Diophantine equations of second degree

In this project we study some properties of Diophantine equations of second degree. Those who advance in the project will develop a theory allowing one to solve a large (and interesting) class of problems. Some exciting examples are presented below.

We start with second degree equations in rational numbers. We will elaborate an algorithm which effectively determines whether an equation has a solution. As an application of this theory, we prove the following theorem by Carl Friedrich Gauss.

Theorem (Gauss)**.** A positive integer number d can be written as a sum of three squares if and only if d cannot be represented in the form $4^n(8m-1)$.

After the semifinal, we will focus on integral solutions of degree 2 equations in two variables. To investigate the solutions of these equations, we will introduce the *maps* of quadratic forms. We will also prove the following statement.

Theorem (J. Conway)**.** There exists a unique[1] homogeneous polynomial $f(x,y,z)$ of degree 2 such that all the equations $f(x,y,z)=m$ with $m=1,\dots,30$ admit integral solutions, but any equation of the form $f(x,y,z)=m$ with $m<0$ has no integral solutions.

Introductory problems

In this subsection we collect several easy problems on (integral) quadratic forms. These problems may be solved using a general algorithm of solution of such equations; we believe that some participants will construct such an algorithm. Nevertheless, all these introductory problems may be solved in a direct way.

Notice that there is no such algorithm for Diophantine equations of an arbitrary degree; the fact that it cannot exist was proved by Yu. Matiyasevich in 1970; by proving this fact he has solved the 10th Hilbert problem.

See Problems 1–9.

If you are stuck on some of these problems, you may proceed to the next sections and return to these problem later, after obtaining some technical background.

The quadratic forms

Definition 1. A *quadratic form* is a homogeneous polynomial of degree 2. Here are two examples: $2x^2+2xy-y^2$ and $x^2-xz+y^2-2z^2$.

For every positive integer d, we denote by $\mathbb{Z}^d$ the set of d-tuples of integers. E.g., the set of pairs of integers is denoted by $\mathbb{Z}^2$. Any quadratic form in two variables x and y determines a function on $\mathbb{Z}^2$ mapping a pair (x,y) to the number $f(x,y)$. Hereafter we will frequently denote a pair $(x,y)\in\mathbb{Z}^2$ by one letter (say, v) and write $f(v)$ for $f(x,y)$.

Definition 2. We say that a quadratic form f *represents an integer* m if there exists a pair $v\in\mathbb{Z}^2$ with $v\neq(0,0)$ and $f(v)=n$. In other words, f represents m if the equation $f(x,y)=n$ has a nonzero integer solution (thus not any quadratic form represents 0).

See Problems 10–11.

Definition 3. We say that two quadratic forms are *equivalent* if each number represented by one of these forms can also be represented by the other one.

[1]Formally, this statement is wrong; this polynomial is unique up to some equivalence which will be described later.

See Problem 12.

It appears that some quadratic forms are easier to deal with than some other ones. One of our aims is the following: Given a quadratic form f, we wish to find some convenient form equivalent to it (e.g., a form like $ax^2 + by^2$). For that, we need to work out some necessary and sufficient conditions on the two quadratic forms to be equivalent. In particular, we will find some explicitly computable invariants of quadratic forms.

Definition 4. We say that a quadratic form f is
a) *positive definite* if $f(v) > 0$ for all $v \neq 0$,
b) *non-negative definite* if $f(v) \geq 0$ for all $v \in \mathbb{Z}^2$,
c) *indefinite* if $f(u) > 0$ for some $u \in \mathbb{Z}^2$ and $f(v) < 0$ for some $v \in \mathbb{Z}^2$.

See Problem 13.

Extended arithmetics: p-adic numbers

The main goal of this section is to impart some sense to the following Metatheorem.

Theorem (Metatheorem)**.** A quadratic equation has a solution in rational numbers if and only if there are no obstacles modulo any prime p.

Using this Metatheorem, one can prove, for instance, the Gauss theorem and the following theorem by Legendre.

Theorem (Legendre)**.** Every positive integer is a sum of four squares of integers.

In our project we split the proof of Metatheorem (as well as of theorems by Gauss and Legendre) into several problems. Any such problem can be solved independently. To start with, we need to impart a formal sense to our Metatheorem (in the previous formulation, it is ambiguous; moreover, it remains wrong after any easy attempt to formalize it). Let us present some example.

Definition 5. We say that m is a *quadratic residue* modulo n if there exists an integer t such that $m \equiv t^2 \pmod{n}$.

See Problems 14–15.

In the case $\gcd(m, n) = 1$, the conditions a)–c) of Problem 14 imply that the equation

$$ax^2 + by^2 = c$$

has a rational solution. On the other hand, in the case $\gcd(m, n) \neq 1$ one needs to introduce additional conditions on m and n which are related to prime divisors of $\gcd(m, n)$. If one writes them down directly, these conditions would look a bit long, although simple.

An elegant (and short) way to present such conditions is based on the notion of *p-adic numbers*. We follow this approach.

For any prime p, a *p-adic integer* is defined as any formal sum of the form

$$a_0 + a_1 p + \ldots + a_n p^n + \ldots \qquad (a_i \in \mathbb{Z}) \tag{1}$$

where the number of summands may be infinite. Two p-adic integers are assumed to be equal if they coincide modulo p^n for any n. For example,

$$1 = (p+1) - (p+1)p + (p+1)p^2 - (p+1)p^3 + \ldots.$$

The set of p-adic integers is denoted by $\mathbb{Z}_p$.

One may add, subtract, and multiply p-adic integers in an obvious way. Therefore, given an equation $f = 0$ with integer coefficients, one may consider its solutions in $\mathbb{Z}_p$. The following problem provides a connection between the sets of solutions of $f = 0$ in integers and in p-adic integers.

See Problem 16.

The notion of a p-adic integer is an extension of a notion of an integer. A similar extension exists for the rational numbers. Namely, for any prime p we define a *p-adic number* (or a *p-adic rational*) as a formal expression of the form

$$a_{-k}p^{-k} + a_{-k+1}p^{-k+1} + \ldots + a_n p^n + \ldots \qquad (k \in \mathbb{Z}, \quad a_i \in \mathbb{Z}); \tag{2}$$

the equality of two p-adic numbers is defined as above. The set of all p-adic numbers is denoted by $\mathbb{Q}_p$. Obviously, any p-adic integer can be represented in the form (2) with $a_{-k} = ... = a_{-1} = 0$ (or with $k \leq 0$).

In order to get acquainted with the notion of p-adic numbers, it is useful to solve the following problems.

See Problems 17–26.

Now we are ready to present a formal version of Metatheorem.

Theorem (the Hasse–Minkowski principle)**.** A quadratic equation $f = 0$ has a rational solution if and only if it simultaneously has solutions

- in real numbers,
- in p-adic numbers for every prime p.

See Problem 27.

The Hasse–Minkowski principle reduces solving an equation in rational numbers to solving the same equation in p-adic numbers. The advantage is that equations in p-adic numbers are much easier to solve. To show this, we first describe an algorithm which allows one to check whether an equation in two variables has a rational solution. Let us first deal with an equation of the form

$$z^2 - ax^2 - by^2 = 0. \tag{3}$$

Definition 6. Consider a prime p and a pair of integers (a, b). Let us define the *Hilbert symbol* $(a, b)_p$ of a pair (a, b) with respect to p as follows: If the equation (3) has a nonzero solution in p-adic integers, then we set $(a, b)_p = 1$; otherwise we set $(a, b)_p = -1$.

Thus, for finding the solutions of (3) it is helpful to learn how to find $(a, b)_p$.

See Problems 28–29.

To write down an expression for the Hilbert symbol in a compact form, we will use the *Legendre symbol* $\left(\frac{x}{p}\right)$ defined for any integer x and prime p. It equals to 1, -1, or 0 depending on whether x is a nonzero quadratic residue, a quadratic non-residue, or zero. For an odd prime p, one may calculate it using the formula

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}} \pmod{p}$$

See Problems 30–39.

Two variables: maps of quadratic forms

In this section of the project we develop a technique which allows us to solve the equation

$$E_m : ax^2 + bxy + cy^2 = m \tag{4}$$

effectively, here x and y are integer variables and a, b, c, m are integer parameters. To do this, we assign a *map* to any quadratic form in two variables and show how to read properties of the form out of this map. We believe that using this approach the participants will be able to solve the following (super)problems. By a solution in this section we always mean a nonzero integer solution if not mentioned otherwise.

Problem 46 (Superproblem)**.** Assume that the equation E_m has a solution for some positive m, for some negative m and has no solutions for $m = 0$. Prove that in this case either E_m has no solutions, or E_m has infinitely many solutions for any m.

Problem 47 (Superproblem)**.** Is it true that if the equation E_m has solutions for

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

then in this case E_m has solutions for any integer m?

Problem 48 (Superproblem)**.** Assume that the equations E_1, E_2, E_3, E_5 have solutions. Show that in this case the equation E_m has solutions for some $m < 0$.

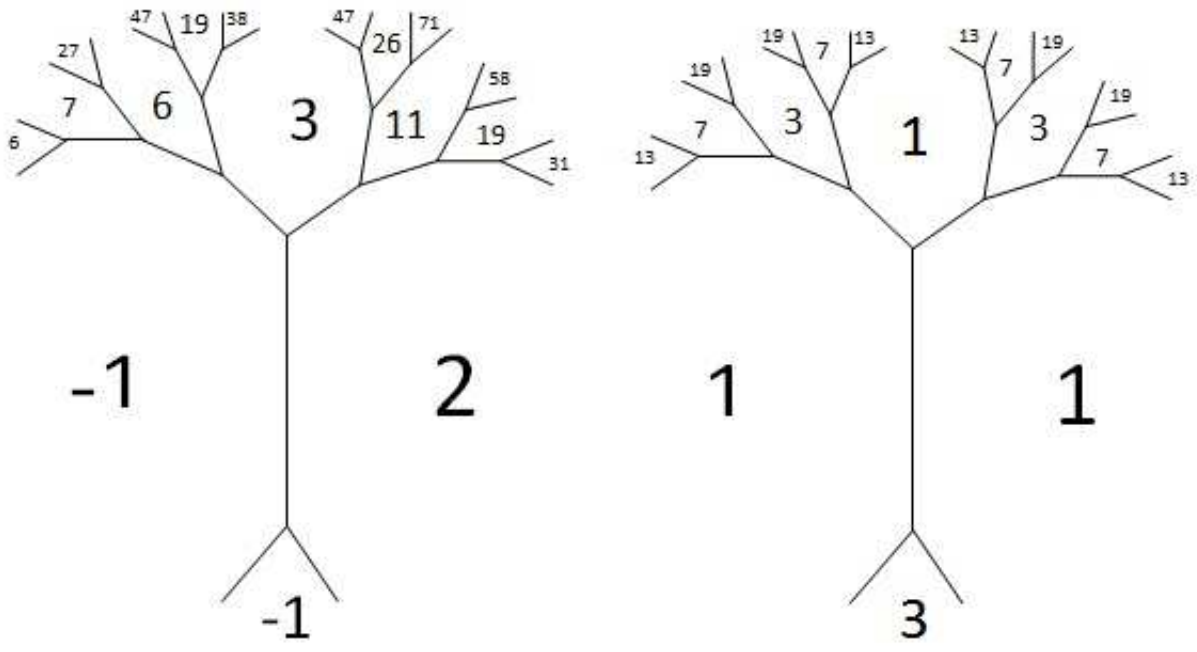
Now we treat two examples to show how the map of a quadratic form may help to solve equations.

Examples of maps

The goal of this subsection is to show that it might be interesting to consider maps of quadratic forms. Given two polynomials

$$2x^2 + 2xy - y^2 = 1 \text{ and } x^2 - xy + y^2 = 2, \tag{5}$$

we assign the following pictures to them, they are called *maps*:



From these maps we see that equations (5) have no integer solutions.

Drawing a map [2]

To find something common in a variety of something very different it was a good idea (from time to time) to consider all this different (some)things simultaneously and providing this "all" by some additional structure. Following this approach, we consider all forms which are linearly equivalent (see definition below) to a form f and provide this set with an oriented graph structure (we put points of quadratic forms on the plane and connect them by edges in some way). To do this we need a notion of basis/superbasis of $\mathbb{Z}^2$.

Definition 7. A *basis* of $\mathbb{Z}^2$ is a pair $w_1, w_2 \in \mathbb{Z}^2$ such that for any $v \in \mathbb{Z}^2$ there exist $m, n \in \mathbb{Z}$, for which

$$v = mw_1 + nw_2.$$

Before semifinal, we had the notion of equivalent forms. Unfortunately, if we work with maps of quadratic forms, it is more natural to use the following notion.

Definition 8. Two forms f_1, f_2 are called *linearly equivalent*, if $\exists$ a, b, c, d, such that $ad - bc = 1$ and

$$f_1(x, y) = f_2(ax + by, cx + dy).$$

See/solve Problems 49–52.

Definition 9. A *superbasis* of $\mathbb{Z}^2$ is a collection $\{\pm w_1, \pm w_2, \pm(w_1 + w_2)\}$, where $\{w_1, w_2\}$ is a basis of $\mathbb{Z}^2$. We say that a basis $\{w_1, w_2\}$ is a *specialization* of a superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$. We say that a superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$ is an *extension* of a basis $\{w_1, w_2\}$.

Example 1. Write down all the extensions of a given basis $\{w_1, w_2\}$. Write down all the specializations of a given superbasis $\{\pm w_1, \pm w_2, \pm(w_1 + w_2)\}$.

Now we are able to describe the map f. We start from a part of this map which does not depend on f at all:

(1) to any superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$, we assign a point on the plane (the vertex of the graph),

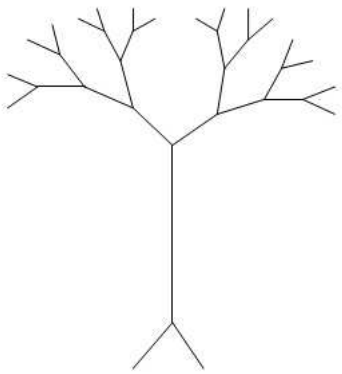
(2) to any basis $\{w_1, w_2\}$, we assign a segment on the plane (the edge of the graph), which connects

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\} \text{ and } \{\pm w_1,\ \pm w_2,\ \pm(w_1 - w_2)\}$$

(we assign the same edge to $\{w_1, w_2\}$, $\{-w_1, w_2\}$, $\{w_1, -w_2\}$, and $\{-w_1, -w_2\}$);

(3) to any $w \in \mathbb{Z}^2$, we assign the region on the plane such that its border consists of edges corresponding to bases containing w (we assign the same region to w and $-w$).

It turns out that it is possible to draw the following picture without self-intersections on the plane.



(6)

[2] If you wish to see a much shorter way to draw a map of a form you could go to the appendix of this section. Try to prove why a map defined in this way satisfies the desired properties.

Note that (6) does not depend on f. Now we will mark the graph with integers depending on f. Integers will be assigned to every region and to every edge of (6) in such a way that it will be possible to restore the class of f up to linear equivalence in the unique way. We use the following rules.

(1) If a region corresponds to $w \in \mathbb{Z}^2$, then $f(w)$ will be assigned to it.

(2) If an edge I corresponds to a basis $\{w_1, w_2\}$, then we assign to it the positive integer

$$|f(w_1 + w_2) - f(w_1) - f(w_2)|.$$

Also, we make I directed: if $f(w_1 + w_2) > f(w_1 - w_2)$, then edge I starts at vertex-superbasis

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 - w_2)\}$$

and ends in

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\};$$

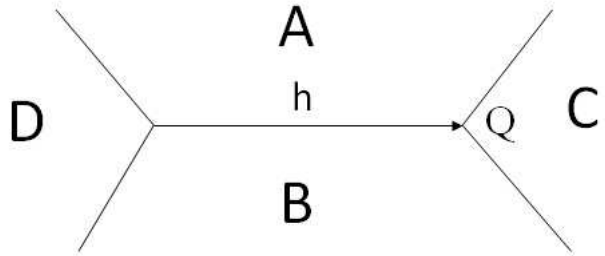
if $f(w_1+w_2) < f(w_1-w_2)$, then otherwise. If $f(w_1+w_2) = f(w_1-w_2)$, then we do not determine the direction of I (and usually omit 0 at I).

The resulting picture will be called the *oriented map* of a quadratic form f. If we omit numbers attached to the edges in this picture, then it will be called the *map* of a quadratic form. For example, the maps of the forms $2x^2 + 2xy - y^2$ and $x^2 - xy + y^2$ are presented on page 4.

Example 2. Draw the (oriented) maps for the quadratic forms

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

In two following problems, the integers A, B, C, D, h are related to the picture



See/solve Problems 53–55.

For positive definite quadratic forms, the following definition plays a key role.

Definition 10. A *well* is a vertex Q of the oriented map of a quadratic form such that all the edges which are incident to Q go out of Q.

See/solve Problems 56–60.

We want to give you an advice:

1) The ideas of proofs of Superproblems 1, 2, and even of every equation of type (4), is very close to Problems 59, 60.

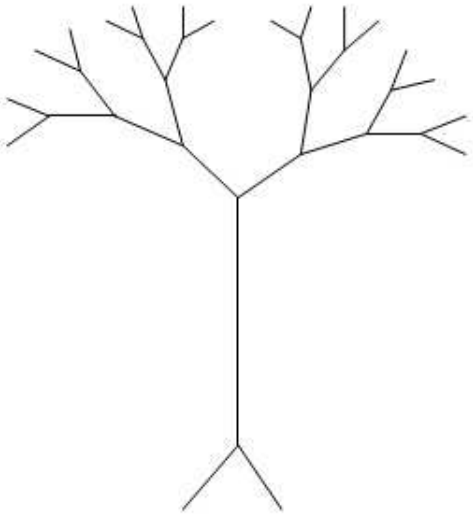
2) In real mathematical life, no one (except yourself) would give you a sequence of (relatively simple) exercises which lead to a proof of any mathematical Problem. You will be very lucky if you learn (most probably, occasionally) a significant piece of the desired methods and ideas somewhere.

3) The goal of this conference is to let you know something about real mathematical life.

If you did not guess, we end up with problems helping you to solve Superproblems 1, 2, 3 and determine when equations (4) have solutions. To simplify your life, we have prepared several pictures which can help you to solve or to guess something.

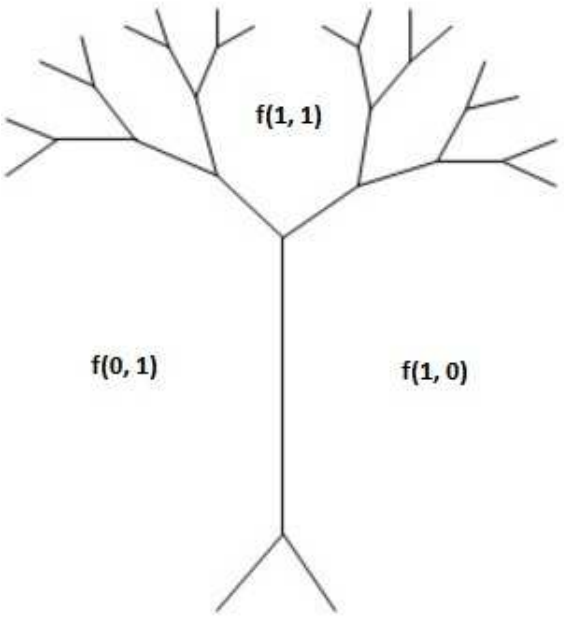
A quick way to describe the map of a quadratic form

Algorithm $f \to \Gamma_f$ **:** We consider an infinite tree (a connected graph without cycles) on the plane such that every its vertex is incident to exactly 3 edges. A part of such a tree is presented below



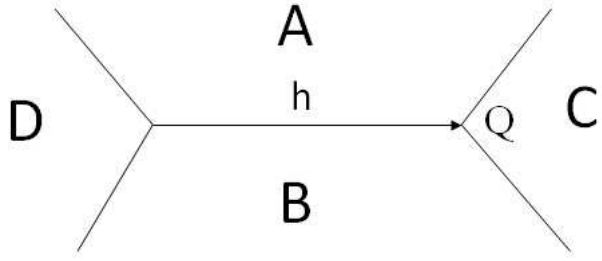
. (7)

We take any vertex of this graph and write integers $f(1,0)$, $f(0,1)$, and $f(1,1)$ on three regions which meet at this vertex.



. (8)

The values at all the other regions are determined by the following rule



. (9)

Rule 1: For a given edge, if 3 values which are adjacent to it (see Figure (9)) are already known, then the fourth one is determined by the formula $2(A+B) = C+D$.

It is easy to see that Rule 1 determines the map Γ_f. Now we need to construct the oriented map $\vec{\Gamma}_f$. We use the following rules (see Figure (9)):

Rule 2: We write $|2(A+B)-C| = |2(A+B)-D|$ at the edge h.

Rule 3: If $C < D$, then the edge h is replaced by an arrow from C to D; if $C < D$ then h is replaced by an arrow from D to C; if $C = D$ then the edge h is unoriented.

Some properties of Γ_f:

1) the points of Γ_f correspond to quadratic forms which are linearly equivalent to f;

2) the regions Γ_f are in one-to-one correspondence with the nonzero rational numbers $\frac{m}{n}$;

3) if the region corresponds to a reduced fraction $\frac{m}{n}$, then integer $f(m,n)$ is written in it;

4) two quadratic forms f and g are linearly equivalent if and only if their maps Γ_f and Γ_g coincide.