(A, B) - 1}{d_A d_B}.$$

Докажем эту формулу. Пусть C — одна из точек пересечения окружностей A и B. Обозначим $\vec{C} = \overrightarrow{OC}$. Заметим, что векторы $[\vec{A}, \vec{C}]$ и $[\vec{B}, \vec{C}]$ параллельны касательным к окружностям A и B соответственно, проведенным в точке C. Поэтому

$$\cos\gamma = \frac{([\vec{A}, \vec{C}], [\vec{B}, \vec{C}])}{|[\vec{A}, \vec{C}]| \cdot |\vec{B}, \vec{C}|}.$$

Приведем правую часть данного равенства к нужному виду. Легко убедиться, что $|[\vec{A}, \vec{C}]| = |\vec{A}| \cdot |\vec{C}| \cdot \sin\angle(\vec{A}, \vec{C}) = d_A$. Поэтому знаменатель данного выражения имеет вид $d_A d_B$. Теперь преобразуем числитель, пользуясь тождеством из второго решения задачи 15:

$$([\vec{A}, \vec{C}], [\vec{B}, \vec{C}]) = (\vec{A}, \vec{B})(\vec{C}, \vec{C}) - (\vec{A}, \vec{C})(\vec{B}, \vec{C}) = (\vec{A}, \vec{B}) - 1.$$

Подставляя в наше равенство выражения для числителя и знаменателя, получим требуемую формулу.
32. Предположим, что вектор $\vec{P} = \frac{[\vec{A}, \vec{B}] + [\vec{B}, \vec{C}] + [\vec{C}, \vec{A}]}{(\vec{A}, \vec{B}, \vec{C})}$ соответствует некоторой окружности P (такой окружности может и не существовать, в этом случае данный вектор не имеет геометрической интерпретации). Заметим, что

$$(\vec{P}, \vec{A}) = \frac{0 + (\vec{A}, [\vec{B}, \vec{C}]) + 0}{(\vec{A}, \vec{B}, \vec{C})} = 1.$$

Значит, по задаче 31, окружность P перпендикулярна окружности A. Аналогично, окружность P перпендикулярна окружностям B и C. Иными словами, P — *общий перпендикуляр* к трем окружностям A, B и C.
Замечание.* Окружности на сфере более естественно ставить в соответствие не вектор, а пару $(\vec{A}, h)$, где $\vec{A}$ — единичный вектор, перпендикулярный плоскости окружности, а h — расстояние от этой плоскости до точки O. Иными словами, каждой окружности ставится в соответствие некоторый кватернион. Будем обозначать его той же буквой, что и саму окружность. Будем считать, что все кватернионы, получающиеся из данного умножением на действительное число, соответствуют одной и той же окружности. Тогда, например, кватернион $AB - BA$ соответствует прямой, проходящей через центры окружностей A и B, а кватернион $ABC - ACB + BCA - BAC + CAB - CBA$ соответствует окружности, перпендикулярной трем окружностям A, B и C. А вот коммутатор четверки кватернионов $ABCD - ABDC + \dots$ всегда равен нулю, и это тождество позволяет получить, например, такую геометрическую теорему:

Утверждение. Пусть D — центр описанной окружности треугольника ABC, а A', B' и C' — центры описанных окружностей треугольников BCD, CAD и ABD соответственно. Тогда прямые AA', BB' и CC' пересекаются в одной точке.

33. Пусть $\vec{A}-\vec{B}+\vec{C}-\vec{D}=0$. Рассмотрим вектор $\vec{P}=\frac{[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]}{(\vec{A},\vec{B},\vec{C})}$. Предположим, что он соответствует некоторой окружности P. Тогда по задаче 32 P — общий перпендикуляр к A, B и C. Заметим, что

$$(\vec{P},\vec{D})=(\vec{P},\vec{A}-\vec{B}+\vec{C})=1-1+1=1,$$

то есть по задаче 31 окружность P перпендикулярна окружности D. Иными словами, четыре окружности A, B, C и D имеют общий перпендикуляр.

34. Докажем вначале одно вспомогательное утверждение:

Лемма. Если найдутся ненулевые числа x, y и z, такие что $x+y+z=0$ и $x\vec{A}+y\vec{B}+z\vec{C}=0$, то окружности A, B и C проходят через одну точку.

Доказательство. Условие данной леммы означает, что концы векторов $\vec{A}$, $\vec{B}$ и $\vec{C}$ лежат на одной прямой. Проведем плоскость π через эту прямую и точку O. Пусть ω — окружность, получающаяся в сечении сферы плоскостью π. Пусть H — полюс нашей прямой в плоскости π относительно окружности γ. Проведем через точку H прямую h, перпендикулярную плоскости π. Тогда можно проверить, что все три окружности A, B и C проходят через точку пересечения сферы и прямой h.

Перейдем к решению нашей задачи. Рассмотрим окружность C, соответствующую вектору $\vec{C}=\frac{d_B}{d_B-d_A}\vec{A}-\frac{d_A}{d_B-d_A}\vec{B}$. Тогда по нашей лемме окружность C проходит через обе точки пересечения окружностей A и B. Обозначим углы между парами окружностей B и C, C и A, A и B через α, β и γ соответственно. Тогда по формуле из решения задачи 31 имеем

$$\cos\beta=\frac{(\vec{A},\vec{C})-1}{d_Ad_C}=\frac{\frac{d_B}{d_B-d_A}(\vec{A},\vec{A})-\frac{d_A}{d_B-d_A}(\vec{A},\vec{B})-1}{d_Ad_C}=\frac{d_Ad_B(1-\cos\gamma)}{d_C(d_B-d_A)}.$$

Записывая аналогичное выражение для $\cos\alpha$, получаем, что $\cos\alpha=\cos\beta$. Иными словами, C — окружность, проходящая через обе точки пересечения окружностей A и B и делящая угол между этими окружностями пополам ('*биссектриса*').

Благодарности. Авторы благодарны В. Арнольду за важные замечания, В. Дремову, которому принадлежит множество ценных идей, реализованных в данном цикле задач, а также О. Карпенкову, И. Лосеву и членам жюри XVIII Летней конференции Турнира городов за полезные обсуждения.

[1] В.И. Арнольд, Теорема о высотах треугольника в геометрии Лобачевского как тождество Якоби в алгебре Ли квадратичных форм на симплектической плоскости, Математическое Просвещение, Третья Серия 9(2005), стр. 93–99.
[2] О.Я. Виро, Ю.В. Добротухина, Сплетения скрещивающихся прямых, Квант 3(1988), стр. 12–19.
[3] Ю.П. Соловьев, А.Б. Сосинский, Геометрия скользящих векторов, Квант 8(1985), стр. 9–17.
[4] J. Conant, R. Schneiderman, P. Teichner, Jacobi identities in low-dimensional topology, preprint (2006).

Theorem on altitudes and the Jacobi identity

S. Dorichenko and M. Skopenkov

The purpose of this problem set is to introduce a nice idea due to academician V. I. Arnold on a relationship between the theorem on altitudes of a triangle and the Jacobi identity [1].

First let us give an interesting application of the idea — a generalization of the theorem on altitudes to the case of the three-dimensional space:

Theorem on altitudes of a 'triside' [1]. Let a, b and c be pairwise non-parallel lines in 3-dimensional space. Let a', b' and c' be the three common perpendiculars to the pairs of lines b and c, c and a, a and b. Finally, let a'', b'' and c'' be the three common perpendiculars to the pairs of lines a and a', b and b', c and c' (it is given that the pairs are pairwise non-parallel). Then the lines a'', b'' and c'' have a common perpendicular (that is, there is a line crossing all of them and perpendicular to all of them).

0. Check that if the lines a, b and c belong to one plane then this theorem becomes to the theorem on altitudes of a triangle.

We start with an introduction to spherical geometry, in which the Arnold idea reveals most clearly.

Subject One. Spherical geometry and vector product.

Consider a unit sphere in three-dimensional space. By a *big circle* (*spherical line*) we mean a section of the sphere by a plane containing the center of the sphere. We identify the antipodes of the sphere.

To each point of the sphere we assign the vector looking from the center of the sphere to this point. By definition let us assume that all the vectors obtained from the given one by multiplication by a number (positive or negative) correspond to the same point.

To each spherical line we assign an arbitrary vector perpendicular to the plane containing the given line. Let us assume that all the vectors obtained from this one by multiplication by a number (positive or negative) correspond to the same line.

In what follows by a point we mean a point on the sphere, and by a line, a spherical line. The definitions of other natural objects of spherical geometry such as segment, triangle, perpendicular etc. are left as an exercise. The problems where we ask to find a geometrical sense of an object or an identity we do not suppose a unique answer, there can be a number of such interpretations.

We denote points by uppercase letters and lines by lowercase ones. We denote the vector corresponding to a point (line) by the same letter as the point (line) itself with the vector symbol.

If $\vec{A}$ and $\vec{B}$ are two vectors then by $[\vec{A}, \vec{B}]$ denote their vector product. Recall that *the vector product* of two vectors $\vec{A}$ and $\vec{B}$ is a vector perpendicular to both of the given vectors, with absolute value equal to the area of the parallelogram spanned by the vectors. The direction of this vector is given by the right hand law. By (A, B) denote the inner product of the vectors.

1. Let A and B be two points in the sphere. Prove that the vector $[\vec{A}, \vec{B}]$ corresponds to the spherical line passing through the points A and B.

2. Let a and b be two spherical lines. Prove that the vector $[\vec{a}, \vec{b}]$ corresponds to their intersection point.

3. Let A be a point and let b be a line. Prove that the vector $[\vec{A}, \vec{b}]$ (if nonzero) corresponds to the perpendicular dropped from the point A to the line b.

4. Let A, B and C be three points. What is the geometrical meaning of the condition $\vec{A} + \vec{B} + \vec{C} = 0$?

5. Let a, b and c be three lines. What is the meaning of the condition $\vec{a} + \vec{b} + \vec{c} = 0$?

6. Prove the identity $[\vec{A}, [\vec{B}, \vec{C}]] = \vec{B}(\vec{A}, \vec{C}) - \vec{C}(\vec{A}, \vec{B})$.

7. Prove *the Jacobi identity*:

$$[\vec{A}, [\vec{B}, \vec{C}]] + [\vec{B}, [\vec{C}, \vec{A}]] + [\vec{C}, [\vec{A}, \vec{B}]] = 0.$$

8. Let A, B and C be the vertices of a spherical triangle. What is the geometrical meaning of the Jacobi identity for the vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$?

[1] This statement was suggested as a problem in a selection competition in the Depatment of Mechanics and Mathematics of Moscow State University for this year's international student contest

9. Prove that in Theorem on the altitudes of a 'triside' formulated above the lines a'', b'' and c'' are parallel to one plane.
10. Let A and B be two points. Let $\vec{A}$ и $\vec{B}$ be the *unit* vectors facing to these points from the center of the sphere. To which point does the vector $\vec{A}+\vec{B}$ correspond?
11. What is the geometrical meaning of the identity $[\vec{A},\vec{B}+\vec{C}]+[\vec{B},\vec{C}+\vec{A}]+[\vec{C},\vec{A}+\vec{B}]=0$?
12. Let A, B and C be three points in the sphere. Let $\vec{A}$, $\vec{B}$ and $\vec{C}$ be the unit vectors pointing from the center of the sphere to A, B and C. To which point does the vector $[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]$ correspond?
13. Let a and b be two lines. Let $\vec{a}$ and $\vec{b}$ be the corresponding vectors such that $|\vec{a}|=|\vec{b}|=1$. What is the geometrical sense of the inner product $(\vec{a},\vec{b})$ and the absolute value of the vector product $|[\vec{a},\vec{b}]|$? Answer the similar question also for two points A and B; for a point A and a line b.
14. What is the geometric sense of the mixed product $([\vec{A},\vec{B}],\vec{C})$ and the identity $(\vec{A},[\vec{B},\vec{C}])=(\vec{B},[\vec{A},\vec{C}])=(\vec{C},[\vec{A},\vec{B}])$?
15. Prove that in spherical geometry there are no distinct similar triangles — the angles of a triangle uniquely define the sides of the triangle.
16. Try to find the geometrical meaning of more algebraic objects and identities. Obtain the proofs of the corresponding theorems of spherical geometry (or, vice versa, the initial algebraic identities, if the corresponding geometrical theorems are already known).
17*. By *a circle* we mean a section of the sphere by an arbitrary plane. A point in the sphere and a spherical line are specific cases of a circle. Investigate how our correspondence between points and lines can be extended to include arbitrary circles, and thus obtain new theorems of spherical geometry. (We shall discuss this subject more after the intermediate consideration of the problems.)

Spherical geometry is an example of a non-Euclidean geometry. It is based on the same axioms as the Euclidean geometry except the axiom on parallel lines. In this geometry many theorems not using the fifth axiom remain true.

18*. Try to prove the theorems of spherical geometry you obtained above (for example, in problems 8 and 11) directly, in a purely geometrical way, starting from the axioms of geometry except the axiom on parallel lines.

The next subject is devoted to the proof of Theorem on altitudes of a 'triside'.

Subject Two. Lines in space and bivectors.

Fix a point O in the three-dimensional Euclidean space. To each line in space assign an ordered pair of vectors $(u;v)$ as follows: take two points A and B on the line and put $u=AB$, $v=[OA,OB]$. In the sequel an ordered pair of vectors is called *a bivector*.

19. Check that, up to multiplication of both vectors u and v by the same number, the constructed bivector does not depend on the choice of points A and B on the line.
20. Check that the initial line is uniquely defined by the constructed bivector.
21. Check that in this way one can obtain any bivector $(u;v)$ such that $u\neq 0$ and $v\perp u$.

Let us extend our correspondence between the lines and the bivectors to bivectors $(u;v)$ such that $v\not\perp u$. Denote by $\operatorname{pr} v$ the projection of the vector v to the plane perpendicular to the vector u. By definition we assume that the bivector $(u;v)$ corresponds to the same line as the bivector $(u;\operatorname{pr} v)$. We shall denote a bivector and the corresponding line by the same letter, and to distinguish them, we are going to put the symbol 'hat' ('$\hat{\ }$') above the notation of the bivector.

Let $\hat{a}=(u;v)$ and $\hat{b}=(u';v')$ be two bivectors. Define their *sum* componentwise: $\hat{a}+\hat{b}=(u+u';v+v')$. Define their *product* by the formula

$$[\hat{a},\hat{b}]=([u,u'];[u,v']+[v,u']).$$

A sum and a product of two bivectors are again bivectors.

Remark*. This definition comes from the formula for the commutator in Lie algebra of the transformation group of the 3-space. One of the geometrical interpretations of our bivectors are *slide vectors* [3].

22. Let the lines a and b be non-parallel. Prove that the bivector $[\hat{a},\hat{b}]$ corresponds to the common perpendicular to the lines a and b.

23. Let the lines a and b be non-parallel. Prove that the line corresponding to the bivector $\hat{a}+\hat{b}$ crosses the common perpendicular to the lines a and b.
24. Prove that the multiplication of bivectors satisfies the Jacobi identity.
25. Prove Theorem on altitudes of a 'triside'.
26. Try to obtain more stereometrical theorems this way.

Theorem on altitudes and the Jacobi identity

A. Zaslavskiy and M. Skopenkov

In this subject we shall show how to use the Arnold idea to obtain theorems not only on points and lines but also on circles. Let us give one example:

Theorem on bisectors of a curved triangle. Let a, b and c be three pairwise intersecting circles in the plane. Draw a circle c' passing through both intersection points of the circles a and b such that the angle between c' and a is equal to the angle between c' and b. (There are two circles satisfying these conditions, choose the one having common points with the interior of the intersection of the areas bounded by a and b.) Define the circles a' and b' analogously. Then the three circles a', b' and c' have a common point.

27. Prove the theorem on bisectors of a curved triangle.
28*. Fix a circle I in the plane. By a *common perpendicular* to a pair of circles a and b we mean a circle perpendicular to all the three circles a, b and I. Prove that Theorem on altitudes of a 'triside' remains true, if one replaces the word 'line' (in space) by 'circle' (in the plane) everywhere in the statement.

Let us start again with spherical geometry, in which the Arnold idea is revealed most clearly. The problems of this subject can be solved independently from the others.

Subject Three. Summing of circles.

By a *circle* we mean the section of the sphere by an arbitrary plane, not necessarily passing trough the center of the sphere. To each circle assign a vector perpendicular to this plane, looking toward the plane and having length $1/h$, where h is the distance between the plane and the center of the sphere.

In what follows uppercase letters always denote circles. The vector corresponding to a circle is denoted by the same letter as the circle, and, to distinguish them, with the symbol of the vector. For a vector $\vec{A}$ denote by d_A the length of a tangent to our sphere, dropped from the end of the vector $\vec{A}$.

29. What is the geometric sense of the condition $\vec{A}+\vec{B}+\vec{C}=0$?
30. Prove the formula $(\vec{A},\vec{A})=1+d_A^2$.
31. What is the geometric sence of the condition $(\vec{A},\vec{B})=1$?
32. To what circle does the vector

$$\frac{[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]}{(\vec{A},\vec{B},\vec{C})}$$

correspond?
33. What is the geometric sense of the condition $\vec{A}-\vec{B}+\vec{C}-\vec{D}=0$?
34. To what circle does the vector

$$\frac{d_B}{d_B-d_A}\vec{A}-\frac{d_A}{d_B-d_A}\vec{B}$$

correspond?
35. Use these facts and algebraic identities to obtain more geometric theorems!

Appendix. Presentation of circles by points in space.

Take a circle in the plane be given by the equation $x^2+y^2+ax+by+c=0$. To this circle we assign the point in space having the coordinates (a,b,c).

36. For which numbers a, b and c there is a circle corresponding to the point (a,b,c)?
37. Consider a circle corresponding to the point P in space. What is the set of all the points corresponding to the circles orthogonal to our one?
38. What pairs of points correspond to a pair of
a)intersecting;
b)non-intersecting;
c)tangent circles?
39. Prove the following theorem (V. Yu. Protasov).
Take three circles such that the first is inside the second and the second is inside the third. Consider chains of circles tangent to both the first and the third one and such that one of the intersection points of any two consequent circles in the chain belong to the second given circle. If this chain is closed for some starting circle, then it is closed for any choice of starting circle.
40. Take a sphere, a plane tangent to the sphere and four points A, B, C and D in the plane. Let D' be the intersection point of three planes passing through the lines AB, BC and CA respectively and tangent to the sphere. Define the points A', B' and C' analogously. Prove that the points A', B', C' and D' lie in one plane tangent to the sphere.

Theorem on altitudes and the Jacobi identity

A. Zaslavskiy and M. Skopenkov

Solutions of the problems suggested before the intermediate finish.

1. Denote by O the center of our sphere. Let γ be the plane passing through the points O, A and B. Let c be the spherical line obtained as the intersection of the sphere and γ. Then c is precisely the spherical line passing through the points A and B.
On the other hand, by our definition, the vectors $\vec{A} = \overrightarrow{OA}$ and $\vec{B} = \overrightarrow{OB}$ correspond to the points A and B, respectively, and the vector $[\vec{A}, \vec{B}]$ is orthogonal to both $\overrightarrow{OA}$ and $\overrightarrow{OB}$. Therefore the vector $[\vec{A}, \vec{B}]$ is also orthogonal to the plane γ. Finally, by our definition of correspondence between vectors and spherical lines, we obtain that the vector $[\vec{A}, \vec{B}]$ corresponds to the spherical line c.
2. Recall that the intersection of two spherical lines is a pair of two diametrically opposite points on the sphere. Let C be one of the intersection points of the spherical lines a and b. We shall now show that the vector $\overrightarrow{OC}$ is parallel to the vector $[\vec{a}, \vec{b}]$. Let α be the plane containing the point O and the spherical line a. By definition of the correspondence between spherical lines and vectors, the vector a is orthogonal to the plane α. The segment OC lies in the plane α, therefore $\overrightarrow{OC} \perp \vec{a}$. Similarly, $\overrightarrow{OC} \perp \vec{b}$. Therefore the vectors $\overrightarrow{OC}$ and $[\vec{a}, \vec{b}]$ are collinear, whence the vector $[\vec{a}, \vec{b}]$ corresponds to the point C (recall here that all vectors collinear to $\overrightarrow{OC}$ correspond to the point C).
3. Let c be the line passing through the point A and orthogonal to the line b. We shall say that c is the *perpendicular* dropped from A onto b. It suffices to prove that the vector $\vec{c}$ corresponding to the line c is orthogonal to both the vector $\vec{A}$ and the vector $\vec{b}$ (indeed, in this case the vector $\vec{c}$ is collinear to the vector $[\vec{A}, \vec{b}]$ and we know that collinear vectors correspond to the same line).
First we show that $\vec{c} \perp \vec{A}$. Consider the plane γ passing through the point O and containing the line c. Since the point A belongs to c, the point A also belongs to γ. Therefore $\vec{c}$ is orthogonal to $\overrightarrow{OA} = \vec{A}$.
Now we show that $\vec{c} \perp \vec{b}$. Denote by β the plane passing through the point O and containing the line b. Since spherical lines b and c are orthogonal, the planes β and γ are also orthogonal. But then any vector orthogonal to β must also be orthogonal to any vector orthogonal to γ, and our proof is complete.
4. *Answer:* the points A, B and C lie on the same line.
Solution. The condition $A + B + C = 0$ implies that the vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$ are parallel to some plane π. We may assume here that the plane π passes through the point O and therefore defines a spherical line p. Then the points A, B, and C lie on the spherical line p.
5. *Answer:* the three lines a, b and c all pass through the same point.
Solution. The condition $\vec{a} + \vec{b} + \vec{c} = 0$ implies that the vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ are parallel to some plane π. Take the point P on our sphere such that $\overrightarrow{OP} \perp \pi$. Then the lines a, b and c all pass through P. Indeed, since $\overrightarrow{OP}$ is orthogonal to π, it is also orthogonal to the vector $\vec{a}$, which, by Problem 3, implies that the point P lies on the line a. Similarly, the point P lies on the lines b and c.
6. Consider the identity $[\vec{A}, [\vec{B}, \vec{C}]] = \vec{B}(A, C) - \vec{C}(A, B)$. If to the vector $\vec{A}$ one adds an arbitrary vector collinear to $[\vec{B}, \vec{C}]$ then neither the left nor the right part of the identity changes. Therefore it suffices to consider the case when the vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$ are all parallel to the same plane.
Now observe that neither part of our identity changes if to the vector $\vec{B}$ one adds a vector collinear to $\vec{C}$. Since $\vec{A}$, $\vec{B}$ and $\vec{C}$ are parallel to the same plane, we may assume that either $\vec{B}$ or $\vec{C}$ is collinear to $\vec{A}$.
For definiteness, assume that $\vec{B} \parallel \vec{A}$. Adding to the vector $\vec{C}$ a vector collinear to $\vec{B}$ we may assume that $\vec{C} \perp \vec{B}$. But if $\vec{B} \parallel \vec{A}$ and $\vec{B} \perp \vec{C}$, then it is easy to check directly that both parts of our identity are equal to $|\vec{A}| \cdot |\vec{B}| \cdot \vec{C}$.
Remark. Our identity may also be proven by using linearity of both its parts.
7. The Jacobi identity is obtained by cyclically permuting the variables in the identity of Problem 6 and summing the resulting identities.

Theorem on altitudes and the Jacobi identity

A. Zaslavskiy and M. Skopenkov

Solutions.

First let us give a table containing the answers to all the problems:

Algebraic object	Geometric sense
$\vec{A}$	a point A
$\vec{a}$	a spherical line a
$[\vec{A},\vec{B}]$	the line passing through A and B
$[\vec{a},\vec{b}]$	the intersection point of a and b
$[\vec{A},\vec{b}]$	the perpendicular dropped from the point A to the line b
$\vec{A}+\vec{B}$	the middle of the segment AB
$(\vec{a},\vec{b})$	cosine of the angle between a and b
$\lvert[\vec{a},\vec{b}]\rvert$	sine of the angle between a and b
$(\vec{A},\vec{B})$	cosine of the length of the segment AB (i. e. the arc of a big circle between A and B)
$\lvert[\vec{A},\vec{B}]\rvert$	sine of the length of the segment AB
$(\vec{A},\vec{b})$	sine of the distance between A and b
$\lvert[\vec{A},\vec{b}]\rvert$	cosine of the distance between A and b
$\vec{A}+\vec{B}+\vec{C}=0$	1) the points A, B and C are collinear 2) the centers of the circles A, B and C are collinear
$\vec{a}+\vec{b}+\vec{c}=0$	the lines a, b and c have a common point
$[\vec{A},[\vec{B},\vec{C}]]+[\vec{B},[\vec{C},\vec{A}]]+[\vec{C},[\vec{A},\vec{B}]]=0$	the altitudes of a triangle ABC have a common point
$[\vec{A},\vec{B}+\vec{C}]+[\vec{B},\vec{C}+\vec{A}]+[\vec{C},\vec{A}+\vec{B}]=0$	the medians of a triangle ABC have a common point
$[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]$	the center of the circumscribed circle of the triangle ABC
$(\vec{A},\vec{B},\vec{C})$	sine of the length of the segment AB $\cdot$ sine of the distance between C and AB
$(\vec{A},[\vec{B},\vec{C}])=(\vec{B},[\vec{C},\vec{A}])=(\vec{C},[\vec{A},\vec{B}])$	The sines theorem for the triangle ABC
$(\vec{A},\vec{B})=1$	the circles A and B are orthogonal
$\frac{[\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]}{(\vec{A},\vec{B},\vec{C})}$	the circle orthogonal to the circles A, B and C
$\vec{A}-\vec{B}+\vec{C}-\vec{D}=0$	there exists a circle orthogonal to the circles A, B, C and D
$\frac{d_B}{d_B-d_A}\vec{A}-\frac{d_A}{d_B-d_A}\vec{B}$	bisector of the circles A and B

8. Let $\vec{A}$, $\vec{B}$ and $\vec{C}$ be the vectors corresponding to the vertices of the spherical triangle ABC. By Problem 1 the vector $[\vec{A},\vec{B}]$ corresponds to the spherical line passing through the points A and B. Then by Problem 3 the vector $[\vec{A},[\vec{B},\vec{C}]]$ corresponds to the perpendicular dropped from the point A to the line BC, that is, it corresponds to the line containing the altitude h_A of the triangle ABC. Analogously the vectors $[\vec{B},[\vec{C},\vec{A}]]$ and $[\vec{C},[\vec{A},\vec{B}]]$ correspond to the lines containing two other altitudes of the triangle ABC. So by Problem 5 the Jacobi identity means that the three constructed lines have a common point, i. e. we get the theorem on altitudes of a spherical triangle.

9. Denote by $\vec{a}$ *the direction vector* of the line a, that is, a vector parallel to the line a. Analogously, let $\vec{b}$ and $\vec{c}$ be the direction vectors of the lines b and c. Then obviously the vector $[\vec{b},\vec{c}]$ is parallel to the common perpendicular to the lines b and c, that is, to the line a'. Thus the vector $[\vec{a},[\vec{b},\vec{c}]]$ is parallel to the common perpendicular to the lines a and a', that is, to the line a''. Analogously, the vectors $[\vec{b},[\vec{c},\vec{a}]]$ and $[\vec{c},[\vec{a},\vec{b}]]$ are parallel to the lines b'' and c'' respectively. By the Jacobi identity the last three vectors are parallel to one plane. Thus the three lines a'', b'' и c'' are parallel to one plane.

10. By symmetry the vector $\vec{A}+\vec{B}$ corresponds to the middle of the arc of the big circle passing through the points A and B. It is natural to say that this point is *the middle* of the spherical segment with ends A and B.

Remark. In fact a pair of points A and B on the sphere determines two distinct segments. If *one* of the two vectors $\vec{A}$ and $\vec{B}$ is replaced by the opposite vector, then their sum corresponds to the middle of the other segment with ends A and B.

11. Let $\vec{A}$, $\vec{B}$ and $\vec{C}$ be unit vectors pointing to A, B and C respectively from the center of the sphere. Then by Problem 10 the vector $\vec{B}+\vec{C}$ corresponds to the middle of the segment BC. Then by Problem 1 the vector $[\vec{A},\vec{B}+\vec{C}]$ corresponds to *the median* of the triangle ABC. Then by Problem 5 the identity $[\vec{A},\vec{B}+\vec{C}]+[\vec{B},\vec{C}+\vec{A}]+[\vec{C},\vec{A}+\vec{B}]=0$ means that the medians of the triangle ABC have a common point.

The theorem on medians can also be proved by consideration of the vector $\vec{A}+\vec{B}+\vec{C}$.

Remark. (A. Mafusalov) Unlike Euclidean geometry, the theorem on medians of a spherical triangle has an external analog: two 'external' medians m_A, m_B and the opposite 'internal' median m_C of a triangle have a common point.

12. First let us prove that the vectors $\vec{A}-\vec{B}$, $\vec{B}-\vec{C}$ and $\vec{C}-\vec{A}$ correspond to the *middle perpendiculars* to the sides of a spherical triangle ABC. Indeed, since

$$(\vec{A}-\vec{B},\vec{A}+\vec{B}) = (\vec{A},\vec{A}) - (\vec{B},\vec{B}) = 1-1 = 0,$$

it follows that the line corresponding to $\vec{A}-\vec{B}$ passes through the middle of AB. And since

$$(\vec{A}-\vec{B},[\vec{A},\vec{B}]) = (\vec{A},[\vec{A},\vec{B}]) - (\vec{B},[\vec{A},\vec{B}]) = 0,$$

it follows that this line is orthogonal to the segment AB. Thus the vector $\vec{A}-\vec{B}$ corresponds to the middle perpendicular to the segment AB.
Denote by $\vec{V} = [\vec{A},\vec{B}]+[\vec{B},\vec{C}]+[\vec{C},\vec{A}]$. Let V be the point on the sphere corresponding to this vector. Let us check that the point V belongs to all the three constructed middle perpendiculars. Indeed, since

$$(\vec{V},\vec{A}-\vec{B}) = ([\vec{B},\vec{C}],\vec{A}) - ([\vec{C},\vec{A}],\vec{B}) = 0,$$

it follows that V belongs to the line corresponding to the vector $\vec{A}-\vec{B}$. Analogously, V belongs to two other middle perpendiculars.
Remark. The identity $(\vec{A}-\vec{B})+(\vec{B}-\vec{C})+(\vec{C}-\vec{A}) = 0$ and Problem 5 also imply that the three middle perpendiculars to the sides of a triangle have a common point.
13. The solution follows directly from the formulas $(\vec{a},\vec{b}) = |\vec{a}|\cdot|\vec{b}|\cdot\cos\gamma$ and $|[\vec{a},\vec{b}]| = |\vec{a}|\cdot|\vec{b}|\cdot\sin\gamma$, where γ is the angle between the vectors $\vec{a}$ and $\vec{b}$.
14. Assume that the lengths of the vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$ are equal to 1. Then by Problem 13 we obtain $|([\vec{A},\vec{B}],\vec{C})| = \sin AB\sin h_C$, where h_C is the length of the altitude of the triangle ABC passing through the vertex C. The identity $(\vec{A},[\vec{B},\vec{C}]) = (\vec{B},[\vec{C},\vec{A}]) = (\vec{C},[\vec{A},\vec{B}])$ means that $\sin AB\sin h_C = \sin BC\sin h_A = \sin CA\sin h_B$ (1). Writing an analogous identity for the unit vectors corresponding to the lines containing the sides of the triangle ABC, we get $\sin\angle C\sin h_C = \sin\angle A\sin h_A = \sin\angle B\sin h_B$ (2). Dividing (1) by (2), we get spherical *sine theorem*:

$$\frac{\sin AB}{\sin\angle C} = \frac{\sin BC}{\sin\angle A} = \frac{\sin CA}{\sin\angle B}.$$

Remark. Mixed product of three vectors has a clear meaning — it is the volume of the parallelepiped spanned by these three vectors, with a sign. We do not know such a natural interpretation of this number in terms of spherical geometry.
15. *First solution.* It is proved analogously to the case Euclidean geometry that *the sides* of a spherical triangle uniquely define the angles of the triangle. Now it remains to consider the triangle dual to the initial one (that is, a triangle such that the vectors pointing to its vertices are orthogonal to the planes containing the sides of the initial triangle).
*Second solution.*Analogously to the solution of Problem 6 one can prove the identity $([\vec{a},\vec{c}],[\vec{b},\vec{c}]) = (\vec{a},\vec{b})(\vec{c},\vec{c}) - (\vec{a},\vec{c})(\vec{b},\vec{c})$. Applying the results of Problem 13, we obtain the formula (*the spherical cosine theorem*)

$$\sin\angle A\sin\angle B\cos AB = \cos\angle C - \cos\angle A\cos\angle B.$$

This implies directly that the angles of a spherical triangle uniquely determine the sides of the triangle.
Remark.* The obtained results remain true also in the Lobachevskiy (hyperbolic) geometry, because it can be transformed to spherical by multiplication of all coordinates by i. Then the words 'the lines have a common point' in all the formulations above should be replaced by 'the lines belong to one sheaf'. Formally, our above proofs remain valid in this case, if one takes the unit sphere in the quasi-euclidean space and puts $[\vec{A},\vec{B}] := *(\vec{A}\wedge\vec{B})$ [1].

19. It is easy to check that the directions of both vectors u and v do not depend on the choice of points A and B: the vector u is parallel to our plane, and the vector v is orthogonal to the plane, passing through our line and the point O. The length of the vector u is equal to AB, and the length of the vector v is equal to two areas of the triangle ABO, that is, to $AB\cdot h$, where h is the distance between O and our line. Thus for another choice of the points A and B the vectors u and v are multiplied by the same number.
20. Indeed, our line should lie in the plane passing through the point O and orthogonal to the vector v. Our line should also be parallel to the vector v, and the distance between our line and the point O is equal to $|v|/|u|$. The vector $[u,v]$ points, in which 'direction' with respect to the point O our line lies. Thus our line is uniquely defined by the pair of vectors u and v.
21. The construction of the desired line, starting from the vectors $\vec{u}$ and $\vec{v}$, is in fact given in the solution of the previous problem.
22. (This is one of the most hard problems of this set) Define *the inner product* of bivectors $\hat{a} = (u,v)$ and $\hat{b} = (u',v')$ by the formula

$$(\hat{a},\hat{b}) = ((u,v);(u,v')+(u',v)).$$

The inner product of two bivectors is a *pair* of numbers. It is easy to check that $(\hat{a},[\hat{a},\hat{b}]) = (\hat{b},[\hat{a},\hat{b}]) = (0;0)$. So our problem is a direct corollary of the following lemma:

Lemma. If $(\hat{a},\hat{b}) = (0;0)$, then the lines a and b intersect, and the angle between them is right.

Proof of the lemma. The lines a and b are orthogonal because the first component of the inner product of $\hat{a}$ and $\hat{b}$ is zero: $(u,u') = 0$. It remains to shows that if $(\hat{a},\hat{b}) = (0;0)$, then the lines a and b intersect.
First consider the case when $u\perp v$ and $u'\perp v'$.
Let us check that the inner product does not depend on the choice of the point O. Assume that in the beginning we fix another point O_1. Let $\hat{a}_1 = (u_1,v_1)$, $\hat{b}_1 = (u_1',u_1')$ be the bivectors obtained from the lines a and b by our construction, if we

assume that the fixed point is O_1. Then one can check directly that $u_1 = u$, $v_1 = v + [u, \overrightarrow{OO_1}]$ и $u'_1 = u'$, $v'_1 = v' + [u', \overrightarrow{OO_1}]$. So

$$(\hat{a}_1, \hat{b}_1) = ((u, u'); (u, v' + [u', \overrightarrow{OO_1}]) + (v + [u, \overrightarrow{OO_1}], u')) = ((u, u'); (u, v') + (v, u')) = (\hat{a}, \hat{b}),$$

because $(u, [u', \overrightarrow{OO_1}]) + ([u, \overrightarrow{OO_1}], u') = 0$. Thus indeed the inner product of bivectors does not depend on the choice of the point O.

Therefore we may assume without loss of generality that the point O belongs to the line a. Then $v = 0$. Thus the condition $(\hat{a}, \hat{b}) = (0; 0)$ means that the vector u is orthogonal to the vector v'. If $v' = 0$, then we are done, because the point O is a common point for a and b in this case. If $v' \neq 0$, then both lines lie in the plane passing through the point O orthogonal to the vector v', and hence these two lines intersect each other. Thus the case $u \perp v$, $u' \perp v'$ has been considered.

Now consider the case when not necessarily $u \perp v$ and $u' \perp v'$. (This case is important for what follows, because even if this condition is satisfied for the bivectors $\hat{a}$ and $\hat{b}$, it may not be satisfied for the bivector $[\hat{a}, \hat{b}]$.)

In this case, by definition, the bivector $\hat{a} = (u, v)$ corresponds to the same line as the bivector $(u, pr\, v)$. Denote the last bivector by $\hat{a}_\perp$. Notice that $pr\, v = v + \alpha u$ for some number α. Define analogously the bivector $b_\perp$ and the number β. Since $(\hat{a}, \hat{b}) = (0; 0)$, it follows that $(\hat{a}_\perp, \hat{b}_\perp) = ((u, u'); (u, v') + \beta(u, u') + (v, u') + \alpha(u, u')) = (\hat{a}, \hat{b}) = (0; 0)$, because $(u, u') = 0$. Then the second case of the lemma is reduced to the first one, and we are done.

*Remark**. The definition of the product of bivectors comes from the formula for the commutator in the Lie algebra of the movements group of the 3-space. *Sliding vector* is another geometric interpretation of our bivectors [3].

23. Notice that $(\hat{a}+\hat{b}, [\hat{a}, \hat{b}]) = (\hat{a}, [\hat{a}, \hat{b}]) + (\hat{b}, [\hat{a}, \hat{b}]) = 0$. Then by lemma from the solution of Problem 22 the line corresponding to the vector $\hat{a} + \hat{b}$ intersects the line corresponding to the vector $[\hat{a}, \hat{b}]$. And by Problem 22 the last line is the common perpendicular to the lines a and b.

24. The Jacobi identity for the product of bivectors follows directly from the Jacobi identity for the product of vectors.

25. Applying Problem 22 several times, we obtain that the bivectors $[\hat{a}, [\hat{b}, \hat{c}]]$, $[\hat{b}, [\hat{c}, \hat{a}]]$ and $[\hat{c}, [\hat{b}, \hat{a}]]$ correspond to the lines a'', b'' and c'' respectively. By the Jacobi identity it follows that $[\hat{c}, [\hat{b}, \hat{a}]] = [\hat{a}, [\hat{b}, \hat{c}]] + [\hat{b}, [\hat{c}, \hat{a}]]$. Then by Problem 23 the line c'' intersects the common perpendicular to the lines a'' and b''. Denote this common perpendicular by h. By Problem 9 the angle between c'' and h is right. So h is a common perpendicular for a'', b'' and c'', and the Theorem on altitudes of a triangle follows.

27. In this solution we will use some facts which are proved in the further solutions.

It is sufficient to prove the statement of the problem for circles on a sphere (since a plane can be mapped into a sphere by a stereographic projection). For some circle, we will denote the corresponding vector by the same letter with the arrow over it (e.g. the vector corresponding to a circle a is denoted by $\vec{a}$). By problem 34, $\vec{c}' = \frac{d_B}{d_B - d_A}\vec{a} - \frac{d_A}{d_B - d_A}\vec{b}$; one can write down the analogous formulae for $\vec{a}'$ and $\vec{b}'$. Then we have

$$\frac{d_C - d_B}{d_C d_B}\vec{a}' + \frac{d_A - d_C}{d_A d_C}\vec{b}' + \frac{d_B - d_A}{d_B d_A}\vec{c}' = 0.$$

Applying the lemma from the solution of problem 34, we obtain that the circles a, b, and c have a common point.

*Note**. Three circles in this statement can form a "convex" or an "concave" triangle. According to this, our theorem is equivalent to the theorem on bisectors in spherical geometry or in Lobachevskiy geometry.

28*. (Solution by A. Mafusalov) We will present a plan of the solution, omitting some technical details. It is based on the following elegant theorem.

Theorem on the altitudes of a curved triangle. Let A, B, C be pairwise intersecting circles. Let C' be a circle perpendicular to C and passing through two points of intersection of A and B. The circles A' and B' are defined analogously. Then three circles A', B' and C' belong to one sheaf. (In particular, if two of them intersect, then the third passes through their common point(s)).

One can check that the statement of problem 28 is essentially the theorem on altitudes for the curved triangle $a'b'c'$.

Proof of the theorem. Let D be the radical center of A, B, and C. Denote by d the degree of the point D with respect to our circles. Consider the sphere of radius $\sqrt{|d|}/2$ touching the initial plane in D. Let us perform the stereographic projection of the plane into this sphere. Three cases are possible.

(i) $d < 0$. One can check that A, B and C are mapped into three big circles on the sphere. In this case our theorem follows from the theorem on altitudes of a spherical triangle (problem 8).

(ii) $d = 0$. Let us perform an inversion by a point D. Then our theorem is essentially the theorem on altitudes of euclidean triangle.

(iii) $d > 0$. One may check that the circles A, B and C map into three circles on the sphere, their centers lying on one spherical line. Denote this line by p. Consider a section of the sphere by the plane containing p, and perform an orthogonal projection from the sphere into this plane. Three circles constructed will map into three chords of p; denote these chords by KL, MN, PQ. Applying consequently our stereographic projection and orthogonal projection to C', we obtain some chord $P'Q'$. One may check that the line XY should pass through the intersection point of KL and MN, as well as through that of the lines tangent to p at P and Q. The chords $K'L'$ and $M'N'$ are defined similarly. We need to prove that the lines $K'L'$, $M'N'$, and $P'Q'$ are concurrent. One can prove this statement using the Ceva theorem in a trigopometric form. (This statement is in fact the theorem on altitudes in Lobachevskiy geometry, being considered in the Klein model.)

29. Let A_1, B_1, and C_1 be the *'centers'* of circles, corresponding to the vectors $\vec{A}$, $\vec{B}$, and $\vec{C}$ (i.e. A_1 is the point of the sphere such that the plane that touches the sphere in this point is parallel to the circle A, and so on). Then the vectors $\overrightarrow{OA_1}$, $\overrightarrow{OB_1}$, and $\overrightarrow{OC_1}$ are parallel to $\vec{A}$, $\vec{B}$, and $\vec{C}$ respectively. Arguments similar to those in the solution of problem 4 show that the points A_1, B_1, C_1, and O lie in one plane, i.e. the 'centers' of the circles A, B, and C lie on one spherical line.

30. Let A_1 be the point such that $\overrightarrow{OA_1} = \vec{A}$. Denote by d_A the length of a tangent from A_1 to our sphere. Consider an arbitrary point C on the circle A. By the Pythagorean theorem for the triangle OA_1C, we obtain $OA_1^2 = OC^2 + A_1C^2$, i.e. $(\vec{A}, \vec{A}) = 1 + d_A^2$, QED.
31. This solution is obtained from the following (more general) formula for the angle γ between A and B:

$$\cos\gamma = \frac{(A, B) - 1}{d_A d_B}.$$

Let us prove this formula. Let C be one of the intersection points of A and B. Let $\vec{C} = \overrightarrow{OC}$. Note that the vectors $[\vec{A}, \vec{C}]$ and $[\vec{B}, \vec{C}]$ are parallel to the lines passing through C and tangent to the circles A and B respectively. Hence

$$\cos\gamma = \frac{([\vec{A}, \vec{C}], [\vec{B}, \vec{C}])}{|[\vec{A}, \vec{C}]| \cdot |\vec{B}, \vec{C}|}.$$

Let us show that the right hand part coincides with the required expression. It is easy to check that $|[\vec{A}, \vec{C}]| = |\vec{A}| \cdot |\vec{C}| \cdot \sin\angle(\vec{A}, \vec{C}) = d_A$. Hence the denominator of the right hand part is equal to $d_A d_B$. Now we will rearrange the numerator using the identity from the second solution of problem 15. We have

$$([\vec{A}, \vec{C}], [\vec{B}, \vec{C}]) = (\vec{A}, \vec{B})(\vec{C}, \vec{C}) - (\vec{A}, \vec{C})(\vec{B}, \vec{C}) = (\vec{A}, \vec{B}) - 1.$$

Substituting the expressions for the numerator and the denominator, we obtain the desired formula.
32. Assume that the vector $\vec{P} = \frac{[\vec{A}, \vec{B}] + [\vec{B}, \vec{C}] + [\vec{C}, \vec{A}]}{(\vec{A}, \vec{B}, \vec{C})}$ corresponds to a circle P (such a circle may not exist, in which case the given vector has no geometric interpretation). Notice that

$$(\vec{P}, \vec{A}) = \frac{0 + (\vec{A}, [\vec{B}, \vec{C}]) + 0}{(\vec{A}, \vec{B}, \vec{C})} = 1.$$

Then by Problem 31, the circle P is orthogonal to the circle A. Analogously, the circle P is orthogonal to the circles B and C. In other words, P is *the common perpendicular* to the circles A, B and C.
Remark.* It is more natural to each circle to assign a pair $(\vec{A}, h)$, where $\vec{A}$ is the unit vector orthogonal to the plane containing the circle, and h is the distance between the plane and the point O. In other words, to each circle we assign a quaternion. Denote this quaternion by the same letter as the circle itself. We assume that all the quaternions, obtained from the given one by the multiplication by a real number, correspond to the same circle. Then, for example, the quaternion $AB - BA$ corresponds to the line passing through the centers of the circles A and B, and the quaternion $ABC - ACB + BCA - BAC + CAB - CBA$ corresponds to the circle orthogonal to three circles A, B and C. And the commutator of a quadruple of quaternions $ABCD - ABDC + ...$ always equals to zero. This identity implies, for example, the following geometrical theorem:

Proposition. Let D be the center of the circumscribed circle of the triangle ABC, and let A', B' and C' be the centers of circumscribed circles of the triangles BCD, CAD and ABD respectively. Then the lines AA', BB' and CC' have a common point.

33. Assume that $\vec{A} - \vec{B} + \vec{C} - \vec{D} = 0$. Consider the vector $\vec{P} = \frac{[\vec{A}, \vec{B}] + [\vec{B}, \vec{C}] + [\vec{C}, \vec{A}]}{(\vec{A}, \vec{B}, \vec{C})}$. Assume that it corresponds to some circle P. Then by Problem 32 the circle P is the common perpendicular to A, B and C. Notice that

$$(\vec{P}, \vec{D}) = (\vec{P}, \vec{A} - \vec{B} + \vec{C}) = 1 - 1 + 1 = 1,$$

hence by Problem 31 the circle P is orthogonal to the circle D. In other words, the four circles A, B, C and D have a common perpendicular.
34. First let us prove a Lemma:

Lemma. If there are nonzero numbers x, y and z, such that $x + y + z = 0$ and $x\vec{A} + y\vec{B} + z\vec{C} = 0$, then the circles A, B and C have a common point.

Proof of the lemma. The assumption of the lemma means that the ends of the vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$ lie on one plane. Draw a plane π passing through this line and the point O. Let ω be the circle, which is the intersection of the sphere and the plane π. Let H be the pole of our line in the plane π with respect to the circle γ. Draw a line h passing through the point H and orthogonal to the plane π. Then one can check that all the circles A, B and C pass through the intersection point of the sphere and the line h.

Now let us solve our problem. Consider a circle C corresponding to the vector $\vec{C} = \frac{d_B}{d_B - d_A}\vec{A} - \frac{d_A}{d_B - d_A}\vec{B}$. Then by our lemma the circle C passes through both intersection points of the circles A and B. Denote by α, β and γ the angles between the pairs of circles B and C, C and A, A and B respectively. Then by the formula from the solution of the problem 31 we have

$$\cos\beta = \frac{(\vec{A}, \vec{C}) - 1}{d_A d_C} = \frac{\frac{d_B}{d_B - d_A}(\vec{A}, \vec{A}) - \frac{d_A}{d_B - d_A}(\vec{A}, \vec{B}) - 1}{d_A d_C} = \frac{d_A d_B (1 - \cos\gamma)}{d_C (d_B - d_A)}.$$

Writing down the analogous formula for $\cos\alpha$, we get $\cos\alpha = \cos\beta$. In other words, C is a circle passing through both intersection points of A and B and dividing the angle between this two circles into two equal parts ('*bisector*').

Acknowledgements. The authors are grateful to V. Arnold for important remarks, to V. Dremov, whose valueable ideas have been realized in this project, and also to O. Karpenkov, I. Losev, and to the jury of the XVIII Summer conference of the Tournament of towns for useful discussions.

[1] V.I. Arnold, Theorem on altitudes of a triangle in the Lobachevsky geometry as the Jacobi identity in the Lie algebra of quadratic forms in the simplectic plane, Math. Prosveschenie, Third Series 9(2005), pages 93–99 (in Russian).
[2] O.Ya. Viro, Yu.V. Dobdotukhina, Interlacements of crossing lines, Quant 3(1988), pages 12–19 (in Russian).
[3] Yu.P. Solov'ev, A.B. Sossinsky, Geometry of sliding vectors, Quant 8(1985), pages 9–17.
[4] J. Conant, R. Schneiderman, P. Teichner, Jacobi identities in low-dimensional topology, preprint (2006).

БАЗИСНОСТЬ ПЛОСКИХ МНОЖЕСТВ

А. Скопенков и И. Шнурников

ЗАДАЧИ ДО ПРОМЕЖУТОЧНОГО ФИНИША.

Мотивировки и соглашения.

При решении 13-й проблемы Гильберта появилось понятие *базисного вложения.* Основной результат настоящего цикла задач (задача 8b) — элементарное решение 'половины' проблемы Арнольда о характеризации базисных подмножеств плоскости. Важнейшие нерешенные задачи данного цикла посвящены попыткам характеризации *гладко базисных* подмножеств плоскости. Мы благодарим В. И. Арнольда за полезные обсуждения.

Трудные задачи отмечены звездочкой, а нерешенные — двумя. Если условие задачи является утверждением, то задача состоит в том, чтобы это утверждение доказать.

Разрывная базисность.

1. (a) Для любых ли четырех чисел $f_{11}, f_{12}, f_{21}, f_{22}$ существуют такие четыре числа g_1, g_2, h_1, h_2, что $f_{ij} = g_i + h_j$ при любых $i, j = 1, 2$?

(b) Андрей Николаевич и Владимир Игоревич играют в игру 'А ну-ка, разложи!'. На шахматной доске отмечено несколько клеток. А. Н. расставляет числа в отмеченных клетках, как хочет. В. И. смотрит на расставленные числа и берет 16 чисел $a_1, \dots, a_8, b_1, \dots, b_8$, т.е. 'весов' столбцов и строк, как хочет. Если число в каждой отмеченной клетке оказалось равным сумме весов строки и столбца этой клетки, то выигрывает В. И., а иначе (т.е. если число хотя бы в одной отмеченной клетке оказалось не равным сумме весов строки и столбца этой клетки) выигрывает А. Н.

Докажите, что при правильной игре А. Н. выигрывает тогда и только тогда, когда не существует замкнутого маршрута ладьи, начальная клетка и клетки поворота которого являются отмеченными (не обязательно все отмеченные клетки задействованы).

Обозначим через $\mathbb{R}^2$ плоскость с фиксированной системой координат. Обозначим через $x(a)$ и $y(a)$ координаты точки $a \in \mathbb{R}^2$. Последовательность (конечная или бесконечная) точек плоскости $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ называется *молнией*, если для каждого i выполнено $a_i \neq a_{i+1}$, и при этом $x(a_i) = x(a_{i+1})$ для четных i и $y(a_i) = y(a_{i+1})$ для нечетных i. Не обязательно все точки молнии различны. Конечная молния $\{a_1, \dots, a_n, \dots\}$ называется *замкнутой*, если $a_1 = a_n$.

2. Рассмотрим замкнутую молнию $\{a_1, \dots, a_n = a_1\}$. Назовем *разложением* расстановку чисел в проекциях точек этой молнии на ось Ox и в проекциях точек этой молнии на ось Oy. Можно ли так расставить в точках молнии числа $f_1, \dots, f_n \in \mathbb{R}$ с $f_1 = f_n$, чтобы для любого разложения некоторое число f_i не было бы равно сумме двух чисел, стоящих в $x(a_i)$ и в $y(a_i)$?

Подмножество $K \subset \mathbb{R}^2$ плоскости называется *разрывно базисным*, если для любой функции $f : K \to \mathbb{R}$ существуют такие функции $g, h : \mathbb{R} \to \mathbb{R}$, что $f(x, y) = g(x) + h(y)$ для каждой точки $(x, y) \in K$.

3. (a) Отрезок $K = 0 \times [0, 1] \subset \mathbb{R}^2$ является разрывно базисным.

(b) Крест $K = 0 \times [-1, 1] \cup [-1, 1] \times 0 \subset \mathbb{R}^2$ является разрывно базисным.

4. (a) *Критерий разрывной базисности плоских множеств.* Подмножество плоскости является разрывно базисным тогда и только тогда, когда оно не содержит замкнутых молний.

(b)** Как по набору отмеченных клеток в кубе $8 \times 8 \times 8$ узнать, кто выигрывает в пространственный аналог игры 'А ну-ка разложи!', в котором В. И. пытается поставить 24 числа $a_1, \dots, a_8, b_1, \dots, b_8, c_1, \dots, c_8$ так, чтобы число в каждой клетке (i, j, k) равнялось сумме $a_i + b_j + c_k$ трех 'весов'?

(c)** Определите разрывную базисность подмножеств трехмерного пространства. Сформулируйте и докажите пространственный аналог приведенного критерия.

Непрерывная базисность.

Через $|z,z_0| = |(x,y),\ (x_0,y_0)| = \sqrt{(x-x_0)^2+(y-y_0)^2}$ обозначается обычное расстояние между точками $z=(x,y)$ и $z_0=(x_0,y_0)$ плоскости. Пусть K — подмножество плоскости $\mathbb{R}^2$. Функция $f: K\to\mathbb{R}$ называется *непрерывной*, если для любых точки $z_0\in K$ и числа $\varepsilon>0$ существует такое число $\delta>0$, что для любой точки $z\in K$ с условием $|z,z_0|<\delta$ выполнено $|f(z)-f(z_0)|<\varepsilon$. Иногда удобно обозначать точки (x,y) вместо z.

5. (a) Функция $f(x,y)=\sqrt{x^2+y^2}$ является непрерывной на плоскости.

(b) Функция $f(x,y)$, равная целой части от $x+y$, не является непрерывной на плоскости.

(c) Если $a_1,\dots,a_n$ — различные точки множества $K\subset\mathbb{R}^2$, то существует непрерывная функция $f:K\to\mathbb{R}$, для которой $f(a_i)=(-1)^i$ и $|f(x)|\le 1$ для любого $x\in K$.

(d) Пусть $K=\{a_1,\dots,a_{4n+4}\}$ — молния из $4n+4$ различных точек на плоскости и $f_1,\dots,f_{4n+4}$ — числа, для которых $|(-1)^i-f_i|<1/2n$. Пусть $g(x(a_i)),h(y(a_i))$, $i=1,\dots,4n+4$, — такие числа, что $f_i=g(x(a_i))+h(y(a_i))$ для любого i (при этом если $x(a_i)=x(a_j)$, то $g(x(a_i))=g(x(a_j))$, и аналогично для y и h). Докажите, что $\max_i\{g(x(a_i))\}>n$.

В дальнейшем все функции предполагаются непрерывными.

Подмножество $K\subset\mathbb{R}^2$ называется *(непрерывно) базисным*, если для любой непрерывной функции $f:K\to\mathbb{R}$ существуют такие непрерывные функции $g,h:\mathbb{R}\to\mathbb{R}$, что $f(x,y)=g(x)+h(y)$ для каждой точки $(x,y)\in K$.

6. (a) Замкнутая молния не базисна.

(b) Отрезок $K=0\times[0,1]\subset\mathbb{R}^2$ является базисным.

(c) Крест $K=0\times[-1,1]\cup[-1,1]\times 0\subset\mathbb{R}^2$ является базисным.

7. (a) Если подмножество плоскости базисно, то оно разрывно базисно.

(b) *Пополненной молнией* называется объединение точки $a_0\in\mathbb{R}^2$ с бесконечной молнией $\{a_1,\dots,a_n,\dots\}\subset\mathbb{R}^2$ из различных точек, *сходящейся* к точке a_0 (т.е. для любого $\varepsilon>0$ найдется такое натуральное N, что для любого $i>N$ выполнено $|a_i,a_0|<\varepsilon$). Докажите, что никакая пополненная молния не является базисной. (Заметим, что она является разрывно базисной).

(c) Через $[a,b]$ обозначим отрезок, соединяющий точки a и b. Докажите, что крест $K=[(-1,-2),(1,2)]\cup[(-1,1),(1,-1)]$ не является базисным.

(d) Пусть $m_{i,j}=2-3\cdot 2^{-i}+j\cdot 2^{-2i}$. Рассмотрим множество, состоящее из точек $(m_{i,2l},m_{i,2l})$ и точек $(m_{i,2l-1},m_{i,2l-2})$, где i от 1 до ∞ и $l=1,2,3,\dots,2^{i-1}$. Докажите, что это подмножество плоскости не содержит бесконечной молнии, но содержит сколь угодно длинные молнии.

(e) Объединение множества из предыдущего пункта с точкой $(2,2)$ не базисно.

8. Пусть $K\subset\mathbb{R}^2$ — образ отрезка $[0,1]$ при непрерывном отображении $[0,1]\to\mathbb{R}^2$.

(a) Любая непрерывная функция $K\to\mathbb{R}$ достигает своего наибольшего и наименьшего значений. Указание: сведите к аналогичной теореме для непрерывных функций $[0,1]\to\mathbb{R}$.

(b)* Если K содержит сколь угодно длинные молнии, то K не базисно.

Указание. Предположим, что K содержит сколь угодно длинные молнии и базисно. Можно считать, что все точки каждой молнии различны. Для каждого n возьмем молнию $\{a_1^n,\dots,a_{4n+4}^n\}$ из $4n+4$ различных точек в K. Тогда существует непрерывная функция $f_n:K\to\mathbb{R}$, для которой $f_n(a_i^n)=(-1)^i$ и $|f_n(x)|\le 1$ для любого $x\in K$. Для функции $G:K\to\mathbb{R}$ ее максимум обозначается через $\|G\|:=\max_{x\in K}|G(x)|$. Пусть $f:K\to\mathbb{R}$ и $g,h:\mathbb{R}\to\mathbb{R}$ — функции такие, что $\|f-f_n\|<1/2n$ и $f(x,y)=g(x)+h(y)$ для любой $(x,y)\in K$. Тогда $\|g\|>n$ ...

БАЗИСНОСТЬ ПЛОСКИХ МНОЖЕСТВ

А. Скопенков и И. Шнурников

задачи до промежуточного финиша.

Мотивировки и соглашения.

При решении 13-й проблемы Гильберта появилось понятие *базисного вложения*. Основной результат настоящего цикла задач (задача 8b) — элементарное решение 'половины' проблемы Арнольда о характеризации базисных подмножеств плоскости. Важнейшие нерешенные задачи данного цикла посвящены попыткам характеризации *гладко базисных* подмножеств плоскости. Мы благодарим В. И. Арнольда за полезные обсуждения.

Трудные задачи отмечены звездочкой, а нерешенные — двумя. Если условие задачи является утверждением, то задача состоит в том, чтобы это утверждение доказать.

Разрывная базисность.

1. (a) Для любых ли четырех чисел $f_{11}, f_{12}, f_{21}, f_{22}$ существуют такие четыре числа g_1, g_2, h_1, h_2, что $f_{ij} = g_i + h_j$ при любых $i, j = 1, 2$?

(b) Андрей Николаевич и Владимир Игоревич играют в игру 'А ну-ка, разложи!'. На шахматной доске отмечено несколько клеток. А. Н. расставляет числа в отмеченных клетках, как хочет. В. И. смотрит на расставленные числа и берет 16 чисел $a_1, \dots, a_8, b_1, \dots, b_8$, т.е. 'весов' столбцов и строк, как хочет. Если число в каждой отмеченной клетке оказалось равным сумме весов строки и столбца этой клетки, то выигрывает В. И., а иначе (т.е. если число хотя бы в одной отмеченной клетке оказалось не равным сумме весов строки и столбца этой клетки) выигрывает А. Н.

Докажите, что при правильной игре В. И. выигрывает тогда и только тогда, когда не существует замкнутого маршрута ладьи, начальная клетка и клетки поворота которого являются отмеченными (не обязательно все отмеченные клетки задействованы).

Обозначим через $\mathbb{R}^2$ плоскость с фиксированной системой координат. Обозначим через $x(a)$ и $y(a)$ координаты точки $a \in \mathbb{R}^2$. Последовательность (конечная или бесконечная) точек плоскости $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ называется *молнией*, если для каждого i выполнено $a_i \neq a_{i+1}$, и при этом $x(a_i) = x(a_{i+1})$ для четных i и $y(a_i) = y(a_{i+1})$ для нечетных i. Не обязательно все точки молнии различны. Конечная молния $\{a_1, \dots, a_{2l+1}\}$ называется *замкнутой*, если $a_1 = a_{2l+1}$.

2. Рассмотрим замкнутую молнию $\{a_1, \dots, a_n = a_1\}$. Назовем *разложением* расстановку чисел в проекциях точек этой молнии на ось Ox и в проекциях точек этой молнии на ось Oy. Можно ли так расставить в точках молнии числа $f_1, \dots, f_n \in \mathbb{R}$ с $f_1 = f_n$, чтобы для любого разложения некоторое число f_i не было бы равно сумме двух чисел, стоящих в $x(a_i)$ и в $y(a_i)$?

Подмножество $K \subset \mathbb{R}^2$ плоскости называется *разрывно базисным*, если для любой функции $f : K \to \mathbb{R}$ существуют такие функции $g, h : \mathbb{R} \to \mathbb{R}$, что $f(x, y) = g(x) + h(y)$ для каждой точки $(x, y) \in K$.

3. (a) Отрезок $K = 0 \times [0, 1] \subset \mathbb{R}^2$ является разрывно базисным.

(b) Крест $K = 0 \times [-1, 1] \cup [-1, 1] \times 0 \subset \mathbb{R}^2$ является разрывно базисным.

4. (a) *Критерий разрывной базисности плоских множеств.* Подмножество плоскости является разрывно базисным тогда и только тогда, когда оно не содержит замкнутых молний.

(b)** Как по набору отмеченных кубиков в кубе $8 \times 8 \times 8$ узнать, кто выигрывает в пространственный аналог игры 'А ну-ка разложи!', в котором В. И. пытается поставить

24 числа $a_1, \dots, a_8, b_1, \dots, b_8, c_1, \dots, c_8$ так, чтобы число в каждом кубике (i, j, k) равнялось сумме $a_i + b_j + c_k$ трех 'весов'?

(c)** Определите разрывную базисность подмножеств трехмерного пространства. Сформулируйте и докажите пространственный аналог приведенного критерия.

Непрерывная базисность.

Через $|z, z_0| = |(x, y),\ (x_0, y_0)| = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ обозначается обычное расстояние между точками $z = (x, y)$ и $z_0 = (x_0, y_0)$ плоскости. Пусть K — подмножество плоскости $\mathbb{R}^2$. Функция $f : K \to \mathbb{R}$ называется *непрерывной*, если для любых точки $z_0 \in K$ и числа $\varepsilon > 0$ существует такое число $\delta > 0$, что для любой точки $z \in K$ с условием $|z, z_0| < \delta$ выполнено $|f(z) - f(z_0)| < \varepsilon$. Иногда удобно обозначать точки (x, y) вместо z.

5. (a) Функция $f(x, y) = \sqrt{x^2 + y^2}$ является непрерывной на плоскости.

(b) Функция $f(x, y)$, равная целой части от $x+y$, не является непрерывной на плоскости.

(c) Если $a_1, \dots, a_n$ — различные точки множества $K \subset \mathbb{R}^2$, то существует непрерывная функция $f : K \to \mathbb{R}$, для которой $f(a_i) = (-1)^i$ и $|f(x)| \le 1$ для любого $x \in K$.

(d) Пусть $K = \{a_1, \dots, a_{4n+4}\}$ — молния из $4n + 4$ различных точек на плоскости и $f_1, \dots, f_{4n+4}$ — числа, для которых $|(-1)^i - f_i| < 1/2n$. Пусть $g(x(a_i)), h(y(a_i))$, $i = 1, \dots, 4n+4$, — такие числа, что $f_i = g(x(a_i)) + h(y(a_i))$ для любого i (при этом если $x(a_i) = x(a_j)$, то $g(x(a_i)) = g(x(a_j))$, и аналогично для y и h). Докажите, что $\max_i |g(x(a_i))| > n$.

В дальнейшем все функции предполагаются непрерывными.

Подмножество $K \subset \mathbb{R}^2$ называется *(непрерывно) базисным*, если для любой непрерывной функции $f : K \to \mathbb{R}$ существуют такие непрерывные функции $g, h : \mathbb{R} \to \mathbb{R}$, что $f(x, y) = g(x) + h(y)$ для каждой точки $(x, y) \in K$.

6. (a) Замкнутая молния не базисна.

(b) Отрезок $K = 0 \times [0, 1] \subset \mathbb{R}^2$ является базисным.

(c) Крест $K = 0 \times [-1, 1] \cup [-1, 1] \times 0 \subset \mathbb{R}^2$ является базисным.

7. (a) Если подмножество плоскости базисно, то оно разрывно базисно.

(b) *Пополненной молнией* называется объединение точки $a_0 \in \mathbb{R}^2$ с бесконечной молнией $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ из различных точек, *сходящейся* к точке a_0 (т.е. для любого $\varepsilon > 0$ найдется такое натуральное N, что для любого $i > N$ выполнено $|a_i, a_0| < \varepsilon$). Докажите, что никакая пополненная молния не является базисной. (Заметим, что она является разрывно базисной).

(c) Через $[a, b]$ обозначим отрезок, соединяющий точки a и b. Докажите, что крест $K = [(-1, -2), (1, 2)] \cup [(-1, 1), (1, -1)]$ не является базисным.

(d) Пусть $m_{i,j} = 2 - 3 \cdot 2^{-i} + j \cdot 2^{-2i}$. Рассмотрим множество, состоящее из точек $(m_{i,2l}, m_{i,2l})$ и точек $(m_{i,2l}, m_{i,2l-2})$, где i от 1 до ∞ и $l = 1, 2, 3, \dots, 2^{i-1}$. Докажите, что это подмножество плоскости не содержит бесконечной молнии, но содержит сколь угодно длинные молнии.

(e) Объединение множества из предыдущего пункта с точкой $(2, 2)$ не базисно.

8. Пусть $K \subset \mathbb{R}^2$ — образ отрезка $[0, 1]$ при непрерывном отображении $[0, 1] \to \mathbb{R}^2$.

(a) Любая непрерывная функция $K \to \mathbb{R}$ достигает своего наибольшего и наименьшего значений. Указание: сведите к аналогичной теореме для непрерывных функций $[0, 1] \to \mathbb{R}$.

(b)* Если K содержит сколь угодно длинные молнии, то K не базисно.

Указание. Предположим, что K содержит сколь угодно длинные молнии и базисно. Можно считать, что все точки каждой молнии различны. Для каждого n возьмем молнию $\{a_1^n, \dots, a_{4n+4}^n\}$ из $4n + 4$ различных точек в K. Тогда существует непрерывная функция $f_n : K \to \mathbb{R}$, для которой $f_n(a_i^n) = (-1)^i$ и $|f_n(x)| \le 1$ для любого $x \in K$. Для функции

$G : K \to \mathbb{R}$ ее максимум обозначается через $|G| := \max_{x\in K} |G(x)|$. Пусть $f : K \to \mathbb{R}$ и $g, h : \mathbb{R} \to \mathbb{R}$ — функции такие, что $|f - f_n| < 1/2n$ и $f(x, y) = g(x) + h(y)$ для любой $(x, y) \in K$. Тогда $|g| > n$ ...

БАЗИСНОСТЬ ПЛОСКИХ МНОЖЕСТВ

А. Скопенков и И. Шнурников (представляется ими и В. Гориным)

задачи после промежуточного финиша.

Критерий базисности.

Продолжение указаний к задаче 8b. Определим по индукции последовательность чисел s_n и функций $F_n : K \to \mathbb{R}$. Положим $s_0 = 1$ и $F_0 = 0$. Предположим, что s_{n-1} и F_{n-1} уже определены. Возьмем функции $G_{n-1}, H_{n-1} : \mathbb{R} \to \mathbb{R}$, для которых $F_{n-1}(x, y) = G_{n-1}(x) + H_{n-1}(y)$ (если таких функций нет, то все доказано). Берем

$$s_n > s_{n-1}! \cdot (|G_{n-1}| + n) \quad \text{и} \quad F_n = F_{n-1} + \frac{f_{s_n}}{s_{n-1}!}$$

Тогда функция $F = \sum\limits_{n=1}^{\infty} \frac{f_{s_n}}{s_{n-1}!}$ не представима в виде $G(x) + H(y)$ (что Вам и остается доказать).

Последовательность точек a_i плоскости называется *сходящейся к точке* a, если для любого $\varepsilon > 0$ найдется такое целое N, что для любого $i > N$ выполнено $|a, a_i| < \varepsilon$.

Подмножество $K \subset \mathbb{R}^2$ плоскости называется *замкнутым*, если для любой бесконечной последовательности точек $a_i \in K$, сходящейся к точке a, выполнено $a \in K$.

9. (a) Подмножество $K \subset \mathbb{R}^2$ плоскости является замкнутым тогда и только тогда, когда для любой точки $a \notin K$ найдется такое $\varepsilon > 0$, что любая точка плоскости с расстоянием менее ε до a не принадлежит K.

(b) Образ отрезка при непрерывном отображении $[0, 1] \to \mathbb{R}^2$ в плоскость является замкнутым подмножеством плоскости.

Геометрический критерий базисности Штернфельда. Замкнутое ограниченное подмножество $K \subset \mathbb{R}^2$ плоскости базисно тогда и только тогда, когда K не содержит сколь угодно длинных молний.

10. (a) Условие замкнутости в критерии действительно необходимо.

(b) Условие ограниченности в критерии действительно необходимо.

(c) Докажите часть 'только тогда' ($\Rightarrow$) критерия.

Пусть K — подмножество плоскости $\mathbb{R}^2$. Для каждой точки $v \in K$ нарисуем две прямые, проходящие через v параллельно координатным осям. Если хотя бы одна из этих двух прямых пересекает K только в точке v, то покрасим v в белый цвет. Обозначим через $E(K)$ множество всех точек K, не являющихся белыми:

$$E(K) = \{v \in K : \ |K \cap (x = x(v))| \geq 2 \text{ и } |K \cap (y = y(v))| \geq 2\}.$$

Пусть $E^2(K) = E(E(K))$, $E^3(K) = E(E(E(K)))$ и т.д.

11. (a) Если $K \subset \mathbb{R}^2$ не содержит сколь угодно длинных молний, то $E^n(K) = \emptyset$ для некоторого n.

(b) Докажите обратное.

(c)* Если $K \subset \mathbb{R}^2$ замкнуто и ограничено, причем $E(K) = \emptyset$, то K базисно.

(d)* Докажите часть 'тогда' ($\Leftarrow$) критерия Штернфельда. Замечание. Это можно пытаться делать, доказав сначала, что разложение $f(x,y) = g(x) + h(y)$ можно получить для кусочно-линейных функций f, причем $|g| + |h| < C_n|f|$, где C_n зависит только от того n, для которого $E^n(K) = \emptyset$.

12. (a) Определите (непрерывную) базисность подмножеств трехмерного пространства. Докажите, что еж $0 \times 0 \times [-1,1] \cup 0 \times [-1,1] \times 0 \cup [-1,1] \times 0 \times 0 \subset \mathbb{R}^3$ является базисным.

(b) Подмножество пространства $\mathbb{R}^3$, состоящее из четырех точек $(0,0,0)$; $(1,1,0)$; $(0,1,1)$; $(1,0,1)$, не базисно. (Но $E^n(K) \neq \emptyset$ для любого n, см. ниже.)

(c)* Пусть подмножество $K \subset \mathbb{R}^3$ пространства замкнуто и ограничено. Аналогично определим $E(K)$, используя вместо прямых плоскости, перпендикулярные осям координат:

$$E(K) = \{v \in K:\ |K \cap (x = x(v))| \geq 2,\ |K \cap (y = y(v))| \geq 2 \text{ и } |K \cap (z = z(v))| \geq 2\}.$$

Докажите, что если $E^n(K) = \emptyset$ для некоторого n, то K базисно.

Гладкая базисность.

Пусть K — подмножество плоскости $\mathbb{R}^2$. Функция $f : K \to \mathbb{R}$ называется *дифференцируемой*, если для любой точки $z_0 \in K$ существуют такие вектор $a \in \mathbb{R}^2$ и бесконечно малая функция $\alpha : \mathbb{R}^2 \to \mathbb{R}$, что для любой точки $z \in K$ выполнено

$$f(z) = f(z_0) + a \cdot (z - z_0) + \alpha(z - z_0)|z, z_0|.$$

Здесь точка означает знак скалярного произведения векторов $a = (f_x, f_y)$ и $z - z_0 = (x, y)$, и т.е. $a \cdot (z - z_0) = x f_x + y f_y$. Функция $\alpha : \mathbb{R}^2 \to \mathbb{R}$ называется *бесконечно малой* , если для любого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что для любой точки $(x, y) \in \mathbb{R}^2$

$$\text{если} \quad \sqrt{x^2 + y^2} < \delta, \quad \text{то} \quad |\alpha(x, y)| < \varepsilon.$$

Подмножество $K \subset \mathbb{R}^2$ плоскости называется *дифференцируемо базисным*, если для любой дифференцируемой функции $f : K \to \mathbb{R}$ существуют такие дифференцируемые функции $g : \mathbb{R} \to \mathbb{R}$ и $h : \mathbb{R} \to \mathbb{R}$, что для любой точки $(x, y) \in K$ выполняется $f(x, y) = g(x) + h(y)$.

13. (a) (b) (c) Решите аналоги задачи 6 для дифференцируемой базисности.

14. (a) График функции $|x|$ на отрезке $[-1, 1]$ является дифференцируемо базисным.

(b) Ломаная с последовательными вершинами $(-2, 0)$, $(-1, 1)$, $(0, 0)$, $(1, 1)$ и $(2, 0)$ не является дифференцируемо базисной. (Заметим, что она непрерывно базисна.)

(c) Пополненная молния $\{([\frac{n+1}{2}]^{-1/2}, [\frac{n}{2}]^{-1/2})\}_{n=1}^{\infty} \cup \{(0, 0)\}$ не является дифференцируемо базисной. (Заметим, что она не является также непрерывно базисной.)

(d) Пополненная молния $\{(2^{-[\frac{n+1}{2}]}, 2^{-[\frac{n}{2}]})\}_{n=1}^{\infty} \cup \{(0, 0)\}$ является дифференцируемо базисной. (Заметим, что она не является непрерывно базисной.)

(e)** Существует ли непрерывное отображение отрезка в плоскость, образ которого является дифференцируемо базисным, но не непрерывно базисным?

15. (a) Крест $K = [(-1, -2), (1, 2)] \cup [(-1, 1), (1, -1)]$ не является дифференцируемо базисным.

(b) Является ли подмножество $\{(t^2, \frac{t^2}{(1+t)^2})\}_{t \in [-\frac{1}{2}, \frac{1}{2}]}$ плоскости дифференцируемо базисным?

(c)** Найдите критерий дифференцируемой базисности для графов, лежащих в плоскости.

16. Пусть K — подмножество плоскости $\mathbb{R}^2$ и $r \geq 0$. Функция $f : K \to \mathbb{R}$ называется *r раз дифференцируемой*, если для любой точки $z_0 \in K$ существуют такие многочлен $\overline{f}(z) = \overline{f}(x, y)$ степени не выше r от двух переменных x и y и бесконечно малая функция $\alpha : \mathbb{R}^2 \to \mathbb{R}$, что $f(z) = \overline{f}(z - z_0) + \alpha(z - z_0)|z, z_0|^r$ для любой точки $z \in K$. (Это определение отличается от общепринятого.)

(a) Ноль раз дифференцируемые функции — это в точности непрерывные, а один раз дифференцируемые — это в точности дифференцируемые.

(b) Для любого целого положительного r определите r-дифференцируемую базисность подмножеств плоскости.

(c) Для любого целого $k \geq 0$ найдется подмножество плоскости, r-дифференцируемо базисное для любого $r = 0, 1, \dots k$, но не r-дифференцируемо базисное ни для какого $r > k$.

(d)** Найдите критерий r-дифференцируемой базисности для графов, лежащих в плоскости.

БАЗИСНОСТЬ ПЛОСКИХ МНОЖЕСТВ

А.Скопенков, И.Шнурников (представляется ими и В. Гориным)

Решения задач, предложенных до промежуточного финиша.

Разрывная базисность.

В дальнейшем мы будем использовать определения, предложенные в задачах после промежуточного финиша.

1. (a) Это неверно. Если $f_{ij} = g_i + h_j$ для $i, j = 1, 2$, то $f_{11} + f_{22} = f_{12} + f_{21}$, но это соотношение не имеет места для некоторых наборов чисел f_{ij}.

(b) "Только тогда" следует из задачи 2. Докажем утверждение "тогда" индукцией по количеству отмеченных клеток. Если отмечена только одна клетка, утверждение задачи тривиально. Обозначим за K множество центров отмеченных клеток. По условию, K не содержит замкнутых молний, следовательно $\#E(K) < \#K$. Значит, по индуктивному предположению В.И. может выиграть на множестве $E(K)$. Все оставшиеся клетки являются единственными отмеченными в своей строке или в своем столбце. Следовательно, В.И. сможет выбрать и оставшиеся веса для K.

2. Да. Если каждое из чисел f_i представимо в виде суммы двух чисел, расположенных в точках $x(a_i)$ и $y(a_i)$, то $f_1 - f_2 + f_3 - \dots - f_{n-1} = 0$, но можно легко подобрать набор чисел f_i, для которого это неверно.

3. (a) Положим $h(y) = f(0, y)$ и $g(x) = 0$.

(b) Положим $g(x) = f(x, 0)$ и $h(y) = f(0, y) - f(0, 0)$.

4. (a) "Только тогда" следует из задачи 2. Докажем утверждение "тогда". Рассмотрим произвольную функцию $f : K \to \mathbb{R}$ и построим по ней функции g и h такие, что $f(x, y) = g(x) + h(y)$. Назовём две точки $a, b \in K$ *эквивалентными*, если существует молния $\{a = a_1, \dots, a_n = b\} \subset K$. Возьмём один из классов эквивалентности $K_1 \subset K$ и определим функции $g : x(K_1) \to \mathbb{R}$ и $h : y(K_1) \to \mathbb{R}$ следующим образом. Зафиксируем произвольную точку $a_1 \in K_1$. Положим $g(x(a_1)) = f(a_1)$ и $h(y(a_1)) = 0$. Если $\{a_1, a_2, \dots, a_{2l}\}$ — молния из точек множества K, то положим

$$h(y(a_{2l})) := f(a_{2l}) - f(a_{2l-1}) + \dots - f(a_1) \quad \text{и} \quad g(x(a_{2l})) = f(a_{2l-1}) - f(a_{2l-2}) + \dots + f(a_1).$$

Если $\{a_1, a_2, \dots, a_{2l+1}\}$ — молния из точек множества K, то положим

$$g(x(a_{2l+1})) = f(a_{2l+1}) - f(a_{2l}) + \dots + f(a_1)$$

(значение $h(y(a_{2l+1}))$ уже определено). Сделаем это построение для всех классов эквивалентности одновременно. Для всех же прочих точек положим $g(x) = 0$ и $h(y) = 0$.

Непрерывно базисные множества.

5. (a) Можно положить $\delta = \varepsilon$, тогда утверждение следует из неравенства треугольника $|f(z) - f(z_0)| \le |z, z_0|$.

(b) Для $x = 1$, $y = 0$ и $\varepsilon = \frac{1}{2}$ такого δ не существует, т.к. $|f(1, 0) - f(1 - \frac{\delta}{2}, 0)| = 1 > \frac{1}{2}$.

(c) Построим сначала непрерывную функцию $f : \mathbb{R}^2 \to \mathbb{R}$, удовлетворяющую условию задачи. Обозначим $s = \min_{i<j} |a_i, a_j|$. Рассмотрим n дисков с центрами в точках a_i и радиусами $\frac{s}{3}$. Вне этих дисков положим $f = 0$. Внутри i-го диска сделаем f линейной функцией от радиуса, равной $(-1)^i$ в центре a_i и нулю на границе. Теперь ограничим построенную функцию на $K \subset \mathbb{R}^2$ и получим требуемую непрерывную функцию $K \to \mathbb{R}$.

(d) Имеем

$$|(f_2 - f_3 + f_4 - f_5 + \dots - f_{4n+3}) - (4n+2)| \le \frac{4n+2}{2n} \le 3.$$

Это означает, что $g(a_2) - g(a_{4n+3}) \ge (4n+2) - 3 > 2n$, из чего немедленно следует требуемое неравенство.

6. (a) Если бы молния $A = \{a_1, \dots, a_{2l+1}\}$ была базисной, то $f(a_1) - f(a_2) + \dots + f(a_{n-2}) - f(a_{2l}) = 0$, но легко подобрать функцию f, для которой это не выполнено. Сравните с задачей 2.

(b),(c) Аналогично задчам 3a, 3b.

7. (a) Если множество не является разрывно базисным, то оно содержит замкнутую молнию. Тогда утверждение задач следует из 6a, т.к. функция f может быть продолжена с замкнутой молнии на всё множество.

(b) Рассмотрим функцию f, для которой $f(a_i) = \frac{(-1)^i}{i}$. Предположим, что $f(x,y) = g(x) + h(y)$ для некоторых непрерывных g и h, тогда

$$f(a_1) - f(a_2) + f(a_3) - f(a_4) + \dots - f(a_{2l}) = h(y(a_1)) - h(y(a_{2l})).$$

Так как $\lim_{l\to\infty} h(y_{2l})$ существует и равен $h(y(a_0))$, то ряд $\sum_{i=1}^{2l} (-1)^i f(a_i)$ сходится при $l \to \infty$. Но это противоречит расходимости гармонического ряда.

(c) Крест содержит замкнутую молнию

$$a_{4k+1} = (\frac{-1}{4^k}, \frac{1}{4^k}),\ a_{4k+2} = (\frac{1}{2\cdot 4^k}, \frac{1}{4^k}),\ a_{4k+3} = (\frac{1}{2\cdot 4^k}, \frac{-1}{2\cdot 4^k}),\ a_{4k+4} = (\frac{-1}{4^{k+1}}, \frac{-1}{2\cdot 4^k})$$

Определим функцию f на этой молнии, используя задачу 7(b), и продолжим её кусочно-линейно на весь крест. Не существует таких функций g и f, что $f(x,y) = g(x) + h(y)$.

(d) Для любого i точки $(x_{i,2l}, x_{i,2l})_{l=1}^{2^{i-1}}$ и $(x_{i,2l}, x_{i,2l-2})_{l=1}^{2^{i-1}}$ образуют молнию из 2^i элементов.

(e) Определим функцию $f(x,y)$ соотношениями

$$f((x_{i,2l}, x_{i,2l})) := \frac{1}{2^i} \quad \text{и} \quad f((x_{i,2l}, x_{i,2l-2})) := -\frac{1}{2^i}$$

Предположим, что $f(x,y) = g(x) + h(y)$ для некоторых непрерывных $g(x)$ и $h(y)$. Теперь для каждого i, используя молнии $(x_{i,2l}, x_{i,2l})$ и $(x_{i,2l}, x_{i,2l-2})$, где $l = 1, 2, 3, \dots 2^{i-1}$, получаем $h(2 - \frac{3}{2^i}) - h(2 - \frac{2}{2^i}) = 1$. Это противоречит непрерывности h в точке $y = 2$.

Функция $f : K \to \mathbb{R}$ называется *ограниченной*, если найдется число M такое, что $|f(x,y)| < M$ для любой точки $(x,y) \in K$.

8. *Лемма.* Непрерывная функция $f : [0,1] \to \mathbb{R}$ ограничена.

Доказательство. Предположим противное. Тогда для каждого целого n существует такая точка $a_n \in [0,1]$, что $|f(a_n)| > n$. Выберем из последовательности a_n подпоследовательность a_{n_i}, сходящуюся к некоторой точке $a \in [0,1]$. Из непрерывности функции f следует, что $|f(a_{n_i})|$ стремится к $|f(a)|$. Но в то же время эта последовательность стремится к бесконечности по построению! Из полученного противоречия следует лемма.

(a) Возьмём такое отображение $k : [0,1] \to \mathbb{R}^2$, что $K = k([0,1])$ и рассмотрим композицию $f \circ k : [0,1] \to \mathbb{R}$. Пусть s есть минимальное число, для которого $f(k(t)) \le s$ для всех $t \in [0,1]$. Если бы не существовало такого t, что $f(k(t)) = s$, то непрерывная функция $\frac{1}{s - f(k(t))} : [0,1] \to \mathbb{R}$ оказалась бы неограниченной. Но это невозможно по Лемме. Следовательно, f достигает своего наибольшего значения. Аналогично доказывается утверждение для наименьшего значения функции f.

БАЗИСНОСТЬ ПЛОСКИХ МНОЖЕСТВ

А.Скопенков и И.Шнурников (представляется ими и В. Гориным)

Решения задач, предложенных после промежуточного финиша.

Критерий базисности.

Для функции $G : K \to \mathbb{R}$ положим $|G| := \max_{x \in K} |G(x)|$.

8. (b) Предположим, что множество K содержит сколь угодно длинные молнии и при этом является базисным. Можно считать, что все точки молний различны. Тогда для любого n существует молния $\{a_1^n, \dots, a_{4n+4}^n\}$ из $(4n+4)$-х различных точек множества K. В таком случае существует непрерывная функция $f_n : K \to \mathbb{R}$ такая, что $f_n(a_i^n) = (-1)^i$ и $|f_n(x)| \le 1$ для любого $x \in K$. Пусть $f : K \to \mathbb{R}$ и $g, h : \mathbb{R} \to \mathbb{R}$ — непрерывные функции, для которых $|f - f_n| < 1/2n$ и $f(x, y) = g(x) + h(y)$ для любой точки $(x, y) \in K$. Тогда $|g| > n$ по задаче 5(d). Теперь достаточно предположить, что $F(x, y) = G(x) + H(y)$ и доказать, что $|G| > n$ для каждого n. Имеем:

$$F - F_n = F - F_{n-1} - \frac{f_{s_n}}{s_{n-1}!} = \frac{s_{n-1}!(F - F_{n-1}) - f_{s_n}}{s_{n-1}!}.$$

Положим $f = s_{n-1}!(F - F_{n-1})$. Тогда $s_n - 1 > s_{n-1}$ при $n > 2$ и

$$|f - f_{s_n}| = s_{n-1}!|F - F_n| < \frac{1}{(s_n - 1) \cdot s_n} \sum_{k=0}^{\infty} \frac{1}{s_{n+1} \cdot \ldots \cdot s_{n+k}} < \frac{1}{(s_n - 1) \cdot s_n} \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{2s_n}.$$

Но в то же время

$$f = s_{n-1}!(G(x) - G_{n-1}(x)) + s_{n-1}!(H(y) - H_{n-1}(y)).$$

По задаче 5(d) получаем $s_{n-1}!|G - G_{n-1}| > s_n$. Следовательно,

$$|G| + |G_{n-1}| \ge |G - G_{n-1}| > \frac{s_n}{s_{n-1}!} > |G_{n-1}| + n. \quad \Box$$

9. (a) Докажем утверждение "только тогда". Пусть K — замкнутое подмножество плоскости. Предположим, что для некоторой точки $a = (x, y) \notin K$ и для произвольного $\varepsilon = \frac{1}{n} > 0$ существует хотя бы одна точка $a_n \in K$, для которой $|a, a_n| \le \frac{1}{n}$. Но тогда последовательность точек $a_n \in K$ сходится к точке a, поэтому $a \in K$. Противоречие.

Теперь докажем утверждение "тогда". Пусть некоторая последовательность a_n сходится к точке a, не лежащей в множестве K. По условию существует $\varepsilon > 0$ такое, что для любой точки $a_n \in K$ расстояние $|a, a_n| > \varepsilon$. Но это противоречит сходимости последовательности.

(b) Рассмотрим непрерывное отображение $f : [0, 1] \to \mathbb{R}^2$. Предположим, что некоторая последовательность точек $\{a_i\}$ из образа f сходится к точке a. Для каждого i выберем $t_i \in f^{-1}(a_i)$. Теперь выделим из последовательности $\{t_i\}$ сходящуюся подпоследовательность $\{t_{i_k}\}$. Её предел обозначим за $t_0 \in [0, 1]$. Отображение f непрерывно, значит, последовательность $f(t_{i_k})$ сходится к $f(t_0)$. Тогда $a = f(t_0)$, следовательно, $f([0, 1])$ замкнуто.

10. (a) Любая бесконечная молния A, не содержащая замкнутых молний и сходящаяся к точке $a \notin A$, является базисной. Это следует из того, что любая функция, определённая на A, непрерывна.

(b) Контрпримером является множество $\{(k,k)\}_{k=1}^{\infty} \cup \{(k,k-1)\}_{k=1}^{\infty}$ точек плоскости.

(c) Доказательство повторяет решение задачи 8(b), используя следущую Лемму.

Лемма. Пусть K — произвольное замкнутое ограниченное подмножество плоскости. Тогда любая непрерывная функция $f : K \to \mathbb{R}$ ограничена.

11. (a) Предположим, что $E^n(K) \neq \emptyset$ для всех n. Для каждого n рассмотрим точку $a_0 \in E^n(K)$. Выберем точки $a_{-1}, a_1 \in E^{n-1}(K)$ такие, что $x(a_{-1}) = x(a_0)$ и $y(a_1) = y(a_0)$. Теперь можно выбрать точки $a_{-2}, a_2 \in E^{n-2}(K)$, для которых $\{a_{-2}, a_{-1}, a_0, a_1, a_2\}$ — молния. Аналогично можно сконструировать молнию из $2n+1$ точек, лежащую целиком в множестве K. Что и требовалось доказать.

(b) Пусть теперь множество K содержит молнию из $2n+1$ точки $\{a_{-n}, \dots, a_0, \dots, a_n\}$. Тогда в множестве $E(K)$ содержится молния из $2n-1$ точки $\{a_{-n+1}, \dots, a_{n-1}\}$. Продолжая, получим, что $a_0 \in E^n(K)$. Следовательно, если $E^n(K) = \emptyset$, то K не содержит молнии из $2n+1$ точек.

(c)* Смотри решение задачи 11(d).

(d)* Приведём неэлементарное доказательство, основанное на переформулировке свойства базисности в терминах *ограниченных линейных операторов* в *банаховых пространствах функций.* Обозначим через $C(X)$ пространство непрерывных функций на X с нормой $|f| = \sup\{|f(x)| \ : \ x \in X\}$. Обозначим $I := [0,1]$. Обозначим через $pr_x(a)$ и $pr_y(a)$ проекции точки $a \in K$ на оси координат.

Для подмножества $K \subset I^2$ определим отображение (*линейный оператор суперпозиции*)

$$\varphi : C(I) \oplus C(I) \to C(K) \quad \text{формулой} \quad \varphi(g,h)(x,y) = g(x) + h(y).$$

Ясно, что подмножество $K \subset I^2$ базисно тогда и только тогда, когда φ эпиморфно. Обозначим через $C^*(X)$ пространство ограниченных линейных функций $C(X) \to \mathbb{R}$ с нормой $|\mu| = \sup\{|\mu(f)| \ : \ f \in C(X), \ |f| = 1\}$. Для подмножества $K \subset I^2$ определим отображение (*двойственный линейный оператор суперпозиции*)

$$\varphi^* : C^*(K) \to C^*(I) \oplus C^*(I) \quad \text{как} \quad \varphi^*\mu(g,h) = (\mu(g \circ pr_x), \mu(h \circ pr_y)).$$

Так как $|\varphi^*\mu| \le 2|\mu|$, то φ^* ограничен. По двойственности, φ эпиморфен тогда и только тогда, когда φ^* мономорфен.

(При этом φ^* может быть инъективным, но не мономорфным. Другими словами, не только линейные соотношения на $\operatorname{im}\varphi$ заставляют его быть строго меньше чем $C(K)$, как показывает пример небазисной пополненной молнии.)

Понятно, что φ^* мономорфен тогда и только тогда, когда *существует* $\varepsilon > 0$ *такое, что* $|\varphi^*\mu| > \varepsilon|\mu|$ *для каждого ненулевого* $\mu \in C^*(K)$.

Остаётся доказать, что последнее условие следует из того, что $E^n(K) = \emptyset$. Приведем доказательство для $n \in \{1,2\}$ (для произвольного n оно аналогично). Мы используем следующий нетривиальный факт: $C^*(K)$ *совпадает с пространством σ-аддитивных регулярных вещественнозначных борелевских мер на* K (далее мы будем называть их просто 'мерами'; используется также термин 'заряды'). Имеем

$$\varphi^*\mu = (\mu_x, \mu_y), \quad \text{где} \quad \mu_x(U) = \mu(pr_x^{-1}U) \quad \text{и} \quad \mu_y(U) = \mu(pr_y^{-1}U).$$

Если $\mu = \mu^+ - \mu^-$ есть разложение меры μ на положительные и отрицательные части, то $|\mu| = \bar{\mu}(X)$, где $\bar{\mu} = \mu^+ + \mu^-$ есть абсолютное значение меры μ.

Обозначим через D_x (и D_y) множество тех точек из K, которые не затеняются любой другой точкой из K в x- (и y-) направлении. Возьмем любую меру μ на K с нормой 1.

Если $n = 1$, то

$$E(K) = \emptyset, \quad \text{поэтому} \quad D_x \cup D_y = K, \quad \text{значит,} \quad 1 = \bar{\mu}(K) \leq \bar{\mu}(D_x) + \bar{\mu}(D_y).$$

Тогда, не уменьшая общности, $\bar{\mu}(D_x) \geq 1/2$. Так как проекция на ось x инъективна на D_x, то $|\mu_x| \geq 1/2$. Из этого вытекает необходимое утверждение для $\varepsilon = \frac{1}{2}$.

Если $n = 2$, то

$$E(E(K)) = \emptyset, \quad \text{поэтому} \quad D_x \cup D_y = K - E(K) \quad \text{и} \quad E(D_x \cup D_y) = \emptyset.$$

Тогда при $\bar{\mu}(E(K)) < 3/4$ имеем $\bar{\mu}(D_x \cup D_y) > 1/4$ и, не уменьшая общности, $\bar{\mu}(D_x) > 1/8$, значит, как и в случае $n = 1$, имеем $|\mu_x| > 1/8$, из чего вытекает необходимое утверждение для $\varepsilon = \frac{1}{8}$.

При $\bar{\mu}(E(K)) \geq 3/4$ имеем $\bar{\mu}(K - E(K)) \leq 1/4$. Как и в случае $n = 1$, не уменьшая общности, $\bar{\mu}_x(pr_x(E(K))) \geq \bar{\mu}(E(K))/2$. Следовательно $|\mu_x| \geq \frac{1}{2} \cdot \frac{3}{4} - \frac{1}{4} = \frac{1}{8}$, из чего и вытекает необходимое утверждение для $\varepsilon = \frac{1}{8}$.

Заметим, что если $K \subset \mathbb{R}^2$ - базисное подмножество, то мы можем доказать без использования φ, что φ^* мономорфно. Определим линейный оператор $\Psi : C^*(I) \oplus C^*(I) \to C^*(K)$ формулой $\Psi(\mu_x, \mu_y)(f) = \mu_x(g) + \mu_y(h)$, где $g, h \in C(I)$ таковы, что $g(0) = 0$ и $f(x, y) = g(x) + h(y)$ для $(x, y) \in K$. Ясно, что $\Psi\varphi^* = \mathrm{id}$ и Ψ ограничено, следовательно φ^* мономорфно.

12. (a) Множество $K \subset \mathbb{R}^3$ называется *(непрерывно) базисным*, если для любой непрерывной функции $f : K \to \mathbb{R}$ существуют непрерывные функции $g, h, l : \mathbb{R} \to \mathbb{R}$, такие, что $f(x, y, z) = g(x) + h(y) + l(z)$ для всех точек $(x, y, z) \in K$.

Для произвольной функции $f : K \to \mathbb{R}$ на кресте K определим $g(x) := f(x, 0, 0)$, $h(y) := f(0, y, 0) - f(0, 0, 0)$ и $l(z) := f(0, 0, z) - f(0, 0, 0)$.

(b) Положим $g(0) = f(0, 0, 0)$, $h(0) = 0$, $l(0) = 0$,

$$2g(1) = f(0, 0, 0) + f(1, 1, 0) + f(1, 0, 1) - f(0, 1, 1),$$

$$2h(1) = -f(0, 0, 0) + f(1, 1, 0) - f(1, 0, 1) + f(0, 1, 1)) \quad \text{и}$$

$$2l(1) = -f(0, 0, 0) - f(1, 1, 0) + f(1, 0, 1) + f(0, 1, 1).$$

(c)* Аналогично задаче 11(d) [St89, §2, Лемма 23.ii].

Гладкая базисность.

13. (a), (b), (c) Аналогично задачам 6(a), 3(a) и 3(b).

14. (a) Пусть $f(x, y)$ - дифференцируемая функция. Тогда $f(x, |x|) - f(0, 0) = ax + b|x| + \alpha(x, |x|)|(x, |x|)|$. Положим $h(y) = by$, $g(x) = f(x, |x|) - b|x|$. Более подробно решение изложено вместе с задачей 16c.

(b) Предположим, что данная ломаная дифференцируемо базисна. Мы знаем, что функция f дифференцируема в точках $(-1, 1)$ и $(1, 1)$. Следовательно, для достаточно малого $d > 0$ имеют место следующие соотношения:

$$f(-1 + d, 1 - d) - f(-1, 1) = f_1 d - f_2 d + \alpha_{(-1,1)}(d, -d)|(d, -d)|,$$

$$f(-1-d,1-d)-f(-1,1)=-f_1d-f_2d+\alpha_{(-1,1)}(-d,-d)|(-d,-d)|,$$

$$f(1+d,1-d)-f(1,1)=f_3d-f_4d+\alpha_{(1,1)}(d,-d)|(d,-d)| \quad \text{и}$$

$$f(1-d,1-d)-f(1,1)=-f_3d-f_4d+\alpha_{(1,1)}(-d,-d)|(-d,-d)|.$$

Кроме того $f(x,y)=g(x)+h(y)$, где $g(x)$ и $h(y)$ дифференцируемы. Значит,

$$f(-1+d,1-d)-f(-1,1)=g(-1+d)-g(-1)+h(1-d)-h(1)=g'(-1)d-h'(1)d+\alpha(d)d \quad \text{и}$$

$$f(-1-d,1-d)-f(-1,1)=g(-1-d)-g(-1)+h(1-d)-h(1)=-g'(-1)d-h'(1)d+\alpha(d)d.$$

Таким образом, $h'(1)=f_2$ (и $g'(-1)=f_1$). Аналогично $h'(1)=f_4$. Следовательно, $h'(1)=f_2=f_4$. Рассмотрим теперь функцию $f(x,y)=xy$, для неё $f_2=-1\neq f_4=1$.

(c) Предположим, что эта пополненная молния дифференцируемо базисна. Положим $a_n=([\frac{n+1}{2}]^{-1/2},[\frac{n}{2}]^{-1/2})$, $f(a_n):=\frac{(-1)^n}{n}$, $n=2,3,\ldots$. Если $f(x,y)=g(x)+h(y)$ для некоторых функций $g(x)$ и $h(y)$, тогда $f(a_2)-f(a_3)+f(a_4)-\ldots$ сходится к $g(1)-g(0)$ (аналогично задаче 7d). Но это противоречит расходимости ряда $\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$

(d) Не ограничивая общности можно считать, что $f(0,0)=0$, тогда возьмём $g(0)=0$ и $h(0)=0$. Положим

$$h(2^{-k})=f(2^{-(k+1)},2^{-k})-f(2^{-(k+1)},2^{-(k+1)})+f(2^{-(k+2)},2^{-(k+1)})-\ldots,$$

$$g(2^{-k})=f(2^{-k},2^{-k})-f(2^{-(k+1)},2^{-k})+f(2^{-(k+1)},2^{-(k+1)})-\ldots,$$

где правые части суть суммы знакопеременных рядов.

Теперь $g(x)$ и $h(y)$ могут быть продолжены до дифференцируемых функций $\mathbb{R}\to\mathbb{R}$.

15. (a) Определим

$$w(0)=0, \quad w(4^{-i}+4^{-3i})=w(4^{-i})=0 \quad \text{и} \quad w(4^{-i}+4^{-3i-1})=2^{3i} \quad \text{для} \quad i=1,2,3,\ldots.$$

Продолжим теперь эту функцию кусочно-линейно до функции $w:[0,1]\to\mathbb{R}$. Для каждого $x\in[0,1]$ определим $f(x,-x)$ как площадь под графиком функции w на отрезке $[0,x]$. На остальном кресте положим $f(x,y)=0$.

Предположим, что $f(x,y)=g(x)+h(y)$ для некоторых дифференцируемых g и h. Не ограничивая общности, будем считать, что $g(0)=h(0)=0$. Докажем, что g не дифференцируема в точке $x=1/4$. (Таким же способом можно доказать, что g не дифференцируема в любой точке вида $x=4^{-i}$.)

Рассматривая две бесконечные молнии из точек креста, начинающиеся в точках $(\frac{1}{4}+d,-\frac{1}{4}-d)$ и $(\frac{1}{4},-\frac{1}{4})$ и сходящиеся к точке $(0,0)$, заключаем, что

$$g\left(\frac{1}{4}+d\right)-g\left(\frac{1}{4}\right)=f\left(\frac{1}{4}+d,-\frac{1}{4}-d\right)-f\left(\frac{1}{4},-\frac{1}{4}\right)+f\left(\frac{1}{4^2}+\frac{d}{4},-\frac{1}{4^2}-\frac{d}{4}\right)-f\left(\frac{1}{4^2},-\frac{1}{4^2}\right)+\ldots$$

Для произвольного положительного $d<\frac{1}{4}$ существует k такое, что $4^{-3i}<d/4^{i-1}$ для всех $i>k$ и $4^{-3k}\geq d/4^{k-1}$. В частности $4^{-2k}\geq 4d>4^{-2(k+1)}$. Таким образом,

$$2(g\left(\frac{1}{4}+d\right)-g\left(\frac{1}{4}\right))>2^{-3(k+1)}+2^{-3(k+2)}+2^{-3(k+3)}\cdots>2^{-3(k+1)}\geq\frac{(4d)^{3/4}}{8}.$$

А это противоречит дифференцируемости функции g в точке $\frac{1}{4}$.

(b) **Гипотеза.** Ответ — нет. Доказательство аналогично задаче 15(a).

(c)** **Гипотеза.** Кусочно-линейный граф в $\mathbb{R}^2$ является гладко базисным тогда и только тогда, когда он не содержит сколь угодно длинных молний, и для любых двух *сингулярных* точек a и b выполнено $x(a) \neq x(b)$ и $y(a) \neq y(b)$. Точка $a \in K$ называется *сингулярной*, если пересечение K с любым диском с центром в a не является прямолинейным отрезком.

16. См. отрывок из [RZ02] на отдельной странице.

Мотивировки.

При решении 13-й проблемы Гильберта [Ar58] появилось понятие *базисного вложения* (ссылки даются по возможности не на оригинальные работы, а на обзоры). Основной результат настоящего цикла задач (задача 8b) — элементарное решение [MT03] 'половины' проблемы Арнольда [Ar58'] о характеризации базисных подмножеств плоскости [St89]. См. также [Vo81, Vo82, Sk95, Ku00, Ku03]. Данный цикл задач пересекается с [KS97, KS98] только по нескольким тривиальным задачам. Важнейшие нерешенные задачи данного цикла посвящены попыткам характеризации *гладко базисных* подмножеств плоскости [RZ02].

Мы благодарим В.И. Арнольда и С.М. Воронина за полезные обсуждения.

Литература

[Ar58] В. И. Арнольд, *О представлении функций нескольких переменных в виде суперпозиции функций меньшего числа переменных,* Мат. Просвещение, 3 (1958), 41–61.

[Ar58'] В. И. Арнольд, *Проблема 6,* Мат. Просвещение, 3 (1958), 273-274.

[Vo81] С. М. Воронин, Функциональный анализ, 15 (1981).

[Vo82] С. М. Воронин, Функциональный анализ, 16 (1982).

[KS97] В. Курлин и А. Скопенков, *Базисные вложения графов в плоскость*, Мат. Образование, 3 (1997), 105–113.

[KS98] В. Курлин и А. Скопенков, *Базисные вложения графов в плоскость,* в кн.: 9-я летняя конференция Турнира Городов, изд-во МЦНМО (1998), 34–44, 106–113.

[Ku00] V. Kurlin, *Basic embeddings into products of graphs,* Topol.Appl. 102 (2000), 113–137.

[Ku03] V. A. Kurlin, *Basic embeddings of graphs and the Dynnikov method of three-pages embeddings (in Russian),* Uspekhi Mat. Nauk, 58:2 (2003), 163–164. English transl.: Russian Math. Surveys, 58:2 (2003).

[MK03] N. Mramor-Kosta and E. Trenklerova, *On basic embeddings of compacta into the plane,* Bull. Austral. Math. Soc. 68 (2003), 471–480.

[RZ02] D. Repovs and M. Zeljko, *On basic embeddings into the plane,* preprint (2002)

[Sk95] A. Skopenkov, *A description of continua basically embeddable in* $\mathbb{R}^2$*,* Topol. Appl. 65 (1995), 29–48.

[St89] Y. Sternfeld, *Hilbert's 13th problem and dimension,* Lect. Notes Math. 1376 (1989), 1–49.

HILBERT'S 13-TH PROBLEM AND BASIC PLANAR SETS

A. Skopenkov and I. Shnurnikov

PROBLEMS PROPOSED BEFORE THE INTERMEDIATE FINISH.

Motivation.

In the course of the solution of the Hilbert's 13-th problem the notion of *basic embedding* appeared. The main result of the present sequence of problems (problem 8b) is an elementary solution of 'half' of the Arnold problem on the characterization of basic subsets of the plane. The most important unsolved problems here concern the characterization of *smoothly basic* subsets of the plane.

The more difficult problems are marked by a star, and unsolved problems by two stars. If the statement of a problem is an assertion, then it is required to prove this assertion.

Discontinuously basic subsets.

1. (a) Is it true that for any four numbers $f_{11}, f_{12}, f_{21}, f_{22}$ there exist four numbers g_1, g_2, h_1, h_2 such that $f_{ij} = g_i + h_j$ for each $i, j = 1, 2$?

(b) Andrey Nikolaevich and Vladimir Igorevich play the 'Dare you to decompose!' game. Some cells of chessboard are marked. A. N. writes numbers in the marked cells as he wishes. V. I. looks at the written numbers and chooses (as he wishes) 16 numbers $a_1, \dots, a_8, b_1, \dots, b_8$ as 'weights' of the columns and the lines. If each number in a marked cell turns out to be equal to the sum of weights of the line and the row (of the cell), then V. I. wins, and in the opposite case (i.e., when the number in at least one marked cell is not equal to the sum of weights of the line and the row) A. N. wins.

Prove that A. N. can win no matter how V. I. plays if and only if there does not exist a closed route of a rook turning only at marked cells (the route is not required to pass through each marked cell).

Let $\mathbb{R}^2$ be the plane with a fixed coordinate system. Let $x(a)$ and $y(a)$ be the coordinates of a point $a \in \mathbb{R}^2$. An ordered set (either finite or infinite) $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ is called an *array* if for each i we have $a_i \neq a_{i+1}$ and $x(a_i) = x(a_{i+1})$ for even i and $y(a_i) = y(a_{i+1})$ for odd i. It is not assumed that points of an array are distinct. An array is called *closed* if $a_1 = a_n$.

2. Consider a closed array $\{a_1, \dots, a_n = a_1\}$. A *decomposition* for such an array is an assignment of numbers at the projections of the points of the array on the x-axis and on the y-axis. Is it possible to put numbers $f_1, \dots, f_n \in \mathbb{R}$, where $f_1 = f_n$, at the points of the array so that for each decomposition there exists an f_i that is not equal to the sum of the two numbers at $x(a_i)$ and $y(a_i)$?

A subset $K \subset \mathbb{R}^2$ is called *discontinuously basic* if for each function $f : K \to \mathbb{R}$ there exist functions $g, h : \mathbb{R} \to \mathbb{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

3. (a) The segment $K = 0 \times [0, 1] \subset \mathbb{R}^2$ is discontinuously basic.

(b) The cross $K = 0 \times [-1, 1] \cup [-1, 1] \times 0 \subset \mathbb{R}^2$ is discontinuously basic.

4. (a) *A criterion for a subset of the plane to be discontinuously basic.* A subset of the plane is discontinuously basic if and only if it does not contain any closed arrays.

(b) Given a set of marked cells in the cube $8 \times 8 \times 8$, how can we see who wins in the 3D analogue of the 'Dare you to decompose!' game? In this analogue V. I. tries to choose 24 numbers $a_1, \dots, a_8, b_1, \dots, b_8, c_1, \dots, c_8$ so that the number at the cell (i, j, k) would be equal to the sum $a_i + b_j + c_k$ of the three weights.

(c)** Define discontinuous basic subsets of the 3-space. Discover and prove the 3D analogue of the above criterion.

Continuously basic subsets.

Denote by $|z,z_0| = |(x,y),\ (x_0,y_0)| = \sqrt{(x-x_0)^2+(y-y_0)^2}$ the ordinary distance between points $z=(x,y)$ и $z_0=(x_0,y_0)$ of the plane. Let K be a subset of $\mathbb{R}^2$. A function $f:K\to\mathbb{R}$ is called *continuous* if for each point $z_0\in K$ and number $\varepsilon>0$ there exists a number $\delta>0$ such that for each point $z\in K$ if $|z,z_0|<\delta$, then $|f(z)-f(z_0)|<\varepsilon$. It is sometimes convenient to write (x,y) instead of z.

5. (a) The function $f(x,y)=\sqrt{x^2+y^2}$ is continuous on the plane.

(b) The function $f(x,y)$ equal to the integer part of $x+y$, is not continuous on the plane.

(c) Let $a_1,\dots,a_n$ be distinct points of $K\subset\mathbb{R}^2$. Prove that there exists a continuous function $f:K\to\mathbb{R}$ such that $f(a_i)=(-1)^i$ and $|f(x)|\le 1$ for each $x\in K$.

(d) Let $K=\{a_1,\dots,a_{4n+4}\}$ be an array of $4n+4$ distinct points and $f_1,\dots,f_{4n+4}$ numbers such that $|(-1)^i-f_i|\le\frac{1}{2n}$. Let $g(x(a_i)),h(y(a_i))$, $i=1,\dots,4n+4$, be numbers such that $f_i=g(x(a_i))+h(y(a_i))$ for each i. Prove that $\max_i\{g(x(a_i))\}>n$.

In the sequel all functions are assumed to be continuous.

A subset $K\subset\mathbb{R}^2$ is called *(continuously) basic* if for each continuous function $f:K\to\mathbb{R}$ there exist continuous functions $g,h:\mathbb{R}\to\mathbb{R}$ such that $f(x,y)=g(x)+h(y)$ for each point $(x,y)\in K$.

6. (a) A closed array is not basic.

(b) The segment $K=0\times[0,1]\subset\mathbb{R}^2$ is basic.

(c) The cross $K=0\times[-1,1]\cup[-1,1]\times 0\subset\mathbb{R}^2$ is basic.

7. (a) If a subset of the plane is basic, then it is discontinuously basic.

(b) A *completed array* is the union of a point $a_0\in\mathbb{R}^2$ with an infinite array $\{a_1,\dots,a_n,\dots\}\subset\mathbb{R}^2$ of distinct points which *converges* to the point a_0 (i.e. for each $\varepsilon>0$ there exists a positive integer N such that for each $i>N$ we have $|a_i,a_0|<\varepsilon$). Prove that any completed array is not basic. (Note that it is discontinuously basic).

(c) Let $[a,b]$ be the rectilinear arc which connects points a and b. Prove that the cross $K=[(-1,-2),(1,2)]\cup[(-1,1),(1,-1)]$ is not basic.

(d) Let $m_{ij}=2-3\cdot2^{-i}+j\cdot2^{-2i}$. Consider the set of points $(x_{i,2l},x_{i,2l})$ and $(x_{i,2l-1},x_{i,2l-2})$, where i varies from 1 to ∞ and $l=1,2,3,\dots,2^{i-1}$. Prove that this subset of the plane does not contain any infinite arrays but contains arbitrary long arrays.

(e) The union of the set from the previous problem and the point $(2,2)$ is not basic.

8. Let $K\subset\mathbb{R}^2$ be the image of an arc $[0,1]$ under a continuous map $[0,1]\to\mathbb{R}^2$.

(a) Each continuous function $f:K\to\mathbb{R}$ assumes its lowest value and greatest value. Hint: reduce this problem to an analogous theorem on continuous functions $[0,1]\to\mathbb{R}$.

(b)* If K contains arbitrary long arrays, then K is not basic.

Hint. Assume that K contains arbitrary long arrays and is basic. We may assume that points of each array are distinct. Therefore for each n there is an array $\{a_1^n,\dots,a_{4n+4}^n\}$ of $4n+4$ distinct points in K. Then there exists continuous function $f_n:K\to\mathbb{R}$ such that $f_n(a_i^n)=(-1)^i$ and $|f_n(x)|\le 1$ for each $x\in K$. For each function $G:K\to\mathbb{R}$ its *maximum* is $\|G\|:=\max_{x\in K}|G(x)|$. Let $f:K\to\mathbb{R}$ and $g,h:\mathbb{R}\to\mathbb{R}$ be functions such that $\|f-f_n\|<1/2n$ and $f(x,y)=g(x)+h(y)$ for each $(x,y)\in K$. Then $\|g\|>n$...

BASIC PLANAR SETS

A. Skopenkov and I. Shnurnikov

problems proposed before the intermediate finish.

Motivation.

In the course of the solution of the Hilbert's 13-th problem the notion of *basic embedding* appeared. The main result of the present sequence of problems (problem 8b) is an elementary solution of 'half' of the Arnold problem on the characterization of basic subsets of the plane. The most important unsolved problems here concern the characterization of *smoothly basic* subsets of the plane.

The more difficult problems are marked by a star, and unsolved problems by two stars. If the statement of a problem is an assertion, then it is required to prove this assertion.

Discontinuously basic subsets.

1. (a) Is it true that for any four numbers $f_{11}, f_{12}, f_{21}, f_{22}$ there exist four numbers g_1, g_2, h_1, h_2 such that $f_{ij} = g_i + h_j$ for each $i, j = 1, 2$?

(b) Andrey Nikolaevich and Vladimir Igorevich play the 'Dare you to decompose!' game. Some cells of chessboard are marked. A. N. writes numbers in the marked cells as he wishes. V. I. looks at the written numbers and chooses (as he wishes) 16 numbers $a_1, \dots, a_8, b_1, \dots, b_8$ as 'weights' of the columns and the lines. If each number in a marked cell turns out to be equal to the sum of weights of the line and the row (of the cell), then V. I. wins, and in the opposite case (i.e., when the number in at least one marked cell is not equal to the sum of weights of the line and the row) A. N. wins.

Prove that V. I. can win no matter how A. N. plays if and only if there does not exist a closed route of a rook starting and turning only at marked cells (the route is not required to pass through each marked cell).

Let $\mathbb{R}^2$ be the plane with a fixed coordinate system. Let $x(a)$ and $y(a)$ be the coordinates of a point $a \in \mathbb{R}^2$. An ordered set (either finite or infinite) $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ is called an *array* if for each i we have $a_i \neq a_{i+1}$ and $x(a_i) = x(a_{i+1})$ for even i and $y(a_i) = y(a_{i+1})$ for odd i. It is not assumed that points of an array are distinct. An array is called *closed* if $a_1 = a_{2l+1}$.

2. Consider a closed array $\{a_1, \dots, a_n = a_1\}$. A *decomposition* for such an array is an assignment of numbers at the projections of the points of the array on the x-axis and on the y-axis. Is it possible to put numbers $f_1, \dots, f_n \in \mathbb{R}$, where $f_1 = f_n$, at the points of the array so that for each decomposition there exists an f_i that is not equal to the sum of the two numbers at $x(a_i)$ and $y(a_i)$?

A subset $K \subset \mathbb{R}^2$ is called *discontinuously basic* if for each function $f : K \to \mathbb{R}$ there exist functions $g, h : \mathbb{R} \to \mathbb{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

3. (a) The segment $K = 0 \times [0, 1] \subset \mathbb{R}^2$ is discontinuously basic.

(b) The cross $K = 0 \times [-1, 1] \cup [-1, 1] \times 0 \subset \mathbb{R}^2$ is discontinuously basic.

4. (a) *A criterion for a subset of the plane to be discontinuously basic.* A subset of the plane is discontinuously basic if and only if it does not contain any closed arrays.

(b)** Given a set of marked unit cubes in the cube $8 \times 8 \times 8$, how can we see who wins in the 3D analogue of the 'Dare you to decompose!' game? In this analogue V. I. tries to choose 24 numbers $a_1, \dots, a_8, b_1, \dots, b_8, c_1, \dots, c_8$ so that the number at the unit cube (i, j, k) would be equal to the sum $a_i + b_j + c_k$ of the three weights.

(c)** Define discontinuous basic subsets of the 3-space. Discover and prove the 3D analogue of the above criterion.

Continuously basic subsets.

Denote by $|z, z_0| = |(x, y),\ (x_0, y_0)| = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ the ordinary distance between points $z = (x, y)$ и $z_0 = (x_0, y_0)$ of the plane. Let K be a subset of $\mathbb{R}^2$. A function $f : K \to \mathbb{R}$ is called *continuous* if for each point $z_0 \in K$ and number $\varepsilon > 0$ there exists a number $\delta > 0$ such that for each point $z \in K$ if $|z, z_0| < \delta$, then $|f(z) - f(z_0)| < \varepsilon$. It is sometimes convenient to write (x, y) instead of z.

5. (a) The function $f(x, y) = \sqrt{x^2 + y^2}$ is continuous on the plane.

(b) The function $f(x, y)$ equal to the integer part of $x + y$ is not continuous on the plane.

(c) Let $a_1, \dots, a_n$ be distinct points of $K \subset \mathbb{R}^2$. Prove that there exists a continuous function $f : K \to \mathbb{R}$ such that $f(a_i) = (-1)^i$ and $|f(x)| \le 1$ for each $x \in K$.

(d) Let $K = \{a_1, \dots, a_{4n+4}\}$ be an array of $4n + 4$ distinct points and $f_1, \dots, f_{4n+4}$ be numbers such that $|(-1)^i - f_i| \le \frac{1}{2n}$. Let $g(x(a_i)), h(y(a_i))$, $i = 1, \dots, 4n + 4$, be numbers such that $f_i = g(x(a_i)) + h(y(a_i))$ for each i. Prove that $\max_i |g(x(a_i))| > n$.

In the sequel all functions are assumed to be continuous.

A subset $K \subset \mathbb{R}^2$ is called *(continuously) basic* if for each continuous function $f : K \to \mathbb{R}$ there exist continuous functions $g, h : \mathbb{R} \to \mathbb{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

6. (a) A closed array is not basic.

(b) The segment $K = 0 \times [0, 1] \subset \mathbb{R}^2$ is basic.

(c) The cross $K = 0 \times [-1, 1] \cup [-1, 1] \times 0 \subset \mathbb{R}^2$ is basic.

7. (a) If a subset of the plane is basic, then it is discontinuously basic.

(b) A *completed array* is the union of a point $a_0 \in \mathbb{R}^2$ with an infinite array $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ of distinct points which *converges* to the point a_0 (i.e. for each $\varepsilon > 0$ there exists a positive integer N such that for each $i > N$ we have $|a_i, a_0| < \varepsilon$). Prove that any completed array is not basic. (Note that it is discontinuously basic).

(c) Let $[a, b]$ be the rectilinear arc which connects points a and b. Prove that the cross $K = [(-1, -2), (1, 2)] \cup [(-1, 1), (1, -1)]$ is not basic.

(d) Let $x_{ij} = 2 - 3 \cdot 2^{-i} + j \cdot 2^{-2i}$. Consider the set of points $(x_{i,2l}, x_{i,2l})$ and $(x_{i,2l}, x_{i,2l-2})$, where i varies from 1 to ∞ and $l = 1, 2, 3, \dots, 2^{i-1}$. Prove that this subset of the plane does not contain any infinite arrays but contains arbitrary long arrays.

(e) The union of the set from the previous problem and the point $(2, 2)$ is not basic.

8. Let $K \subset \mathbb{R}^2$ be the image of an arc $[0, 1]$ under a continuous map $[0, 1] \to \mathbb{R}^2$.

(a) Each continuous function $f : K \to \mathbb{R}$ assumes its lowest value and greatest value. Hint: reduce this problem to an analogous theorem on continuous functions $[0, 1] \to \mathbb{R}$.

(b)* If K contains arbitrary long arrays, then K is not basic.

Hint. Assume that K contains arbitrary long arrays and is basic. We may assume that points of each array are distinct. Therefore for each n there is an array $\{a^n_1, \dots, a^n_{4n+4}\}$ of $4n+4$ distinct points in K. Then there exists continuous function $f_n : K \to \mathbb{R}$ such that $f_n(a^n_i) = (-1)^i$ and $|f_n(x)| \le 1$ for each $x \in K$. For each function $G : K \to \mathbb{R}$ its *maximum* is $|G| := \max_{x \in K} |G(x)|$. Let $f : K \to \mathbb{R}$ and $g, h : \mathbb{R} \to \mathbb{R}$ be functions such that $|f - f_n| < 1/2n$ and $f(x, y) = g(x) + h(y)$ for each $(x, y) \in K$. Then $|g| > n$...

BASIC PLANAR SETS

A. Skopenkov and I. Shnurnikov (presented by them and V. Gorin)

Problems proposed after the intermediate finish.

The Sternfeld criterion for being a basic subset.

Further hints to problem 8b. Define integers s_n and functions $F_n : K \to \mathbb{R}$ inductively as follows. Set $s_0 = 1$ and $F_0 = 0$. Suppose now that F_{n-1} and s_{n-1} are defined. If F_{n-1} is not representable as $G_{n-1}(x) + H_{n-1}(y)$, then we are done. If it is representable in this way, then take

$$s_n > s_{n-1}!(|G_{n-1}| + n) \quad \text{and} \quad F_n = F_{n-1} + \frac{f_{s_n}}{s_{n-1}!}$$

It remains to prove that if we can construct in this way an infinite number of s_n and F_n, then the function

$$F = \lim_{n\to\infty} F_n = \sum_{n=1}^{\infty} \frac{f_{s_n}}{s_{n-1}!}$$

is not representable as $G(x) + H(y)$.

A sequence of points $\{a_1, \dots, a_n, \dots\} \subset \mathbb{R}^2$ *converges to a point* $a \in \mathbb{R}^2$ if for each $\varepsilon > 0$ there exists an integer N such that for each $i > N$ we have $|a_i, a| < \varepsilon$.

A subset $K \subset \mathbb{R}^2$ of the plane is called *closed*, if for each infinite sequence of points $a_i \in K$ converging to a point a this point belongs to K.

9. (a) A subset $K \subset \mathbb{R}^2$ of the plane is closed if and only if for each point $a \notin K$ there exists $\varepsilon > 0$ such that if for a point b of the plane we have $|a, b| \le \varepsilon$, then b does not belong to K.

(b) The image of an arc under a continuous map $[0, 1] \to \mathbb{R}^2$ is a closed subset of $\mathbb{R}^2$.

The Sternfeld criterion for being a basic subset. A closed bounded subset $K \subset \mathbb{R}^2$ of the plane is basic if and only if K does not contain arbitrary long arrays.

10. (a) The criterion is false without the assumption that K closed.

(b) The criterion is false without the assumption that K bounded.

(c) Prove the 'only if' part ($\Rightarrow$) of the criterion.

Suppose that K is a subset of $\mathbb{R}^2$. For every point $v \in K$ consider the pair of lines passing through v and parallel to the x-axis and the y-axis. If one of these two lines intersects K only at point v, we colour v in white. Define $E(K)$ as the set of noncoloured points of K:

$$E(K) = \{v \in K : \ |K \cap (x = x(v))| \ge 2 \text{ and } |K \cap (y = y(v))| \ge 2\}.$$

Let $E^2(K) = E(E(K))$, $E^3(K) = E(E(E(K)))$ etc.

11. (a) If a subset K of the plane does not contain arbitrary long arrays, then $E^n(K) = \emptyset$ for some n.

(b) Prove the converse statement.

(c)* If K is a closed bounded subset of $\mathbb{R}^2$ and $E(K) = \emptyset$, then K is basic.

(d)* Prove the 'if' part ($\Leftarrow$) of the criterion. Remark. It can be proven first that for piecewise-linear maps f there is a decomposition $f(x, y) = g(x) + h(y)$ with $|g| + |h| < C_n|f|$, where C_n depends only on that n for which $E^n(K) = \emptyset$.

12. (a) Define the property of being a (continuously) basic subset of $\mathbb{R}^3$. Prove that the hedgehog $0 \times 0 \times [-1, 1] \cup 0 \times [-1, 1] \times 0 \cup [-1, 1] \times 0 \times 0 \subset \mathbb{R}^3$ is basic.

(b) The set of 4 points $(0,0,0)$; $(1,1,0)$; $(0,1,1)$; $(1,0,1)$ is basic. (But $E^n(K)\neq\emptyset$ for each n, see below.)

(c)* Let K be a closed bounded subset of $\mathbb{R}^3$. Define $E(K)$ analogously to the above, only instead of lines we use planes orthogonal to the axes:

$$E(K)=\{v\in K:\ |K\cap(x=x(v))|\geq 2,\ |K\cap(y=y(v))|\geq 2 \text{ and } |K\cap(z=z(v))|\geq 2\}.$$

Prove that if $E^n(K)=\emptyset$ for some n, then K is basic.

Smoothly basic subsets of the plane.

Let K be a subset of the plane $\mathbb{R}^2$. A function $f:K\to\mathbb{R}$ is called *differentiable* if for each point $z_0\in K$ there exist a vector $a\in\mathbb{R}^2$ and infinitesimal function $\alpha:\mathbb{R}^2\to\mathbb{R}$ such that for each point $z\in K$

$$f(z)=f(z_0)+a\cdot(z-z_0)+\alpha(z-z_0)|z,z_0|.$$

Here the dot denotes scalar product of vectors $a=(f_x,f_y)$ and $z-z_0=(x,y)$, i.e. $a\cdot(z-z_0)=xf_x+yf_y$. A function $\alpha:\mathbb{R}^2\to\mathbb{R}$ is *infinitesimal*, if for each number $\varepsilon>0$ there exists a number $\delta>0$ such that for each point $(x,y)\in\mathbb{R}^2$

$$\text{if}\quad \sqrt{x^2+y^2}<\delta,\quad \text{then}\quad |\alpha(x,y)|<\varepsilon.$$

A subset $K\subset\mathbb{R}^2$ of the plane is called *differentiably basic* if for each differentiable function $f:K\to\mathbb{R}$ there exist differentiable functions $g:\mathbb{R}\to\mathbb{R}$ and $h:\mathbb{R}\to\mathbb{R}$ such that $f(x,y)=g(x)+h(y)$ for each point $(x,y)\in K$.

13. (a) (b) (c) Solve the analogues of problem 6 for differentiably basic sets.

14. (a) The graph of the function $y=|x|$, where $x\in[-1,1]$, is differentiably basic.

(b) The broken line whose consecutive vertices are $(-2,0)$, $(-1,1)$, $(0,0)$, $(1,1)$ and $(2,0)$ is not differentiably basic. (Note that it is continuously basic).

(c) The completed array $\{([\frac{n+1}{2}]^{-1/2},[\frac{n}{2}]^{-1/2})\}_{n=2}^{\infty}\cup\{(0,0)\}$ is not differentiably basic. (Note that it is also not continuously basic.)

(d) The completed array $\{(2^{-[\frac{n+1}{2}]},2^{-[\frac{n}{2}]})\}_{n=1}^{\infty}\cup\{(0,0)\}$ is differentiably basic. (Note that it is not continuously basic.)

(e)** Is there a continuous map of arc $[0,1]$ to $\mathbb{R}^2$, whose image is differentiably basic but not continuously basic?

15. (a) The cross $K=[(-1,-2),(1,2)]\cup[(-1,1),(1,-1)]$ is not differentiably basic.

(b) Is the subset $\{(t^2,\frac{t^2}{(1+t)^2})\}_{t\in[-\frac12,\frac12]}$ of the plane differentiably basic?

(c)** Find a criterion for graphs in $\mathbb{R}^2$ to be differentiably basic.

16. Let $r\geq 0$ be an integer and $K\in\mathbb{R}^2$ a subset. A function $f:K\to\mathbb{R}$ is called r *times differentiable* if for each point $z_0\in K$ there exist a polynomial $\overline{f}(z)=\overline{f}(x,y)$ degree at most r of 2 variables x and y and infinitesimal function $\alpha:\mathbb{R}^2\to\mathbb{R}$ such that $f(z)=\overline{f}(z-z_0)+\alpha(z-z_0)|z,z_0|^r$ for each point $z\in K$. (This definition differs from the one generally accepted.)

(a) Functions differentiable zero times are exactly continuous functions, and functions differentiable one time are exactly differentiable functions.

(b) For each positive integer r define the property of being an r times differentiably basic subset of the plane $\mathbb{R}^2$.

(c) For each integer $k\geq 0$ there is a subset of the plane which is r times differentiably basic for $r=0,1\ldots k$ but is not r times differentiably basic for each $r>k$.

(d)** Find a criterion for graphs in $\mathbb{R}^2$ to be r times differentiably basic.

BASIC PLANAR SETS

A. Skopenkov and I. Shnurnikov (presented by them and V. Gorin)

Solutions of problems proposed before the intermediate finish.

Discontinuously basic subsets.

We use definitions given in problems proposed after the intermediate finish.

1. (a) It is not true. If $f_{ij} = g_i + h_j$ for each $i, j = 1, 2$, then $f_{11} + f_{22} = f_{12} + f_{21}$, but this is false for some numbers f_{ij}.

(b) The statement 'only if' follows from the problem 2. Let us prove the 'if' part by induction on the number of the marked cells. If only one cell is marked then we are done. Let K be the set of centres of the marked cells. The set K does not contain any closed array, therefore $\#E(K) < \#K$. So by the induction hypothesis V. I. can win for $E(K)$. Each cell from $K - E(K)$ is the only marked cell on its line or column, thus V. I. can choose the remaining weights for K.

2. Yes, it is. If every f_i is equal to the sum of two numbers at $x(a_i)$ and $y(a_i)$, then $f_1 - f_2 + f_3 - \dots - f_{n-1} = 0$, but this is false for some numbers f_i.

3. (a) Set $h(y) = f(0, y)$ and $g(x) = 0$.

(b) Set $g(x) = f(x, 0)$ and $h(y) = f(0, y) - f(0, 0)$.

4. (a) The statement 'only if' follows from the problem 2. Let us prove the 'if' part. Consider a function $f : K \to \mathbb{R}$. Our aim is to construct functions g and h so that $f(x, y) = g(x) + h(y)$. Two points $a, b \in K$ are called *equivalent* if there is an array $\{a = a_1, \dots, a_n = b\} \subset K$. Now take an equivalence class $K_1 \subset K$. Define function $g : x(K_1) \to \mathbb{R}$ and $h : y(K_1) \to \mathbb{R}$ in the following way. Take any point $a_1 \in K_1$ and set $g(x(a_1)) = f(a_1)$ and $h(y(a_1)) = 0$. If $\{a_1, a_2, \dots, a_{2l}\}$ is an array, then set

$$h(y(a_{2l})) := f(a_{2l}) - f(a_{2l-1}) + \dots - f(a_1) \quad \text{and} \quad g(x(a_{2l})) := f(a_{2l-1}) - f(a_{2l-2}) + \dots + f(a_1).$$

If $\{a_1, a_2, \dots, a_{2l+1}\}$ is an array, then set

$$g(x(a_{2l+1})) := f(a_{2l+1}) - f(a_{2l}) + \dots + f(a_1)$$

($h(y(a_{2l+1}))$ is already defined). Make this construction for each equivalence class. Then set $g = 0$ and $h = 0$ at all other points of $\mathbb{R}$.

Continuously basic subsets.

5. (a) We can set $\delta = \varepsilon$. Then the statement follows from the triangle inequality: $|f(z) - f(z_0)| \le |z, z_0|$.

(b) For the point $(1, 0)$ and $\varepsilon = \frac{1}{2}$ there is no such δ because $|f(1, 0) - f(1 - \frac{\delta}{2}, 0)| = 1 > \frac{1}{2}$.

(c) First define a continuous function $f : \mathbb{R}^2 \to \mathbb{R}$. Denote $s = \min_{i<j} |a_i, a_j|$. Take n disks with centers a_i and radii $\frac{s}{3}$. Outside of these disks set $f = 0$. Inside the i-th disk take f to be $(-1)^i$ in the center a_i, 0 on the boundary and extend it linearly in the distance to a_i. Then restriction of f to $K \subset \mathbb{R}^2$ is a continuous function $K \to \mathbb{R}$.

(d) We have $|(f_2 - f_3 + f_4 - f_5 + \dots - f_{4n+3}) - (4n+2)| \le \frac{4n+2}{2n} \le 3$. Therefore $g(a_2) - g(a_{4n+3}) \ge (4n+2) - 3 > 2n$, which implies the required inequality.

6. (a) If an array $A = \{a_1, \dots, a_{2l+1}\}$ is basic, then $f(a_1) - f(a_2) + \dots + f(a_{n-2}) - f(a_{2l}) = 0$. But this is false for some functions f. Cf. problem 2.

(b),(c) Analogously to problems 3a,3b.

7. (a) If the subset is not discontinuously basic, then it contains a closed array. Hence the statement follows by extension of f on the subset and using problem 6a.

(b) Define function f by $f(a_n) = \frac{(-1)^n}{n}$. Suppose that $f(x, y) = g(x) + h(y)$ for some g and h. Then

$$f(a_1) - f(a_2) + f(a_3) - f(a_4) + \cdots - f(a_{2l}) = h(y(a_1)) - h(y(a_{2l})).$$

Since $\lim_{l\to\infty} h(y_{2l})$ exists and equals to $h(y(a_0))$, it follows that $\sum_{i=1}^{2l}(-1)^i f(a_i)$ converges when $l \to \infty$, which is a contradiction.

(c) The cross contains a completed array

$$a_{4k+1} = (\frac{-1}{4^k}, \frac{1}{4^k}),\ a_{4k+2} = (\frac{1}{2\cdot 4^k}, \frac{1}{4^k}),\ a_{4k+3} = (\frac{1}{2\cdot 4^k}, \frac{-1}{2\cdot 4^k}),\ a_{4k+4} = (\frac{-1}{4^{k+1}}, \frac{-1}{2\cdot 4^k})$$

Define a function f on this array using problem 7(b) and then extend it (e.g. piecewise linearly) to the cross. Then there are no functions g and h such that $f(x, y) = g(x) + h(y)$.

(d) For every i the set $(x_{i,2l}, x_{i,2l})_{l=1}^{2^{i-1}} \cup (x_{i,2l}, x_{i,2l-2})_{l=1}^{2^{i-1}}$ is an array of 2^i points.

(e) Define a function f by

$$f((x_{i,2l}, x_{i,2l})) := \frac{1}{2^i} \quad \text{and} \quad f(x_{i,2l}, x_{i,2l-2}) := -\frac{1}{2^i}.$$

If $f(x, y) = g(x) + h(y)$ for some g and h, then for every i using array of points $(x_{i,2l}, x_{i,2l})$ and $(x_{i,2l}, x_{i,2l-2})$, where $l = 1, 2, 3, \dots 2^{i-1}$, we obtain $h(2 - \frac{3}{2^i}) - h(2 - \frac{2}{2^i}) = 1$. This contradicts to the continuity of h.

A function $f : K \in \mathbb{R}$ is called *bounded*, if there exists a number M such that $|f(x)| < M$ for every $x \in K$.

8. *Lemma.* Every continuous function $f : [0, 1] \to \mathbb{R}$ is bounded.

Proof. Suppose that for each integer n there exists a point $a_n \in [0, 1]$ such that $|f(a_n)| > n$. In the sequence a_n take a subsequence a_{n_i} converging to a point $a \in [0, 1]$. Since the function f is continuous, the sequence $f(a_{n_i})$ converges to the $f(a)$ when $i \to \infty$. But $|f(a_{n_i})| > n_i$ and $f(a_{n_i}) \to \infty$. This contradiction implies the Lemma.

(a) Take a map $k : [0, 1] \to \mathbb{R}^2$ such that $K = k([0, 1])$. Consider a composition $f \circ k : [0, 1] \to \mathbb{R}$. Let s be a minimal number for which $f(k(t)) \le s$ for each $t \in [0, 1]$. If there were no t such that $f(k(t)) = s$ then continuous function $\frac{1}{s - f(k(t))} : [0, 1] \to \mathbb{R}$ would be unbounded. This contradiction shows that f assume its greatest value. Analogously f assume its lowest value.

BASIC PLANAR SETS

A. Skopenkov and I. Shnurnikov (presented by them and V. Gorin)

Solutions of problems proposed after the intermediate finish.

The Sternfeld criterion for being a basic subset.

For function $G : K \to \mathbb{R}$ set $|G| := \max_{x\in K} |G(x)|$.

8. (b) Assume that K contains arbitrary long arrays and is basic. We may assume that points of each array are distinct. Therefore for each n there is an array $\{a_1^n, \dots, a_{4n+4}^n\}$ of $4n+4$ distinct points in K. Then there exists continuous function $f_n : K \to \mathbb{R}$ such that $f_n(a_i^n) = (-1)^i$ and $|f_n(x)| \le 1$ for each $x \in K$. Let $f : K \to \mathbb{R}$ and $g, h : \mathbb{R} \to \mathbb{R}$ be functions such that $|f - f_n| < 1/2n$ and $f(x, y) = g(x) + h(y)$ for each $(x, y) \in K$. Then by the problem 5(d) $|g| > n$. It suffices to assume that $F(x, y) = G(x) + H(y)$ and prove that $|G| > n$ for each n. We have

$$F - F_n = F - F_{n-1} - \frac{f_{s_n}}{s_{n-1}!} = \frac{s_{n-1}!(F - F_{n-1}) - f_{s_n}}{s_{n-1}!}.$$

Set $f = s_{n-1}!(F - F_{n-1})$. Then for $n > 2$ we have $s_n - 1 > s_{n-1}$ and

$$|f - f_{s_n}| = s_{n-1}!|F - F_n| < \frac{1}{(s_n-1)\cdot s_n}\sum_{k=0}^{\infty}\frac{1}{s_{n+1}\cdot\ldots\cdot s_{n+k}} < \frac{1}{(s_n-1)\cdot s_n}\sum_{k=0}^{\infty}\frac{1}{2^k} = \frac{1}{2s_n}.$$

We have

$$f = s_{n-1}!(G(x) - G_{n-1}(x)) + s_{n-1}!(H(y) - H_{n-1}(y)).$$

So by the above problem 5(d) it follows that $s_{n-1}!|G - G_{n-1}| > s_n$. Therefore

$$|G| + |G_{n-1}| \ge |G - G_{n-1}| > \frac{s_n}{s_{n-1}!} > |G_{n-1}| + n. \quad \square$$

9. (a) Let us prove the 'only if' part. Let K be a closed subset of the plane. Suppose that for some point $a = (x, y) \notin K$ and for each $\varepsilon = \frac{1}{n} > 0$ there exists a point $a_n \in K$ (at least one) such that $|a, a_n| \le \frac{1}{n}$. The sequence of points $a_n \in K$ converges to the point a, thus $a \in K$. Contradiction.

Now let us prove the 'if' part. Suppose that a sequence a_n converges to a point a and the point $a = (x, y)$ is not in K. There exists $\varepsilon > 0$ such that for every point $a_n \in K$ the distance $|a, a_n| > \varepsilon$. This is a contradiction.

(b) Take a continuous map $f : [0, 1] \to \mathbb{R}^2$. Suppose that a sequence of points $\{a_i\}_{i=1}^{\infty}$ in the image of f converges to a point a. Select a point $t_i \in f^{-1}(a_i)$. Now take a subsequence of points $\{t_{i_k}\}$ of $\{t_i\}$ converging to a point $t_0 \in [0, 1]$. Since the map f is continuous, the subsequence $f(t_{i_k})$ converges to the point $f(t_0)$. Therefore $a = f(t_0)$, so $f([0, 1])$ is closed.

10. (a) Any infinite array A not containing closed arrays and converging to a point $a \notin A$ is basic. This follows because each function defined on A is continuous.

(b) A counterexample is $\{(k, k)\}_{k=1}^{\infty} \cup \{(k, k-1)\}_{k=1}^{\infty}$.

(c) The proof is the same as that of problem 8(b) by the following Lemma.

Lemma. Let K be a closed bounded subset of the plane. Then every continuous function $f : K \to \mathbb{R}$ is bounded.

11. (a) Suppose that $E^n(K) \neq \emptyset$ for each n. For each n take a point $a_0 \in E^n(K)$. Then there exist points $a_{-1}, a_1 \in E^{n-1}(K)$ such that $x(a_{-1}) = x(a_0)$ and $y(a_1) = y(a_0)$. Analogously there exist points $a_{-2}, a_2 \in E^{n-2}(K)$ such that $\{a_{-2}, a_{-1}, a_0, a_1, a_2\}$ is an array. Analogously we construct an array of $2n+1$ points in K, which is a contradiction.

(b) Suppose that K contains an array of $2n+1$ points $\{a_{-n}, \dots, a_0, \dots, a_n\}$. Then there is an array of $2n-1$ points $\{a_{-n+1}, \dots, a_{n-1}\}$ in $E(K)$. Analogously $a_0 \in E^n(K)$. Thus if $E^n(K) = \emptyset$, then K does not contain an array of $2n+1$ points.

(c)* See solution of the problem 11(d).

(d)* We present a non-elementary solution based on a reformulation of the property of being a basic subset in terms of *bounded linear operators* in *Banach functional spaces.* Denote by $C(X)$ the space of continuous functions on X with the norm $|f| = \sup\{|f(x)| \ : \ x \in X\}$. We may assume that $K \subset I^2 := [0,1] \times [0,1]$. Define a map (*linear superposition operator*)

$$\varphi\colon C(I) \oplus C(I) \to C(K) \quad \text{by} \quad \varphi(g,h)(x,y) := g(x) + h(y).$$

Clearly, the subset $K \subset I^2$ is basic if and only if φ is surjective, or equivalently, epimorphic.

Denote by $C^*(X)$ the space of *bounded linear functions* $C(X) \to \mathbb{R}$ with the norm $|\mu| = \sup\{|\mu(f)| \ : \ f \in C(X),\ |f| = 1\}$. Denote by $pr_x(a)$ and $pr_y(a)$ the projections of a point $a \in K$ on the coordinate axes.

For a subset $K \subset I^2$ define a map (*dual linear superposition operator*)

$$\varphi^*\colon C^*(K) \to C^*(I) \oplus C^*(I) \quad \text{by} \quad \varphi^*\mu(g,h) := (\mu(g \circ pr_x), \mu(h \circ pr_y)).$$

Since $|\varphi^*\mu| \le 2|\mu|$, it follows that φ^* is bounded. By duality, φ is epimorphic if and only if φ^* is monomorphic.

(We remark that φ^* can be injective but not monomorphic. In other words not only some linear relation on $\operatorname{im}\varphi$ can force it to be strictly less than $C(K)$.)

It is clear that φ^* is monomorphic if and only if *there exist* $\varepsilon > 0$ *such that* $|\varphi^*\mu| > \varepsilon|\mu|$ *for each unzero* $\mu \in C^*(K)$.

So it remains to prove that $E^n(K) = \emptyset$ implies the latter condition. We present the proof for $n \in \{1,2\}$. The proof for arbitrary n is analogous. We use the following non-trivial fact: $C^*(X)$ *is the space of σ-additive regular real valued Borel measures on* X (in the sequel we call them simply 'measures'). We have

$$\varphi^*\mu = (\mu_x, \mu_y), \quad \text{where} \quad \mu_x(U) = \mu(pr_x^{-1}U) \quad \text{and} \quad \mu_y(U) = \mu(pr_y^{-1}U) \quad \text{for each Borel set} \quad U \subset I.$$

If $\mu = \mu^+ - \mu^-$ is the decomposition of a measure μ into its positive and negative parts, then $|\mu| = \bar\mu(X)$, where $\bar\mu = \mu^+ + \mu^-$ is the absolute value of μ.

Let D_x (D_y) be the set of points of K which are not shadowed by some other point of K in x- (y-) direction. Take any measure μ on K of the norm 1.

If $n = 1$, then

$$E(K) = \emptyset, \quad \text{then} \quad D_x \cup D_y = K, \quad \text{so} \quad 1 = \bar\mu(K) \le \bar\mu(D_x) + \bar\mu(D_y).$$

Therefore without loss of generality, $\bar\mu(D_x) \ge 1/2$. Since the projection onto the x-axis is injective over D_x, it follows that $|\mu_x| \ge 1/2$, thus the required assertion holds for $\varepsilon = \frac{1}{2}$.

If $n = 2$, then

$$E(E(K)) = \emptyset, \quad \text{hence} \quad D_x \cup D_y = K - E(K), \quad \text{so} \quad E(D_x \cup D_y) = \emptyset.$$

Therefore in the case when $\bar\mu(E(K)) < 3/4$ we have $\bar\mu(D_x \cup D_y) > 1/4$ and without loss of generality $\bar\mu(D_x) > 1/8$. Then as for $n = 1$ we have $|\mu_x| > 1/8$, thus the required assertion holds for $\varepsilon = \frac{1}{8}$.

In the case when $\bar\mu(E(K)) \geq 3/4$ we have $\bar\mu(K - E(K)) \leq 1/4$. By the case $n = 1$ above without loss of generality $\bar\mu_x(pr_x(E(K))) \geq \bar\mu(E(K))/2$. Hence $|\mu_x| \geq \frac{1}{2}\cdot\frac{3}{4} - \frac{1}{4} = \frac{1}{8}$, thus the required assertion holds for $\varepsilon = \frac{1}{8}$.

If an embedding $K \subset \mathbb{R}^2$ is basic, then we can prove that φ^* is monomorphic without use of φ as follows. Define a linear operator

$$\Psi\colon C^*(I) \oplus C^*(I) \to C^*(K) \quad \text{by} \quad \Psi(\mu_x, \mu_y)(f) = \mu_x(g) + \mu_y(h),$$

where $g, h \in C(I)$ are such that $g(0) = 0$ and $f(x, y) = g(x) + h(y)$ for $(x, y) \in K$. Clearly, $\Psi\varphi^* = \mathrm{id}$ and Ψ is bounded, hence φ^* is monomorphic.

12. (a) A subset $K \subset \mathbb{R}^3$ is called *(continuously) basic* if for each continuous function $f : K \to \mathbb{R}$ there exist continuous functions $g, h, l : \mathbb{R} \to \mathbb{R}$ such that $f(x, y, z) = g(x)+h(y)+l(z)$ for each point $(x, y, z) \in K$.

For each functuion $f : K \to \mathbb{R}$ on the cross K define $g(x) := f(x, 0, 0)$, $h(y) := f(0, y, 0) - f(0, 0, 0)$ and $l(z) := f(0, 0, z) - f(0, 0, 0)$.

(b) Set $g(0) = f(0, 0, 0)$, $h(0) = 0$, $l(0) = 0$,

$$2g(1) = f(0, 0, 0) + f(1, 1, 0) + f(1, 0, 1) - f(0, 1, 1)$$

$$2h(1) = -f(0, 0, 0) + f(1, 1, 0) - f(1, 0, 1) + f(0, 1, 1))$$

$$\text{and} \quad 2l(1) = -f(0, 0, 0) - f(1, 1, 0) + f(1, 0, 1) + f(0, 1, 1).$$

(c)* Analogously to problem 11(d) [St89, §2, Lemma 23.ii].

Smoothly basic subsets of the plane.

13. (a), (b), (c) Analogously to problems 6(a), 3(a) and 3(b).

14. (a) See solution of 16c.

(b) Suppose the broken line is differentiably basic. Since f is differentiable at points $(-1, 1)$ and $(1, 1)$, the following relations hold for sufficiently small $d > 0$:

$$f(-1 + d, 1 - d) - f(-1, 1) = f_1 d - f_2 d + \alpha_{(-1,1)}(d, -d)|(d, -d)|,$$

$$f(-1 - d, 1 - d) - f(-1, 1) = -f_1 d - f_2 d + \alpha_{(-1,1)}(-d, -d)|(-d, -d)|,$$

$$f(1 + d, 1 - d) - f(1, 1) = f_3 d - f_4 d + \alpha_{(1,1)}(d, -d)|(d, -d)| \quad \text{and}$$

$$f(1 - d, 1 - d) - f(1, 1) = -f_3 d - f_4 d + \alpha_{(1,1)}(-d, -d)|(-d, -d)|.$$

Also we have $f(x, y) = g(x) + h(y)$ and both $g(x)$, $h(y)$ are differentiable. Hence

$$f(-1+d, 1-d)-f(-1, 1) = g(-1+d)-g(-1)+h(1-d)-h(1) = g'(-1)d-h'(1)d+\alpha(d)d \quad \text{and}$$

$$f(-1 - d, 1 - d) - f(-1, 1) = g(-1 - d) - g(-1) + h(1 - d) - h(1) = -g'(-1)d - h'(1)d + \alpha(d)d.$$

Therefore $h'(1) = f_2$ (and $g'(-1) = f_1$). Analogously $h'(1) = f_4$. Thus $h'(1) = f_2 = f_4$. But for function $f(x, y) = xy$ we have $f_4 = 1 \neq f_2 = -1$.

(c) Suppose that this completed array is differentiably basic. Set $a_n = ([\frac{n+1}{2}]^{-1/2}, [\frac{n}{2}]^{-1/2})$, $f(a_n) := \frac{(-1)^n}{n}$, $n = 2, 3, \ldots$. If $f(x, y) = g(x) + h(y)$ for some functions $g(x)$ and $h(y)$, then the series $f(a_2) - f(a_3) + f(a_4) - \ldots$ converges to $g(1) - g(0)$ (analogously to Problem 7d). This is a contradiction because the series $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$ diverges.

(d) Without loss of generality assume that $f(0, 0) = 0$, then take $g(0) = 0$ and $h(0) = 0$. Set

$$h(2^{-k}) = f(2^{-(k+1)}, 2^{-k}) - f(2^{-(k+1)}, 2^{-(k+1)}) + f(2^{-(k+2)}, 2^{-(k+1)}) - \ldots,$$

$$g(2^{-k}) = f(2^{-k}, 2^{-k}) - f(2^{-(k+1)}, 2^{-k}) + f(2^{-(k+1)}, 2^{-(k+1)}) - \ldots,$$

where the right-hand sides are sums of alternating series. Now $g(x)$ and $h(y)$ may be extended to differentiable functions $\mathbb{R} \to \mathbb{R}$.

15. (a) Define

$$w(0) = 0, \quad w(4^{-i} + 4^{-3i}) = w(4^{-i}) = 0 \quad \text{and} \quad w(4^{-i} + 4^{-3i-1}) = 2^{3i} \quad \text{for} \quad i = 1, 2, 3, \ldots.$$

Extend piecewice-linearly to obtain a function $w : [0, 1] \to \mathbb{R}$. For every $x \in [0, 1]$ define $f(x, -x)$ as the area under the graph of w on $[0, x]$. Define $f(x, y) = 0$ on the rest of the cross.

Suppose that $f(x, y) = g(x) + h(y)$ for some differentiable functions g and h. Without loss of generality we assume that $g(0) = h(0) = 0$. Let us prove that g is not differentiable at $x = 1/4$. (In the same way one can prove that g is not differentiable at $x = 4^{-i}$ for each i.)

Using two infinite arrays starting at points $(\frac{1}{4} + d, -\frac{1}{4} - d)$ and $(\frac{1}{4}, -\frac{1}{4})$ and converging to the point $(0, 0)$ we obtain that

$$g\left(\frac{1}{4}+d\right) - g\left(\frac{1}{4}\right) = f\left(\frac{1}{4}+d, -\frac{1}{4}-d\right) - f\left(\frac{1}{4}, -\frac{1}{4}\right) + f\left(\frac{1}{4^2}+\frac{d}{4}, -\frac{1}{4^2}-\frac{d}{4}\right) - f\left(\frac{1}{4^2}, -\frac{1}{4^2}\right) + \ldots$$

For every positive $d < \frac{1}{4}$ there is k such that $4^{-3i} < d/4^{i-1}$ for each $i > k$ and $4^{-3k} \geq d/4^{k-1}$. In particular, $4^{-2k} \geq 4d > 4^{-2(k+1)}$. Then

$$2(g\left(\frac{1}{4}+d\right) - g\left(\frac{1}{4}\right)) > 2^{-3(k+1)} + 2^{-3(k+2)} + 2^{-3(k+3)} \cdots > 2^{-3(k+1)} \geq \frac{(4d)^{3/4}}{8}.$$

This contradicts to the differentiability of g at $\frac{1}{4}$.

(b) **Conjecture.** The answer is no, the proof is analogous to that of problem 15(a).

(c)** **Conjecture.** A piecewise-linear graph in $\mathbb{R}^2$ is differentiably basic if and only if it does not contain arbitrary long arrays and for each two singular points a and b we have $x(a) \neq x(b)$ and $y(a) \neq y(b)$. A point $a \in K$ is *singular* if the intersection of K with each disk centered at a is not a rectilinear arc.

16. See an extract of [RZ02] on a separate sheet.

Motivations.

In the course of solution of the Hilbert 13-th problem [Ar58] there appeared the notion of *basic embedding* (we give references to surveys not to original papers). The main result of this sequence of problems (problem 8b) is an elementary solution [MT03] of the 'half' of the Arnold problem [Ar58'] on characterization of basic subsets of the plane [St89]. See also [Vo81, Vo82, Sk95, Ku00, Ku03]. This sequence of problems has only some trivial problems in common with [KS97, KS98]. The most important unsolved problems here are on characterization of *smoothly* *basic* subsets of the plane [RZ02]. We would like to acknowledge V. I. Arnold and S. M. Voronin for useful discussions.

References

[Ar58] V. I. Arnold, *Representation of functions of some number of variables as superposition of functions of less number of variables (in Russian)*, Math. Prosveshenie, 3 (1958), pages 41-61.

[Ar58'] V. I. Arnold, *Problem 6 (in Russian)*, Math Prosveshenie, 3 (1958), pages 273-274.

[Vo81] S.M. Voronin, Functional Analysis, 15 (1981).

[Vo82] S.M. Voronin, Functional Analysis, 16 (1982).

[KS97] V. Kurlin and A. Skopenkov, *Basic embeddings of graphs into the plane (in Russian)*, journal Math. Obrazovanie, 3 (1997), pages 105–113

[KS98] V. Kurlin and A. Skopenkov *Basic embeddings of graphs into the plane (in Russian)*, publ. in: 9-th summer conference of Tournament of Towns, MCCME (1998), pages 34–44, 106–113

[Ku00] V. Kurlin, *Basic embeddings into products of graphs,* Topol.Appl. 102 (2000), 113–137.

[Ku03] V. A. Kurlin, *Basic embeddings of graphs and the Dynnikov method of three-pages embeddings (in Russian),* Uspekhi Mat. Nauk, 58:2 (2003), 163–164. English transl.: Russian Math. Surveys, 58:2 (2003).

[MK03] N. Mramor-Kosta and E. Trenklerova, *On basic embeddings of compacta into the plane,* Bull. Austral. Math. Soc. 68 (2003), 471–480.

[RZ02] D. Repovs and M. Zeljko, *On basic embeddings into the plane,* preprint (2002)

[Sk95] A. Skopenkov, *A description of continua basically embeddable in* $\mathbb{R}^2$*,* Topol. Appl. 65 (1995), 29–48.

[St89] Y. Sternfeld, *Hilbert's 13th problem and dimension,* Lect. Notes Math. 1376 (1989), 1–49.

BASIC PLANAR SETS

A. Skopenkov and I. Shnurnikov (presented by them and V. Gorin)

An extract of [RZ02].

14. (a) Take a differentiable function $f : V \to \mathbb{R}$. Since f is differentiable at $(0,0)$, it follows that there exist $a, b \in \mathbb{R}$ such that

$$f(x,|x|) = f(0,0) + ax + b|x| + o(\sqrt{x^2+|x|^2}), \quad \text{where} \quad x \to 0.$$

Take $h(y) = by$ and $g(x) = f(x,|x|) - h(|x|)$. Clearly, h is differentiable and g is differentiable outside 0. Since $g(x) = f(0,0) + ax + o(x)$ when $x \to 0$, it follows that g is differentiable also at 0.

16. (a) It is clear.

(b) A subset $K \subset \mathbb{R}^2$ is called r *times differentiably basic* if for each r times differentiable function $f : K \to \mathbb{R}$ there exist r times differentiable functions $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ such that $f(x,y) = g(x) + h(y)$ for each point $(x,y) \in K$.

(c) We can take the graph V_k of the function $y = |x|^k$, $x \in [-1,1]$ for k odd, and $W_{k+1} = (V_{k+1} - (2,0)) \cup (V_{k+1} + (2,0))$ for k even.

Proof for k even. Let us prove that W_{k+1} is r times differentiably basic for each $0 \le r \le k$. Given an r times differentiable function $f : W_{k+1} \to \mathbb{R}$, take functions $h(y) = 0$ and $g(x) = f(x, |x - 2\operatorname{sign} x|^{k+1})$. Clearly, h is r times differentiable and $f(x,y) = g(x) + h(y)$ for each $(x,y) \in W_{k+1}$. Since the function $p(t) = |t|^{k+1}$ is k times differentiable and $r \le k$, it follows that g is r times differentiable.

Let us prove that W_{k+1} is not r times differentiably basic for k even and each $k < r$. Define an differentiable function $f : W_{k+1} \to \mathbb{R}$ by $f(x,y) = y \operatorname{sign} x$. If W_{k+1} is r times differentiably basic, then there are r times differentiable functions g and h such that $f(x,y) = g(x) + h(y)$. For $x \in [-1,1]$ we have

$$g(\pm 2 + x) + h(|x|^{k+1}) = f(\pm 2 + x, |x|^{k+1}) = \pm |x|^{k+1}.$$

Since g is $(k+1)$ times differentiable and $k+1$ is odd, it follows that $\frac{dh}{dx}|_{x=0} = +1$ and $\frac{dh}{dx}|_{x=0} = -1$, which is a contradiction. $\square$

Proof for k odd.

Now we prove that V_k is r times differentiably basic for each $0 \le r \le k$. Take an r times differentiable function $f : V_k \to \mathbb{R}$. Since f is r times differentiable at $(0,0)$, it follows that there exist $\{a_{ij}\}_{i,j=0}^r \subset \mathbb{R}$ such that

$$a_{00} = f(0,0) \quad \text{and} \quad f(x,|x|^k) = \sum_{i,j=0}^{r} a_{ij} x^i |x|^{kj} \ + \ o([x^2 + x^{2r}]^{r/2}), \quad \text{where} \quad x \to 0.$$

Since

$$o([x^2 + x^{2r}]^{r/2}) = o_1(x^r), \quad \text{we have} \quad f(x,|x|^k) = a_{00} + a_{01}|x|^k + a_{10}x + \dots + a_{r0}x^r + o_2(x^r).$$

Take $h(y) = a_{01}y$ and $g(x) = f(x,|x|^k) - h(|x|^k)$. Clearly, h is r times differentiable and g is r times differentiable outside 0. We also have $g(x) = a_{00} + a_{10}x + \dots + a_{r0}x^r + o_2(x^r)$ when $x \to 0$. So g is r times differentiable also at 0.

Next we prove that $V = V_1$ is not r times differentiably basic for each $1 < r$. Define a differentiable function $f : V \to \mathbb{R}$ by $f(x, y) = xy$, where $y = |x|$. If V is r times differentiably basic for some $r \geq 2$, then there are r times differentiable functions

$$g, h : \mathbb{R} \to \mathbb{R} \quad \text{such that} \quad f(x, |x|) = x|x| = g(x) + h(|x|) \quad \text{for each} \quad x \in [0, 1].$$

Hence $g(x) - g(-x) = 2x^2$. But this is impossible because g is 2 times differentiable, hence

$$g(x) = g(0) + ax + bx^2 + o(x^2) \quad \text{and so} \quad g(-x) = g(0) - ax + bx^2 + o(x^2) \quad \text{for} \quad x \to +0.$$

At last we prove that V_k is not r times differentiably basic for k odd and each $k < r$. Define a differentiable function $f : V_k \to \mathbb{R}$ by $f(x, y) = xy$, where $y = |x|^k$. If V is r times differentiably basic for some $r > k$, then there are r times differentiable functions

$$g, h : \mathbb{R} \to \mathbb{R} \quad \text{such that} \quad f(x, |x|^k) = x|x|^k = g(x) + h(|x|^k) \quad \text{for each} \quad x \in [0, 1].$$

Hence $g(x) - g(-x) = 2x|x|^k$. But this is impossible for k odd because g is $(k + 1)$ times differentiable, hence

$$g(x) = g_0 + g_1 x + \dots + g_{k+1} x^{k+1} + o(x^{k+1}) \quad \text{and so} \quad g(-x) = g_0 - g_1 x + \dots + g_{k+1} x^{k+1} + o(x^{k+1})$$

$$\text{for} \quad x \to +0. \qquad \square$$

БАЗИСНЫЕ ВЛОЖЕНИЯ И 13-Я ПРОБЛЕМА ГИЛЬБЕРТА [1]

А. Скопенков

В этой статье рассказано, как при решении 13-й проблемы Гильберта о суперпозициях непрерывных функций появилось понятие базисного подмножества и базисного вложения. Приводятся красивые результаты об этих понятиях, большая часть которых доступна старшекласснику. Три части статьи можно читать независимо друг от друга (в небольшом количестве мест, в которых одна часть использует другую, приведены точные ссылки).

В первой части приводится элементарное изложение идеи решения 13-й проблемы Гильберта о суперпозициях А. Н. Колмогоровым и В. И. Арнольдом (по мотивам [Ar58]). При этом показывается, как естественно появилось понятие базисного вложения, и остается без доказательства важнейшее место (лемма Колмогорова о деревьях).

Подмножество $K \subset \mathbf{R}^2$ называется *базисным*, если для любой непрерывной функции $f : K \to \mathbf{R}$ существуют такие непрерывные функции $g, h : \mathbf{R} \to \mathbf{R}$, что $f(x, y) = g(x) + h(y)$ для каждой точки $(x, y) \in K$.

Во второй части приводится характеризация графов, которые можно вложит в плоскость в качестве базисных подмножеств [St89, Sk95] (решение проблемы Штернфельда), а также ее обобщения [St89, Sk95, Ku00, Ku03, Ku03'].

Третья часть наиболее элементарна (см., например, задачи 1b и 4a). Приводится характеризация базисных подмножеств плоскости (решение проблемы Арнольда) [Ar58', St89, MKT03, Mi09], а также ее обобщения [St89, Vo81, Vo82, RZ07].

В тексте много задач, которые нумеруются жирными цифрами и к большинству которых приводятся решения (в конце соответствующего пункта). Если условие задачи является утверждением, то задача состоит в том, чтобы это утверждение доказать. Трудные задачи отмечены звездочкой, а нерешенные — двумя.

Если некоторое замечание в сноске или условие задачи Вам непонятно, то его нужно просто игнорироваить. Это не приведет к непониманию дальнейшего материала.

Благодарю В.И. Арнольда, Ю.М. Бурмана, С.М. Воронина, М.Н. Вялого, А.Р. Сафина и И.Н. Шнурникова за полезные обсуждения, а также М.Н. Вялого за подготовку рисунков.

О РЕШЕНИИ 13-Й ПРОБЛЕМЫ ГИЛЬБЕРТА

13-я проблема Гильберта.

Пусть дано некоторое множество функций $F = \{f_\alpha(x_1, \dots, x_{n_\alpha})\}_{\alpha\in A}$ (не обязательно конечное). Определим *суперпозицию* функций из F (или *формулу* над F) индуктивно:

(1) сами функции f_α и все переменные x_j являются суперпозициями функций из F.

(2) если функции $f(x_1, \dots, x_n)$, $g_1(\dots)$, $\dots$, $g_n(\dots)$ являются суперпозициями функций из F (не обязательно различными), то и функция $f(g_1(\dots), \dots, g_n(\dots))$ является суперпозицией функций из F. Здесь в качестве аргументов функций g_i можно подставлять любые наборы переменных (эти переменные могут идти в любом порядке, а некоторые из переменных могут совпадать). [2]

Например, полином $\sum a_{k_1\dots k_n} x_1^{k_1}\cdots x_n^{k_n}$ есть суперпозиция функций $f(x, y) = x + y$ и $g(x, y) = xy$ и констант. Причем если можно использовать функции одной переменной, то для $x, y > 0$ произведение можно и не использовать, так как $xy = 2^{\log_2 x+\log_2 y}$.

Рассмотрим следующие вопросы.

[1] Отдельные части работы представлялись на летних конференциях Турнира Городов 1997 и 2006 годов (В.А. Гориным, В.А. Курлиным, И.Н. Шнурниковым и автором), на семинарах мехмата МГУ и на семинаре по геометрии в МЦНМО. Некоторые задачи о гладкой базисности представлялись А.А. Бараном на международной конференции школьников Intel ISEF в 2003 году. Ссылки даются по возможности не на оригинальные работы, а на обзоры.

[2] Определение суперпозиции можно также сформулировать графически, на языке схем.

1. Можно ли каждую функцию нескольких аргументов записать в виде суперпозиции функций не более чем двух аргументов?

2. Можно ли каждую функцию двух аргументов записать как суперпозицию функций одного аргумента и простейшей функции двух аргументов (например, сложения)? [3]

Поскольку плоскость и прямая равномощны, то любую функцию трех и более переменных можно выразить в виде суперпозиции (вообще говоря, *разрывных*) функций двух переменных (см. детали в [Ar58]). Поэтому указанные вопросы интересны только для *непрерывных* функций.

Через
$$|x,y| = |(x_1,\dots,x_n),\ (y_1,\dots,y_n)| = \sqrt{(x_1-y_1)^2+\dots+(x_n-y_n)^2}$$
обозначается обычное расстояние между точками $x=(x_1,\dots,x_n)$ и $(y_1,\dots,y_n)$ пространства $\mathbf{R}^n$. Пусть K — подмножество пространства $\mathbf{R}^n$. Функция $f:K\to\mathbf{R}$ называется *непрерывной*, если для любых точки $x_0\in K$ и числа $\varepsilon>0$ существует такое число $\delta>0$, что для любой точки $x\in K$ с условием $|x,x_0|<\delta$ выполнено $|f(x)-f(x_0)|<\varepsilon$. Например, функция $f(x_1,x_2)=\sqrt{x_1^2+x_2^2}$ является непрерывной на плоскости, а функция $f(x_1,x_2)$, равная целой части от x_1+x_2, — нет. В дальнейшем все функции предполагаются непрерывными, если явно не оговорено противное.

Ясно, что любая элементарная функция представляется в виде суперпозиции функций двух переменных. Простейшие неэлементарные функции — корни алгебраических уравнений. К 1900 году было известно, что любое алгебраическое уравнение n-й степени сводится (при помощи радикалов, сложения, вычитания, умножения и деления) к такому, у которого коэффициенты при x^n и x^0 равны 1, а при x^{n-1}, x^{n-2} и x^{n-3} равны 0. Таким образом, остается $n-4$ переменных коэффициента. Поэтому 'простейшей' функцией, для которой не было известно выражение через функции двух переменных, была функция $x(a,b,c)$, выражающая решение уравнения $x^7+ax^3+bx^2+cx+1=0$ седьмой степени. Поэтому Гильберт сформулировал свою 13-ю проблему так:

Доказать, что уравнение седьмой степени $x^7+ax^3+bx^2+cx+1=0$ *неразрешимо без использования функций трех переменных.*

Гильберту удалось показать, что некоторые *аналитические* функции трех переменных не являются суперпозициями аналитических же функций двух переменных [Ar58]. В 1954 Витушкин доказал, что некоторые *r раз непрерывно дифференцируемые* функции не являются суперпозициями r раз непрерывно дифференцируемых функций двух переменных [Ar58, Vi04]. Для *непрерывных* же функций гипотеза Гильберта была опровергнута в 1957 году Колмогоровым и Арнольдом.

Теорема Колмогорова-Арнольда. *Любая непрерывная функция представляется в виде суперпозиции непрерывных функций одного и двух аргументов.*

Теорема Колмогорова: к суперпозициям функций трех переменных.

Сначала в 1956 г. Колмогорову удалось доказать, что *произвольная непрерывная функция более чем трех переменных является суперпозицией непрерывных функций трех переменных.* Он использовал следующее понятие. (Если это понятие или последующий текст до леммы Колмогорова об универсальных функциях покажутся Вам сложными, Вы можете сразу перейти к этой лемме. Другой вариант — прочитать этот текст для $n=2$, а в этой лемме снова вернуться к произвольному n.)

[3] Для функций алгебры логики ответы на оба вопроса положительны (поскольку любую функцию алгебры логики можно выразить через 'и' и 'не'). Аппроксимационная теорема Вейерштраса показывает, что функция нескольких аргументов может быть *равномерно приближена* на компактном множестве полиномами, т.е. суперпозициями констант, сложения и умножения.

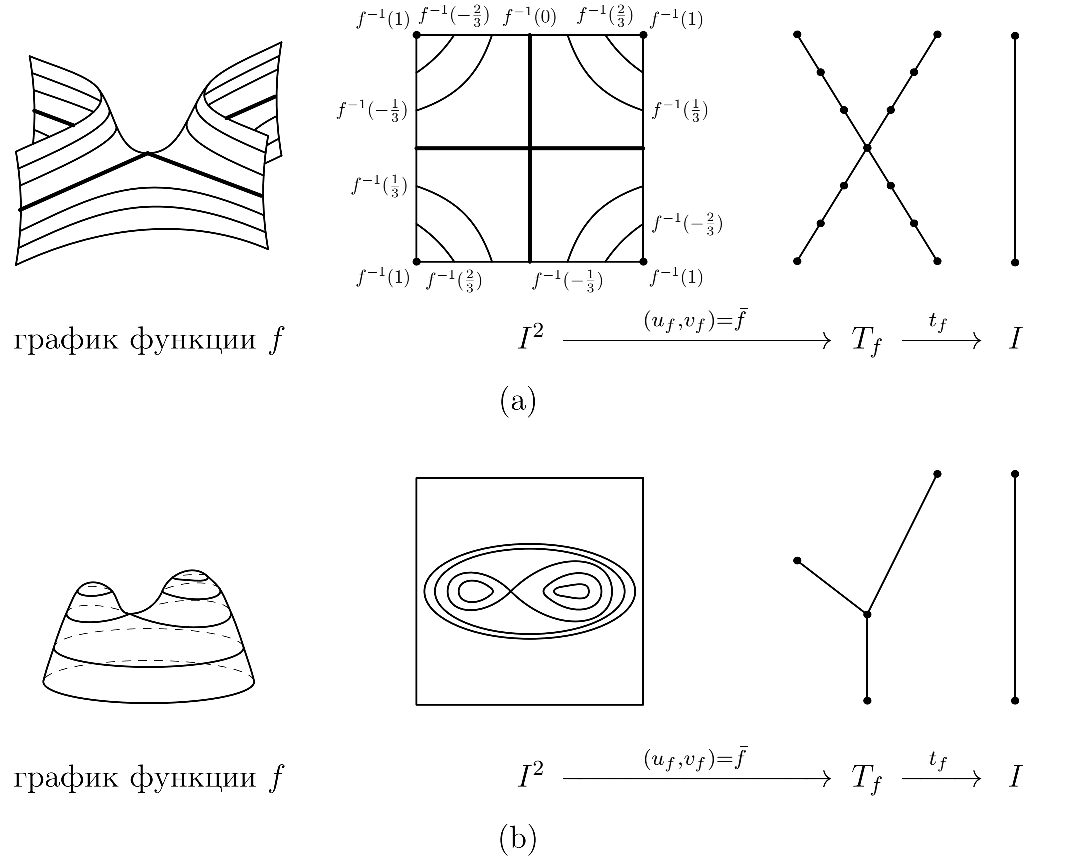


Рис. 1: Примеры деревьев компонент множеств уровня и разложений Кронрода

Обозначим $I = [-1; 1]$. *Деревом T_f компонент множеств уровня* функции $f : I^n \to I$ называется метрическое пространство, точками которого являются компоненты связности множеств $f^{-1}(c)$, $c \in I$, а метрика определена в http://en.wikipedia.org/wiki/Hausdorf_distance

Например, дерево компонент множеств уровня функции $f : I^2 \to I$, $f(x, y) = xy$, гомеоморфно букве X. Другие примеры приведены на рисунке (см. детали в [Ar58]).

Очевидно, что любая функция $f : I^n \to I$ от n переменных представляется в виде композиции $I^n \stackrel{t_f}{\to} T_f \stackrel{\overline{f}}{\to} I$ для некоторых отображений $\overline{f}$ и t_f. Пространство T_f можно считать лежащим без самопересечений в квадрате I^2. [4] Поэтому t_f можно считать парой функций $u_f, v_f : I^n \to I$. Функцию $\overline{f}$ можно продолжить на весь квадрат I^2 (по теореме Урысона о продолжении), т.е. считать функцией двух переменных. Итак, имеем (рис. 1) *разложение Кронрода*

$$f(x_1, \ldots, x_n) = \overline{f}(u_f(x_1, \ldots, x_n), v_f(x_1, \ldots, x_n)).$$

Лемма Колмогорова об универсальных деревьях. *Для любого $n \geq 2$ существуют такие деревья $T_1, \ldots, T_{n+1}$ и функции $t_i : I^n \to T_i$, что для любой непрерывной функции $f : I^n \to I$ от n переменных существуют непрерывные функции $g_1, \ldots, g_{n+1} : I^n \to I$ от n переменных, для которых*

$$T_{g_i} = T_i, \quad t_{g_i} = t_i \quad \text{и} \quad f = g_1 + \cdots + g_{n+1}.$$

Важно, что деревья T_{g_i} компонент множеств уровня функций g_i (и соответствующие отображения t_{g_i}) не зависят от f, хотя сами функции g_i могут зависеть от f.

[4] Для доказательства нужно показать, что T_f является *деревом*, т. е. одномерным стягиваемым локально связным компактом. Примеры деревьев находятся на рис. 3,4,5 ниже. Любое дерево планарно [Ku68].

По-видимому, лемма верна и для $n = 1$ (но это нетривиально).

Доказательства мы не приводим. Хотя оно являются важным шагом в доказательстве теоремы Колмогорова-Арнольда, но наша цель — осветить именно те шаги, в которых появилось понятие базисного вложения. Кроме того, проблему Гильберта можно решить намного проще: см. ниже суперпозиционную теорему Колмогорова и ее доказательство в [Ar58].

Из леммы Колмогорова об универсальных деревьях и разложения Кронрода вытекает следующий результат (докажите!).

Лемма Колмогорова об универсальных функциях. *Для любого $n \geq 3$*

существует такой набор из $2n+2$ непрерывных функции $u_i, v_i : I^n \to I$ $(i = 1, \dots, n+1)$ от n переменных,

что для любой непрерывной функции $f : I^n \to I$ от n переменных

существуют непрерывные функции $f_i : I^2 \to I$ $(i = 1, \dots, n+1)$ двух переменных, для которых

$$f(x_1, \dots, x_n) = \sum_{i=1}^{n+1} f_i(u_i(x_1, \dots, x_n), v_i(x_1, \dots, x_n)).$$

Важно, что функции u_i, v_i не зависят от f (при фиксированном n), хотя функции f_i могут зависеть от f.

Эта лемма тривиальна для $n = 1$ и $n = 2$ (подумайте, почему).

Набросок доказательства теоремы Колмогорова о выразимости через функции трех переменных. Для функции $f(x_1, x_2, x_3, x_4)$ четырех переменных имеем

$$f(x_1, x_2, x_3, x_4) \;=\; f_{x_4}(x_1, x_2, x_3) \;\overset{(*)}{=}\; \sum_{i=1}^{4} f_{x_4,i}(u_i(x_1, x_2, x_3), v_i(x_1, x_2, x_3)) \;=$$

$$= \sum_{i=1}^{4} F_i(u_i(x_1, x_2, x_3), v_i(x_1, x_2, x_3), x_4), \quad \text{где} \quad F_i(a, b, c) = f_{c,i}(a, b).$$

Функция f_{x_4} непрерывно зависит от параметра x_4. Равенство (*) получается применением леммы Колмогорова об универсальных функциях. Мы используем усиленный вариант этой леммы, утверждающий, что каждая из $n+1$ функций f_i непрерывно зависит от исходной функции f. Из этого варианта вытекает непрерывная зависимость функций $f_{x_4,i}$ от параметра x_4 $(i = 1, 2, 3, 4)$. А это влечет непрерывность функций F_i $(i = 1, 2, 3, 4)$.

Для функции большего количества переменных рассуждение аналогично. QED

1. За одну копейку автомат выдает значение заданной Вами непрерывной функции трех переменных на заданной Вами тройке чисел. За какую сумму Вы заведомо сможете вычислить заданную непрерывную функцию n переменных (при условии наличия у Вас неограниченной памяти)?

Теорема Арнольда: к суперпозициям функций двух переменных.

Для доказательства теоремы Колмогорова-Арнольда осталось произвольную непрерывную функцию трех переменных выразить через суперпозицию непрерывных функций двух переменных. Для этого полезно следующее понятие.

Подмножество $T \subset I^3$ назовем *базисным*, если любая непрерывная функция на T может быть разложена в сумму трех функций, каждая из которых зависит только от одной координаты. Или, формально, если для любой непрерывной функции $f : T \to I$ существуют такие непрерывные функции $g_1, g_2, g_3 : I \to I$, что

$$f(x, y, z) = g_1(x) + g_2(y) + g_3(z) \quad \text{для} \quad (x, y, z) \in T.$$

Лемма Арнольда о деревьях. *Любое дерево реализуется как базисное подмножество в I^3 (т.е. топологически эквивалентно некоторому базисному подмножеству в I^3).* [5]

Идея доказательства видна на примере доказательства либо базисной вложимости в плоскость конечного дерева, из каждой вершины которого выходит либо одно, либо три ребра [Ar58], либо более сильного утверждения (c) в предпоследнем пункте второй части.

Лемма Арнольда об универсальных функциях. *Существует такой набор из девяти непрерывных функции $u_i : I^2 \to I$ $(i = 1, 2, \dots, 9)$ двух переменных, что*
для любой непрерывной функции $f : I^2 \to I$ двух переменных
существуют непрерывные функции $f_i : I \to I$ $(i = 1, 2, \dots, 9)$ одной переменной,
для которых $f(x, y) = f_1(u_1(x, y)) + \dots + f_9(u_9(x, y))$.

Важно, что функции u_i не зависят от f, хотя функции f_i могут зависеть от f.

Доказательство. Возьмем деревья T_1, T_2, T_3 из леммы Колмогорова об универсальных деревьях для $n = 2$. По лемме Арнольда о деревьях существуют базисные вложения $(u_{i1}, u_{i2}, u_{i3}) : T_i \to I^3$. Положим $u_{3(i-1)+j} := u_{ij}$. Возьмем функции $g_1, g_2, g_3 : I^2 \to I$ (зависящие от f) из леммы Колмогорова об универсальных деревьях для $n = 2$. Применяя аналог разложения Кронрода и определение базисности, получаем

$$g_i(x, y) = \overline{g}_i(u_{i1}(x, y), u_{i2}(x, y), u_{i3}(x, y)) = g_{i1}(u_{i1}(x, y)) + g_{i2}(u_{i2}(x, y)) + g_{i3}(u_{i3}(x, y))$$

для некоторых функций $g_{ij} : I^2 \to I$. Остается положить $f_{3(i-1)+j} := g_{ij}$. QED

Теперь теорема Колмогорова-Арнольда вытекает из

$$f(x, y, z) = f_z(x, y) = \sum_{i=1}^{9} f_{i,z}(u_i(x, y)) = \sum_{i=1}^{9} F_i(u_i(x, y), z), \quad \text{где} \quad F_i(t, z) = f_{i,z}(t).$$

Непрерывность функций F_i доказывается аналогично предыдущему пункту (с использованием соответствующего усиления леммы Арнольда об универсальных функциях).

2. За одну копейку автомат выдает значение заданной Вами непрерывной функции двух переменных на заданной Вами паре чисел. За какую сумму Вы заведомо сможете вычислить заданную непрерывную функцию n переменных (при наличии неограниченной памяти)?

Теорема Колмогорова: к функциям одной переменной и сложению.

В том же 1957 году Колмогоров доказал еще более сильный результат, из которого также вытекает решение проблемы Гильберта.

Суперпозиционная теорема Колмогорова. *Любая непрерывная функция представляется в виде суперпозиции сложения и непрерывных функций одной переменной.*

Более точно, для каждого $n > 1$ существует набор $n(2n+1)$ таких непрерывных функций $u_{ij} : I \to I$ $(i = 1, \dots, 2n+1,\ j = 1, \dots, n)$ одной переменной, что
для любой непрерывной функции $f : I^n \to I$ от n переменных
существуют непрерывные функции $f_1, \dots, f_{2n+1} : I \to I$ одной переменной, для которых

$$f(x_1, \dots, x_n) \ = \ \sum_{i=1}^{2n+1} f_i\left(\sum_{j=1}^{n} u_{ij}(x_j)\right).$$

Здесь важно, что функции u_{ij} не зависят от f, хотя функции f_i могут зависеть от f. Элементарное изложение доказательства приведенной теоремы можно найти в [Ar58].

[5] Доказательство леммы Арнольда использует теорему Менгера о существовании универсального дерева. На самом деле, Арнольд доказал эту лемму для деревьев с точками ветвления третьего порядка. Этого было достаточно для решения проблемы Гильберта. Общий случай леммы доказан Острандом в 1965 [St89].

3. За одну копейку автомат либо складывает два заданных Вами числа, либо выдает значение заданной Вами непрерывной функции одной переменной в заданной Вами точке. За какую сумму Вы заведомо сможете вычислить заданную непрерывную функцию n переменных (при наличии неограниченной памяти)?

Об аналитических проблемах, связанных с этим выдающимся результатом Колмогорова, см. [St89, Vi04]. О топологических проблемах написано далее.

Базисные вложения в многомерные пространства. [6]

Подмножество $K \subset \mathbf{R}^n$ называется *базисным*, если для любой непрерывной функции $f : K \to \mathbf{R}$ существуют такие непрерывные функции $f_1, \dots, f_n : \mathbf{R} \to \mathbf{R}$, что

$$f(x_1, \dots, x_n) = f_1(x_1) + \dots + f_n(x_n) \quad \text{для} \quad (x_1, \dots, x_n) \in K.$$

Пространство K называется *базисно вложимым* в $\mathbf{R}^n$, если существует вложение $K \to \mathbf{R}^n$, образ которого базисный.

Функции на произвольном n-мерном компакте уже нельзя представлять себе как функции n переменных. Однако понятие базисной вложимости доставляет аналог разложимости функций на компактах в суперпозицию фиксированных функций и сложения.

Из суперпозиционной теоремы Колмогорова следует, что *n-мерный куб базисно вложим в $\mathbf{R}^{2n+1}$*. Действительно, функции $u_i = \sum\limits_{j=1}^{n} u_{ij}$ $(i = 1, \dots, 2n+1)$ из теоремы Колмогорова определяют базисное вложение $I^n \to I^{2n+1}$. Остранд заметил в 1965 г., что этот факт можно обобщить.

Теорема Остранда. *Любой n-мерный компакт базисно вложим в $\mathbf{R}^{2n+1}$* [St89].

Эта теорема усиливает теорему Неблинга–Менгера–Понтрягина о вложимости любого n-мерного компакта в $\mathbf{R}^{2n+1}$ [Ku68].

На самом деле Остранд доказал следующий более сильный результат, обобщающий суперпозиционную теорему Колмогорова (а не только ее следствие).

Пусть $X_1, \dots, X_m$ — конечномерные метрические пространства. Положим $n = \dim X_1 + \dots + \dim X_m$ и $X = X_1 \times \dots \times X_m$. Тогда существуют такие непрерывные функции $u_{ij} : X_j \to \mathbf{R}$, $(i = 1, \dots, 2n+1,\ j = 1, \dots, m)$, что для функций $u_i(x_1, \dots, x_m) = u_{i1}(x_1) + \dots + u_{im}(x_m)$ и любой непрерывной функции $f : X \to \mathbf{R}$ существуют непрерывные функции $f_1, \dots, f_{2n+1} : \mathbf{R} \to \mathbf{R}$, для которых

$$f(x_1, \dots, x_m) = f_1(u_1(x_1, \dots x_m)) + \dots + f_{2n+1}(u_{2n+1}(x_1, \dots, x_m)).$$

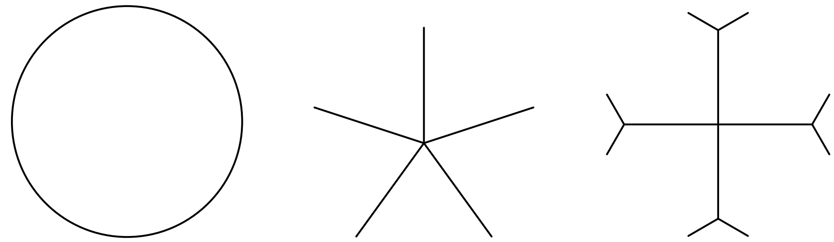
Имеются n-мерные полиэдры, не вложимые в $\mathbf{R}^{2n}$ [Pr04, Sk].

Теорема Штернфельда. *Для любого $n \geq 2$ любой n-мерный компакт (например, n-мерный куб) не вложим базисно в $\mathbf{R}^{2n}$* [St89].

Интересно, что теорема Штернфельда редуцируется к комбинаторно-геометрическому утверждению при помощи многомерного аналога критерия базисности, приведенного в третьей части настоящей статьи.

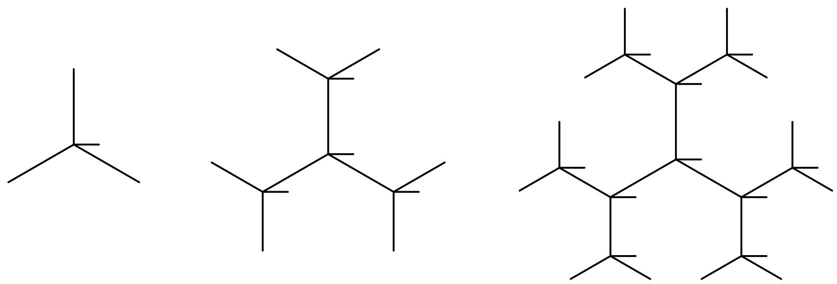
Очевидно, что K базисно вложим в $\mathbf{R}$ тогда и только тогда, когда K топологически вложим в $\mathbf{R}$. Из теорем Остранда и Штернфельда следует, что *для $m > 2$ компакт K базисно вложим в $\mathbf{R}^m$ тогда и только тогда, когда* $\dim K < m/2$. Таким образом, в 1989 г. оставалось неизвестным лишь описание компактов, базисно вложимых *в плоскость*.

[6] Этот пункт неэлементарен, формально не используется в дальнейшем и может быть опущен читателем. Однако мы приводим его, поскольку он дает четкую картину для разных размерностей.

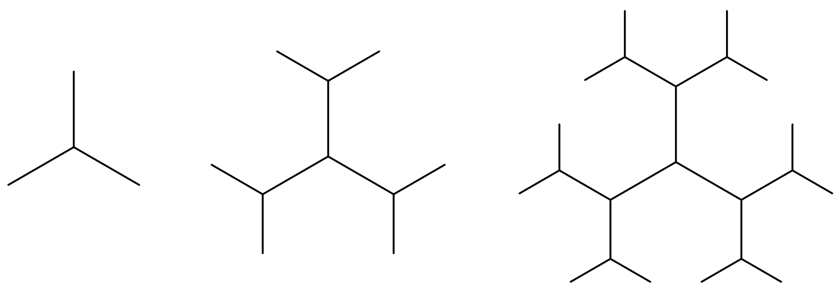


S^1 T_5 C

Рис. 2:



R_1 R_2 R_3



F_1 F_2 F_3

Рис. 3:

БАЗИСНЫЕ ВЛОЖЕНИЯ В ПЛОСКОСТЬ

Базисная вложимость в плоскость.

Граф (или компакт) K называется *базисно вложимым* в плоскость, если существует такое вложение $\varphi : K \to \mathbf{R}^2$, что для любой непрерывной функции $f : \varphi(K) \to \mathbf{R}$ существуют такие непрерывные функции $g, h : \mathbf{R} \to \mathbf{R}$, что $f(x,y) = g(x) + h(y)$ для любой точки $(x,y) \in \varphi(K)$. (Определение непрерывной функции напомнено в начале части 1.)

Проблема описания графов (и компактов), базисно вложимых в плоскость, поставлена Штернфельдом [St89]. Критерий базисной вложимости линейно-связных компактов в плоскость получен в [Sk95]. Для конечных графов он формулируется особенно просто.

Критерий базисной вложимости графов. [Sk95, ср. Sk05] *Конечный граф K базисно вложим в плоскость тогда и только тогда, когда выполнено одно из следующих двух эквивалентных условий:*

(S) K не содержит подграфов, гомеоморфных окружности S^1, пентоду $T_5 = C_1$ или кресту с разветвленными концами $C = C_2$ (рис. 2);

(U) K содержится в одном из графов R_n (рис. 3).

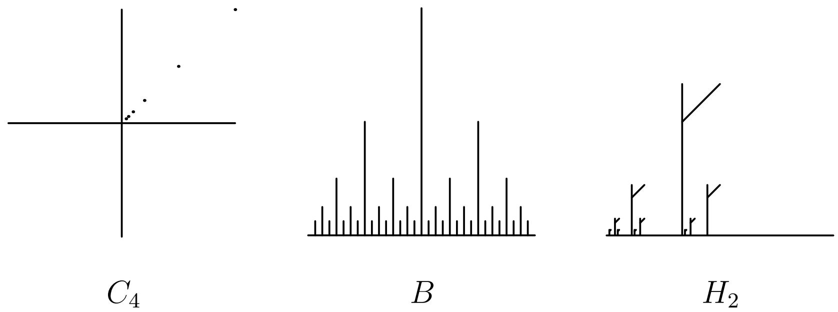


Рис. 4:

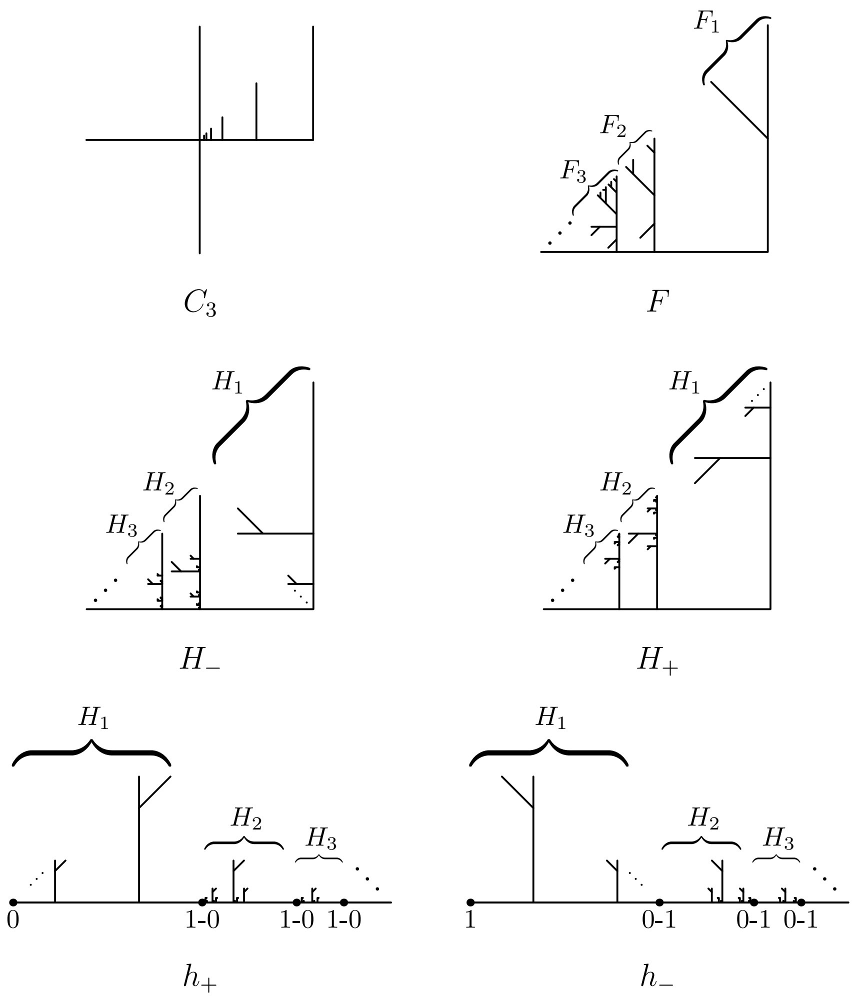


Рис. 5:

Определение графов F_n и R_n (рис. 3). Пусть F_1 — триод, и для любого n граф F_{n+1} получен из F_n разветвлением каждого висячего ребра графа F_n. Граф R_n получается из графа F_n добавлением висячего ребра к каждой точке ветвления графа F_n.

Доказательство приводится в следующем пункте.

Приведем без доказательства решение проблемы Штернфельда для более общего случая (его можно пропустить без ущерба для понимания дальнейшего).

Критерий базисной вложимости линейно-связных компактов. [Sk95, ср. Sk05] *Линейно-связный компакт K базисно вложим в плоскость тогда и только тогда, когда он является локально связным (т.е. пеановским) и выполнено одно из следующих двух эквивалентных условий:*

(1) K не содержит подкомпактов S^1, C_2, C_4, B и, для некоторого n, подкомпактов F_n и H_n (рис. 2, 3, 4);

(2) K не содержит подконтинуумов $S^1, C_1, C_2, C_3, B, F, H_+,\ H_-, h_+, h_-$ (рис. 2, 4, 5).

Введем использованные обозначения (рис. 4, 5). *Нуль-последовательностью* множеств называется последовательность множеств, диаметры которых стремятся к нулю.

Обозначим через C_3 крест с нуль-последовательностью дуг, сходящихся к его 'центру' и приклеенных к одной из его 'ветвей'.

Обозначим через C_4 крест с последовательностью точек, сходящихся к его 'центру'.

Каждый из континуумов B, H_n, F, H_+, H_- является объединением отрезка $I = [0;1]$ и нуль-последовательности

• дуг, приклеенных за один конец к $(0,1)$ в двоично-рациональных точках (для B; очевидно, что топологический тип пространства B не зависит от вариаций в этом построении);

• триодов, приклеенных к I за один конец в точках множества $\{3^{-l_1} + \dots + 3^{-l_s} \mid s \le n,\ 0 < l_1 < \dots < l_s$ – целые$\}$ (для H_n);

• графов F_n, приклеенных к точкам $1/n$ за одну из висячих вершин (для F);

• континуумов H_n, соединенных с точками $1/n$ дугами, пересекающими H_n в $1 \in I \subset H_n$ (для H_+) или в $0 \in I \subset H_n$ (для H_-).

Континуум h_+ (соответственно h_-) получен из нуль-последовательности континуумов H_n склеиванием точек $1 \in I \subset H_n$ и $0 \in I \subset H_{n-1}$ (соответственно точек $0 \in I \subset H_n$ и $1 \in I \subset H_{n-1}$).

Гипотеза о базисной вложимости (не обязательно линейно-связных) континуумов в плоскость еще более громоздка. Она сформулирована в [Sk95, RS99].

Набросок доказательства критерия базисной вложимости графов.

Достаточно доказать следующие три утверждения.

(a) Окружность S^1, пентод T_5 и крест с разветвленными концами C (рис. 2) не вложимы базисно в $\mathbf{R}^2$.

(b) Если конечный граф K не содержит ни одного из графов S^1, T_5 и C (рис. 2), то K содержится в R_n (рис. 3) для некоторого n.

(c) Каждый граф R_n (рис. 3) базисно вложим в $\mathbf{R}^2$.

Набросок доказательства утверждения (b). Докажем, что *дерево K с n невисячими вершинами, не содержащее графов T_5 и C, содержится в R_n*. Возьмем произвольную вершину $a \in K$. Поскольку K не содержит T_5 и C, то $\deg a \le 4$, причем если $\deg a = 4$, то a имеет висячее ребро. Значит, окрестность точки a из K можно вложить в R_n так, что a попадет в центр R_n, а эта окрестность перейдет в окрестность T_4 центра R_n. С каждой вершиной, соседней с a, поступаем аналогично. Поскольку 'глубина' графа R_n (т.е. количество невисячих вершин на самой длинной ветви от центра) равна n, а невисячих вершин в K ровно n, то, продолжая этот процесс дальше, мы сможем вложить в R_n весь граф K. QED

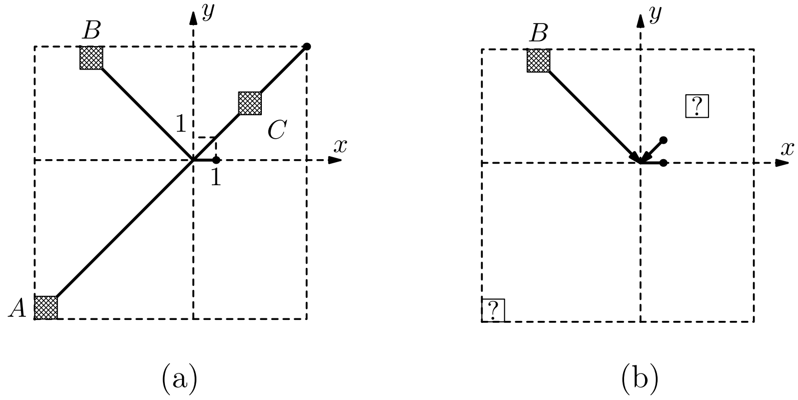


Рис. 6:

Набросок доказательства утверждения (c). Обозначим через R_0 отрезок. Базисные вложения $R_n \to \mathbf{R}^2$ строятся по индукции для $n \geq 0$. Вложим R_0 в $[-7;5]^2$ как диагональ, соединяющую точки $(-7,-7)$ и $(5,5)$.

Вложение $R_n \to \mathbf{R}^2$ получается из вложения на рис.6a добавлением трех вложений графа R_{n-1} в квадратики A, B, C. Проекции A_x, B_x, C_x квадратиков на ось Ox не пересекаются друг с другом и с проекцией отрезка в R_n, параллельного горизонтальной оси. Аналогичное утверждение справедливо и для проекций A_y, B_y, C_y квадратиков на ось Oy.

Докажем, что построенные вложения базисные, при помощи индукции по n. База индукции $n = 0$ очевидна. Докажем шаг индукции.

Пусть $n \geq 1$ и $f : R_n \to \mathbf{R}$ — непрерывная функция. Для $t \in [0,1]$ положим $g(t) := f(t,0)$,

$$h(t) := f(t,t) - g(t) = f(t,t) - f(t,0) \quad \text{и} \quad g(-t) := f(-t,t) - h(t) = f(-t,t) - f(t,t) + f(t,0).$$

По предположению индукции существуют такие функции

$$g : A_x \cup B_x \cup C_x \to \mathbf{R} \quad \text{и} \quad h : A_y \cup B_y \cup C_y \to \mathbf{R}, \quad \text{что}$$

$$f(x,y) = g(x) + h(y) \quad \text{для} \quad (x,y) \in (A \cup B \cup C) \cap R_n.$$

Положим

$$g(-t) := f(-t,t) - h(t) \quad \text{для} \quad t \in C_y \quad \text{и} \quad h(t) = f(t,t) - g(t) \quad \text{для} \quad t \in (-C_y) \cup B_x \to \mathbf{R}.$$

Продолжим построенную функцию $g : A_x \cup B_x \cup (-C_y) \cup [-1;1] \cup C_x \to \mathbf{R}$ до непрерывной функции $g : [-7;1] \cup C_x \to \mathbf{R}$ (например, линейно). Положим

$$h(t) = f(-|t|,t) - g(t) \quad \text{для} \quad t \in [-6;4] \quad \text{и} \quad g(t) := f(t,t) - h(t) \quad \text{для} \quad t \in [1;2] \cup [3;5]$$

(это определение совпадает с прежним для $t \in [-1;1]$). После этого продолжим g и h до непрерывных функций $\mathbf{R} \to \mathbf{R}$. Ясно, что полученные функции g и h — искомые. QED

При доказательстве утверждения (a) мы используем критерий базисности из части 3 (а также приведенное перед ним определение молнии и приведенное после него определение операции E).

Доказательство базисной невложимости окружности. [St89] Пусть задано вложение окружности $S \subset \mathbf{R}^2$. На первом шаге применения E в S закрашивается в белый цвет не более четырех точек (это точки, в которых существуют опорные прямые, параллельные координатным осям и пересекающие K ровно в одной точке). Если после n-го шага применения E закрашено белым цветом k точек, то на следующем шаге закрашивается не более $2k$ точек.

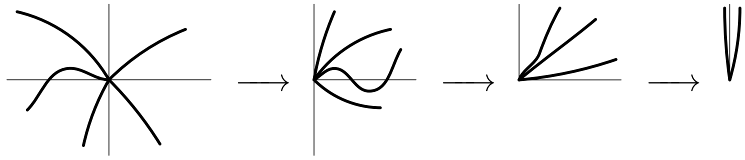


Рис. 7: Доказательство базисной невложимости пентода

В самом деле, если закрашивается точка $a \in E^n(S)$, то хотя бы на одной из двух прямых, проходящих через a и параллельных координатным осям, есть закрашенная ранее точка. В противном случае каждая из этих прямых высекает в $E^n(S)$ не менее двух точек, т.е. a не может быть закрашена на $(n+1)$-м шаге. Итак, после конечного числа шагов будет закрашено конечное число точек, т.е. $E^n(S) \neq \emptyset$ для любого n. QED

Доказательство базисной невложимости пентода. Предположим, что пентод T_5 базисно вложен в плоскость. Пусть d — вершина пентода. Так как $E^n(T_5) = \emptyset$ для некоторого n, то существует максимальное k такое, что $E^k(T_5)$ содержит проколотую окрестность U вершины d в T_5. Тогда на $(k+1)$-м шаге в *одном* из ребер $A \subset T_5$ закрашивается в белый цвет некоторая последовательность точек a_n, сходящаяся к d. Значит (при необходимости переходя к подпоследовательности в a_n и меняя направление оси x), мы можем считать, что одна из прямых $x = x(a_i)$ или $y = y(a_i)$ не содержит других точек из $E^k(T_5)$, кроме a_i. Поскольку окрестность $(U - A) \cup d \cong T_4$ связна, она лежит по одну сторону от всех этих прямых, т.е., в полуплоскости $\mathbf{R}_+ \times \mathbf{R}$. При этом $E^n(T_4) = \emptyset$, т.е.

некоторый крест $T_4 \subset T_5$ базисно вложен в $\mathbf{R}_+ \times \mathbf{R}$ так, что $d = (0,0)$.

Теперь аналогично доказывается, что

некоторый триод $T_3 \subset T_4$ базисно вложен в $\mathbf{R}_+ \times \mathbf{R}_+$ или в $0 \times \mathbf{R}$ так, что $d = (0,0)$.

Второй случай невозможен. Теперь аналогично доказывается, что

некоторый диод $T_2 \subset T_3$ базисно вложен в луч $\mathbf{R}_+ \times 0$ так, что $d = (0,0)$.

Получили противоречие. QED

Доказательство базисной невложимости креста с разветвленными концами. Предположим, что C базисно вложен в плоскость. Аналогично доказательству базисной невложимости пентода получаем, что *если крест T_4 базисно вложен в плоскость $\mathbf{R}^2$, то одна из его ветвей содержит прямолинейный отрезок с концом в вершине креста, параллельный одной из координатных осей.*

Теперь базисная невложимость графа C вытекает из следующей леммы. QED

Лемма о схлопывании. *Пусть K — базисное подмножество плоскости. Определим отображение*

$$q : \mathbf{R}^2 \to \mathbf{R}^2 \quad \textit{формулой} \quad q(x,y) = \begin{cases} (x,y), & x < a \\ (a,y), & a \le x \le b \\ (x-(b-a),y), & x > b \end{cases}.$$

(a) $q|_{K-[a;b]\times c}$ инъективно;

(b) $q(K) \subset \mathbf{R}^2$ базисно.

Доказательство. (a) Пусть, напротив, две точки из $K - [a;b] \times c$ склеиваются при схлопывании q. Тогда они лежат в полосе $[a;b] \times \mathbf{R}$ и имеют одинаковую ординату d. Тогда эти точки (x_1, d), (x_2, d) вместе с (x_1, c), (x_2, c) являются вершинами прямоугольника со сторонами, параллельными координатным осям. Это множество не базисно в $\mathbf{R}^2$. Противоречие.

(b) Достаточно доказать, что если $q(K)$ содержит молнию $A = \{a_1, \dots, a_n\}$ длины n, то и K содержит молнию длины n. Если $q^{-1}(A)$ — молния в K, то нужное утверждение доказано. Иначе найдутся точки a_i, a_{i+1} — вершины вертикального звена — такие, что $p_x(a_i) = p_x(a_{i+1})$.

Тогда добавим к $q^{-1}(A)$ точки $(x(q^{-1}(a_i)), c)$ и $(x(q^{-1}(a_{i+1})), c)$ (здесь полагаем $q^{-1}(a, c) = (a, c)$). Проделав такую операцию несколько раз, получим молнию в K длины больше n. QED

Базисная вложимость в произведение графов.

Декартовым произведением $X \times Y$ двух множеств X и Y называется множество всех пар (a, b) таких, что $a \in X$ и $b \in Y$. Определение базисного вложения может быть очевидно обобщено на вложения в произвольное декартово произведение $X \times Y$. Если X и Y — графы, то мы можем представлять себе произведение $X \times Y$ как двумерный объект (в некоторых случаях можно считать, что этот объект расположен в трехмерном пространстве). Обозначим через T_n звезду с n лучами. Например, для триода T_3 пространство $T_3 \times I$ является 'книжкой с тремя страницами', $S^1 \times I$ — цилиндром и $S^1 \times S^1$ — тором. Произведение $T_n \times I$ назовем *книжкой с n страницами.*

Замечание. *Для любых конечных графов X и Y найдется конечный граф, который нельзя базисно вложить в произведение $X \times Y$.*

Действительно, обозначим через k максимальную степень вершин графов X и Y. Докажем, что звезда T_{4k^2+1} не вложима базисно и *кусочно-линейно* в $X \times Y$. Предположим, противное. Малая окрестность центра v звезды в произведении $X \times Y$ состоит из не более чем k^2 прямоугольников $I \times J$, являющихся произведениями частей ребер I и J графов X и Y. По принципу Дирихле среди $4k^2 + 1$ ребер звезды найдется по крайней мере пять, начала которых выходят в один прямоугольник. Поэтому существует подзвезда $T_5 \subset T_{4k^2+1}$, базисно вложенная в один из таких прямоугольников. Это противоречит критерию базисной вложимости графов. QED

Теорема универсальности. [Ku03] *Любой конечный граф базисно вложим в произведение $X \times I$ для некоторого букета X окружностей и отрезков.*

Число окружностей в букете можно взять равным $E-V+C$, где E, V, C — количества вершин, ребер и компонент связности графа. В частности, любое дерево базисно вложимо в книжку с некоторым числом страниц.

Результаты этого пункта приводятся без доказательства.

Следующая гипотеза навеяна теоремой Робертсона-Симора о вложимости графов в поверхности [Sk05].

Гипотеза. *(a) Существует лишь конечное число 'запрещенных' подграфов для базисной вложимости конечного графа в данное произведение графов.*

(b) Существует алгоритм проверки базисной вложимости конечного графа в данное произведение графов.

Критерий базисной вложимости графов в книжку с n страницами. [Ku00] Дефектом *графа K называется сумма*

$$\delta(K) = (degA_1 - 2) + \cdots + (degA_k - 2),$$

где $A_1, \ldots, A_k$ — все вершины графа K, либо имеющие степень больше четырех, либо степени 4, не имеющие висячих ребер. Конечный граф K базисно вложим в $I \times T_n$ тогда и только тогда, когда он является деревом и

- *либо $\delta(K) < n$,*
- *либо $\delta(K) = n$ и K содержит вершину степени больше четырех, имеющую висячее ребро.*

Из этого результата вытекает положительное решение гипотезы для произведения $I \times T_n$. В [Ku03] доказаны также обобщения этого результата.

БАЗИСНОСТЬ ПЛОСКИХ МНОЖЕСТВ [7]

Разрывная базисность.

0. Представьте функцию $f : [(-1,-1),(1,1)] \cup [(0,0),(1,-1)] \to \mathbf{R}$, $f(x,y) = xy$ в виде суммы $g(x) + h(y)$ двух функций, каждая из которых зависит только от одной координаты.

1. (a) Для любых ли четырех чисел $f_{11}, f_{12}, f_{21}, f_{22}$ существуют такие четыре числа g_1, g_2, h_1, h_2, что $f_{ij} = g_i + h_j$ при любых $i, j = 1, 2$?

(b) Андрей Николаевич и Владимир Игоревич играют в игру 'А ну-ка, разложи!'. На шахматной доске отмечено несколько клеток. А. Н. расставляет числа в отмеченных клетках, как хочет. В. И. смотрит на расставленные числа и берет 16 чисел $a_1, \dots, a_8, b_1, \dots, b_8$, т.е. 'весов' столбцов и строк, как хочет. Если число в каждой отмеченной клетке оказалось равным сумме весов строки и столбца этой клетки, то выиграл В. И., а иначе (т.е. если число хотя бы в одной отмеченной клетке оказалось не равным сумме весов строки и столбца этой клетки) выиграл А. Н.

Докажите, что при правильной игре В. И. выигрывает тогда и только тогда, когда не существует замкнутого маршрута ладьи, начальная клетка и клетки поворота которого являются отмеченными (не обязательно все отмеченные клетки задействованы).

Определение молнии. Обозначим через $\mathbf{R}^2$ плоскость с фиксированной системой координат. Обозначим через $x(a)$ и $y(a)$ координаты точки $a \in \mathbf{R}^2$. Последовательность (конечная или бесконечная) точек плоскости $\{a_1, \dots, a_n, \dots\} \subset \mathbf{R}^2$ называется *молнией*, если для каждого i выполнено $a_i \neq a_{i+1}$, и при этом $x(a_i) = x(a_{i+1})$ для четных i и $y(a_i) = y(a_{i+1})$ для нечетных i. (Не обязательно все точки молнии различны.)

Конечная молния $\{a_1, \dots, a_{2l+1}\}$ называется *замкнутой*, если $a_1 = a_{2l+1}$.

2. Рассмотрим замкнутую молнию $\{a_1, \dots, a_n = a_1\}$. Назовем *разложением* расстановку чисел в проекциях точек этой молнии на ось Ox и в проекциях точек этой молнии на ось Oy. Можно ли так расставить в точках молнии числа $f_1, \dots, f_n \in \mathbf{R}$ с $f_1 = f_n$, чтобы для любого разложения некоторое число f_i не было бы равно сумме двух чисел, стоящих в $x(a_i)$ и в $y(a_i)$?

Подмножество $K \subset \mathbf{R}^2$ плоскости называется *разрывно базисным*, если для любой функции $f : K \to \mathbf{R}$ существуют такие функции $g, h : \mathbf{R} \to \mathbf{R}$, что $f(x,y) = g(x) + h(y)$ для каждой точки $(x,y) \in K$.

3. (a) Отрезок $K = 0 \times [0;1] \subset \mathbf{R}^2$ является разрывно базисным.

(b) Крест $K = 0 \times [-1;1] \cup [-1;1] \times 0 \subset \mathbf{R}^2$ является разрывно базисным.

(c) *Критерий разрывной базисности.* Подмножество плоскости разрывно базисно тогда и только тогда, когда оно не содержит замкнутых молний.

4.** Андрей Николаевич и Владимир Игоревич играют в 3D-игру 'А ну-ка, разложи!'. В кубе $n \times n \times n$, разбитом на n^3 единичных кубиков, отмечено несколько кубиков. А. Н. расставляет числа в отмеченных кубиках, как хочет. В. И. смотрит на расставленные числа и берет $3n$ чисел $a_1, \dots, a_n, b_1, \dots, b_n, c_1, \dots, c_n$ — 'весов' колонок, продольных строк и поперечных строк (т.е. рядов, параллельных оси z, оси x и оси y) — как хочет. Если число в каждом отмеченном кубике (i,j,k) (поставленное А. Н.) оказалось равным сумме $a_i + b_j + c_k$ трех весов колонки, продольной строки и поперечной строки этого кубика, то выиграл В. И., а иначе (т.е. если число хотя бы в одном отмеченном кубике оказалось не равным сумме трех весов) выиграл А. Н.

Как по набору отмеченных кубиков узнать, кто выигрывает?

(Ясно, что *алгоритм* распознавания выигрышности данного набора отмеченных кубиков существует. Желательно найти простой критерий типа того, который имеется для плоского аналога этой игры. Интересны даже ответы для маленьких n.)

---

[7] Первый и остальные пункты этой части независимы друг от друга.

5.** (a) Определите разрывную базисность подмножеств трехмерного пространства. Сформулируйте и докажите пространственный аналог приведенного критерия.

(b) То же для многомерного случая.

Решения задач.

1. (a) Это неверно. Если $f_{ij} = g_i + h_j$ для $i, j = 1, 2$, то $f_{11} + f_{22} = f_{12} + f_{21}$, но это соотношение не имеет места для некоторых наборов чисел f_{ij}.

(b) "Только тогда"следует из задачи 2. Докажем утверждение "тогда"индукцией по количеству отмеченных клеток. Если отмечена только одна клетка, утверждение задачи тривиально. Обозначим через K множество центров отмеченных клеток. Множество $E(K)$ определено в следующем пункте после задачи 9. По условию, K не содержит замкнутых молний, следовательно $\#E(K) < \#K$. Значит, по индуктивному предположению В.И. может выиграть на множестве $E(K)$. Все оставшиеся клетки являются единственными отмеченными в своей строке или в своем столбце. Следовательно, В.И. сможет выбрать и оставшиеся веса для K.

2. Да. Если каждое из чисел f_i представимо в виде суммы двух чисел, расположенных в точках $x(a_i)$ и $y(a_i)$, то $f_1 - f_2 + f_3 - \dots - f_{n-1} = 0$, но можно легко подобрать набор чисел f_i, для которого это неверно.

3. (a) Положим $h(y) = f(0, y)$ и $g(x) = 0$.

(b) Положим $g(x) = f(x, 0)$ и $h(y) = f(0, y) - f(0, 0)$.

(c) "Только тогда"следует из задачи 2. Докажем утверждение "тогда". Рассмотрим произвольную функцию $f : K \to \mathbf{R}$ и построим по ней функции g и h такие, что $f(x, y) = g(x) + h(y)$. Назовчм две точки $a, b \in K$ *эквивалентными*, если существует молния $\{a = a_1, \dots, a_n = b\} \subset K$. Возьмчм один из классов эквивалентности $K_1 \subset K$ и определим функции $g : x(K_1) \to \mathbf{R}$ и $h : y(K_1) \to \mathbf{R}$ следующим образом. Зафиксируем произвольную точку $a_1 \in K_1$. Положим $g(x(a_1)) = f(a_1)$ и $h(y(a_1)) = 0$. Если $\{a_1, a_2, \dots, a_{2l}\}$ — молния из точек множества K, то положим

$$h(y(a_{2l})) := f(a_{2l}) - f(a_{2l-1}) + \dots - f(a_1) \quad \text{и} \quad g(x(a_{2l})) := f(a_{2l-1}) - f(a_{2l-2}) + \dots + f(a_1).$$

Если $\{a_1, a_2, \dots, a_{2l+1}\}$ — молния из точек множества K, то положим

$$g(x(a_{2l+1})) = f(a_{2l+1}) - f(a_{2l}) + \dots + f(a_1)$$

(значение $h(y(a_{2l+1}))$ уже определено). Сделаем это построение для всех классов эквивалентности одновременно. Для всех же прочих точек положим $g(x) = 0$ и $h(y) = 0$.

Непрерывная базисность.

Подмножество $K \subset \mathbf{R}^2$ называется *(непрерывно) базисным*, если для любой непрерывной функции $f : K \to \mathbf{R}$ существуют такие непрерывные функции $g, h : \mathbf{R} \to \mathbf{R}$, что $f(x, y) = g(x) + h(y)$ для каждой точки $(x, y) \in K$. (Определение непрерывной функции напомнено в начале части 1.) Слово 'непрерывно' далее опускается.

Проблема Арнольда. *Какие подмножества плоскости являются базисными?*

Чтобы подойти к ответу, рассмотрим несколько примеров.

6. (a) Замкнутая молния не базисна.

(b) Отрезок $K = 0 \times [0; 1] \subset \mathbf{R}^2$ является базисным.

(c) Крест $K = 0 \times [-1; 1] \cup [-1; 1] \times 0 \subset \mathbf{R}^2$ является базисным.

7. (a) Если подмножество плоскости базисно, то оно разрывно базисно. (Определение и критерий разрывной базисности см. в предыдущем пункте.)

(b) *Пополненной молнией* называется объединение точки $a_0 \in \mathbf{R}^2$ с бесконечной молнией $\{a_1, \dots, a_n, \dots\} \subset \mathbf{R}^2$ из различных точек, *сходящейся* к точке a_0 (т.е. для любого $\varepsilon > 0$ найдется такое натуральное N, что для любого $i > N$ выполнено $|a_i, a_0| < \varepsilon$). Докажите, что никакая пополненная молния не является базисной. (Заметим, что она является разрывно базисной).

(c) Через $[a,b]$ обозначим отрезок, соединяющий точки a и b. Докажите, что крест $[(-1,-2),(1,2)]\cup[(-1,1),(1,-1)]$ не является базисным.

(d) Пусть $m_{i,j}=2-3\cdot 2^{-i}+j\cdot 2^{-2i}$. Рассмотрим множество, состоящее из точек $(m_{i,2l},m_{i,2l})$ и точек $(m_{i,2l},m_{i,2l-2})$, где $i=1,2,\dots$ и $l=1,2,3,\dots,2^{i-1}$. Докажите, что это подмножество плоскости не содержит бесконечной молнии, но содержит сколь угодно длинные молнии.

(e) Объединение множества из предыдущего пункта с точкой $(2,2)$ не базисно.

Последовательность точек a_i плоскости называется *сходящейся к точке* a, если для любого $\varepsilon>0$ найдется такое целое N, что для любого $i>N$ выполнено $|a,a_i|<\varepsilon$.

Подмножество $K\subset\mathbf{R}^2$ плоскости называется *замкнутым*, если для любой бесконечной последовательности точек $a_i\in K$, сходящейся к точке a, выполнено $a\in K$.

8. Подмножество $K\subset\mathbf{R}^2$ плоскости является замкнутым тогда и только тогда, когда для любой точки $a\notin K$ найдется такое $\varepsilon>0$, что любая точка плоскости с расстоянием менее ε до a не принадлежит K.

Критерий базисности. *Замкнутое ограниченное подмножество плоскости базисно тогда и только тогда, когда оно не содержит сколь угодно длинных молний* [St89].

Приведем здесь замечания и переформулировку (используемую в доказательстве). Само доказательство приводится в следующем пункте.

9. (a) Условие замкнутости в критерии действительно необходимо (т.е. если в формулировке теоремы опустить это условие, то получится неверное утверждение).

(b) Условие ограниченности в критерии действительно необходимо (т.е. если в формулировке теоремы опустить это условие, то получится неверное утверждение).

(c)** Найдите критерий базисности для замкнутых (но неограниченных) подмножеств плоскости.

Пусть K — подмножество плоскости $\mathbf{R}^2$. Для каждой точки $v\in K$ нарисуем две прямые, проходящие через v параллельно координатным осям. Если хотя бы одна из этих двух прямых пересекает K только в точке v, то покрасим v в белый цвет. Обозначим через $E(K)$ множество всех точек K, не являющихся белыми:

$$E(K)=\{v\in K:\ |K\cap(x=x(v))|\geq 2 \text{ и } |K\cap(y=y(v))|\geq 2\}.$$

Например, рис. 6b получается из рис. 6a операцией E. Пусть $E^2(K)=E(E(K))$, $E^3(K)=E(E(E(K)))$ и т.д.

10. Подмножество $K\subset\mathbf{R}^2$ не содержит сколь угодно длинных молний тогда и только тогда, когда $E^n(K)=\emptyset$ для некоторого n.

12. (a)* Докажите элементарно (т.е. без использования описания пространства $C^*(K)$ в терминах мер, см. следующий пункт), что если $K\subset\mathbf{R}^2$ замкнуто и ограничено, причем $E(K)=\emptyset$, то K базисно [Mi09].

Указание. Получите сначала разложение $f(x,y)=g(x)+h(y)$ для *кусочно-линейных* функций f, причем $|g|+|h|<5|f|$.

(b)** Докажите элементарно часть 'тогда' критерия базисности.

Указание. То же, $|g|+|h|<C_n|f|$, где C_n зависит только от того n, для которого $E^n(K)=\emptyset$.

11. Базисность подмножеств трехмерного пространства определена выше перед леммой Арнольда о деревьях.

(a) Докажите, что еж $0\times0\times[-1;1]\cup0\times[-1;1]\times0\cup[-1;1]\times0\times0\subset\mathbf{R}^3$ является базисным.

(b) Подмножество пространства $\mathbf{R}^3$, состоящее из четырех точек $(0,0,0)$; $(1,1,0)$; $(0,1,1)$; $(1,0,1)$, базисно. (Но $E^n(K)\neq\emptyset$ для любого n, см. ниже.)

(c)* Для $K \subset \mathbf{R}^3$ аналогично определим $E(K)$, используя вместо прямых плоскости, перпендикулярные осям координат:

$$E(K) := \{v \in K : \ |K \cap (x = x(v))| \geq 2, \ |K \cap (y = y(v)) \geq 2| \text{ и } |K \cap (z = z(v))| \geq 2\}.$$

Докажите, что если K замкнуто, ограничено и $E^n(K) = \emptyset$ для некоторого n, то K базисно [St89, Lemma 23.ii].

(d)* Никакое подмножество пространства $\mathbf{R}^3$ (или даже $\mathbf{R}^4$), гомеоморфное двумерному диску, не является базисным. Указание: определение многомерной молнии см. в [St89, 6.12, p.39].

Решения задач.

6. (a) Если бы молния $A = \{a_1, \dots, a_{2l+1} = a_1\}$ была базисной, то $f(a_1) - f(a_2) + \dots + f(a_{2l-1}) - f(a_{2l}) = 0$, но легко подобрать функцию f, для которой это не выполнено. Сравните с задачей 2.

(b),(c) Аналогично задачам 3a, 3b.

7. (a) Если множество не является разрывно базисным, то по критерию разрывной базисности из предыдущего пункта оно содержит замкнутую молнию. Тогда утверждение задачи следует из 6a, так как функция f может быть продолжена с замкнутой молнии на всч множество.

(b) Рассмотрим функцию f, для которой $f(a_i) = \frac{(-1)^i}{i}$. Предположим, что $f(x,y) = g(x) + h(y)$ для некоторых непрерывных g и h, тогда

$$f(a_1) - f(a_2) + f(a_3) - f(a_4) + \dots - f(a_{2l}) = h(y(a_1)) - h(y(a_{2l})).$$

Так как $\lim_{l \to \infty} h(y_{2l})$ существует и равен $h(y(a_0))$, то ряд $\sum_{i=1}^{2l} (-1)^i f(a_i)$ сходится при $l \to \infty$. Но это противоречит расходимости гармонического ряда.

(c) Крест содержит замкнутую молнию

$$a_{4k+1} = (\frac{-1}{4^k}, \frac{1}{4^k}), \ a_{4k+2} = (\frac{1}{2 \cdot 4^k}, \frac{1}{4^k}), \ a_{4k+3} = (\frac{1}{2 \cdot 4^k}, \frac{-1}{2 \cdot 4^k}), \ a_{4k+4} = (\frac{-1}{4^{k+1}}, \frac{-1}{2 \cdot 4^k})$$

Определим функцию f на этой молнии, используя задачу 7(b), и продолжим еч кусочно-линейно на весь крест. Не существует таких функций g и f, что $f(x,y) = g(x) + h(y)$.

(d) Для любого i точки $(m_{i,2l}, m_{i,2l})_{l=1}^{2^{i-1}}$ и $(m_{i,2l}, m_{i,2l-2})_{l=1}^{2^{i-1}}$ образуют молнию из 2^i элементов.

(e) Определим функцию $f(x,y)$ соотношениями

$$f((m_{i,2l}, m_{i,2l})) := \frac{1}{2^i} \quad \text{и} \quad f((m_{i,2l}, m_{i,2l-2})) := -\frac{1}{2^i}$$

Предположим, что $f(x,y) = g(x) + h(y)$ для некоторых непрерывных $g(x)$ и $h(y)$. Теперь для каждого i, используя молнии $(m_{i,2l}, m_{i,2l})$ и $(m_{i,2l}, m_{i,2l-2})$, где $l = 1, 2, 3, \dots 2^{i-1}$, получаем $h(2 - \frac{3}{2^i}) - h(2 - \frac{2}{2^i}) = 1$. Это противоречит непрерывности h в точке $y = 2$.

8. Докажем утверждение "только тогда". Пусть K — замкнутое подмножество плоскости. Предположим, что для некоторой точки $a = (x,y) \notin K$ и для произвольного $\varepsilon = \frac{1}{n} > 0$ существует хотя бы одна точка $a_n \in K$, для которой $|a, a_n| \leq \frac{1}{n}$. Но тогда последовательность точек $a_n \in K$ сходится к точке a, поэтому $a \in K$. Противоречие.

Теперь докажем утверждение "тогда". Пусть некоторая последовательность a_n сходится к точке a, не лежащей в множестве K. По условию существует $\varepsilon > 0$ такое, что для любой точки $a_n \in K$ расстояние $|a, a_n| > \varepsilon$. Но это противоречит сходимости последовательности.

9. (a) Любая бесконечная молния A, не содержащая замкнутых молний и сходящаяся к точке $a \notin A$, является базисной. Это следует из того, что любая функция, определчнная на A, непрерывна.

(b) Контрпримером является множество $\{(k,k)\}_{k=1}^{\infty} \cup \{(k, k-1)\}_{k=1}^{\infty}$ точек плоскости.

10. Докажем часть 'только тогда'. Предположим, что $E^n(K) \neq \emptyset$ для всех n. Для каждого n рассмотрим точку $a_0 \in E^n(K)$. Выберем точки $a_{-1}, a_1 \in E^{n-1}(K)$ такие, что $x(a_{-1}) = x(a_0)$ и $y(a_1) = y(a_0)$. Теперь можно выбрать точки $a_{-2}, a_2 \in E^{n-2}(K)$, для которых $\{a_{-2}, a_{-1}, a_0, a_1, a_2\}$ —

молния. Аналогично можно сконструировать молнию из $2n+1$ точек, лежащую целиком в множестве K. Что и требовалось доказать.

Докажем часть 'тогда'. Пусть множество K содержит молнию из $2n+1$ точки $\{a_{-n}, \dots, a_0, \dots, a_n\}$. Тогда в множестве $E(K)$ содержится молния из $2n-1$ точки $\{a_{-n+1}, \dots, a_{n-1}\}$. Продолжая, получим, что $a_0 \in E^n(K)$. Следовательно, если $E^n(K) = \emptyset$, то K не содержит молнии из $2n+1$ точек.

11. (a) Для произвольной функции $f : K \to \mathbf{R}$ на еже K определим $g(x) := f(x,0,0)$, $h(y) := f(0,y,0) - f(0,0,0)$ и $l(z) := f(0,0,z) - f(0,0,0)$.

(b) Положим $g(0) = f(0,0,0)$, $h(0) = 0$, $l(0) = 0$,

$$2g(1) = f(0,0,0) + f(1,1,0) + f(1,0,1) - f(0,1,1),$$

$$2h(1) = -f(0,0,0) + f(1,1,0) - f(1,0,1) + f(0,1,1)) \quad \text{и}$$

$$2l(1) = -f(0,0,0) - f(1,1,0) + f(1,0,1) + f(0,1,1).$$

Доказательство критерия базисности.

Пусть K — произвольное замкнутое ограниченное подмножество плоскости. Известно, что тогда любая непрерывная функция $f : K \to \mathbf{R}$ ограничена. Функция $f : K \to \mathbf{R}$ называется *ограниченной*, если найдется число M такое, что $|f(x)| < M$ для любой точки $x \in K$. Для ограниченной функции $G : K \to \mathbf{R}$ положим $|G| := \sup_{x\in K} |G(x)|$.

Начало доказательства части 'только тогда' критерия базисности. Предположим, напротив, что K содержит сколь угодно длинные молнии и базисно. Выбирая подпоследовательности, можно добиться того, чтобы в каждой молнии точки попарно различны. Поэтому будем считать, что это выполнено. Тогда для любого $n \ge 4$ существует молния $\{a_1^n, \dots, a_{2n+5}^n\}$ из $(2n+5)$-и различных точек множества K.

Существует непрерывная функция

$$f_n : K \to \mathbf{R} \quad \text{такая, что} \quad f_n(a_i^n) = (-1)^i \quad \text{и} \quad |f_n(x)| \le 1 \quad \text{для любого} \quad x \in K.$$

(Действительно, построим сначала непрерывную функцию $f : \mathbf{R}^2 \to \mathbf{R}$, удовлетворяющую этим условиям. Обозначим $s = \min_{i<j} |a_i, a_j|$. Рассмотрим n дисков с центрами в точках a_i и радиусами $\frac{s}{3}$. Вне этих дисков положим $f = 0$. Внутри i-го диска сделаем f линейной функцией от радиуса, равной $(-1)^i$ в центре a_i и нулю на границе. Теперь ограничим построенную функцию на $K \subset \mathbf{R}^2$ и получим требуемую непрерывную функцию $K \to \mathbf{R}$.)

Определим по индукции последовательность чисел s_n и функций $F_n : K \to \mathbf{R}$. Положим $s_0 = 1$ и $F_0 = 0$. Предположим, что s_{n-1} и F_{n-1} уже определены. Возьмем функции $G_{n-1}, H_{n-1} : \mathbf{R} \to \mathbf{R}$, для которых $F_{n-1}(x,y) = G_{n-1}(x) + H_{n-1}(y)$ (если таких функций нет, то все доказано). Берем

$$s_n > s_{n-1}! \cdot (|G_{n-1}| + n) \quad \text{и} \quad F_n = F_{n-1} + \frac{f_{s_n}}{s_{n-1}!}$$

Достаточно доказать, что функция

$$F = \sum_{n=1}^{\infty} \frac{f_{s_n}}{s_{n-1}!} = \lim_{n\to\infty} F_n$$

не представима в виде $G(x) + H(y)$.

Предположим, что, напротив $F(x,y) = G(x) + H(y)$ для некоторых G и H. Для получения противоречия достаточно доказать, что $|G| > n$ для каждого n. А для этого достаточно показать, что $s_{n-1}!|G - G_{n-1}| > s_n$: тогда будет

$$|G| + |G_{n-1}| \ge |G - G_{n-1}| > \frac{s_n}{s_{n-1}!} > |G_{n-1}| + n.$$

Лемма. *Пусть* $m \geq 4$,

• $K = \{a_1, \dots, a_{2m+5}\}$ *— молния из* $2m+5$ *различных точек на плоскости,*

• $f(a_1), \dots, f(a_{2m+5})$ *— числа, для которых* $|(-1)^i - f(a_i)| < 1/m$ *и*

• $g(x(a_i)), h(y(a_i))$, $i = 1, \dots, 2m+5$, *— такие числа, что* $f(a_i) = g(x(a_i)) + h(y(a_i))$ *для любого* i *(при этом если* $x(a_i) = x(a_j)$, *то* $g(x(a_i)) = g(x(a_j))$, *и аналогично для* y *и* h*).*

Тогда $\max_i |g(x(a_i))| > m$.

Доказательство. Можно считать, что прямая a_1a_2 параллельна оси Ox (если это не так, то увеличим все индексы на 1 в последующих формулах). Имеем

$$|(f(a_1) - f(a_2) + f(a_3) - f(a_4) + \dots - f(a_{2m+4})) - (2m+4)| \leq \frac{2m+4}{m} \leq 3.$$

Это означает, что $|g(x(a_1)) - g(x(a_{2m+4}))| \geq (2m+4) - 3 > 2m$. Отсюда следует требуемое неравенство. QED

Окончание доказательства части 'только тогда' критерия непрерывной базисности. Имеем:

$$F - F_n = F - F_{n-1} - \frac{f_{s_n}}{s_{n-1}!} = \frac{s_{n-1}!(F - F_{n-1}) - f_{s_n}}{s_{n-1}!}.$$

Применим лемму к

$$m = s_n, \quad a_i = a_i^{s_n}, \quad f = s_{n-1}!(F - F_{n-1}), \quad g = s_{n-1}!(G - G_{n-1}), \quad h = s_{n-1}!(H - H_{n-1}).$$

Это возможно, так как $f(x,y) = g(x) + h(y)$ и (так как $s_n - 1 > s_{n-1}$ при $n > 2$)

$$|f - f_{s_n}| = s_{n-1}!|F - F_n| < \frac{1}{(s_n - 1)\cdot s_n}\sum_{k=0}^{\infty}\frac{1}{(s_n+1)\cdot\dots\cdot s_{n+k}} < \frac{1}{(s_n - 1)\cdot s_n}\sum_{k=0}^{\infty}\frac{1}{2^k} < \frac{1}{s_n}.$$

По лемме получим $s_{n-1}!|G - G_{n-1}| > s_n$. QED

Доказательство критерия базисности [St89, §2, Лемма 23.ii]. [8] Оно основано на переформулировке свойства базисности в терминах *ограниченных линейных операторов* в *банаховых пространствах функций.* Обозначим через $C(X)$ пространство непрерывных функций на X с нормой $|f| = \sup\{|f(x)| \ : \ x \in X\}$. В этом доказательстве обозначим через $pr_x(a)$ и $pr_y(a)$ проекции точки $a \in K$ на оси координат.

Для подмножества $K \subset I^2$ определим отображение (*линейный оператор суперпозиции*)

$$\phi : C(I) \oplus C(I) \to C(K) \quad \text{формулой} \quad \phi(g,h)(x,y) = g(x) + h(y).$$

Норма на пространстве $C(I) \oplus C(I)$ вводится естественным способом. Ясно, что подмножество $K \subset I^2$ базисно тогда и только тогда, когда ϕ эпиморфно.

Обозначим через $C^*(X)$ пространство ограниченных линейных функций $C(X) \to \mathbf{R}$ с нормой $|\mu| = \sup\{|\mu(f)| \ : \ f \in C(X), \ |f| = 1\}$. Для подмножества $K \subset I^2$ определим отображение (*двойственный линейный оператор суперпозиции*)

$$\phi^* : C^*(K) \to C^*(I) \oplus C^*(I) \quad \text{как} \quad \phi^*\mu(g,h) = (\mu(g \circ pr_x), \mu(h \circ pr_y)).$$

Норма на пространстве $C^*(I) \oplus C^*(I)$ вводится естественным способом. Так как $|\phi^*\mu| \leq 2|\mu|$, то ϕ^* ограничен. По двойственности, ϕ эпиморфен тогда и только тогда, когда ϕ^* мономорфен. [9]

[8] Это доказательство неэлементарно, формально не используется в дальнейшем и может быть опущено читателем. Однако мы приводим его, поскольку наше изложение короче и яснее данного в [St89].

[9] При этом ϕ^* может быть инъективным, но не мономорфным. Другими словами, не только линейные соотношения на $\operatorname{im}\phi$ заставляют его быть строго меньше чем $C(K)$, как показывает пример небазисной пополненной молнии.
Заметим, что если $K \subset \mathbf{R}^2$ - базисное подмножество, то мы можем доказать без использования ϕ, что ϕ^* мономорфно. Определим линейный оператор $\Psi : C^*(I) \oplus C^*(I) \to C^*(K)$ формулой $\Psi(\mu_x, \mu_y)(f) = \mu_x(g) + \mu_y(h)$, где $g, h \in C(I)$ таковы, что $g(0) = 0$ и $f(x,y) = g(x) + h(y)$ для $(x,y) \in K$. Ясно, что $\Psi\phi^* = \mathrm{id}$ и Ψ ограничено, следовательно ϕ^* мономорфно.

Понятно, что ϕ^* мономорфен тогда и только тогда, когда

(*) *существует* $\varepsilon > 0$ *такое, что* $|\phi^*\mu| > \varepsilon|\mu|$ *для каждого ненулевого* $\mu \in C^*(K)$.

Чтобы работать с условием (*), используем следующий нетривиальный факт: $C^*(K)$ *совпадает с пространством σ-аддитивных регулярных вещественнозначных борелевских мер на* K (далее мы будем называть их просто 'мерами'; используется также термин 'заряды'). Имеем

$$\phi^*\mu = (\mu_x, \mu_y), \quad \text{где} \quad \mu_x(U) = \mu(pr_x^{-1}U) \quad \text{и} \quad \mu_y(U) = \mu(pr_y^{-1}U).$$

Если $\mu = \mu^+ - \mu^-$ есть разложение меры μ на положительные и отрицательные части, то $|\mu| = \bar{\mu}(X)$, где $\bar{\mu} = \mu^+ + \mu^-$ есть абсолютное значение меры μ.

Доказательство того, что условие (*) влечет отсутствие сколь угодно длинных молний, оставляем в качестве упражнения. (Это доказывает часть 'только тогда', для которой у нас уже есть элементарное доказательство.)

Остаҷтся доказать, что условие (*) следует из $E^n(K) = \emptyset$. (Это доказывает часть 'тогда'.) Приведем доказательство для $n \in \{1, 2\}$ (для произвольного n оно аналогично).

Обозначим через D_x (и D_y) множество тех точек из K, которые не затеняются никакой другой точкой из K в x- (и y-) направлении. Возьмем любую меру μ на K с нормой 1.

Если $n = 1$, то

$$E(K) = \emptyset, \quad \text{поэтому} \quad D_x \cup D_y = K, \quad \text{значит,} \quad 1 = \bar{\mu}(K) \le \bar{\mu}(D_x) + \bar{\mu}(D_y).$$

Тогда, не уменьшая общности, $\bar{\mu}(D_x) \ge 1/2$. Так как pr_x инъективна на D_x, то $|\mu_x| \ge 1/2$. Поэтому условие (*) выполнено для $\varepsilon = \frac{1}{2}$.

Если $n = 2$, то

$$E(E(K)) = \emptyset, \quad \text{поэтому} \quad D_x \cup D_y = K - E(K) \quad \text{и} \quad E(D_x \cup D_y) = \emptyset.$$

При $\bar{\mu}(E(K)) < 3/4$ имеем $\bar{\mu}(D_x \cup D_y) > 1/4$ и, не уменьшая общности, $\bar{\mu}(D_x) > 1/8$, значит, как и в случае $n = 1$, имеем $|\mu_x| > 1/8$. Поэтому условие (*) выполнено для $\varepsilon = \frac{1}{8}$.

При $\bar{\mu}(E(K)) \ge 3/4$ имеем $\bar{\mu}(K - E(K)) \le 1/4$. Как и в случае $n = 1$, не уменьшая общности, $\bar{\mu}_x(pr_x(E(K))) \ge \bar{\mu}(E(K))/2$. Следовательно $|\mu_x| \ge \frac{1}{2} \cdot \frac{3}{4} - \frac{1}{4} = \frac{1}{8}$. Поэтому условие (*) выполнено для $\varepsilon = \frac{1}{8}$. QED

Гладкая базисность.

Пусть K — подмножество плоскости $\mathbf{R}^2$. Функция $f : K \to \mathbf{R}$ называется *дифференцируемой* (по Уитни), если для любой точки $z_0 \in K$ существуют такие вектор $a \in \mathbf{R}^2$ и бесконечно малая функция $\alpha : \mathbf{R}^2 \to \mathbf{R}$, что для любой точки $z \in K$ выполнено

$$f(z) = f(z_0) + a \cdot (z - z_0) + \alpha(z - z_0)|z, z_0|.$$

Здесь точка означает знак скалярного произведения векторов $a =: (f_x, f_y)$ и $z - z_0 =: (x, y)$, т.е. $a \cdot (z - z_0) = xf_x + yf_y$. Функция $\alpha : \mathbf{R}^2 \to \mathbf{R}$ называется *бесконечно малой*, если для любого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что для любой точки $(x, y) \in \mathbf{R}^2$

$$\text{если} \quad \sqrt{x^2 + y^2} < \delta, \quad \text{то} \quad |\alpha(x, y)| < \varepsilon.$$

Подмножество $K \subset \mathbf{R}^2$ плоскости называется *дифференцируемо базисным*, если для любой дифференцируемой функции $f : K \to \mathbf{R}$ существуют такие дифференцируемые функции $g : \mathbf{R} \to \mathbf{R}$ и $h : \mathbf{R} \to \mathbf{R}$, что для любой точки $(x, y) \in K$ выполняется $f(x, y) = g(x) + h(y)$.

13. (a) (b) (c) Решите аналоги задачи 6 для дифференцируемой базисности.

14. (a) График функции $|x|$ на отрезке $[-1; 1]$ является дифференцируемо базисным.

(b) Ломаная с последовательными вершинами $(-2,0)$, $(-1,1)$, $(0,0)$, $(1,1)$ и $(2,0)$ не является дифференцируемо базисной. (Заметим, что она базисна.)

(c) Пополненная молния $\{([\frac{n+1}{2}]^{-1/2},[\frac{n}{2}]^{-1/2})\}_{n=1}^{\infty}\cup\{(0,0)\}$ не является дифференцируемо базисной. (Заметим, что она не является также базисной.)

(d) Пополненная молния $\{(2^{-[\frac{n+1}{2}]},2^{-[\frac{n}{2}]})\}_{n=1}^{\infty}\cup\{(0,0)\}$ является дифференцируемо базисной. (Заметим, что она не является базисной.)

(e)** **Гипотеза И. Шнурникова.** Пополненная молния $\{a_n\}_{n=1}^{\infty}\cup\{(0,0)\}$ дифференцируемо базисна тогда и только тогда, когда последовательность $\frac{\sum\limits_{n=k}^{\infty}|a_n|}{a_k}$ ограничена.

15. (a) Крест $K=[(-1,-2),(1,2)]\cup[(-1,1),(1,-1)]$ не является дифференцируемо базисным.

(b)** **Гипотеза.** Подмножество $\{(t^2,\frac{t^2}{(1+t)^2})\}_{t\in[-\frac{1}{2};\frac{1}{2}]}$ плоскости не является дифференцируемо базисным. Указание: можно пытаться делать аналогично задаче 15(a).

(c)** **Гипотеза.** Кусочно-линейный граф на плоскости является дифференцируемо базисным тогда и только тогда, когда он не содержит сколь угодно длинных молний, и для любых двух *сингулярных* точек a и b выполнено $x(a)\neq x(b)$ и $y(a)\neq y(b)$. Точка $a\in K$ называется *сингулярной*, если пересечение K с любым диском с центром в a не является прямолинейным отрезком.

(d)** Найдите критерий дифференцируемой базисности для графов в плоскости.

(e)** Существует ли непрерывное отображение отрезка в плоскость, образ которого является дифференцируемо базисным, но не базисным?

16. Пусть K — подмножество плоскости $\mathbf{R}^2$ и $r\ge 0$. Функция $f:K\to\mathbf{R}$ называется *r раз дифференцируемой*, если для любой точки $z_0\in K$ существуют такие многочлен $\overline{f}(z)=\overline{f}(x,y)$ степени не выше r от двух переменных x и y и бесконечно малая функция $\alpha:\mathbf{R}^2\to\mathbf{R}$, что $f(z)=\overline{f}(z-z_0)+\alpha(z-z_0)|z,z_0|^r$ для любой точки $z\in K$. (Это определение отличается от общепринятого.)

(a) Ноль раз дифференцируемые функции — это в точности непрерывные, а один раз дифференцируемые — это в точности дифференцируемые.

(b) Для любого целого положительного r определите r-дифференцируемую базисность подмножеств плоскости.

(c)* Для любого целого $k\ge 0$ найдется подмножество плоскости, r-дифференцируемо базисное для любого $r=0,1,\dots k$, но не r-дифференцируемо базисное ни для какого $r>k$ [RZ07].

(d)** Найдите критерий r-дифференцируемой базисности для графов в плоскости.

Решения задач.

13. (a), (b), (c) Аналогично задачам 6(a), 3(a) и 3(b).

14. (a) Пусть $f(x,y)$ - дифференцируемая функция. Тогда $f(x,|x|)-f(0,0)=ax+b|x|+\alpha(x,|x|)|(x,|x|)|$. Положим $h(y)=by$, $g(x)=f(x,|x|)-b|x|$. Подробнее см. [RZ07].

(b) Предположим, что данная ломаная дифференцируемо базисна. Функция $f(x,y)=xy$ дифференцируема. Поэтому $f(x,y)=g(x)+h(y)$ для некоторых дифференцируемых функций g и h. Тогда

$$2-2d=f(1+d,1-d)+f(1-d,1-d)=g(1+d)+g(1-d)+2h(1-d)=2g(1)+2h(1)-2h'(1)d+o(d).$$

Значит, $h'(1)=1$. Аналогично

$$2d-2=f(-1+d,1-d)+f(-1-d,1-d)=g(-1+d)+g(-1-d)+2h(1-d)=2g(-1)+2h(1)-2h'(1)d+o(d).$$

Значит, $h'(1)=-1$. Противоречие.

(c) Предположим, что эта пополненная молния дифференцируемо базисна. Положим $a_n = ([\frac{n+1}{2}]^{-1/2}, [\frac{n}{2}]^{-1/2})$, $f(a_n) := \frac{(-1)^n}{n}$, $n = 2, 3, \ldots$. Если $f(x, y) = g(x) + h(y)$ для некоторых функций $g(x)$ и $h(y)$, тогда $f(a_2) - f(a_3) + f(a_4) - \ldots$ сходится к $g(1) - g(0)$ (аналогично задаче 7b). Но это противоречит расходимости ряда $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$

(d) Не ограничивая общности можно считать, что $f(0,0) = 0$, тогда возьмчм $g(0) = 0$ и $h(0) = 0$. Положим

$$h(2^{-k}) = f(2^{-(k+1)}, 2^{-k}) - f(2^{-(k+1)}, 2^{-(k+1)}) + f(2^{-(k+2)}, 2^{-(k+1)}) - \ldots,$$

$$g(2^{-k}) = f(2^{-k}, 2^{-k}) - f(2^{-(k+1)}, 2^{-k}) + f(2^{-(k+1)}, 2^{-(k+1)}) - \ldots,$$

где правые части суть суммы знакопеременных рядов.

Теперь $g(x)$ и $h(y)$ могут быть продолжены до дифференцируемых функций $\mathbf{R} \to \mathbf{R}$.

15. (a) Определим

$$w(0) = 0, \quad w(4^{-i} + 4^{-3i}) = w(4^{-i}) = 0 \quad \text{и} \quad w(4^{-i} + 4^{-3i-1}) = 2^{3i} \quad \text{для} \quad i = 1, 2, 3, \ldots.$$

Продолжим теперь эту функцию кусочно-линейно до функции $w : [0; 1] \to \mathbf{R}$. Для $x \in [0; 1]$ определим $W(x)$ как площадь под графиком функции w на отрезке $[0; x]$. Определим $f(x, -x) = W(x)$ для $x \in [0; 1]$ and $f(x, y) = 0$ на остальных точках креста.

Ясно, что f дифференцируема вне $(0, 0)$. Можно проверить, что f дифференцируема и в $(0, 0)$.

Предположим, что $f(x, y) = g(x) + h(y)$ для некоторых дифференцируемых g и h. Не ограничивая общности, будем считать, что $g(0) = h(0) = 0$. Функция g не дифференцируема в точке $x = 1/4$, поскольку для $0 < d < \frac{1}{4}$ выполнено

$$g\left(\frac{1}{4} + d\right) - g\left(\frac{1}{4}\right) = W\left(\frac{1}{4} + d\right) - W\left(\frac{1}{4}\right) + W\left(\frac{1}{4^2} + \frac{d}{4}\right) - W\left(\frac{1}{4^2}\right) + \cdots >$$

$$> W\left(\frac{1}{4^{k+1}} + \frac{d}{4^k}\right) - W\left(\frac{1}{4^{k+1}}\right) = \frac{2^{3k} \cdot 4^{-3k}}{2} \geq \frac{(4d)^{3/4}}{2}.$$

Здесь

- первое равенство доказывается с использованием двух бесконечных молний из точек креста, начинающихся в точках $(\frac{1}{4} + d, -\frac{1}{4} - d)$ и $(\frac{1}{4}, -\frac{1}{4})$ и сходящихся к точке $(0, 0)$;
- $k \geq 0$ таково, что $4^{-2k} \geq 4d > 4^{-2(k+1)}$;
- первое неравенство выполнено, поскольку W невозрастающая функция;
- второе неравенство выполнено, поскольку $\frac{d}{4^k} > \frac{1}{4^{3(k+1)}}$;
- второе равенство выполнено по определению числа k.

(Аналогично доказывается, что g не дифференцируема ни в какой точке вида $x = 4^{-i}$.)

ЛИТЕРАТУРА

[Ar58] В. И. Арнольдь, *О представлении функций нескольких переменных в виде суперпозиции функций меньшего числа переменных,* Мат. Просвещение, Сер. 2, вып. 3, (1958), 41–61. http://ilib.mirror1.mccme.ru/djvu/mp2/mp2-3.djvu?, djvuopts&page=43

[Ar58'] В. И. Арнольдь, *Проблема 6,* Мат. Просвещение, Сер. 2, вып. 3, (1958), 273-274. http://ilib.mirror1.mccme.ru/djvu/mp2/mp2-3.djvu?, djvuopts&page=243

[Vi04] А. Г. Витушкин. *13-я проблема Гильберта и смежные вопросы,* УМН, т. 59, вып. 1(355), 2004. С. 11–24.

[Vo81] С. М. Воронин. *Аналитическая классификация ростков конформных отображений $(\mathbb{C}, 0) \to (\mathbb{C}, 0)$ с тождественной линейной частью,* Функц. анализ и его прил. Т. 15, вып. 1, 1981. С. 1–17.

[Vo82] С. М. Воронин. *Аналитическая классификация пар инволюций и ее приложения,* Функц. анализ и его прил. Т. 16, вып. 2, 1982. С. 21–29.

[Ku68] К. Куратовский, Топология, Мир, Москва, 1969, т. 1,2

[Ku00] V. Kurlin, *Basic embeddings into products of graphs,* Topol.Appl. 102 (2000), 113–137.

[Ku03] В. А. Курлин. *Базисные вложения графов и метод трехстраничных вложений Дынникова,* УМН, т. 58, вып. 2(350), 2003. С. 163–164.

[Ku03'] В. А. Курлин, Базисные вложения графов и метод трехстраничных вложений Дынникова, диссертация (2003), http://maths.dur.ac.uk/∼dma0vk/PhD.html

[Mi09] E. Miliczka, *Constructive decomposition of a function of two variables as a sum of functions of one variable*, Proc. AMS, 137:2 (2009), 607-614.

[MKT03] N. Mramor-Kosta and E. Trenklerova [Miliczka], *On basic embeddings of compacta into the plane,* Bull. Austral. Math. Soc. 68 (2003), 471–480.

[Pr04] В. В. Прасолов, Элементы комбинаторной и дифференциальной топологии, Москва, МЦНМО, 2004. http://www.mccme.ru/prasolov

[RS99] Д. Реповш и А. Скопенков, "Новые результаты о вложениях полиэдров и многообразий в евклидовы пространства", УМН 54:6 (1999), р. 61–109

[RZ07] D. Repovs and M. Zeljko, *On basic embeddings into the plane,* Rocky Mountain J. Math., to appear.

[Sk95] A. Skopenkov, *A description of continua basically embeddable in* $\mathbf{R}^2$*,* Topol. Appl. 65 (1995), 29–48.

[Sk05] А. Скопенков, Вокруг критерия Куратовского планарности графов, Мат. Просвещение, 9 (2005), 116–128 и 10 (2006), 276–277. http://www.mccme.ru/free-books/matprosa.html, http://arxiv.org/abs/0802.3820

[Sk] А. Б. Скопенков, Алгебраическая топология с элементарной точки зрения, Москва, МЦНМО, в печати. http://arxiv.org/abs/0808.1395

[St89] Y. Sternfeld, *Hilbert's 13th problem and dimension,* Lect. Notes Math. 1376 (1989), 1–49.

BASIC EMBEDDINGS AND HILBERT'S 13TH PROBLEM [1]

A. Skopenkov [2]

Abstract. This note is purely expository. In the course of the Kolmogorov-Arnold solution of Hilbert's 13th problem on superpositions there appeared the notion of *basic embedding*. A subset K of $\mathbf{R}^2$ is *basic* if for each continuous function $f\colon K \to \mathbf{R}$ there exist continuous functions $g, h\colon \mathbf{R} \to \mathbf{R}$ such that $f(x,y) = g(x)+h(y)$ for each point $(x,y) \in K$. We present descriptions of basic subsets of the plane (with a proof) and description of graphs basically embeddable into the plane (solutions of Arnold's and Sternfeld's problems). We present some results and open problems on the smooth version of the property of being basic. This note is accessible to undergraduates and could be an interesting easy reading for mature mathematicians. The two sections can be read independently on each other.

arXiv:1003.1586v1 [math.FA] 8 Mar 2010

HILBERT'S 13TH PROBLEM AND BASIC EMBEDDINGS

Hilbert's 13th problem

Let us recall informally the concept of *superposition*. Suppose that there is a set of functions of several variables, including all variables considered as functions. Represent each of the functions as an element of a circuit with several entries and one exit. Then a *superposition* of functions of this set is a function that can be repsesented by a circuit constructed from given elements; the circuit should not contain oriented cycles.

For example, a polynomial $a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$ is a superposition of the constant functions and the functions $f(x,y) = x+y$, $g(x,y) = xy$. It is clear that any elementary function can be represented as a superposition of functions of at most two variables. *Is it possible to represent each function of several arguments as a superposition of functions of at most two arguments?*

Since there is a 1–1 correspondence between a segment and a square, any function of three and more variables is superposition of (in general, discontinuous) functions of two variables. So the above question is only interesting for continuous functions. [3] From now on we assume all functions to be continuous, unless the contrary is explicitly specified.

Hilbert's 13th problem. *Can the equation $x^7 + ax^3 + bx^2 + cx + 1 = 0$ of degree seven be solved without using functions of three variables?*

This question was answered affirmatively in 1957 by Kolmogorov and Arnold. They proved that any continuous function of n variables defined on a compact subset of $\mathbf{R}^n$ can be represented as a superposition of continuous functions of one variable and addition. For an exposition accessible to undergraduates see [Ar58]. See also [Vi04].

Basic embeddings into higher-dimensional spaces [4]

Ostrand extended the Kolmogorov-Arnold Theorem this theorem to arbitrary n-dimensional compacta [St89]. It is in the Kolmogorov-Arnold-Ostrand papers that the notion of basic subset

---

[1] This is an English version of the paper in Russian under the same title. The English version has much shorter first section (which corresponds to two sections in Russian version), but contains solutions of problems 14a and 16c from the third section. Whenever possible I give references to surveys not to original papers. I would like to acknowledge V.I.Arnold, Yu.M. Burman, I.N. Shnurnikov, A.R. Safin, S.M. Voronin and M. Vyaliy for useful discussions, and M. Vyaliy for preparation of figures.

[2] skopenko@mccme.ru, http://dfgm.math.msu.su/people/skopenkov/papersc.ps

[3] Denote by

$$|x,y| = |(x_1,\dots,x_n),\ (y_1,\dots,y_n)| = \sqrt{(x_1-y_1)^2+\dots+(x_n-y_n)^2}$$

the ordinary distance between points $x = (x_1,\dots,x_n)$ and $y = (y_1,\dots,y_n)$ of $\mathbf{R}^n$. Let K be a subset of $\mathbf{R}^n$. A function $f : K \to \mathbf{R}$ is called *continuous* if for each point $x_0 \in K$ and number $\varepsilon > 0$ there exists a number $\delta > 0$ such that for each point $x \in K$ if $|x, x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$. E. g. the function $f(x_1,x_2) = \sqrt{x_1^2+x_2^2}$ is continuous on the plane, whereas the function $f(x_1,x_2)$ equal to the integer part of x_1+x_2 is not.

[4] This subsection is not used in the sequel and so can be omitted.

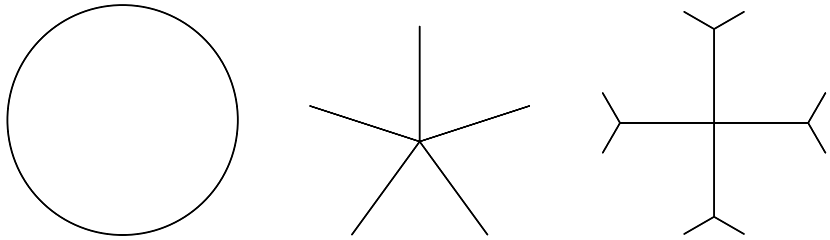


Figure 1:

appeared for the first time. It was explicitly introduced by Sternfeld [St89]. A subset $K \subset \mathbf{R}^m$ is *basic* if for each continuous function $f : K \to \mathbf{R}$ there exist continuous functions $g_1, \dots, g_m : \mathbf{R} \to \mathbf{R}$ such that $f(x_1, \dots, x_m) = g_1(x_1) + \dots + g_m(x_m)$ for each point $(x_1, \dots, x_m) \in K$.

Theorem 1. [St89] *Any n-dimensional compactum is basically embeddable into $\mathbf{R}^{2n+1}$ and, for $n > 1$, is not basically embeddable into $\mathbf{R}^{2n}$.*

It is interesting to compare this theorem with the Nöbeling-Menger-Pontryagin theorem on embeddability of any n-dimensional compact space into $\mathbf{R}^{2n+1}$ and the example of an n-dimensional polyhedron non-embeddable into $\mathbf{R}^{2n}$.

Obviously, K is basically embeddable into $\mathbf{R}$ if and only if K is topologically embeddable into $\mathbf{R}$. It follows from Theorem 1 that a compactum K is basically embeddable into $\mathbf{R}^m$ for $m > 2$ if and only if $\dim K < m/2$. Thus, the only remaining case is $m = 2$ (Sternfeld's problem).

Basic embeddings into the plane

A subset K of $\mathbf{R}^2$ is *basic* if for each continuous function $f\colon K \to \mathbf{R}$ there exist continuous functions $g, h\colon \mathbf{R} \to \mathbf{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

Let us present the characterization of arcwise connected compacta basically embeddable into the plane [Sk95] (this is a partial solution of Sternfeld's problem). We formulate the criterion first for graphs and then for the general case. A conjecture on embeddability of (not necessarily arcwise connected) connected compacta into the plane can be found in [Sk95]. Compacta used in the statements are defined after the statements.

Theorem 2. [Sk95] *A finite graph K is basically embeddable into the plane if and only if any of the following two equivalent conditions holds:*

(a) K does not contain subgraphs homeomorphic to S, C_1, C_2 (fig. 1), that is, a circle, a five-point star, and a cross with branched endpoints;

(b) K is contained in one of the graphs R_n, $n = 1, 2, 3, \dots$ (fig. 2).

Let F_1 be a triod. The graph F_{n+1} is obtained from F_n by branching its endpoints (fig. 2). The graph R_n is obtained from F_n by by adding a hanging edge to each non-hanging vertex.

Theorem 3. [Sk95] *An arcwise-connected compactum K is basically embeddable into the plane if and only if it is locally connected (i.e., is a Peano continuum) and any of the two following (equivalent) conditions hold:*

(1) K does not contain S^1, C_2, C_4, B as subcompacta and contains only finitely many subcontinua F_n, H_n (fig. 1,2,3);

(2) K does not contain any of the continua S^1, C_1, C_2, C_3, B, F, H_+, H_-, h_+, h_- (fig. 1,3,4).

Let $I = [0; 1]$. A sequence of sets is called a *null-sequence* if their diameters tend to zero. Define

• H_n to be the union of I with a null-sequence of triods having endpoints attached to I at points $3^{-l_1} + \dots + 3^{-l_s}$, where $s \le n$ and $0 < l_1 < \dots < l_s$ are integers;

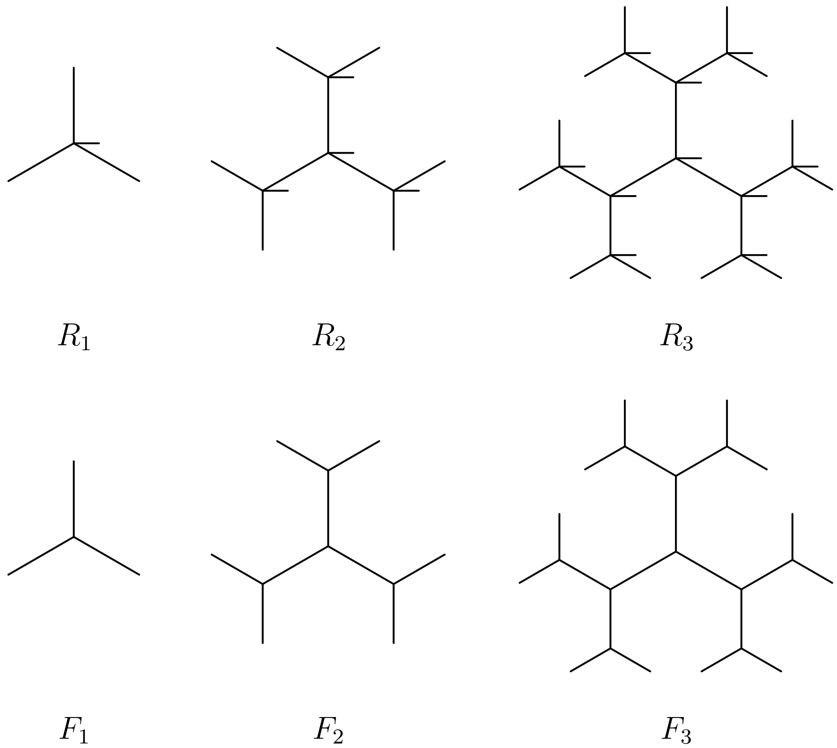


Figure 2:

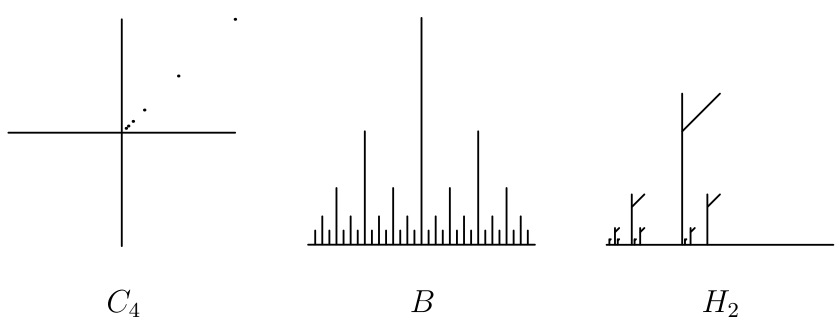


Figure 3:

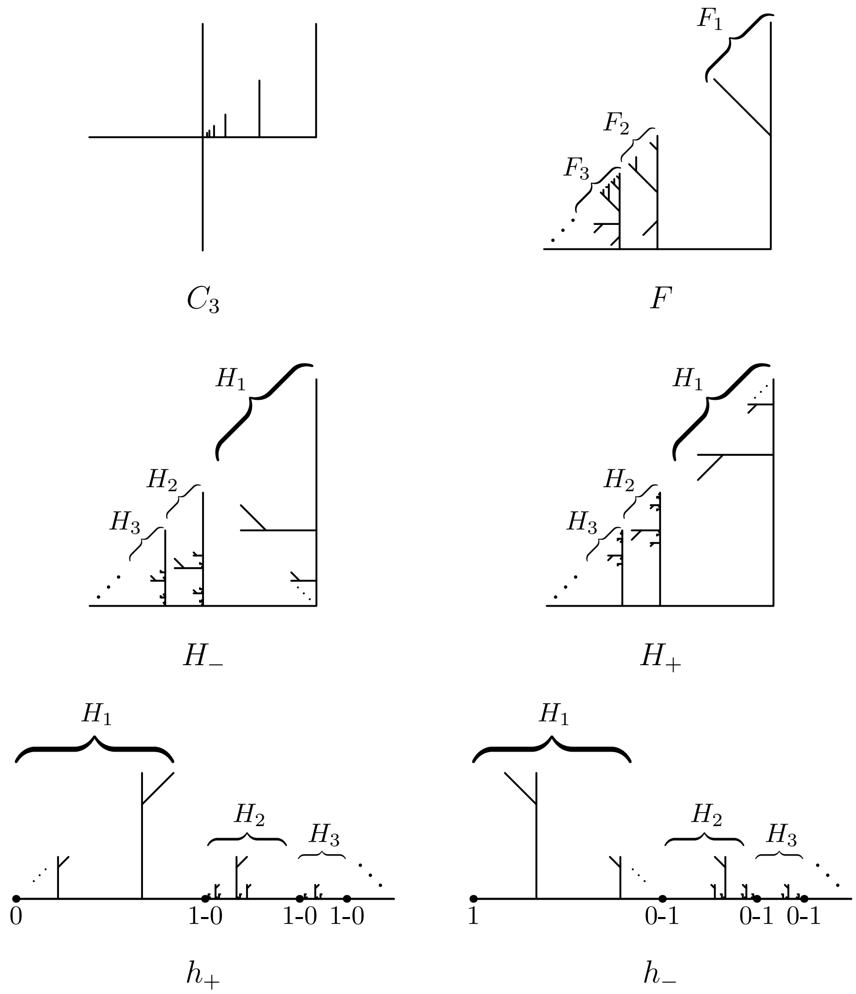

C_3
F_1
F_2
F_3
F
H_1
H_2
H_3
H_-
H_1
H_2
H_3
H_+
H_1
H_2
H_3
0
1-0
1-0
1-0
h_+
H_1
H_2
H_3
1
0-1
0-1
0-1
h_-

Figure 4:

• C_3 to be a cross with a null-sequence of arcs attached to one of its branches and converging to its center;

• C_4 to be a cross with a sequence of points converging to its center;

• B to be the union of the arc I and a null-sequence of arcs attached to $(0;1)$ by their endpoints at rational points;

• F to be the union of I with a null-sequence of sets F_n each having an endpoint attached to the point $1/n \in I$;

• H_+ (H_-) to be the union of I with a null-sequence of continua H_n connected to the points $1/n \in I$ by arcs that intersect H_n at the points $1 \in I \subset H_n$ ($0 \in I \subset H_{n-1}$, respectively);

• h_+ (h_-) to be obtained from a null-sequence of continua H_n by pasting together the points $1 \in I \subset H_n$ and $0 \in I \subset H_{n-1}$ ($0 \in I \subset H_n$ and $1 \in I \subset H_{n-1}$, respectively).

An embedding $K \subset X \times Y$ is *basic* if for any continuous function $f : K \to \mathbf{R}$ there exist continuous functions $g : X \to \mathbf{R}$, $h : Y \to \mathbf{R}$ such that $f(x,y) = g(x) + h(y)$ for any point $(x,y) \in K$.

Denote by T_n an n-od, i.e., an n-pointed star. A vertex of a graph K is called *horrible* if its degree is greater than 4 and *awful* if its degree is equal to 4 and it is not an endpoint of a hanging edge. The *defect* of a graph K is the sum $\delta(K) = (degA_1 - 2) + \dots + (degA_k - 2)$, where $A_1, \dots, A_k$ are all the horrible and awful vertices of K.

Theorem 4. [Ku99] *A finite graph K admits a basic embedding $K \subset \mathbf{R} \times T_n$ if and only if K is a tree and either $\delta(K) < n$ or $\delta(K) = n$ and K has a horrible vertex with a hanging edge.*

BASIC PLANAR SETS

The material is presented as a sequence of problems, which is peculiar not only to Zen monasteries but also to elite mathematical education (at least in Russia). Difficult problems are marked by a star, and unsolved problems by two stars. If the statement of a problem is an assertion, then it is required to prove this assertion.

Discontinuously basic subsets.

1. (a) Is it true that for any four numbers $f_{11}, f_{12}, f_{21}, f_{22}$ there exist four numbers g_1, g_2, h_1, h_2 such that $f_{ij} = g_i + h_j$ for each $i, j = 1, 2$?

(b) Andrey Nikolaevich and Vladimir Igorevich play the 'Dare you to decompose!' game. Some cells of chessboard are marked. A. N. writes numbers in the marked cells as he wishes. V. I. looks at the written numbers and chooses (as he wishes) 16 numbers $a_1, \dots, a_8, b_1, \dots, b_8$ as 'weights' of the columns and the lines. If each number in a marked cell turns out to be equal to the sum of weights of the line and the row (of the cell), then V. I. wins, and in the opposite case (i.e., when the number in at least one marked cell is not equal to the sum of weights of the line and the row) A. N. wins.

Prove that V. I. can win no matter how A. N. plays if and only if there does not exist a closed route of a rook starting and turning only at marked cells (the route is not required to pass through each marked cell).

Let $\mathbf{R}^2$ be the plane with a fixed coordinate system. Let $x(a)$ and $y(a)$ be the coordinates of a point $a \in \mathbf{R}^2$. An ordered set (either finite or infinite) $\{a_1, \dots, a_n, \dots\} \subset \mathbf{R}^2$ is called an *array* if for each i we have $a_i \neq a_{i+1}$ and $x(a_i) = x(a_{i+1})$ for even i and $y(a_i) = y(a_{i+1})$ for odd i. It is not assumed that points of an array are distinct. An array is called *closed* if $a_1 = a_{2l+1}$.

2. Consider a closed array $\{a_1, \dots, a_n = a_1\}$. A *decomposition* for such an array is an assignment of numbers at the projections of the points of the array on the x-axis and on the y-axis. Is it possible to put numbers $f_1, \dots, f_n \in \mathbf{R}$, where $f_1 = f_n$, at the points of the array so that for each decomposition there exists an f_i that is not equal to the sum of the two numbers at $x(a_i)$ and $y(a_i)$?

A subset $K \subset \mathbf{R}^2$ is called *discontinuously basic* if for each function $f : K \to \mathbf{R}$ there exist functions $g, h : \mathbf{R} \to \mathbf{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

3. (a) The segment $K = 0 \times [0; 1] \subset \mathbf{R}^2$ is discontinuously basic.

(b) The cross $K = 0 \times [-1; 1] \cup [-1; 1] \times 0 \subset \mathbf{R}^2$ is discontinuously basic.

(c) *A criterion for a subset of the plane to be discontinuously basic.* A subset of the plane is discontinuously basic if and only if it does not contain any closed arrays.

4.** Given a set of marked unit cubes in the cube $8 \times 8 \times 8$, how can we see who wins in the 3D analogue of the 'Dare you to decompose!' game? In this analogue V. I. tries to choose 24 numbers $a_1, \dots, a_8, b_1, \dots, b_8, c_1, \dots, c_8$ so that the number at the unit cube (i, j, k) would be equal to the sum $a_i + b_j + c_k$ of the three weights.

5.** (a) Define discontinuous basic subsets of the 3-space. Discover and prove the 3D analogue of the above criterion.

(b) The same for higher-dimensional case.

Solutions.

1. (a) It is not true. If $f_{ij} = g_i + h_j$ for each $i, j = 1, 2$, then $f_{11} + f_{22} = f_{12} + f_{21}$, but this is false for some numbers f_{ij}.

(b) The statement 'only if' follows from the problem 2. Let us prove the 'if' part by induction on the number of the marked cells. If only one cell is marked then we are done. Let K be the set of centres of the marked cells. The set $E(K)$ is defined in the following subsection after Problem 9. The set K does not contain any closed array, therefore $\#E(K) < \#K$. So by the induction hypothesis V. I. can win for $E(K)$. Each cell from $K - E(K)$ is the only marked cell on its line or column, thus V. I. can choose the remaining weights for K.

2. Yes, it is. If every f_i is equal to the sum of two numbers at $x(a_i)$ and $y(a_i)$, then $f_1 - f_2 + f_3 - \dots - f_{n-1} = 0$, but this is false for some numbers f_i.

3. (a) Set $h(y) = f(0, y)$ and $g(x) = 0$.

(b) Set $g(x) = f(x, 0)$ and $h(y) = f(0, y) - f(0, 0)$.

(c) The statement 'only if' follows from the problem 2. Let us prove the 'if' part. Consider a function $f : K \to \mathbf{R}$. Our aim is to construct functions g and h so that $f(x, y) = g(x) + h(y)$. Two points $a, b \in K$ are called *equivalent* if there is an array $\{a = a_1, \dots, a_n = b\} \subset K$. Now take an equivalence class $K_1 \subset K$. Define function $g : x(K_1) \to \mathbf{R}$ and $h : y(K_1) \to \mathbf{R}$ in the following way. Take any point $a_1 \in K_1$ and set $g(x(a_1)) = f(a_1)$ and $h(y(a_1)) = 0$. If $\{a_1, a_2, \dots, a_{2l}\}$ is an array, then set

$$h(y(a_{2l})) := f(a_{2l}) - f(a_{2l-1}) + \dots - f(a_1) \quad \text{and} \quad g(x(a_{2l})) := f(a_{2l-1}) - f(a_{2l-2}) + \dots + f(a_1).$$

If $\{a_1, a_2, \dots, a_{2l+1}\}$ is an array, then set $g(x(a_{2l+1})) := f(a_{2l+1}) - f(a_{2l}) + \dots + f(a_1)$ ($h(y(a_{2l+1}))$ is already defined). Make this construction for each equivalence class. Then set $g = 0$ and $h = 0$ at all other points of $\mathbf{R}$.

Continuously basic subsets.

A subset $K \subset \mathbf{R}^2$ is called *(continuously) basic* if for each continuous function $f : K \to \mathbf{R}$ there exist continuous functions $g, h : \mathbf{R} \to \mathbf{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

The Arnold problem. *Which subsets of the plane are basic?* [Ar58]

In order to approach a solution consider some examples.

6. (a) A closed array is not basic.

(b) The segment $K = 0 \times [0; 1] \subset \mathbf{R}^2$ is basic.

(c) The cross $K = 0 \times [-1; 1] \cup [-1; 1] \times 0 \subset \mathbf{R}^2$ is basic.

(d) The graph V of the function $y = |x|$, $x \in [-1; 1]$ is basic.

A sequence of points $\{a_1, \dots, a_n, \dots\} \subset \mathbf{R}^2$ *converges to a point* $a \in \mathbf{R}^2$ if for each $\varepsilon > 0$ there exists an integer N such that for each $i > N$ we have $|a_i, a| < \varepsilon$.

7. (a) If a subset of the plane is basic, then it is discontinuously basic.

(b) A *completed array* is the union of a point $a_0 \in \mathbf{R}^2$ with an infinite array $\{a_1, \dots, a_n, \dots\} \subset \mathbf{R}^2$ of distinct points which converges to the point a_0. Prove that any completed array is not basic. (Note that it is discontinuously basic).

(c) Let $[a, b]$ be the rectilinear arc which connects points a and b. Prove that the cross $K = [(-1, -2), (1, 2)] \cup [(-1, 1), (1, -1)]$ is not basic.

(d) Let $m_{ij} = 2 - 3 \cdot 2^{-i} + j \cdot 2^{-2i}$. Consider the set of points $(m_{i,2l}, m_{i,2l})$ and $(m_{i,2l}, m_{i,2l-2})$, where i varies from 1 to ∞ and $l = 1, 2, 3, \dots, 2^{i-1}$. Prove that this subset of the plane does not contain any infinite arrays but contains arbitrary long arrays.

(e) The union of the set from the previous problem and the point $(2, 2)$ is not basic.

A subset $K \subset \mathbf{R}^2$ of the plane is *closed*, if for each sequence $a_i \in K$ converging to a point a this point belongs to K.

8. A subset $K \subset \mathbf{R}^2$ of the plane is closed if and only if for each point $a \notin K$ there exists $\varepsilon > 0$ such that if for a point b of the plane we have $|a, b| \le \varepsilon$, then b does not belong to K.

The Sternfeld criterion for being a basic subset. *A closed bounded subset $K \subset \mathbf{R}^2$ of the plane is basic if and only if K does not contain arbitrary long arrays.*

9. (a) The criterion is false without the assumption that K closed.

(b) The criterion is false without the assumption that K bounded.

(c)** Find a criterion of being a basic subset for closed (but unbounded) subsets of the plane.

Suppose that K is a subset of $\mathbf{R}^2$. For every point $v \in K$ consider the pair of lines passing through v and parallel to the x-axis and the y-axis. If one of these two lines intersects K only at point v, we colour v in white. Define $E(K)$ as the set of noncoloured points of K:

$$E(K) = \{v \in K : \ |K \cap (x = x(v))| \ge 2 \text{ and } |K \cap (y = y(v))| \ge 2\}.$$

Let $E^2(K) = E(E(K))$, $E^3(K) = E(E(E(K)))$ etc.

10. A subset K of the plane does not contain arbitrary long arrays if and only if $E^n(K) = \emptyset$ for some n.

12. (a)* Give an elementary proof that if K is a closed bounded subset of $\mathbf{R}^2$ and $E(K) = \emptyset$, then K is basic [Mi09].

Hint. It can be proven that for piecewise-linear maps f there is a decomposition $f(x, y) = g(x) + h(y)$ with $|g| + |h| < 5|f|$.

(b)* Prove the 'if' part of the criterion without using the functional spaces as below.

Hint. Same as above with $|g| + |h| < C_n|f|$, where C_n depends only on that n for which $E^n(K) = \emptyset$.

11. A subset $K \subset \mathbf{R}^3$ is called *(continuously) basic* if for each continuous function $f : K \to \mathbf{R}$ there exist continuous functions $g, h, l : \mathbf{R} \to \mathbf{R}$ such that $f(x, y, z) = g(x) + h(y) + l(z)$ for each point $(x, y, z) \in K$.

(a) The 'hedgehog' $0 \times 0 \times [-1; 1] \cup 0 \times [-1; 1] \times 0 \cup [-1; 1] \times 0 \times 0 \subset \mathbf{R}^3$ is basic.

(b) The set of 4 points $(0, 0, 0)$; $(1, 1, 0)$; $(0, 1, 1)$; $(1, 0, 1)$ is basic. (But $E^n(K) \ne \emptyset$ for each n, see below.)

(c)* Define $E(K)$ analogously to the above, only instead of lines use planes orthogonal to the axes:

$$E(K) = \{v \in K : \ |K \cap (x = x(v))| \ge 2, \ |K \cap (y = y(v))| \ge 2 \text{ and } |K \cap (z = z(v))| \ge 2\}.$$

Let K be a closed bounded subset of $\mathbf{R}^3$. Prove that if $E^n(K) = \emptyset$ for some n, then K is basic [St89, Lemma 23.ii].

Solutions.

6. (a) If an array $A = \{a_1, \dots, a_{2l+1}\}$ is basic, then $f(a_1) - f(a_2) + \dots + f(a_{n-2}) - f(a_{2l}) = 0$. But this is false for some functions f. Cf. problem 2.

(b),(c) Analogously to problems 3a,3b.

(d) Take $h(y) = 0$ and $g(x) = f(x, y)$.

7. (a) If the subset is not discontinuously basic, then it contains a closed array. Hence the statement follows by extension of f on the subset and using problem 6a.

(b) Define function f by $f(a_n) = \frac{(-1)^n}{n}$. Suppose that $f(x, y) = g(x) + h(y)$ for some g and h. Then

$$f(a_1) - f(a_2) + f(a_3) - f(a_4) + \dots - f(a_{2l}) = h(y(a_1)) - h(y(a_{2l})).$$

Since $\lim_{l \to \infty} h(y_{2l})$ exists and equals to $h(y(a_0))$, it follows that $\sum_{i=1}^{2l} (-1)^i f(a_i)$ converges when $l \to \infty$, which is a contradiction.

(c) The cross contains a completed array

$$a_{4k+1} = (-2^{-2k}, 2^{-2k}),\ a_{4k+2} = (2^{-2k-1}, 2^{-2k}),\ a_{4k+3} = (2^{-2k-1}, -2^{-2k-1}),\ a_{4k+4} = (-2^{-2k-2}, -2^{-2k-1}).$$

Define a function f on this array using problem 7.b and then extend it (e.g. piecewise linearly) to the cross. Then there are no functions g and h such that $f(x, y) = g(x) + h(y)$.

(d) For every i the set $(m_{i,2l}, m_{i,2l})_{l=1}^{2^{i-1}} \cup (m_{i,2l}, m_{i,2l-2})_{l=1}^{2^{i-1}}$ is an array of 2^i points.

(e) Define a function f by

$$f((m_{i,2l}, m_{i,2l})) := 2^{-i} \quad \text{and} \quad f(m_{i,2l}, m_{i,2l-2}) := -2^{-i}.$$

If $f(x, y) = g(x) + h(y)$ for some g and h, then for every i using array of points $(m_{i,2l}, m_{i,2l})$ and $(m_{i,2l}, m_{i,2l-2})$, where $l = 1, 2, 3, \dots 2^{i-1}$, we obtain $h(2 - \frac{3}{2^i}) - h(2 - \frac{2}{2^i}) = 1$. This contradicts to the continuity of h.

8. Let us prove the 'only if' part. Let K be a closed subset of the plane. Suppose that for some point $a = (x, y) \notin K$ and for each $\varepsilon = \frac{1}{n} > 0$ there exists a point $a_n \in K$ (at least one) such that $|a, a_n| \le \frac{1}{n}$. The sequence of points $a_n \in K$ converges to the point a, thus $a \in K$. Contradiction.

Now let us prove the 'if' part. Suppose that a sequence a_n converges to a point a and the point $a = (x, y)$ is not in K. There exists $\varepsilon > 0$ such that for every point $a_n \in K$ the distance $|a, a_n| > \varepsilon$. This is a contradiction.

9. (a) Any infinite array A not containing closed arrays and converging to a point $a \notin A$ is basic. This follows because each function defined on A is continuous.

(b) A counterexample is $\{(k, k)\}_{k=1}^{\infty} \cup \{(k, k-1)\}_{k=1}^{\infty}$.

10. Let us prove the 'only if' part. Suppose that $E^n(K) \neq \emptyset$ for each n. For each n take a point $a_0 \in E^n(K)$. Then there exist points $a_{-1}, a_1 \in E^{n-1}(K)$ such that $x(a_{-1}) = x(a_0)$ and $y(a_1) = y(a_0)$. Analogously there exist points $a_{-2}, a_2 \in E^{n-2}(K)$ such that $\{a_{-2}, a_{-1}, a_0, a_1, a_2\}$ is an array. Analogously we construct an array of $2n + 1$ points in K, which is a contradiction.

Let us prove the 'if' part. Suppose that K contains an array of $2n + 1$ points $\{a_{-n}, \dots, a_0, \dots, a_n\}$. Then there is an array of $2n - 1$ points $\{a_{-n+1}, \dots, a_{n-1}\}$ in $E(K)$. Analogously $a_0 \in E^n(K)$. Thus if $E^n(K) = \emptyset$, then K does not contain an array of $2n + 1$ points.

11. (a) For each functuion $f : K \to \mathbf{R}$ on K define $g(x) := f(x, 0, 0)$, $h(y) := f(0, y, 0) - f(0, 0, 0)$ and $l(z) := f(0, 0, z) - f(0, 0, 0)$.

(b) Set $g(0) = f(0, 0, 0)$, $h(0) = 0$, $l(0) = 0$,

$$2g(1) = f(0,0,0) + f(1,1,0) + f(1,0,1) - f(0,1,1), \quad 2h(1) = -f(0,0,0) + f(1,1,0) - f(1,0,1) + f(0,1,1)$$

$$\text{and} \quad 2l(1) = -f(0,0,0) - f(1,1,0) + f(1,0,1) + f(0,1,1).$$

Proof of the criterion for being a basic subset.

Let K be a closed bounded subset of the plane. It is known that each continuous function $f : K \to \mathbf{R}$ is bounded. A function $f : K \in \mathbf{R}$ is called *bounded*, if there exists a number M such that $|f(x)| < M$ for every $x \in K$. For a bounded function $G : K \to \mathbf{R}$ denote $|G| := \sup_{x \in K} |G(x)|$.

Beginning of the proof of the 'only if' part of the criterion. Assume to the contrary that K contains arbitrary long arrays and is basic. Choosing subsequences we may assume that points of each array are distinct. Therefore for each n there is an array $\{a_1^n, \dots, a_{2n+5}^n\}$ of $2n+5$ distinct points in K.

Then there exists continuous function

$$f_n : K \to \mathbf{R} \quad \text{such that} \quad f_n(a_i^n) = (-1)^i \quad \text{and} \quad |f_n(x)| \le 1 \quad \text{for each} \quad x \in K.$$

(Indeed, first define such a continuous function $f : \mathbf{R}^2 \to \mathbf{R}$. Denote $s = \min_{i<j} |a_i, a_j|$. Take n disks with centers a_i and radii $\frac{s}{3}$. Outside of these disks set $f = 0$. Inside the i-th disk take f to be $(-1)^i$ in the center a_i, 0 on the boundary and extend it linearly in the distance to a_i. Then restrict f to $K \subset \mathbf{R}^2$.)

Define integers s_n and functions $F_n : K \to \mathbf{R}$ inductively as follows. Set $s_0 = 1$ and $F_0 = 0$. Suppose now that F_{n-1} and s_{n-1} are defined. If F_{n-1} is not representable as $G_{n-1}(x) + H_{n-1}(y)$, then we are done. If it is representable in this way, then take

$$s_n > s_{n-1}!(|G_{n-1}| + n) \qquad \text{and} \qquad F_n = F_{n-1} + \frac{f_{s_n}}{s_{n-1}!}$$

It remains to prove that if we can construct in this way an infinite number of s_n and F_n, then the function

$$F = \lim_{n\to\infty} F_n = \sum_{n=1}^{\infty} \frac{f_{s_n}}{s_{n-1}!}$$

is not representable as $G(x) + H(y)$.

Assume to the contrary that $F(x,y) = G(x) + H(y)$ for some G and H. It suffices to prove that $|G| > n$ for each n. For this it suffices to prove that $s_{n-1}!|G - G_{n-1}| > s_n$: then we would have

$$|G| + |G_{n-1}| \ge |G - G_{n-1}| > \frac{s_n}{s_{n-1}!} > |G_{n-1}| + n.$$

Lemma. *Let $m \ge 4$,*
- $K = \{a_1, \dots, a_{2m+5}\}$ *be an array of $2m+5$ distinct points,*
- $f(a_1), \dots, f(a_{2m+5})$ *numbers such that $|(-1)^i - f(a_i)| \le 1/m$,*
- $g(x(a_i)), h(y(a_i))$, $i = 1, \dots, 2m+5$, *numbers such that $f_i = g(x(a_i)) + h(y(a_i))$ for each i.*

Then $\max_i |g(x(a_i))| > n$.

Proof. We may assume that $a_1a_2 \| Ox$. Then

$$|(f(a_1) - f(a_2) + f(a_3) - f(a_4) + \dots - f(a_{2m+4}) - (2m+4)| \le \frac{2m+4}{m} \le 3.$$

Therefore $g(x(a_1)) - g(x(a_{2m+4})) \ge (2m+4) - 3 > 2m$. This implies the required inequality. $\square$

Completion of the proof of the 'only if' part of the criterion. We have

$$F - F_n = F - F_{n-1} - \frac{f_{s_n}}{s_{n-1}!} = \frac{s_{n-1}!(F - F_{n-1}) - f_{s_n}}{s_{n-1}!}.$$

Apply the Lemma to

$$m = s_n, \quad a_i = a_i^{s_n}, \quad f = s_{n-1}!(F - F_{n-1}), \quad g = s_{n-1}!(G(x) - G_{n-1}(x)), \quad h = s_{n-1}!(H(y) - H_{n-1}(y)).$$

This is possible because $f(x,y) = g(x) + h(y)$ and (since $s_n - 1 > s_{n-1}$ for $n > 2$)

$$|f - f_{s_n}| = s_{n-1}!|F - F_n| < \frac{1}{(s_n - 1)\cdot s_n} \sum_{k=0}^{\infty} \frac{1}{(s_n+1)\cdot \dots \cdot s_{n+k}} < \frac{1}{(s_n - 1)\cdot s_n} \sum_{k=0}^{\infty} \frac{1}{2^k} < \frac{1}{s_n}.$$

By Lemma we obtain $s_{n-1}!|G - G_{n-1}| > s_n$. $\square$

Proof of the criterion. [5] The proof is based on a reformulation of the property of being a basic subset in terms of *bounded linear operators* in *Banach functional spaces.* Denote by $C(X)$ the space of continuous functions on X with the norm $|f| = \sup\{|f(x)| \ : \ x \in X\}$. In this proof denote by $pr_x(a)$ and $pr_y(a)$ the projections of a point $a \in K$ on the coordinate axes.

For $K \subset I^2 := [0;1] \times [0;1]$ define a map (*linear superposition operator*)

$$\phi\colon C(I) \oplus C(I) \to C(K) \quad \text{by} \quad \phi(g,h)(x,y) := g(x) + h(y).$$

Clearly, the subset $K \subset I^2$ is basic if and only if ϕ is surjective, or equivalently, epimorphic.

Denote by $C^*(X)$ the space of *bounded linear functions* $C(X) \to \mathbf{R}$ with the norm $|\mu| = \sup\{|\mu(f)| \ : \ f \in C(X),\ |f| = 1\}$. For a subset $K \subset I^2$ define a map (*dual linear superposition operator*)

$$\phi^*\colon C^*(K) \to C^*(I) \oplus C^*(I) \quad \text{by} \quad \phi^*\mu(g,h) := (\mu(g \circ pr_x), \mu(h \circ pr_y)).$$

Since $|\phi^*\mu| \le 2|\mu|$, it follows that ϕ^* is bounded. By duality, ϕ is epimorphic if and only if ϕ^* is monomorphic. [6]

It is clear that ϕ^* is monomorphic if and only if

(*) there exists $\varepsilon > 0$ such that $|\phi^*\mu| > \varepsilon|\mu|$ for each unzero $\mu \in C^*(K)$.

We leave as an excercise the proof that (*) implies the abcense of arbitrarily large arrows. (This proves the 'only if' part of the criterion, for which we already have an elementary proof.)

So it remains to prove that $E^n(K) = \emptyset$ implies the condition (*). We present the proof for $n \in \{1,2\}$. The proof for arbitrary n is analogous. We use the following non-trivial fact: *$C^*(X)$ is the space of σ-additive regular real valued Borel measures on X* (in the sequel we call them simply 'measures'). We have

$$\phi^*\mu = (\mu_x, \mu_y), \quad \text{where} \quad \mu_x(U) = \mu(pr_x^{-1}U) \quad \text{and} \quad \mu_y(U) = \mu(pr_y^{-1}U) \quad \text{for each Borel set} \quad U \subset I.$$

If $\mu = \mu^+ - \mu^-$ is the decomposition of a measure μ into its positive and negative parts, then $|\mu| = \bar\mu(X)$, where $\bar\mu = \mu^+ + \mu^-$ is the absolute value of μ.

Let D_x (D_y) be the set of points of K which are not shadowed by some other point of K in x- (y-) direction. Take any measure μ on K of the norm 1.

If $n = 1$, then

$$E(K) = \emptyset, \quad \text{then} \quad D_x \cup D_y = K, \quad \text{so} \quad 1 = \bar\mu(K) \le \bar\mu(D_x) + \bar\mu(D_y).$$

Therefore without loss of generality, $\bar\mu(D_x) \ge 1/2$. Since the projection onto the x-axis is injective over D_x, it follows that $|\mu_x| \ge 1/2$, thus the required assertion holds for $\varepsilon = \frac{1}{2}$.

If $n = 2$, then

$$E(E(K)) = \emptyset, \quad \text{hence} \quad D_x \cup D_y = K - E(K), \quad \text{so} \quad E(D_x \cup D_y) = \emptyset.$$

In the case when $\bar\mu(E(K)) < 3/4$ we have $\bar\mu(D_x \cup D_y) > 1/4$ and without loss of generality $\bar\mu(D_x) > 1/8$. Then as for $n = 1$ we have $|\mu_x| > 1/8$, thus (*) holds for $\varepsilon = \frac{1}{8}$.

---

[5] This proof is not elementary, is not used in the sequel and could be omitted.

[6] We remark that ϕ^* can be injective but not monomorphic. In other words not only some linear relation on $\operatorname{im}\phi$ can force it to be strictly less than $C(K)$.
If an embedding $K \subset \mathbf{R}^2$ is basic, then we can prove that ϕ^* is monomorphic without use of ϕ as follows. Define a linear operator

$$\Psi\colon C^*(I) \oplus C^*(I) \to C^*(K) \quad \text{by} \quad \Psi(\mu_x, \mu_y)(f) = \mu_x(g) + \mu_y(h),$$

where $g, h \in C(I)$ are such that $g(0) = 0$ and $f(x,y) = g(x) + h(y)$ for $(x,y) \in K$. Clearly, $\Psi\phi^* = \mathrm{id}$ and Ψ is bounded, hence ϕ^* is monomorphic.

In the case when $\bar\mu(E(K)) \ge 3/4$ we have $\bar\mu(K - E(K)) \le 1/4$. By the case $n = 1$ above without loss of generality $\bar\mu_x(pr_x(E(K))) \ge \bar\mu(E(K))/2$. Hence $|\mu_x| \ge \frac{1}{2}\cdot\frac{3}{4} - \frac{1}{4} = \frac{1}{8}$, thus (*) holds for $\varepsilon = \frac{1}{8}$. $\square$

Smoothly basic subsets of the plane.

Let K be a subset of the plane $\mathbf{R}^2$. A function $f : K \to \mathbf{R}$ is called *differentiable* if for each point $z_0 \in K$ there exist a vector $a \in \mathbf{R}^2$ and infinitesimal function $\alpha : \mathbf{R}^2 \to \mathbf{R}$ such that for each point $z \in K$

$$f(z) = f(z_0) + a \cdot (z - z_0) + \alpha(z - z_0)|z, z_0|.$$

Here the dot denotes scalar product of vectors $a =: (f_x, f_y)$ and $z - z_0 =: (x, y)$, i.e. $a\cdot(z - z_0) = xf_x + yf_y$. A function $\alpha : \mathbf{R}^2 \to \mathbf{R}$ is *infinitesimal*, if for each number $\varepsilon > 0$ there exists a number $\delta > 0$ such that for each point $(x, y) \in \mathbf{R}^2$

$$\text{if} \quad \sqrt{x^2 + y^2} < \delta, \quad \text{then} \quad |\alpha(x, y)| < \varepsilon.$$

Let V be the graph of the function $y = |x|$, where $x \in [-1; 1]$. A function $f : V \to \mathbf{R}$ is differentiable if and only if $f(x, |x|)$ is differentiable on the segments $[-1; 0]$ and $[0; 1]$.

A subset $K \subset \mathbf{R}^2$ of the plane is called *differentiably basic* if for each differentiable function $f : K \to \mathbf{R}$ there exist differentiable functions $g : \mathbf{R} \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$ such that $f(x, y) = g(x) + h(y)$ for each point $(x, y) \in K$.

13. (a) (b) (c) Solve the analogues of problem 6 for differentiably basic sets.

14. (a) The graph V is differentiably basic.

(b) $W := (V - (2, 0)) \cup (V + (2, 0))$ is not differentiably basic.

(c) The broken line whose consecutive vertices are $(-2, 0)$, $(-1, 1)$, $(0, 0)$, $(1, 1)$ and $(2, 0)$ is not differentiably basic. (Note that it is continuously basic).

(d) The completed array $\{([\frac{n+1}{2}]^{-1/2}, [\frac{n}{2}]^{-1/2})\}_{n=2}^{\infty} \cup \{(0, 0)\}$ is not differentiably basic. (Note that it is also not continuously basic.)

(e) The completed array $\{(2^{-[\frac{n+1}{2}]}, 2^{-[\frac{n}{2}]})\}_{n=1}^{\infty} \cup \{(0, 0)\}$ is differentiably basic. (Note that it is not continuously basic.)

(f) (I. Shnurnikov) The cross $K = [(-1, -2), (1, 2)] \cup [(-1, 1), (1, -1)]$ is not differentiably basic. (This assertion and Conjecture 15a imply that the property of being differentably basic is not hereditary.)

(g) If a graph is basically embeddable in the plane, then it is differentiably basically embeddable in the plane. (This is non-trivial because the plane contains graphs which are basic but not differentaibly basic and vice versa.) [RZ06]

15. Conjectures.** (a) (I. Shnurnikov) A completed array $\{a_n\}_{n=1}^{\infty} \cup \{(0, 0)\}$ is differentiably basic if and only if the sequence $\frac{\sum_{n=k}^{\infty} |a_n|}{|a_k|}$ is bounded.

(b) The subset $\{(t^2, \frac{t^2}{(1+t)^2})\}_{t\in[-\frac{1}{2};\frac{1}{2}]}$ of the plane is not differentiably basic.

Hint. One can try to prove this analogously to 14f. Cf. [Vo81, Vo82].

(c) A piecewise-linear graph in $\mathbf{R}^2$ is differentiably basic if and only if it does not contain arbitrary long arrays and for each two singular points a and b we have $x(a) \ne x(b)$ and $y(a) \ne y(b)$. A point $a \in K$ is *singular* if the intersection of K with each disk centered at a is not a rectilinear arc.

It would be interesting to find a criterion of being differentiably basic for closed bounded subsets of the plane. Apparently a simple-to-state criterion (analogous to the Sternfeld criterion) does not exist. Another interesting question: is there a continuous map $[0; 1] \to \mathbf{R}^2$ whose image is differentiably basic but not basic?

16. Let $r\ge 0$ be an integer and $K\in\mathbf{R}^2$ a subset. A function $f:K\to\mathbf{R}$ is called r *times differentiable* if for each point $z_0\in K$ there exist a polynomial $\overline{f}(z)=\overline{f}(x,y)$ of degree at most r of 2 variables x and y and an infinitesimal function $\alpha:\mathbf{R}^2\to\mathbf{R}$ such that $f(z)=\overline{f}(z-z_0)+\alpha(z-z_0)|z,z_0|^r$ for each point $z\in K$. (This definition differs from the one generally accepted.)

(a) Functions differentiable zero times are exactly continuous functions, and functions differentiable one time are exactly differentiable functions.

(b) For each positive integer r define the property of being an r times differentiably basic subset of the plane $\mathbf{R}^2$.

(c) For each integer $k\ge 0$ there is a subset of the plane which is r times differentiably basic for $r=0,1\dots k$ but is not r times differentiably basic for each $r>k$.

(d)** Find a criterion for graphs in $\mathbf{R}^2$ to be r times differentiably basic.

Solutions.

13. (a), (b), (c) Analogously to problems 6(a), 3(a) and 3(b).

14. (a) Take a differentiable function $f:V\to\mathbf{R}$. Since f is differentiable at $(0,0)$, it follows that there exist $a,b\in\mathbf{R}$ such that

$$f(x,|x|)=f(0,0)+ax+b|x|+\alpha(x),\quad\text{where}\quad \alpha(x)=o(\sqrt{x^2+|x|^2})\quad\text{when}\quad x\to 0.$$

Take $h(y):=by$ and $g(x):=f(0,0)+ax+\alpha(x)$. Clearly, h is differentiable and g is differentiable outside 0. Since $\alpha(x)=o(x)$ when $x\to 0$, it follows that g is differentiable also at 0.

(b) See 16c for $k=0$.

(c) Suppose the broken line is differentiably basic. The function $f(x,y)=xy$ is differentiable. We have $f(x,y)=g(x)+h(y)$, where both g and h are differentiable. Then

$$2-2d=f(1+d,1-d)+f(1-d,1-d)=g(1+d)+g(1-d)+2h(1-d)=2g(1)+2h(1)-2h'(1)d+o(d).$$

Hence $h'(1)=1$. Analogously

$$2d-2=f(-1+d,1-d)+f(-1-d,1-d)=g(-1+d)+g(-1-d)+2h(1-d)=2g(-1)+2h(1)-2h'(1)d+o(d).$$

Hence $h'(1)=-1$. A contradiction.

(d) Suppose that this completed array is differentiably basic. Set $a_n=([\frac{n+1}{2}]^{-1/2},[\frac{n}{2}]^{-1/2})$, $f(a_n):=\frac{(-1)^n}{n}$, $n=2,3,\dots$. If $f(x,y)=g(x)+h(y)$ for some functions $g(x)$ and $h(y)$, then the series $f(a_2)-f(a_3)+f(a_4)-\dots$ converges to $g(1)-g(0)$ (analogously to Problem 7b). This is a contradiction because the series $\frac12+\frac13+\frac14+\dots$ diverges.

(e) Without loss of generality assume that $f(0,0)=0$, then take $g(0)=0$ and $h(0)=0$. Set

$$h(2^{-k})=f(2^{-(k+1)},2^{-k})-f(2^{-(k+1)},2^{-(k+1)})+f(2^{-(k+2)},2^{-(k+1)})-\dots,$$

$$g(2^{-k})=f(2^{-k},2^{-k})-f(2^{-(k+1)},2^{-k})+f(2^{-(k+1)},2^{-(k+1)})-\dots,$$

where the right-hand sides are sums of alternating series. Now $g(x)$ and $h(y)$ may be extended to differentiable functions $\mathbf{R}\to\mathbf{R}$.

(f) Define

$$w(0)=w(4^{-i}+4^{-3i})=w(4^{-i})=0\quad\text{and}\quad w(4^{-i}+4^{-3i-1})=2^{3i}\quad\text{for}\quad i=1,2,3,\dots.$$

Extend piecewise-linearly to obtain a function $w:[0;1]\to\mathbf{R}$. For $x\in[0;1]$ define $W(x)$ as the area under the graph of w on $[0;x]$. (This is well-defined because this area is finite.) Define $f(x,-x)=W(x)$ for $x\in[0;1]$ and $f(x,y)=0$ on the rest of the cross.

Clearly, f is differentiable outside $(0,0)$. It is easy to check that f is differentiable at $(0,0)$.

Suppose that $f(x,y)=g(x)+h(y)$ for some differentiable functions g and h. Without loss of generality we assume that $g(0)=h(0)=0$. The function g is not differentiable at $x=1/4$ because for $0<d<\frac{1}{4}$ we have

$$g\left(\frac{1}{4}+d\right)-g\left(\frac{1}{4}\right)=W\left(\frac{1}{4}+d\right)-W\left(\frac{1}{4}\right)+W\left(\frac{1}{4^2}+\frac{d}{4}\right)-W\left(\frac{1}{4^2}\right)+\dots>$$
$$>W\left(\frac{1}{4^{k+1}}+\frac{d}{4^k}\right)-W\left(\frac{1}{4^{k+1}}\right)=\frac{2^{3k}\cdot 4^{-3k}}{2}\ge\frac{(4d)^{3/4}}{2}.$$

Here

- the first equality is proved using two infinite arrays starting at points $(\frac{1}{4}+d,-\frac{1}{4}-d)$ and $(\frac{1}{4},-\frac{1}{4})$ and converging to the point $(0,0)$;
- $k\ge 0$ is such that $4^{-2k}\ge 4d>4^{-2(k+1)}$;
- the first inequality follows because W is a non-decreasing function;
- the second inequality follows because $\frac{d}{4^k}>\frac{1}{4^{3(k+1)}}$;
- the second equality follows by definition of k.

(In the same way one can prove that g is not differentiable at $x=4^{-i}$ for each i.)

15. (a) Hints. For the 'only if' part use the idea of Problem 7b and prove that if $\sum\limits_{n=1}^{\infty}|a_n|=\infty$, then there is a sequence $b_n\to 0$ such that $\sum\limits_{n=1}^{\infty}|a_n|b_n=\infty$.

For the 'if' part we may assume that numbers $x(a_i)$ are distinct, numbers $y(a_i)$ are distinct, $x(a_{2i})=x(a_{2i+1})$, $y(a_{2i})=y(a_{2i-1})$. If $f(0,0)=0$, define

$$g(x(a_{2i})):=f(a_1)-f(a_2)+f(a_3)-\dots+f(a_{2i+1}),\quad g(0):=\sum_{i=1}^{\infty}(-1)^i f(a_i),$$

$$h(y(a_{2i})):=-f(a_1)+f(a_2)-f(a_3)-\dots+f(a_{2i-2})\quad\text{and}\quad h(0):=\sum_{i=1}^{\infty}(-1)^i f(a_i).$$

Prove that g and h are differentiable at 0.

16. (a) It is clear.

(b) A subset $K\subset\mathbf{R}^2$ is called r *times differentiably basic* if for each r times differentiable function $f:K\to\mathbf{R}$ there exist r times differentiable functions $g:\mathbf{R}\to\mathbf{R}$ and $h:\mathbf{R}\to\mathbf{R}$ such that $f(x,y)=g(x)+h(y)$ for each point $(x,y)\in K$.

(c) We can take the graph V_k of the function $y=|x|^k$, $x\in[-1;1]$ for k odd, and $W_{k+1}=(V_{k+1}-(2,0))\cup(V_{k+1}+(2,0))$ for k even.

Proof for k even. Let us prove that W_{k+1} is r times differentiably basic for each $0\le r\le k$. Given an r times differentiable function $f:W_{k+1}\to\mathbf{R}$, take functions $h(y)=0$ and $g(x)=f(x,|x-2\operatorname{sign}x|^{k+1})$. Clearly, h is r times differentiable and $f(x,y)=g(x)+h(y)$ for each $(x,y)\in W_{k+1}$. Since the function $p(t)=|t|^{k+1}$ is k times differentiable and $r\le k$, it follows that g is r times differentiable.

Let us prove that W_{k+1} is not r times differentiably basic for k even and each $k<r$. Define a function $f:W_{k+1}\to\mathbf{R}$ by $f(x,y)=y\operatorname{sign}x$. Clearly, f is r times differentiable. If W_{k+1} is r times differentiably basic, then there are r times differentiable functions g and h such that $f(x,y)=g(x)+h(y)$. For $t\in[-1;1]$ we have

$$g(\pm 2+t)+h(|t|^{k+1})=f(\pm 2+t,|t|^{k+1})=\pm|t|^{k+1}.$$

Since g is $(k+1)$ times differentiable and $k+1$ is odd, it follows that $h'(0)=+1$ and $h'(0)=-1$, which is a contradiction. $\square$

Proof for k odd. First we prove that V_k is r times differentiably basic for each $0\le r\le k$. Take an r times differentiable function $f:V_k\to\mathbf{R}$. Since f is r times differentiable at $(0,0)$, it follows that there exist $\{a_{ij}\}_{i,j=0}^r\subset\mathbf{R}$ such that

$$a_{00}=f(0,0)\quad\text{and}\quad f(x,|x|^k)=\sum_{i,j=0}^{r}a_{ij}x^i|x|^{kj}+o([x^2+x^{2r}]^{r/2})\quad\text{when}\quad x\to 0.$$

Since

$$o([x^2+x^{2r}]^{r/2})=o_1(x^r),\quad \text{we have}\quad f(x,|x|^k)=a_{00}+a_{01}|x|^k+a_{10}x+\dots+a_{r0}x^r+o_2(x^r).$$

Take $h(y)=a_{01}y$ and $g(x)=f(x,|x|^k)-h(|x|^k)$. Clearly, h is r times differentiable and g is r times differentiable outside 0. We also have $g(x)=a_{00}+a_{10}x+\dots+a_{r0}x^r+o_2(x^r)$ when $x\to 0$. So g is r times differentiable also at 0.

Next we prove that $V=V_1$ is not r times differentiably basic for each $1<r$. Define a differentiable function $f:V\to\mathbf{R}$ by $f(x,y)=xy$, where $y=|x|$. If V is r times differentiably basic for some $r\ge 2$, then there are r times differentiable functions

$$g,h:\mathbf{R}\to\mathbf{R}\quad \text{such that}\quad f(x,|x|)=x|x|=g(x)+h(|x|).$$

Hence $g(x)-g(-x)=2x^2$ for $x\in[0;1]$. But this is impossible because g is 2 times differentiable, hence for $x\to +0$

$$g(x)=g(0)+ax+bx^2+o(x^2)\quad \text{and}\quad g(-x)=g(0)-ax+bx^2+o(x^2).$$

At last we prove that V_k is not r times differentiably basic for k odd and each $k<r$. Define a differentiable function $f:V_k\to\mathbf{R}$ by $f(x,y)=xy$, where $y=|x|^k$. If V is r times differentiably basic for some $r>k$, then there are r times differentiable functions

$$g,h:\mathbf{R}\to\mathbf{R}\quad \text{such that}\quad f(x,|x|^k)=x|x|^k=g(x)+h(|x|^k).$$

Hence $g(x)-g(-x)=2x^{k+1}$ for each $x\in[0;1]$. But this is impossible for k odd because g is $(k+1)$ times differentiable, hence for $x\to +0$

$$g(x)=g_0+g_1x+\dots+g_{k+1}x^{k+1}+o(x^{k+1})\quad \text{and}\quad g(-x)=g_0-g_1x+\dots+g_{k+1}x^{k+1}+o(x^{k+1}).\qquad \square$$

References

[Ar58] V.I. Arnold, *Representation of functions of some number of variables as superposition of functions of less number of variables (in Russian)*, Mat. Prosveschenie, 3 (1958), 41–61. http://ilib.mirror1.mccme.ru/djvu/mp2/mp2-3.djvu?, djvuopts&page=43

[Ar58'] V.I. Arnold, *Problem 6 (in Russian)*, Mat. Prosveschenie, 3 (1958), 273-274. http://ilib.mirror1.mccme.ru/djvu/mp2/mp2-3.djvu?, djvuopts&page=243

[Ku00] V. Kurlin, *Basic embeddings into products of graphs,* Topol.Appl. 102 (2000), 113–137.

[Ku03] V. A. Kurlin, *Basic embeddings of graphs and the Dynnikov method of three-pages embeddings (in Russian),* Uspekhi Mat. Nauk, 58:2 (2003), 163–164. English transl.: Russian Math. Surveys, 58:2 (2003).

The full text of dissertation is available at http://maths.dur.ac.uk/∼dma0vk/PhD.html

[Mi09] E. Miliczka, *Constructive decomposition of a function of two variables as a sum of functions of one variable*, Proc. AMS, 137:2 (2009), 607-614.

[MK03] N. Mramor-Kosta and E. Trenklerova, *On basic embeddings of compacta into the plane,* Bull. Austral. Math. Soc. 68 (2003), 471–480.

[RZ06] D. Repovš and M. Željko, *On basic embeddings into the plane,* Rocky Mountain J. Math., 36:5 (2006), 1665-1677.

[Sk95] A. Skopenkov, *A description of continua basically embeddable in* $\mathbf{R}^2$*,* Topol. Appl. 65 (1995), 29–48.

[St89] Y. Sternfeld, *Hilbert's 13th problem and dimension,* Lect. Notes Math. 1376 (1989), 1–49.

[Vi04] A.G. Vitushkin, *Hilbert's 13th problem and related questions,* Russian Math. Surveys, 59:1, (2004), 11–24.

[Vo81] S.M. Voronin, Funkcionalniy Analiz, 15:1 (1981), 1–17.

[Vo82] S.M. Voronin, Funkcionalniy Analiz, 16:2 (1982), 21–29.

Емельянов Л.А., Емельянова Т.Л.

Вокруг оснований биссектрис

Вводная часть

Теоретические сведения

- Прямая Эйлера.
- Окружность 9-ти точек.
- Ортоцентрическая четвёрка. Некоторые свойства ортоцентра.
- Вписанная и вневписанные окружности треугольника, их центры.
- Степень точки относительно окружности, радикальная ось двух окружностей, радикальный центр трёх окружностей.

Задачи

1. Лемма Мансиона в полном варианте. *Середина дуги AC описанной окружности треугольника ABC, не содержащая вершину B, равноудалена от вершин A и C, центра I вписанной окружности и центра* I_2 *вневписанной окружности. Середина дуги AC описанной окружности треугольника ABC, содержащая вершину B, равноудалена от вершин A и C, и центров* I_1 *и* I_3 *вневписанных окружностей.*
2. Формулы Эйлера для вписанной и вневписанной окружностей. *Расстояние между центрами вписанной и описанной окружностей треугольника выражается формулой:* $IO^2 = R^2 - 2Rr$. *Расстояние между центрами вневписанной и описанной окружностей треугольника выражается формулой:* $I_kO^2 = R^2 + 2Rr_k, \quad k = 1,2,3.$
3. Теорема Понселе (внутренняя). *Если окружности* Ω *и* ω *соответственно описанная и вписанная для некоторого треугольника, то треугольников с этими же описанной и вписанной окружностями существует бесконечно много и любая точка* Ω *может быть вершиной такого треугольника*. Докажите, что условие теоремы Понселе эквивалентно формуле Эйлера для вписанной и описанной окружностей.
4. Теорема Понселе (внешняя). *Если окружности* Ω *и* ω_1 *соответственно описанная и вневписанная для некоторого треугольника, то треугольников с этими же описанной и вневписанной окружностями существует бесконечно много*. Докажите, что данное условие теоремы Понселе эквивалентно формуле Эйлера для вневписанной и описанной окружностей.
5. Основания внешних биссектрис треугольника лежат на одной прямой. (Эта прямая называется *осью внешних биссектрис*, будем обозначать её ℓ). Прямая ℓ перпендикулярна прямой IO.
6. Выразите расстояние от точки I до оси внешних биссектрис через радиусы описанной и вписанной окружностей.

7. Пусть A_1, B_1 и C_1 – основания внутренних биссектрис на соответствующих сторонах треугольника. Прямые $A_1B_1 = \ell_3$, $B_1C_1 = \ell_1$, $C_1A_1 = \ell_2$ называются *осями внутренних биссектрис*. Докажите, что каждая из этих прямых проходит через основание внешней биссектрисы, а также, что ℓ_k перпендикулярна прямой I_kO (k = 1, 2, 3).

8. Геометрический аналог внешней формулы Эйлера. *Пусть окружности* Ω *и* ω_1 *пересекаются в точках* P *и* Q. *Тогда* Ω *является описанной, а* ω_1 *– вневписанной окружностью некоторого треугольника тогда и только тогда, касательные к* ω_1, *проведённые в точках* P *и* Q, *вторично пересекают* Ω *в точках касания общих внешних касательных, проведённых к* Ω *и* ω_1.

9. Пусть окружности Ω и ω являются соответственно описанной и вписанной для некоторого треугольника. Тогда геометрическим местом оснований внешних биссектрис множества треугольников с теми же описанной и вписанной окружностями является прямая.

10. Пусть окружности Ω и ω_1 являются соответственно описанной и вневписанной (соответствующей вершине A) для некоторого треугольника, а P и Q – точки касания с Ω общих внешних касательных этих окружностей. Тогда геометрическим местом оснований внутренних биссектрис углов B и C множества треугольников с теми же описанной и вневписанной окружностями является отрезок PQ без точек P и Q.

11. Радиус описанной окружности треугольника равен радиусу его вневписанной окружности тогда и только тогда, когда центр описанной окружности лежит на соответствующей оси внутренних биссектрис.

12. Пусть B_0' – середина дуги AC окружности Ω, содержащей точку B, а B_2 – основание внешней биссектрисы на стороне AC. Докажите, что прямая I_2B_0' перпендикулярна прямой B_2I.

Емельянов Л.А., Емельянова Т.Л.

Вокруг оснований биссектрис

Основная часть

Теоретические сведения

- Точка Нагеля.
- Точка Жергонна.
- Изогональность.
- Симедианы треугольника и точка Лемуана.
- Полный четырёхсторонник. Прямые Гаусса и Обера.
- Теорема о трёх центрах гомотетии.
- Теорема Фейербаха. Точки Фейербаха.

13. Биссектриса угла A треугольника ABC пересекает сторону BC в точке A_1, а описанную окружность в точке A_0. Аналогично определяются точки C_1 и C_0. Прямые A_0C_0 и A_1C_1 пересекаются в точке B_{00}. Тогда $B_{00}I$ параллельна стороне AC.
14. $B_{00}B$ – касательная к описанной окружности треугольника.
15. Окружность с центром в B_{00}, проходящая через B, (назовём её b_{00}) проходит через I.
16. Окружность b_{00} вторично пересекается с Ω в точке B_3, а биссектриса угла B вторично пересекает Ω в точке B_0. Тогда $\angle B_0B_3I = 90^\circ$.
17. Если N – точка Нагеля треугольника ABC, то BN и BB_3 – изогональные прямые угла ABC.
18. Пусть L – точка пересечения A_0C_1 и C_0A_1, тогда прямая LI проходит через середину AC.
19. Если аналогично B_{00} определить точки A_{00} и C_{00} (см. задачу 13), то эти три точки лежат на одной прямой (ℓ_{00}), параллельной оси внешних биссектрис.
20. Биссектрисы углов треугольника вторично пересекают его описанную окружность в точках A_0, B_0 и C_0. Докажите, что точки пересечения прямых AB и A_0B_0, BC и B_0C_0, CA и C_0A_0 лежат на одной прямой (ℓ_0), параллельной оси внешних биссектрис.
21. Докажите, что прямые ℓ_{00} и ℓ_0 делят расстояние между центром вписанной окружности I и осью внешних биссектрис ℓ на три равные части.
22. Пусть ось внутренних биссектрис ℓ_2 пересекает Ω в точках E и D. Проведём окружность через точки I, E и D и назовём её b_2. Тогда b_2 проходит через центры вневписанных окружностей I_1 и I_3.
23. Радиус окружности b_2 в 2 раза больше радиуса Ω.
24. I_2 – центр гомотетии Ω и b_2.

25. Прямая $B_{00}I$ (см. задачу 13) является касательной к b_2.

26. Касательные к вневписанной окружности ω_2, проведённые в точках её пересечения с Ω, являются касательными к b_2.

27. Обозначим одну из точек пересечения оси внутренних биссектрис ℓ_2 с описанной окружностью Ω через D, а через D' – точку, диаметрально противоположную D на Ω. Тогда один из концов общей хорды окружностей Ω и ω_2 лежит на прямой $D'I_2$, а другой – на окружности $(OD'I_2)$.

28. Выразите длину хорды, по которой пересекается ось внутренних биссектрис ℓ_2 с описанной окружностью Ω через радиусы R и r_2.

29. Прямая OB является касательной к окружности, для которой основания внутренней и внешней биссектрис угла ABC – диаметрально противоположные точки.

30. Окружность, построенная на основаниях внутренней и внешней биссектрис угла ABC как на диаметре, пересекается с описанной окружностью в точке, принадлежащей симедиане угла ABC.

31. Центр описанной окружности треугольника лежит на прямой Обера четырёхсторонника, образованного четырьмя осями биссектрис.

32. Точка Лемуана треугольника лежит на прямой Обера четырёхсторонника, образованного четырьмя осями биссектрис. (Примечание: *из задач 31 и 32 следует, что ортоцентр треугольника, образованного основаниями внутренних биссектрис, лежит на прямой, проходящей через центр описанной окружности и точку Лемуана.*)

33. Точка B_5 пересечения касательных (задача 26) лежит на отрезке, соединяющем I с точкой касания ω_2 со стороной AC.

34. Аналогично B_5 определим точки A_5 и C_5. Тогда прямые AA_5, BB_5 и CC_5 пересекаются в одной точке T, лежащей на прямой, проходящей через точку Жергонна G и центр тяжести M так, что $GM:MT = 2:1$.

35. Пусть F, F_1, F_2 и F_3 – соответственно внутренняя и внешние точки Фейербаха. Докажите, что основания внутренних и внешних биссектрис треугольника лежат на шести прямых, определяемых точками F, F_1, F_2 и F_3.

36. (*В. Тебо*) Докажите, что треугольники, образованные основаниями внутренних биссектрис $(\Delta A_1B_1C_1)$ и внешними точками Фейербаха $(\Delta F_1F_2F_3)$, подобны.

37. Докажите, что окружность, проходящая через основания внутренних биссектрис, проходит через внутреннюю точку Фейербаха.

Емельянов Л.А., Емельянова Т.Л.

Вокруг оснований биссектрис

Вводная часть

Решения задач

1. Пусть I_2 – центр вневписанной окружности, соответствующей вершине B (рис. 1*a*). Рассмотрим окружность, построенную на отрезке II_2 как на диаметре. Так как вершины A и C лежат на этой окружности, центр её лежит на серединном перпендикуляре к отрезку AC. Этот серединный перпендикуляр пересекается с диаметром II_2 в точке B_0 – середине дуги AC окружности Ω. Следовательно, B_0 равноудалена от вершин A, C, центра вписанной окружности I и центра вневписанной окружности I_2.

Пусть I_1 и I_3 – центры вневписанных окружностей, соответствующих вершинам A и C (рис. 1*b*). Рассмотрим окружность, построенную на отрезке I_1I_3 как на диаметре. Так как вершины A и C лежат на этой окружности, центр её лежит на серединном перпендикуляре к отрезку AC. Этот серединный перпендикуляр пересекается с диаметром I_1I_3 в точке B_0' – середине дуги AC окружности Ω, содержащей вершину B. Следовательно, B_0' равноудалена от вершин A, C, и центров вневписанных окружностей I_1 и I_3.

2. Пусть C' – точка касания вписанной окружности со стороной AB (рис. 2*a*). Степень точки I относительно окружности Ω равна $IO^2 - R^2 = -BI \cdot IB_0$. Рассмотрим треугольники BIC' и $B_0'CB_0$. Они подобны по двум углам (прямому и $\angle B/2$), следовательно, $BI / IC' = B_0'B_0 / B_0C$. Но, по задаче 1, $B_0C = B_0I$, значит, $BI \cdot B_0I = B_0'B_0 \cdot IC' = 2R \cdot r$. Таким образом, $IO^2 - R^2 = -2R \cdot r$ или $IO^2 = R^2 - 2R \cdot r$.

Пусть C' – точка касания вневписанной окружности ω_2 со стороной AB (рис. 2*b*). Степень точки I_2 относительно окружности Ω равна $I_2O^2 - R^2 = I_2B \cdot I_2B_0$. Рассмотрим треугольники BI_2C' и $B_0'CB_0$. Они подобны по двум углам (прямому и $\angle B/2$), следовательно, $BI_2 / I_2C' = B_0'B_0 / B_0C$. Но, по задаче 1, $B_0C = B_0I_2$, значит, $BI_2 \cdot B_0I_2 = B_0'B_0 \cdot I_2C' = 2R \cdot r$. Таким образом, $I_2O^2 - R^2 = 2R \cdot r$ или $I_2O^2 = R^2 + 2R \cdot r$.

3. Рассмотрим окружности $\Omega = (O, R)$ и $\omega = (I, r)$, являющиеся описанной и вписанной для некоторого треугольника. Из задачи 2 следует, что $IO^2 = R^2 - 2R \cdot r$. Выберем произвольную точку B на Ω и построим хорды BA и BC, касательные к ω (рис. 3). Из подобия треугольников BIC' и $B_0'CB_0$ следует: $B_0C / 2R = r / BI$, то есть $2R \cdot r = BI \cdot B_0C$, но из формулы Эйлера следует, что степень точки I относительно Ω равна $-2R \cdot r = -BI \cdot IB_0$. Таким образом, $BI \cdot B_0C = BI \cdot IB_0$, значит, треугольник B_0CI равнобедренный и $\angle B_0IC = \angle ICB_0$, но $\angle B_0IC = \angle B/2 + \angle ICB$, а $\angle ICB_0 = \angle B/2 + \angle ICA$.

Получаем, что $\angle IBC = \angle ICA$. Это означает, что прямая AC симметрична прямой BC относительно прямой CI, проходящей через центр окружности ω, то есть AC – касательная к этой окружности.

4. Рассмотрим окружности $\Omega = (O, R)$ и $\omega_2 = (I, r_2)$, являющиеся описанной и вневписанной для некоторого треугольника. Из задачи 2 следует, что $I_2O^2 = R^2 + 2R \cdot r_2$. Выберем произвольную точку B на Ω и построим прямые BA и BC, касательные к ω_2 (рис. 4). Из подобия треугольников BI_2C' и $B_0'CB_0$ следует: $B_0C / 2R = r_2 / BI_2$, то есть $2R \cdot r_2 = BI_2 \cdot B_0C$, но из формулы Эйлера следует, что степень точки I_2 относительно Ω равна $2R \cdot r_2 = BI_2 \cdot I_2B_0$. Таким образом, $BI_2 \cdot B_0C = BI_2 \cdot I_2B_0$, значит, треугольник B_0CI_2 равнобедренный и $\angle B_0I_2C = \angle I_2CB_0$, но $\angle B_0IC + \angle I_2CB = \angle BB_0C = \angle A$, то есть $\angle I_2CB_0 = \angle A / 2$. Получаем, что $\angle I_2CA = \angle A / 2 + \angle B_0CA = (\angle A + \angle B) / 2$. Это означает, что прямая I_2C является биссектрисой внешнего угла B,то есть AC – касательная к окружности ω_2.

5. Сначала докажем, что ортоцентрическая ось есть радикальная ось описанной окружности и окружности девяти точек. Рассмотрим две окружности: Ω_B – построенная на AC как на диаметре и ω_B – имеющая диаметром отрезок HB (рис. 5). Сторона ортотреугольника H_1H_3 является их общей хордой, то есть лежит на их радикальной оси. Значит, $H_2'H_3 \cdot H_2'H_1 = H_2'A \cdot H_2'C$. Теперь рассмотрим описанную окружность Ω и окружность девяти точек ω_0. Степень точки H_2' относительно Ω равна $H_2'A \cdot H_2'C$, относительно ω_0 равна $H_2'H_3 \cdot H_2'H_1$, то есть точка пересечения стороны ортотреугольника с соответственной стороной исходного треугольника имеют одинаковые степени относительно окружностей Ω и ω_0. Из этого следует, что ортоцентрическая ось треугольника есть радикальная ось его описанной окружности и окружности девяти точек, то есть перпендикулярна прямой Эйлера этого треугольника.

Теперь рассмотрим треугольник $I_1I_2I_3$, образованный центрами вневписанных окружностей. Для него исходный треугольник ABC является ортотреугольником, а точка I – ортоцентром. Значит, точки пересечения внешних биссектрис треугольника ABC с его сторонами лежат на ортоцентрической оси треугольника $I_1I_2I_3$ то есть на одной прямой, перпендикулярной прямой Эйлера треугольника $I_1I_2I_3$. Но прямая Эйлера треугольника $I_1I_2I_3$ проходит через ортоцентр (I) и центр окружности девяти точек (O) этого треугольника, то есть совпадает с прямой IO.

6. Решим вспомогательную задачу: даны окружности $\omega_1 = (O_1, R_1)$ и $\omega_2 = (O_2, R_2)$, их радикальная ось пересекает линию центров в точке P (рис. 6). Найти длину отрезка PO_1. Так как степени P относительно окружностей равны, имеем $PO_1^2 - R_1^2 = PO_2^2 - R_2^2$, $PO_2^2 - PO_1^2 = R_2^2 - R_1^2$, $O_1O_2 \cdot (2PO_1 + O_1O_2) = R_2^2 - R_1^2$. Откуда легко выразить PO_1 через радиусы окружностей и расстояние между их центрами.

Теперь возьмём в качестве ω_1 описанную окружность треугольника ABC радиуса R, а в качестве ω_2 – окружность $(I_1I_2I_3)$ радиуса $2R$. Тогда расстояние d_1 от точки O до радикальной оси этих окружностей (ℓ) равно $\dfrac{R^2 + Rr}{\sqrt{R^2 - 2Rr}}$. Тогда искомое расстояние равно

$$d = d_1 - IO = \frac{R^2 + Rr}{\sqrt{R^2 - 2Rr}} - \sqrt{R^2 - 2Rr} = \frac{3Rr}{\sqrt{R^2 - 2Rr}}.$$

7. Решение аналогично решению задачи 5 с заменой треугольника $I_1I_2I_3$ на треугольники II_2I_3, I_1II_3, I_1I_2I. Эти треугольники имеют общую окружность девяти точек (Ω), а прямые I_kO являются прямыми Эйлера в соответствующем треугольнике. Оси же внутренних биссектрис треугольника ABC являются радикальными осями описанной окружности соответственного треугольника и окружности Ω.

8. Пусть Ω и ω_2 – описанная и вневписанная окружности треугольника ABC (рис. 8). В семействе треугольников с этими описанной и вневписанной окружностями есть два предельных «треугольника». Рассмотрим случай, когда секущая AB превращается в касательную – это будет общая внешняя касательная к Ω и ω_2. При этом A и B точки сливаются в точке D, а прямые BC и AC сливаются в касательную PD.
Теперь рассмотрим окружности Ω и ω_2, для которых касательная к ω_2, проведённая в точке P их пересечения, проходит через точку D касания общей внешней касательной Ω и ω_2. Тогда $DK^2 + (r_2 - R)^2 = OI_2^2$, $\operatorname{tg}(\angle I_2DK) = \dfrac{r_1}{DK} = t$, так как $\angle PDK = 2 \cdot \angle I_2DK$ имеем $\sin(\angle PDK) = \dfrac{2 \cdot t}{1 + t^2} = \dfrac{2 \cdot r_2 \cdot DK}{DK^2 + r_2^2}$. Длина хорды DP окружности Ω равна $DP = 2R \cdot \sin(\angle DPK)$, но так как DP и DK – касательные к ω_2 $DP = DK$. Из этого следует $2R\dfrac{2r_2 \cdot DK}{DK^2 + r_2^2} = DP$, $4R \cdot r_2 = DK^2 + r_2^2$. Используя выражение для DK^2, приходим к формуле Эйлера $I_2O^2 = R^2 + 2Rr_2$.

9. Так как окружности Ω и ω фиксированы, а из решения задачи 6 следует, что расстояние от центра ω до оси внешних биссектрис выражается только через радиусы этих окружностей, получаем, что основания внешних биссектрис всех треугольников Понселе лежат на фиксированной прямой.
Обратно. Возьмём любую точку (A_2) на этой прямой, проведём через неё касательную к ω и обозначим точки пересечения с Ω через B и C. По теореме Понселе сторона BC порождает треугольник ABC, основание внешней биссектрисы которого и есть A_2.

10. Так как окружности Ω и ω_1 фиксированы, а из решения задачи 8 следует, что ось внутренних биссектрис ℓ_1 проходит через точки касания общих внешних касательных этих окружностей, то есть ℓ_1 – фиксированная прямая.
Обратно. Возьмём любую точку (B_1) на интервале PQ, проведём через неё касательную к ω_1 и обозначим точки пересечения с Ω через A и C. По внешней теореме Понселе сторона AC порождает треугольник ABC, основание внутренней биссектрисы которого и есть B_1.

11. Воспользуемся результатом задачи 8. Пусть $R = r_2$. Это равносильно тому, что общие внешние касательные к Ω и ω_2 параллельны прямой OI_2, значит, DE – диаметр описанной окружности, т.е. $O \in A_1C_1$.

12. Пусть ω' – окружность с диаметром IB_0', ω'' – окружность с диаметром II_2 (рис.12). Прямая $B_0'B_2$ – радикальная ось ω' и Ω, прямая CB_2 – радикальная ось ω'' и Ω. Следовательно, прямая IB_2 – радикальная ось ω' и ω''. Пусть K – вторая точка пересечения этих окружностей. Поскольку $\angle IKB_0' = \angle IKI_2 = 90°$, точка K лежит на прямой $B_0'I_2$, и, таким образом, $B_0'I_2 \perp B_2I$.

Емельянов Л.А., Емельянова Т.Л.

Вокруг оснований биссектрис

Основная часть

Решения задач

С использованием решений Бажова И. и Чекалкина С.

13. Так как биссектриса BB_2 внешнего угла ABC (рис. 13) параллельна прямой A_0C_0, соединяющей середины дуг описанной окружности, $\frac{C_1B_{00}}{B_{00}B_2} = \frac{C_1X}{XB}$. В треугольнике XBY прямая BI является биссектрисой и высотой, значит, BI делит XY пополам. Так как $C_0I = C_0B$ A_0C_0 (а значит и XY) является серединным перпендикуляром отрезка BI. Из этого следует, что $BXIY$ – ромб и $IX \parallel BC$. Значит $\frac{C_1X}{XB} = \frac{C_1I}{IC} \Rightarrow \frac{C_1I}{IC} = \frac{C_1B_{00}}{B_{00}B_2}$, то есть $IB_{00} \parallel AC$.

14. Поскольку B_{00} находится на серединном перпендикуляре к отрезку BI (рис. 13), $\angle B_{00}BI = \angle B_{00}IB \Rightarrow \angle B_{00}IB = \angle B_2B_1 = \frac{1}{2}\overset{\smile}{BB_0}$. Значит, $\angle B_{00}BB_0 = \frac{1}{2}\overset{\smile}{BB_0} \Rightarrow B_{00}B$ – касательная к окружности Ω.

15. Поскольку B_{00} находится на серединном перпендикуляре к отрезку BI (рис. 13), $B_{00}I = B_{00}B$.

16. $\angle IB_3B_0 = \angle B_0B_3B' + \angle B'B_3I = \angle B'BB_0 + \angle B'B_3I$ (рис. 16). Так как $\angle B_0BO = 90^\circ$, $B'B$ – касательная и $\angle B'BI = \frac{1}{2}\overset{\smile}{BI}$. Таким образом, $\angle B_0B_3I = \angle B'B_3I + \angle IB_3B = \angle B'B_3B = 90^\circ$.

17. Рассматривая гомотетию ω и ω_2 с центром в точке B (рис. 17), нетрудно доказать, что прямая BN пересекает ω в точке, диаметрально противоположной точке касания ω со стороной AC. Из этого следует, что прямые BN и B_0B_0' пересекаются в точке K, так что $IK \parallel AC$.
Далее, $\angle IBB_0' = \angle IKB_0'$, следовательно четырёхугольник $IBB_0'K$ – вписанный (рис. 17). Из задачи 16 следует, что точки B_0', I и B_3 лежат на одной прямой. Значит, $\angle IBN = \angle IBK = \angle IB_0'K = \angle B_3B_0'B_0 = \angle B_3BB_0$.

18. Так как IA_0, IL, IC_0, IB_{00} – гармоническая четвёрка лучей (рис. 18), IA, IM, IC, IB_{00} также гармоническая четвёрка лучей. Но (по задаче 13) $IB_{00} \parallel AC$, следовательно, M – середина AC.

19. Рассмотрим гомотетию с центром в точке C_1, переводящую I в C (рис. 19). Так как $IB_{00} \| B_2C$, то при этой гомотетии $B_{00} \to B_2$, $A_{00} \to A_2$. Значит, $A_{00}B_{00} \| A_2B_2$. Аналогично, $B_{00}C_{00} \| B_2C_2$. Таким образом, A_{00}, B_{00}, C_{00} лежат на одной прямой, параллельной ℓ.

20. Рассмотрим гомотетию с центром в точке C_0, переводящую I в C. Так как $IB_{00} \| B_2'C$, то при этой гомотетии $B_{00} \to B_2'$, $A_{00} \to A_2'$. Значит, $A_2'B_2' \|\ A_{00}B_{00} \| A_2B_2$. Аналогично, $B_2'C_2' \| B_2C_2$. Таким образом, A_2', B_2', C_2' лежат на одной прямой, параллельной ℓ.

21. Проведём через точки B, I и середину отрезка BI прямые, перпендикулярные биссектрисе BI (рис. 21). Поскольку $B_{00}I = B_{00}B$ и $B_{00}B$ является касательной к окружности Ω точка B_{00} имеет одинаковые степени относительно описанной окружности Ω и точки I. Так как $\angle B_6IA = \angle BIA - 90^\circ = \angle ICA$, $B_6I^2 = B_6A \cdot B_6C$, значит, B_6 лежит на той же радикальной оси, которая совпадает с прямой ℓ_{00}. Треугольник B_8BI прямоугольный, следовательно $B_{00}B = B_{00}I = B_{00}B_8$. В силу параллельности трёх проведённых прямых, а также прямых $B_{00}B_6$ и B_7B_2, получим $B_{00}I = B_4B_6 = B_{00}B_8 = B_2B_4 = B_7B_8$. Из этого следует, что отрезок IB_7 делится прямыми ℓ_0 и ℓ_{00} на три равные части, откуда следует утверждение задачи.

22. В решении задачи 7 доказано, что ℓ_2 является радикальной осью окружностей Ω и (II_1I_3). Значит, степени каждого из оснований биссектрис A_1 и C_1 относительно этих окружностей равны, то есть точки D и E, лежащие на этой радикальной оси – общие точки окружностей Ω и (II_1I_3).

23. Для треугольника II_1I_3 окружность Ω – окружность девяти точек, а b_2 – описанная, то есть отношение их радиусов равно 2.

24. Поскольку точка I_2 – ортоцентр треугольника II_1I_3 окружности Ω и b_2 гомотетичны относительно I_2.

25. Точка B_{00} лежит на прямой DE, которая является радикальной осью окружностей Ω и b_2, а так как $B_{00}I = B_{00}B$ отрезок $B_{00}I$ – касательная к окружности b_2.

26. При гомотетии с центром в точке I_2 и коэффициентом $\frac{1}{2}$ точка D переходит в точку D_1, лежащую на Ω (рис 26). Поскольку PD – касательная к ω_2 $\angle DPI_2 = 90^\circ$. Значит, $PD_1 = DD_1$, следовательно касательная к Ω, проведённая в точке D_1, параллельна PD и при гомотетии I_2 и коэффициентом 2 перейдёт в прямую PD. Но при такой гомотетии окружность Ω перейдёт в b_2, значит, PD – касательная к b_2.

27. Так как DP – касательная к ω_2, а DD' – диаметр Ω $\angle DPI_2 = \angle DPD' = 90^\circ$, значит, точки D', P и I_2 лежат на одной прямой (рис. 27). Далее

$\angle OD'I_2 = \angle OPD' = 180^\circ - \angle OPI_2 = 180^\circ - \angle OQI_2$. То есть точки O, D', I_2 и Q лежат на одной окружности.

28. Искомая хорда DE в 2 раз больше отрезка DL (рис 28). Но $DL = OD \cdot \sin \angle I_2OD = R \cdot \dfrac{KI_2}{OI_2} = R \cdot \dfrac{DS}{\sqrt{R^2 + 2R \cdot r_2}} = \dfrac{R}{\sqrt{R^2 + 2R \cdot r_2^2}} \cdot \sqrt{OI_2^2 - OK^2} =$

$= R \cdot \sqrt{\dfrac{4Rr_2 - r_2^2}{R^2 + 2R \cdot r_2^2}}$.

Значит, длина искомой хорды $DE = 2\sqrt{R \cdot r_2 \cdot \dfrac{4R - r_2}{R + 2r_2}}$.

Примечание: Из этой формулы нетрудно получить, что если $r_2 = R$, то $DE = 2R$, то есть DE – диметр окружности Ω.(задача 11). А также явно видно неравенство для радиусов вневписанных окружностей: $r_i < 4R$.

29. Так как $\angle BB_1B_2$ – внешний угол треугольника ABB_1 (рис 29) $\angle BB_1B_2 = \angle A + \angle B / 2$. Из прямоугольного треугольника B_1B_2B получаем $\angle BB_2B_1 = 90^\circ - \angle A - \angle B / 2$. С другой стороны $\angle OBB_1 = \angle OBC - \angle B_1BC$, следовательно, $\angle OBB_1 = (90^\circ - \angle A) - \angle B / 2$. Значит, $\angle BB_1B_2 = \angle OBB_1$ то есть OB – касательная к окружности (B_1BB_2).

30. Так как (B_1BB_2) – окружность Аполлония отрезка AC (рис. 30) $\dfrac{AB}{BC} = \dfrac{AK}{KC}$. Значит, $AB \cdot KC = AK \cdot BC$ и четырёхугольник $ABCK$ – гармонический. Из этого следует, что касательные к Ω, проведённые в его вершинах A и C, пересекаются в точке S, лежащей на продолжении диагонали BK. То есть BK – симедиана треугольника ABC.

31. Как известно, прямая Обера полного четырёхсторонника является радикальной осью трёх окружностей, построенных на его диагоналях как на диаметрах. Для четырёхсторонника, образованного осями биссектрис диагоналями являются отрезки сторон, заключённые между основаниями внутренней и внешней биссектрис – отрезки A_1A_2, B_1B_2 и C_1C_2. Из решения задачи 30 следует, что степень точки O относительно окружностей, построенных на этих отрезках как на диаметрах, одинаковые и равны R^2. Следовательно, O лежит на радикальной оси этих окружностей, то есть на прямой Обера.

32. Степень точки Лемуана L относительно окружности, построенной на отрезке B_1B_2 как на диаметре, равна $-BL \cdot LK$ (рис 30). Она же равна степени, относительно описанной окружности Ω. Следовательно, степени точки L относительно трёх окружностей, построенных на диагоналях четырёхсторонника, равны степени точки Лемуана относительно описанной окружности треугольника, то есть точка L лежит на прямой Обера.

33. Из задачи 26 следует, что точка B_5 является точкой пересечения общих внутренних касательных окружностей ω_2 и b_2 следовательно, B_5 – центр гомотетии этих

окружностей. При этой гомотетии касательная AC к ω_2 переходит в параллельную ей касательную к b_2, то есть в прямую IB_{00}. Значит, точка касания вневписанной окружности ω_2 со стороной AC и центр I вписанной окружности – соответственные точки при этой гомотетии. Из этого следует, что точка B_5 лежит на отрезке, соединяющем указанные точки.

34. *Авторы не знают геометрического решения этой задачи. К сожалению его не нашли и школьники, поэтому приводим план вычислительного решения в барицентрческих координатах.*
По барицентрческим координатам точки Жергонна G центра тяжести M можно найти барицентрические координаты точки T, делящей отрезок GM в отношении -3. Теперь можно найти отношение, в котором прямая BT делит сторону AC. Далее, поскольку B_5 делит отрезок IB'' (B'' – точка касания ω_2 со стороной AC) в отношении $2R/r_2$ (по задаче 33), можно найти барицентрические координаты точки B_5. После этого достаточно убедиться, что прямая BB_5 делит AC в том же отношении, что и прямая BT.

35. Основание внутренней биссектрисы является центром внутренней гомотетии вписанной и соответствующей вневписанной окружностей, а основание внешней биссектрисы – центром внешней гомотетии двух вневписанных окружностей. Тока F является внешним центром гомотетии вписанной окружности и окружности девяти точек, а точки F_1, F_2 и F_3 – центрами внутренних гомотетий вневписанных окружностей и окружности девяти точек. Утверждение задачи следует из теоремы о трёх центрах гомотетии.

36. Решение можно найти в книге И. Ф. Шарыгин. «Геометрия 9-11», задача № 586.

37. По теореме о трёх центрах гомотетии, точки F_2, B_1 и F лежат на одной прямой. Аналогично для троек F_3, C_1, F и F_1, A_1, F. Таким образом, треугольники $F_1F_2F_3$ и $A_1B_1C_1$ перспективны относительно точки F. Кроме того, эти треугольники подобны (задача 36). Поэтому, угол $C_1B_1A_1$, равный углу $F_1F_2F_3$, составляет в сумме с углом C_1FA_1 180°, то есть, точки A_1, B_1,C_1, F – лежат на одной окружности.

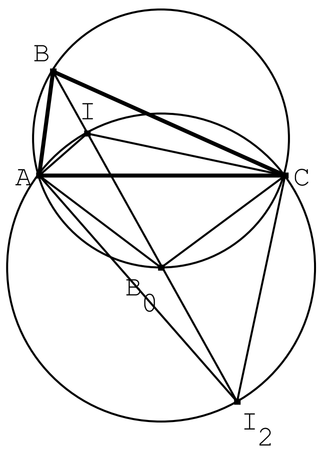


Рис 1 *a*

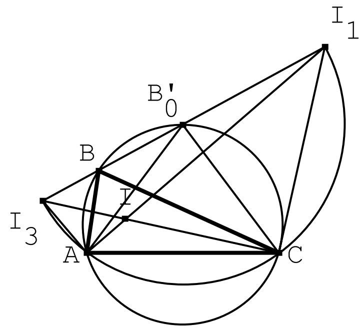


Рис 1 *b*

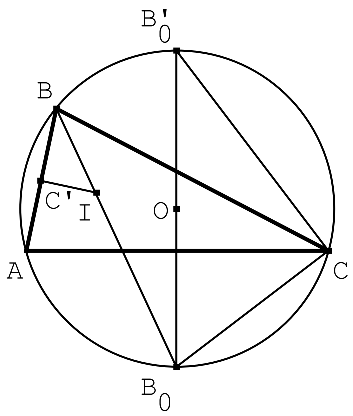


Рис 2 *a*

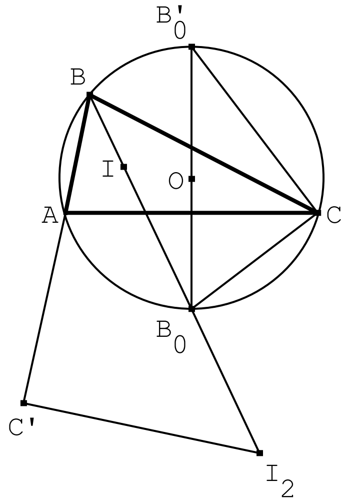


Рис 2 *b*

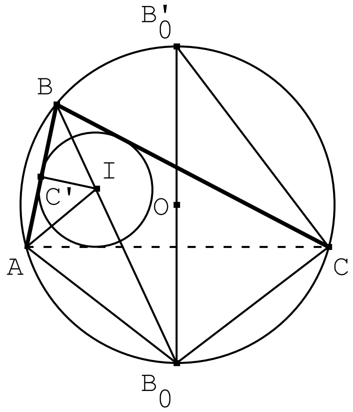


Рис 3

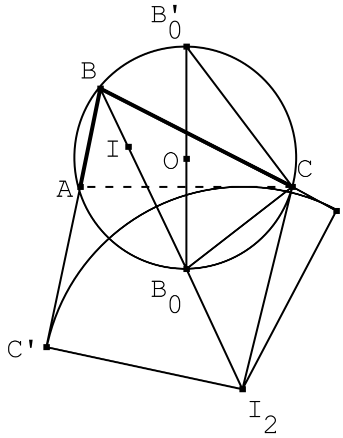


Рис 4

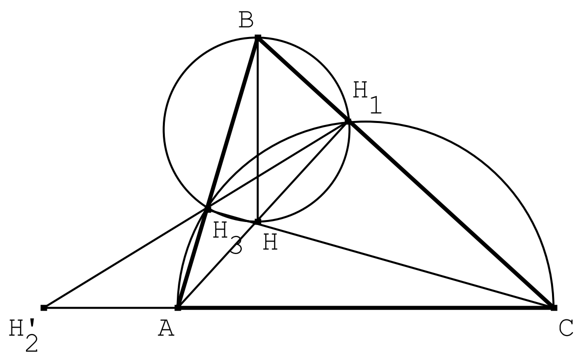


Рис 5

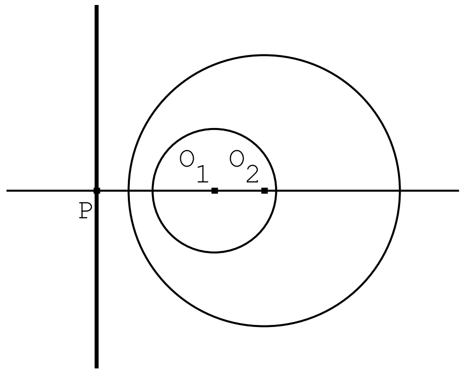


Рис 6

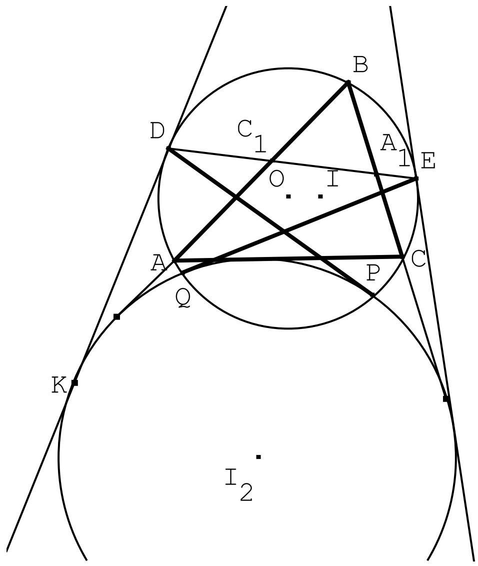


Рис 8

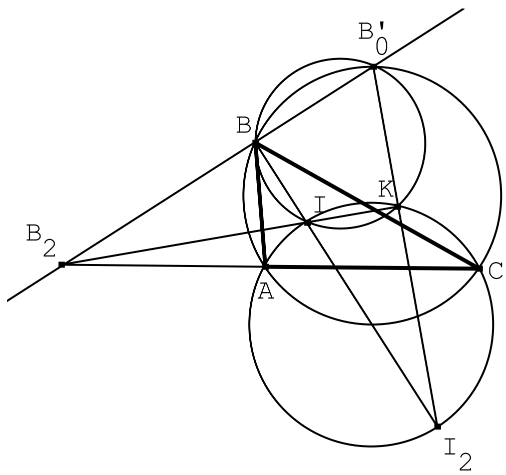


Рис 12

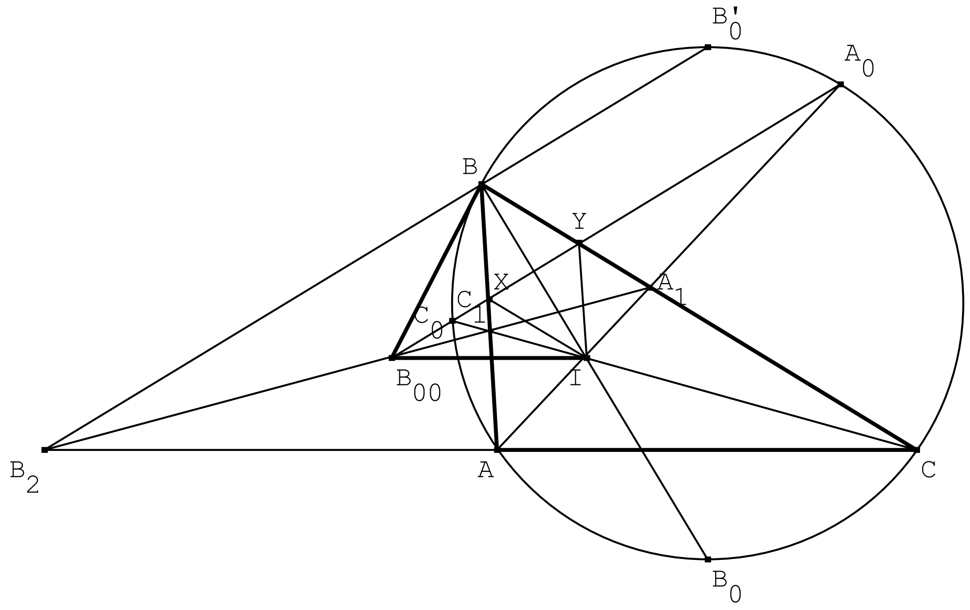


Рис 13

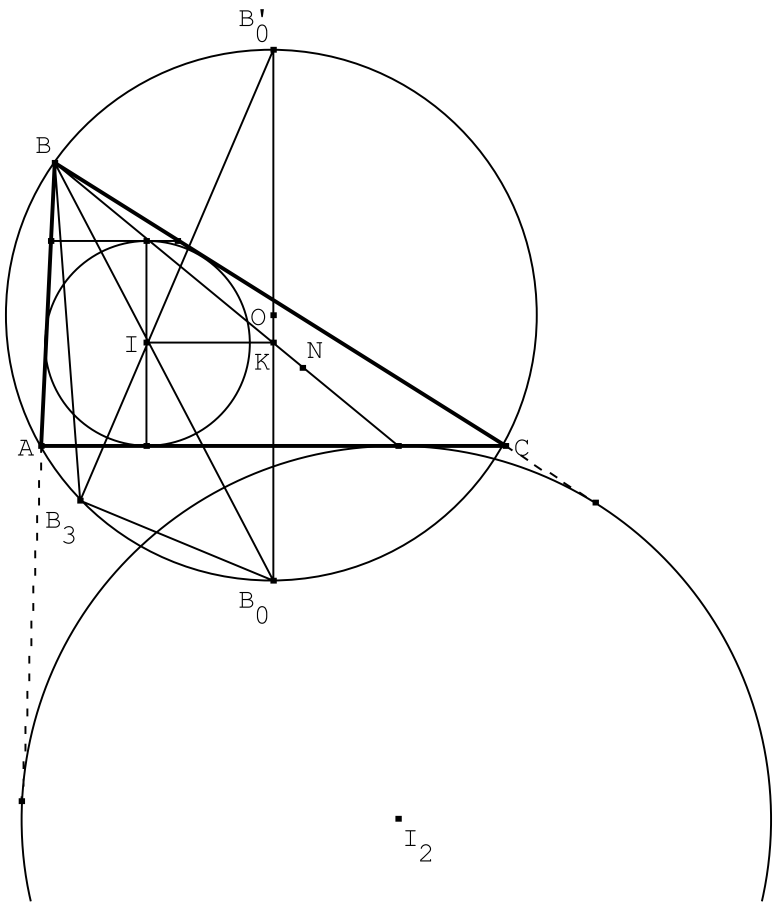


Рис 17

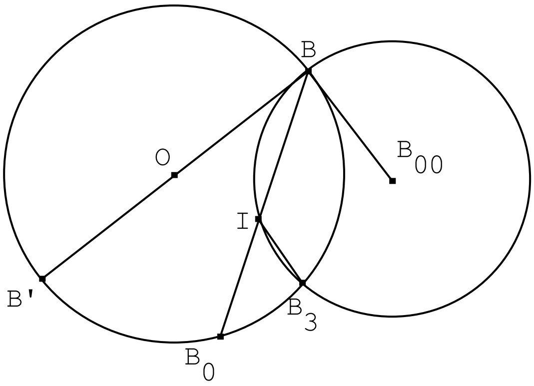


Рис 16

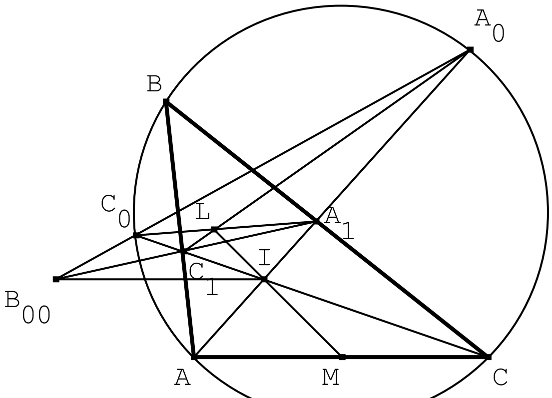


Рис 18

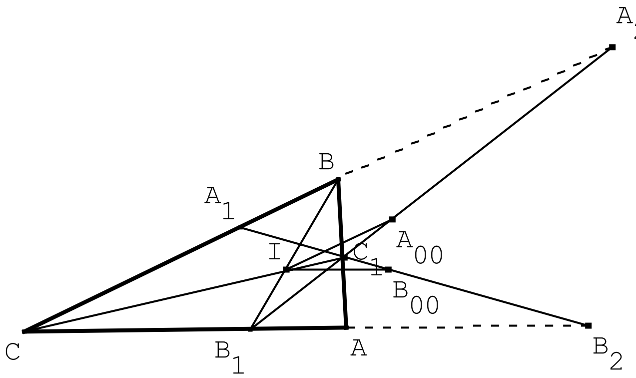


Рис 19

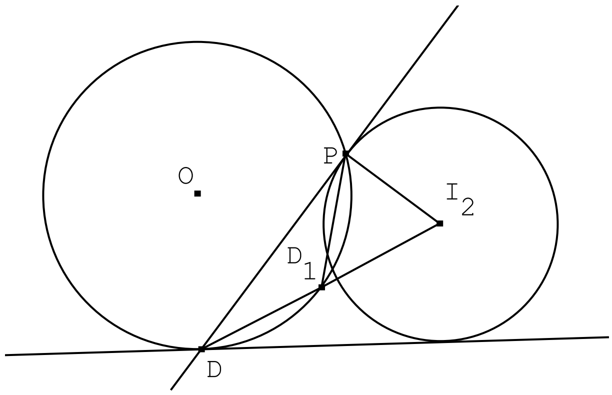


Рис 26

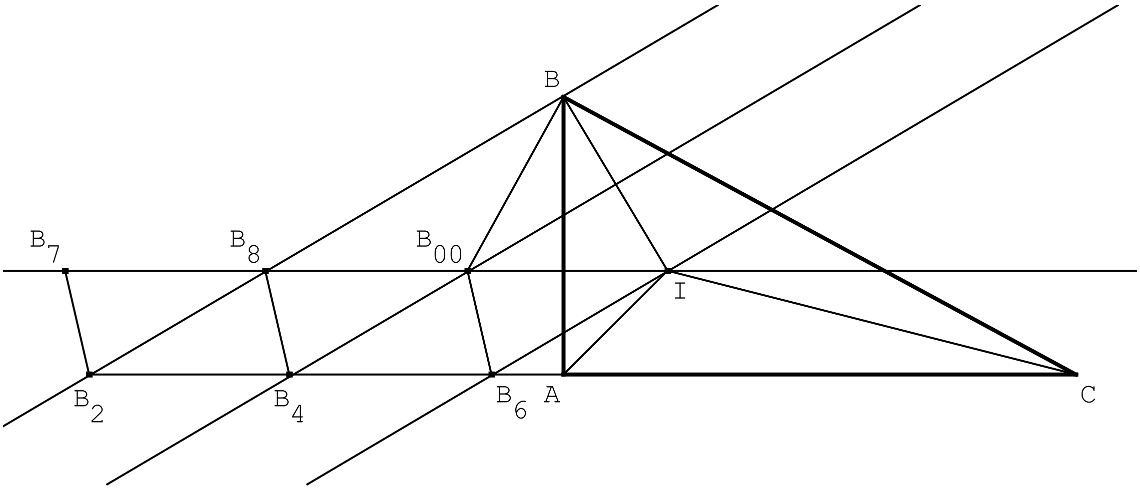


Рис 21

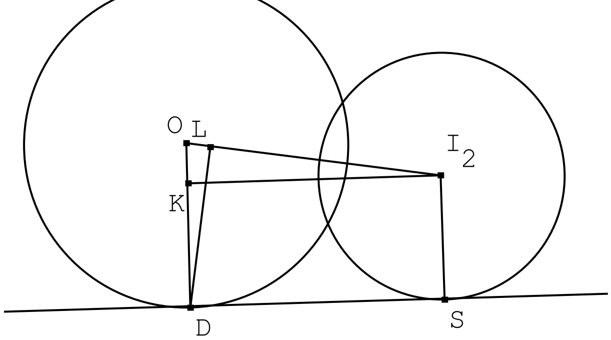


Рис 28

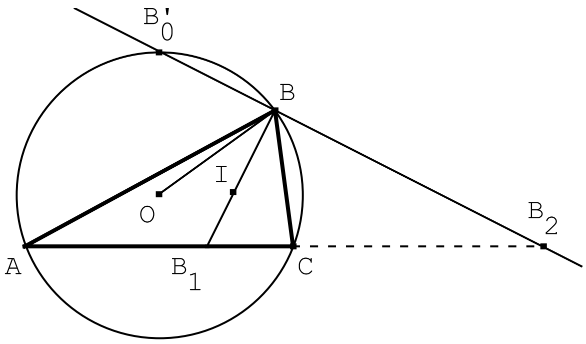


Рис 29

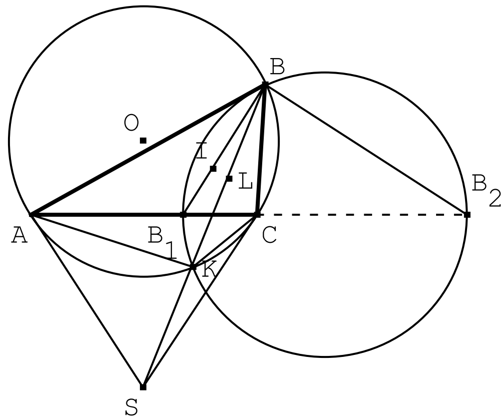


Рис 30

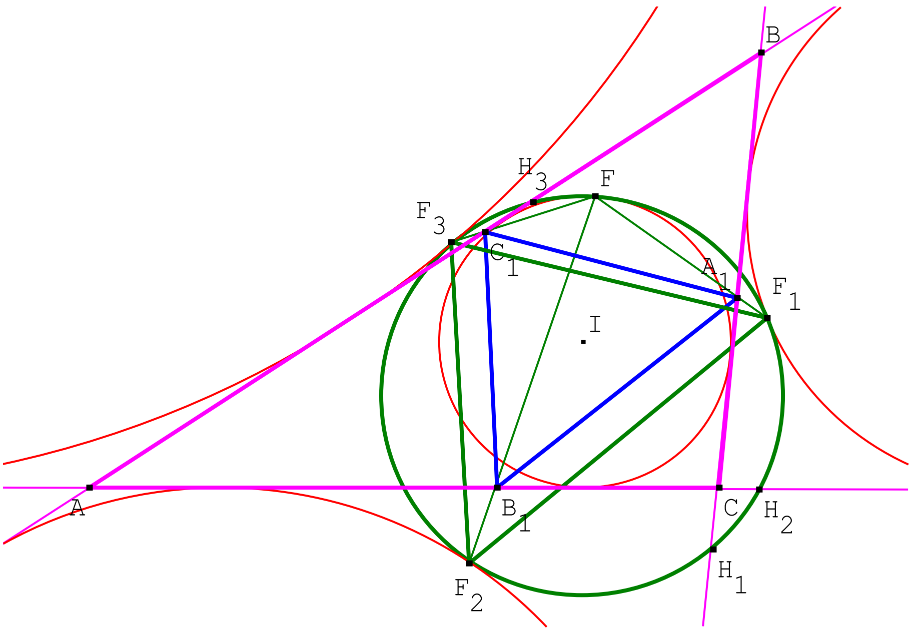


Рис 37

Emelyanov L., Emelyanova T.

Around of Feet of Bisectors

Introduction

Some theoretical facts

- The Euler line.
- The nine-point circle.
- Orthocentric quadruple. Some properties of the orthocenter.
- The incircle and the excircle, their centers.
- Power of a point with respect to a circle, the radical axe of two circles, the radical center of three circles.

Problems

1. Mansion lemma in common case. *The middle of arc AC of circumcircle of triangle ABC non-containing of vertex B is equidistance from vertices A and C, center I of its incircle and center I_2 of its B-excircle. The middle of arc AC of circumcircle of triangle ABC containing of vertex B is equidistance from vertices A and C, centers I_1 and I_3 of its A-excircle and C-excircle.*
2. Euler formula for incircle and excircle. *Distance between the centers of the incircle and the circumcircle satisfies to formula: $IO^2 = R^2 - 2Rr$. Distance between the centers of the excircle and the circumcircle satisfies to formula: $I_kO^2 = R^2 + 2Rr_k$, $k = 1,2,3$.*
3. Poncelet theorem (internal). *If Ω and ω are the circumcircle and the incircle of any triangle respectively then there are endless amount of triangles with the same circumcircle and incircle and any point of Ω may be the vertex of such triangle*. Prove that condition of Poncelet theorem is equivalent to Euler formula for the incircle and the circumcircle.
4. Poncelet theorem (external). *If Ω and ω_1 are the circumcircle and the excircle of any triangle respectively then there are endless amount of triangles with the same circumcircle and excircle*. Prove that the given condition of Poncelet theorem is equivalent to Euler formula for the excircle and the circumcircle.
5. The feet of external bisectors are collinear. (This line is called *the axe of external bisectors*, we note it ℓ). Line ℓ is perpendicular to line IO.
6. Express distance from point I to the axe of external bisectors through the radii of the incircle and the circumcircle.
7. Let A_1, B_1 и C_1 be the feet of internal bisectors on respective sides of triangle. Lines $A_1B_1 = \ell_3$, $B_1C_1 = \ell_1$, $C_1A_1 = \ell_2$ are called the axes of internal bisectors. Prove that each of these lines passes through the foot of the respective external bisector and that ℓ_k is perpendicular to line I_kO (k = 1, 2, 3).

8. Geometrical analog of external Euler formula. *Let circles* Ω *u* ω_1 *intersect in points* P *and* Q. *Then* Ω *is circumcircle and* ω_1 *is excircle for a triangle if and only if tangents to* ω_1 *at points* P *and* Q *secondary intersect* Ω *in tangent points of common extangents to* Ω *and* ω_1.
9. Let Ω and ω be the circumcircle and incircle for a triangle. Then the locus of feet of external bisectors of family of triangles with the same circumcircle and incircle is a line.
10. Let Ω and ω_1 be the circumcircle and A-excircle for a triangle, P and Q are the tangent points of Ω and common extangents of these circles. Then the locus of feet of internal bisectors of angles B and C of family of triangles with the same circumcircle and A-excircle is interval PQ (without P and Q).
11. Circumradius of a triangle equals to exradius if and only if the circumcenter lies on respective axe of internal bisectors.
12. Let B_0' be the middle of arc AC of Ω containing B, B_2 be the foot of external bisector on AC. Prove that $I_2B_0' \perp B_2I$.

Emelyanov L., Emelyanova T.

Around the Feet of Bisectors

Main problems

Theory

- Nagel point.
- Gergonne point.
- Isogonality.
- Symmedians and Lemoine point.
- Quadrilateral complete. The Gauss line and the Aubert line.
- Three homothety centers theorem.
- Feuerbach theorem. Feuerbach points

13. Let the triangle ABC be given. The bisector of angle A intersects the sideline BC in point A_1 and the circumcircle in point A_0. The points C_1 and C_0 are defined similarly. B_{00} is the common point of lines A_0C_0 and A_1C_1. Then the lines $B_{00}I$ and AC are parallel.
14. Prove that the line $B_{00}B$ touches the circumcircle of *ABC*.
15. Let the circle b_{00} with center B_{00} pass through B. Then b_{00} passes through I.
16. Let B_3 be the common points of circle Ω, distinct from B, with b_{00}. Let B_0 be the common points of circle Ω, distinct from B, with the bisector of angle B. Then $\angle B_0B_3I = 90^\circ$.
17. Let N be the Nagel point of ABC. Then the lines BN and BB_3 are isogonal with respect to *ABC*.
18. Let L be the common point of A_0C_1 and C_0A_1. Then the midpoint of the segment AC lies in the line LI.
19. Let the points A_{00} and C_{00} be defined similarly to B_{00} (see problem 13). Then these three points are collinear. Let the line $A_{00}B_{00}$ be denoted by l_{00}. Then l_{00} is parallel to the external bisectors axe.
20. Let the bisectors of triangle *ABC* intersect its circumcircle in points A_0, B_0, C_0, distinct from *A, B, C*. Then the common points of lines AB and A_0B_0, BC and B_0C_0, CA and C_0A_0 are collinear on the line l_0 parallel to external bisectors axe.
21. Prove that the lines ℓ_{00} and ℓ_0 divide the distance between I and the external bisectors axe ℓ into three equal parts.
22. Let the internal bisectors axe ℓ_2 intersect Ω in the points E, D. The circle passing through the points I, E, D is called b_2. Then b_2 passes through the excenters I_1 and I_3.

23. Prove that the radius of b_2 is twice greater than the circumradius *R*.

24. Prove that I_2 is the homothety center of Ω and b_2.

25. Prove that the line $B_{00}I$ (see problem 13) is tangent to the circle b_2.

26. Prove that the tangents to the excircle ω_2 in its common points with Ω are also tangent to the circle b_2.

27. Let D be one of common points of internal bisectors axe ℓ_2 with Ω. D' is the point of Ω opposite to D. Then one of common points of Ω and ω_2 lies on the line $D'I_2$ and the second lies on the circle $(OD'I_2)$.

28. Given the radius R and r_2. Find the distance between the common points of internal bisectors axe ℓ_2 with Ω.

29. Prove that the line OB touches the circle passing through *B* and the feet of the internal and external bisectors of angle *ABC*.

30. Let *K* be distinct from *B* common point of Ω with the circle passing through *B* and the feet of the internal and external bisectors of angle *ABC*. Prove that *K* lies on the symmedian of angle ABC.

31. Prove that the circumcenter of the triangle lies in the Aubert line of quadrilateral formed by four bisectors axes.

32. Prove that the Lemoine point of the triangle lies in the Ober line of quadrilateral formed by four bisectors axes. (Remark: from the problems *31 and 32 the orthocenter of the triangle formed by the feet of internal bisectors is collinear with the circumcenter and the Lemoine point.*)

33. Let B_5 be the common point of tangents defined in problem 26. Prove that B_5 lies in the segment between I and the touching point of ω_2 with the sideline AC.

34. Let the points A_5 and C_5 be defined similarly to B_5. Then the lines AA_5, BB_5, CC_5 are concurrent in the point T. The point T is collinear with Gergonne point G and the centroid M and $GM:MT = 2:1$.

35. Let F, F_1, F_2, F_3 be the internal and external Feuerbach points. Prove that the feet of internal and external bisectors lies in 6 lines defined by F, F_1, F_2, F_3.

36. (*V.Tebault*) Prove that the triangles formed by the feet of external bisectors $(\Delta A_1B_1C_1)$ and external Feuerbach points $(\Delta F_1F_2F_3)$ are similar.

37. Prove that the circumcircle of the triangle formed by the feet of external bisectors $(\Delta A_1B_1C_1)$ pass through the internal Feuerbach point.

Emelyanov L., Emelyanova T.

Around of Feet of Bisectors

Introduction

Solutions

1. Let I_2 be the B-excenter (fig. 1*a*). Let us consider the circle with diameter II_2. The vertices A and C lie on this circle, therefore its center lies on the perpendicular bisector of AC which intersects the diameter II_2 at the midarc B_0 of AC of Ω. Hence B_0 is equidistant from A, C, I and I_2.

Let I_1 and I_3 be the A-excenter and C-excenter (fig. 1*b*). Let us consider the circle with diameter I_1I_3. The vertices A and C lie on this circle, therefore its center lies on the perpendicular bisector of AC which intersects the diameter I_1I_3 at the midarc B_0' of AC of Ω containing B. Hence B_0' is equidistant from A, C, I_1 and I_3.

2. Let C' be the touch point of the incircle and AB (fig. 2*a*). Power of I with respect to Ω is $IO^2 - R^2 = -BI \cdot IB_0$. The triangles BIC' and $B_0'CB_0$ are similar, therefore $BI / IC' = B_0'B_0 / B_0C$. From problem 1 it follows that $B_0C = B_0I$, hence $BI \cdot B_0I = B_0'B_0 \cdot IC' = 2R \cdot r$. Therefore $IO^2 - R^2 = -2R \cdot r$, i.e. $IO^2 = R^2 - 2R \cdot r$.

Let C' be the touch point of the excircle ω_2 and AB (fig. 2*b*). Power of I_2 with respect to Ω is $I_2O^2 - R^2 = I_2B \cdot I_2B_0$. The triangles BI_2C' and $B_0'CB_0$ are similar, therefore $BI_2 / I_2C' = B_0'B_0 / B_0C$. From problem 1 it follows that $B_0C = B_0I_2$, hence $BI_2 \cdot B_0I_2 = B_0'B_0 \cdot I_2C' = 2R \cdot r$. Therefore $I_2O^2 - R^2 = 2R \cdot r$, i.e. $I_2O^2 = R^2 + 2R \cdot r$.

3. Let $\Omega = (O, R)$ be the circumcircle and $\omega = (I, r)$ be the incircle of some triangle. From problem 2 it follows that $IO^2 = R^2 - 2R \cdot r$. Take an arbitrary point on Ω, denote it B and draw the chords BA and BC tangent to ω (fig. 3). From similarity of the triangles BIC' and $B_0'CB_0$ it follows that $B_0C / 2R = r / BI$, i.e. $2R \cdot r = BI \cdot B_0C$. From the Euler formula it follows that power of I with respect to Ω is $-2R \cdot r = -BI \cdot IB_0$. Therefore $BI \cdot B_0C = BI \cdot IB_0$, it means that in the triangle B_0CI $\angle B_0IC = \angle ICB_0$, but $\angle B_0IC = \angle B/2 + \angle ICB$, $\angle ICB_0 = \angle B/2 + \angle ICA$. We obtain that $\angle IBC = \angle ICA$. It means that the lines AC and BC are symmetric with respect to CI, therefor AC is tangent to ω.

4. Consider the circles $\Omega = (O, R)$ and $\omega_2 = (I, r_2)$, which are the circumcircle and the excircle of some triangle. From problem 2 it yields that $I_2O^2 = R^2 + 2R \cdot r_2$. Take any point B in Ω and let the lines BA and BC be tangents to ω_2 (fig. 4). As the triangles BI_2C' and $B_0'CB_0$

are similar $B_0C/2R = r_2/BI_2$, i.e. $2R\cdot r_2 = BI_2\cdot B_0C$, but by Euler formula the degree of point I_2 with respect to Ω is equal to $2R\cdot r_2 = BI_2\cdot I_2B_0$. So, $BI_2\cdot B_0C = BI_2\cdot I_2B_0$. This follows that the triangle B_0CI_2 is isoscelles and $\angle B_0I_2C = \angle I_2CB_0$, but $\angle B_0IC + \angle I_2CB = \angle BB_0C = \angle A$, i.e $\angle I_2CB_0 = \angle A/2$. We obnain that $\angle I_2CA = \angle A/2 + \angle B_0CA = (\angle A + \angle B)/2$. It means that the line I_2C is the external bisector of angle B, therefor AC is tangent to ω_2.

5. Firstly prove that the orthocentric axe is the radical axe of the circumcircle and the nine point circle. Consider two circles: Ω_B with diameter AC and ω_B with diameter HB (fig. 5). The sideline H_1H_3 of orthotriangle is its common chord so lies in its radical axe. Therefor $H_2'H_3\cdot H_2'H_1 = H_2'A\cdot H_2'C$. Now consider the circumcircle Ω and the nine point circle ω_0. The degrees of point H_2' with respect to Ω and ω_0 are equal to $H_2'A\cdot H_2'C$ and $H_2'H_3\cdot H_2'H_1$ respectively, i.e the degrees of the common point of respective sidelines of the triangle and its orthotriangle with respect to Ω and ω_0 are equal. This follows that the orthocentric axe is the radical axe of the circumcircle and nine point circle so it is perpendicular to the Euler line.

   Consider now the triangle $I_1I_2I_3$ formed by three excenters. Original triangle ABC is its orthotriangle, and the point I is its orthocenter. So the common points of external bisectors of the triangle ABC with respecive sidelines lie in the orthocentric axe of the triangle $I_1I_2I_3$ i.e in the line perpendicular to the Euler line of this triangle. But the Euler line of the triangle $I_1I_2I_3$ pass through its orthocenter (I) and nine point center (O), therefof it coincide with the line IO.

6. Firstly consider next problem: given two circles $\omega_1 = (O_1, R_1)$ and $\omega_2 = (O_2, R_2)$, their radical axe and center line intersect in the point P (fig. 6). Find the lenght of the segment PO_1. As the degrees of P with respect to both circles are equal, $PO_1^2 - R_1^2 = PO_2^2 - R_2^2$, $PO_2^2 - PO_1^2 = R_2^2 - R_1^2$, $O_1O_2\cdot(2PO_1 + O_1O_2) = R_2^2 - R_1^2$. So it is easy to express PO_1 through the radius of the circles and the distance O_1O_2..

   Now take the circumcircle of the triangle ABC with radius R as ω_1, and the circle $(I_1I_2I_3)$ with radius $2R$ as ω_2. Then the distance d_1 from the circumcenter O to radical axe (ℓ) is equal to $\dfrac{R^2+Rr}{\sqrt{R^2-2Rr}}$. Therefor the required distance is equal to $d = d_1 - IO = \dfrac{R^2+Rr}{\sqrt{R^2-2Rr}} - \sqrt{R^2-2Rr} = \dfrac{3Rr}{\sqrt{R^2-2Rr}}$.

7. The solution is analogously to the solution of the problem 5 with replacing of the triangle $I_1I_2I_3$ to the triangles II_2I_3, I_1II_3, I_1I_2I. The circumcircle (Ω) is the common nine-point circle of all these triangles and the lines I_kO are the Euler lines of respective triangles. So the internal bisecors axis of the triangle ABC are the the radical axis of Ω and the circumcircles of the respective triangles.

8. Let Ω and ω_2 be the circumcircle and the excircle of the triangle ABC (fig. 8). Let D be the touching point of its common external tangent with Ω. There are two limit "triangles" in the family of triangles with Ω and ω_2 as the circumcircle and the excircle. Consider a case when the secant AB becomes tangent. This will be the common external tangent to Ω and ω_2. Then the points A and B coincide in the point D, and the lines BC and AC coincide in the tangent PD.
Consider now the circles Ω and ω_2, such that the tangent to ω_2 in its common point P passes through the point D. Then $DK^2+(r_2-R)^2=OI_2^2$, $\operatorname{tg}(\angle I_2DK)=\dfrac{r_1}{DK}=t$. As $\angle PDK = 2\cdot\angle I_2DK$ we obnain $\sin(\angle PDK)=\dfrac{2\cdot t}{1+t^2}=\dfrac{2\cdot r_2\cdot DK}{DK^2+r_2^2}$. The lenght of chord DP of circle Ω is equal to $DP=2R\cdot\sin(\angle DPK)$. But as DP and DK are the tangents to ω_2, $DP=DK$. So $2R\dfrac{2r_2\cdot DK}{DK^2+r_2^2}=DP$, $4R\cdot r_2 = DK^2+r_2^2$. Using the expression for DK^2 we obtain the Euler formula $I_2O^2=R^2+2Rr_2$.

9. As the circles Ω and ω are fixed, and by problem 6 the distance from the center of ω to the external bisectors axe can be expressed through the radius of these circles, we obtain that all feet of external bisectors lies in the fixed line.
Inversely. Let A_2 be an arbitrary point of this line. Let B and C be the common points of tangent to ω passing through A_2 with Ω. By Poncelet theorem the sideline BC generate the triangle ABC which have A_2 as the foot of the external bisector.

10. As the circles Ω and ω_1 are fixed, and by problem 8 the internal bisectors axe ℓ_1 passes through the touching points of of common external tangents of these circles, ℓ_1 is the fixed line.
Inversely. Let B_1 be an arbitrary point in the segment PQ. Let A and C be the common points of tangent to ω_1 passing through B_1 with Ω. By external Poncelet theorem the sideline AC generate the triangle ABC which have B_1 as the foot of the internal bisector.

11. Let be $R=r_2$. By problem 8 this is eqivalent that the common external tangents to Ω and ω_2 are parallel to the line OI_2. So DE is the diameter of the circumcenter i.e. $O\in A_1C_1$.

12. Let ω' be the circle with IB_0' as a diameter. Let ω'' be the circle with II_2 as a diameter (fig.12). The line $B_0'B_2$ is the radical axe of ω' and Ω. The line CB_2 is the radical axe of ω'' and Ω. So the line IB_2 is the radical axe of ω' and ω''. Let K be the second common point of these circles. As $\angle IKB_0' = \angle IKI_2 = 90°$, the point K lies in the line $B_0'I_2$, and therefor $B_0'I_2 \perp B_2I$.

Emelyanov L., Emelyanova T.

Around the Feet of Bisectors

Main problems

Solutions

Using solutions by Bazhov I. and Chekalkin S.

13. The external bisector BB_2 (fig. 13) is parallel to A_0C_0, hence $\frac{C_1B_{00}}{B_{00}B_2} = \frac{C_1X}{XB}$. In the triangle XBY the line BI is the bisector and the altitude, it means that BI is the median. Because $C_0I = C_0B$ A_0C_0 (i.e. XY) is the perpendicular bisector of BI. Therefore $IX \parallel BC$. Hence $\frac{C_1X}{XB} = \frac{C_1I}{IC} \Rightarrow \frac{C_1I}{IC} = \frac{C_1B_{00}}{B_{00}B_2}$, i.e. $IB_{00} \parallel AC$.

14. Because B_{00} lies on the perpendicular bisector of BI (fig. 13), $\angle B_{00}BI = \angle B_{00}IB \Rightarrow \angle B_{00}IB = \angle B_2B_1 = \frac{1}{2}\overset{\cup}{BB_0}$. It means that $\angle B_{00}BB_0 = \frac{1}{2}\overset{\cup}{BB_0} \Rightarrow B_{00}B$ is tangent to Ω.

15. Because B_{00} lies on the perpendicular bisector of BI (fig. 13), $B_{00}I = B_{00}B$.

16. $\angle IB_3B_0 = \angle B_0B_3B' + \angle B'B_3I = \angle B'BB_0 + \angle B'B_3I$ (fig. 16). Because $\angle B_0BO = 90^\circ$, $B'B$ is tangent and $\angle B'BI = \frac{1}{2}\overset{\cup}{BI}$. Then $\angle B_0B_3I = \angle B'B_3I + \angle IB_3B = \angle B'B_3B = 90^\circ$.

17. Let us consider the homothety ω and ω_2 with the center B (fig. 17). It is not difficult to show that BN intersects ω in the point opposite to the touching point of ω and AC. This means that the common point of lines BN и B_0B_0' is such point K that $IK \parallel AC$.
Now as $\angle IBB_0' = \angle IKB_0'$ the quadrilateral $IBB_0'K$ is cyclic (fig. 17). By problem 16 the points B_0', I and B_3 are collinear. So $\angle IBN = \angle IBK = \angle IB_0'K = \angle B_3B_0'B_0 = \angle B_3BB_0$.

18. As the quadruple of rays IA_0, IL, IC_0, IB_{00} is harmonic (fig. 18) the quadruple IA, IM, IC, IB_{00} also is harmonic. But by problem 13 $IB_{00} \parallel AC$ so M is the midpoint of AC.

19. Consider the homothety with center C_1 transforming I to C (fig. 19). As $IB_{00} \parallel B_2C$, this homothety transforms $B_{00} \to B_2$, $A_{00} \to A_2$. It follows that $A_{00}B_{00} \parallel A_2B_2$. Similarly $B_{00}C_{00} \parallel B_2C_2$. So A_{00}, B_{00}, C_{00} are collinear on the line parallel to ℓ.

20. Consider the homothety with center C_0 transforming I to C. As $IB_{00} \parallel B_2'C$ this homothety transforms $B_{00} \to B_2'$, $A_{00} \to A_2'$. It follows that $A_2'B_2' \parallel A_{00}B_{00} \parallel A_2B_2$. Similarly $B_2'C_2' \parallel B_2C_2$. So A_2', B_2', C_2' are collinear on the line parallel to ℓ.

21. Consider the perpendiculars to BI passing through B, I and the midpoint of the segment BI (fig. 21). As $B_{00}I = B_{00}B$ and $B_{00}B$ touches the circle Ω, the degrees of B_{00} with respect to Ω and the point I are equal. As $\angle B_6IA = \angle BIA - 90^\circ = \angle ICA$, $B_6I^2 = B_6A \cdot B_6C$, the radical axis of Ω and I coinciding with ℓ_{00} passes through B_6. The triangle B_8BI is recanguler so $B_{00}B = B_{00}I = B_{00}B_8$. As three considered lines are parallel and the lines $B_{00}B_6$ and B_7B_2 are also parallel, $B_{00}I = B_4B_6 = B_{00}B_8 = B_2B_4 = B_7B_8$. So the segment IB_7 is divided by ℓ_0 and ℓ_{00} to three equal parts.

22. By problem 7 ℓ_2 is the radical axis of the circles Ω and (II_1I_3). So the degrees of bisector's feet A_1 and C_1 with respect to these circles are equal. This means that the points D and E lying in this radical axis are the common points of Ω and (II_1I_3).

23. The circle Ω is the nine-point circle of the triangle II_1I_3 and the circle b_2 is its circumcircle. So the ratio of their radius is equal to 2.

24. As I_2 is the orthocenter of the triangle II_1I_3 the circles Ω and b_2 are homothetic with center I_2.

25. The radical axis of circles Ω и b_2 coincide with the line DE and passes through the point B_{00}. As $B_{00}I = B_{00}B$ the segment $B_{00}I$ is the tangent to the circle b_2.

26. The homothety with center I_2 and coefficient $\frac{1}{2}$ transforms the point D to the point D_1 lying on Ω (fig 26). As PD touches ω_2 $\angle DPI_2 = 90^\circ$. It means that $PD_1 = DD_1$, so the tangent to Ω in D_1 is parallel to PD. The homothety with center I_2 and coefficient 2 transforms this tangent to the line PD. But this homothety transforms Ω in b_2. So PD touches b_2.

27. As DP touches ω_2 and DD' is the diameter of Ω, $\angle DPI_2 = \angle DPD' = 90^\circ$. It follows that the points D', P and I_2 are collinear (fig. 27). Now $\angle OD'I_2 = \angle OPD' = 180^\circ - \angle OPI_2 = 180^\circ - \angle OQI_2$. It means that the points O, D', I_2 and Q are on the circle.

28. The chord DE is twice greater than the segment DL (fig 28). But $DL = OD \cdot \sin\angle I_2OD = R \cdot \dfrac{KI_2}{OI_2} = R \cdot \dfrac{DS}{\sqrt{R^2 + 2R \cdot r_2}} = \dfrac{R}{\sqrt{R^2 + 2R \cdot r_2^2}} \cdot \sqrt{OI_2^2 - OK^2} =$

$$= R \cdot \sqrt{\frac{4Rr_2 - r_2^2}{R^2 + 2R \cdot r_2^2}}.$$

So $DE = 2\sqrt{R \cdot r_2 \cdot \frac{4R - r_2}{R + 2r_2}}$.

Remark: Using this formula it is easy to prove that $r_2 = R$ follows $DE = 2R$ i.e. DE is the diameter of Ω (problem 11). Also the inequality for the exradius can be obtained

29. As $\angle BB_1B_2$ is the external angle of the triangle ABB_1 (fig. 29) $\angle BB_1B_2 = \angle A + \angle B/2$. from the right-angled triangle B_1B_2B we obtain that $\angle BB_2B_1 = 90^\circ - \angle A - \angle B/2$. On the other hand, $\angle OBB_1 = \angle OBC - \angle B_1BC$, so $\angle OBB_1 = (90^\circ - \angle A) - \angle B/2$. Thus $\angle BB_1B_2 = \angle OBB_1$, that is, OB touches the circle (B_1BB_2).

30. As (B_1BB_2) is the Apollonius circle of the segment AC (fig. 30) we have $\frac{AB}{BC} = \frac{AK}{KC}$. Thus $AB \cdot KC = AK \cdot BC$ and the quadrilateral $ABCK$ is harmonic. It follows that the tangents to Ω at its vertices A and C intersect in the point S lying in the line BK. It means that BK is the symmedian of the triangle ABC.

31. It is well known that the Aubert line of the complete quadrilateral is the radical axis of three circles having its diagonals as diameters. If the triangle is formed by the bisectors axis than the segments A_1A_2, B_1B_2 и C_1C_2 are its diagonals. By problem 30 the degrees of the point O with respect to three circles having this segments as diameters are equal to R^2. So O lies in the radical axis of these circles.

32. The degree of Lemoine point L with respect to the circle having B_1B_2 as diameter is equal to $-BL \cdot LK$ (fig.30). The degree of this point with respect to Ω is the same. So the degrees of L with respect to three circles having the diagonals of the quadrilateral as diameters are equal and L lies on the Aubert line.

33. By problem 26 B_5 is the common point of two common internal tangents of the circles ω_2 and b_2. So B_5 is the homothety center of these circles. This homothety transforms the line AC touching ω_2 to the parallel line IB_{00} touching b_2. So the touching point of ω_2 with the sideline AC and the incenter I are the respective points of this homothety. It means that the point B_5 lies in the segment between these points

34. *The authors don't know the synthetic solution of this problem. The participants also didn't find it. There is a plan of calculation in barycentric coordinates.*
Using the coordinates of the Gergonne point G and the centroid M we can find the coordinates of the point T dividing the segment GM in the ratio equal to -3. Also we can find the ratio in which the line BT divide the sideline AC. Now as B_5 divide the segment IB'' (B'' is the touching point of ω_2 with AC) in the ratio equal to $2R/r_2$ (problem 33), we can find the coordinates of B_5 and verify that the lines BB_5 and BT divide AC in the equal ratio.

35. The foot of internal bisector is the internal homothety center of the incircle and the respective excircle. The foot of the external bisector is the external homothety center of two excircles. The point F is the external homothety center of the incircle and the nine-point circle and the points F_1, F_2 и F_3 are the internal homothety centers of the nine-point circle and respective excircles. The asertion of the problem follows from three homothety centers theorem.

36. The solution is in the book: I. F. Sharygin. «Geometry 9-11», problem № 586 (in Russian).

37. By three homothety centers theorem the points F_2, B_1 and F are collinear. Similarly the points F_3, C_1, F are collinear and the points F_1, A_1, F are collinear. Thus the triangles $F_1F_2F_3$ and $A_1B_1C_1$ are perspective with center F. Also by problem 36 these triangles are similar. So the sum of the angle $C_1B_1A_1 = F_1F_2F_3$ and the angle C_1FA_1 is equal to 180°. It means that the points A_1, B_1,C_1, F are concyclic.