A combinatorial approach to sums of two squares and related problems

Christian Elsholtz

Abstract In this paper we study elementary approaches to classical theorems on representations of primes of the form ax^2+by^2, in particular the two squares theorem. While most approaches make use of quadratic residues, we study a route initiated by Liouville, and simplified by Heath-Brown and Zagier.

Key words: 11A41, 11E25
Fermat's two squares theorem, binary quadratic forms

Dedicated to Melvyn Nathanson. With many thanks for his beautiful expositions in additive number theory, emphasizing elementary methods.

1 Introduction

In this paper we study elementary approaches to classical theorems on representations of primes of the form ax^2+by^2, in particular the two squares theorem.

1.1 The sums of two squares theorem

Theorem 1. *A positive integer n can be written as a sum of two integer squares, if and only if the canonical prime factorization $n=p_1^{\gamma_1}\cdots p_r^{\gamma_r}$ (where the p_i are distinct primes) satisfies the condition: if $p_i\equiv 3 \bmod 4$, then γ_i is even.*

Christian Elsholtz, Institut für Mathematik A, Technische Universität Graz, Steyrergasse 30, A-8010 Graz, Austria.
email: elsholtz@math.tugraz.at

In order to prove this theorem one proves the following theorem and several minor lemmata.

Theorem 2. *A prime $p \equiv 1 \bmod 4$ can be written as $p = x^2 + y^2$.*

Wells [38] includes Theorem 2 in a list of the 10 most beautiful results in mathematics.

In his "Apology" Hardy [18] writes: "Another famous and beautiful theorem is Fermat's 'two square' theorem... All the primes of the first class" [i.e. 1 mod 4] ... "can be expressed as the sum of two integral squares... This is Fermat's theorem, which is ranked, very justly, as one of the finest of arithmetic. Unfortunately, there is no proof within the comprehension of anybody but a fairly expert mathematician."

In this paper we discuss quite elementary proofs and it would be interesting to know if Hardy would also have written this about the types of proof (and its simplifications), discussed in sections 1.2, 1.3 and 1.6.2.

The history of the theorems above is described in detail in Dickson [8] (volume 2, chapter VI), and also in Edwards [10]. Already Diophant discussed representations of integers as a sum of two squares, and, by slightly altering the text, Jacobi interpreted Diophant's writing in such a way that Diophant possibly essentially knew and was able to prove: if a square-free number n is a sum of two squares, then neither n nor any factor of n is of the form $4k-1$, (see [8], page 236).

The first correct statement of the necessary and sufficient conditions for writing an integer as a sum of two integer squares, without a proof, might have been by Albert Girard. The theorem is also often attributed to Fermat, who wrote he had a proof. His proof is not known to us, even though in this case it is believed he had the right methods to prove the theorem indeed. Euler eventually gave the first proof that has survived.

Since $p = 2 = 1^2 + 1^2$, and since all squares are of the from 0 or 1 mod 4 so that no number $n \equiv 3 \bmod 4$ can be a sum of two squares. Theorem 2 implies

Corollary 1. *A prime p can be written as $p = x^2 + y^2$ if and only if $p = 2$ or $p \equiv 1 \bmod 4$.*

Lemma 1. *If $m = x_1^2 + y_1^2$ and $n = x_2^2 + y_2^2$ can be written as sums of two integer squares, then their product mn can also be written in this form.*

Proof of Lemma: This follows immediately from the identity $mn = (x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2$, an identity which can be motivated by means of complex numbers:

$$\begin{aligned} mn &= ((x_1 + y_1 i)(x_1 - y_1 i))((x_2 + y_2 i)(x_2 - y_2 i)) \\ &= ((x_1 + y_1 i)(x_2 + y_2 i))((x_1 - y_1 i)(x_2 - y_2 i)) \\ &= (x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1))((x_1x_2 - y_1y_2 - i(x_1y_2 + x_2y_1)) \\ &= (x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2. \quad \square \end{aligned}$$

Lemma 2. *If n is divisible by a prime $p \equiv 3 \bmod 4$, and $n = x^2 + y^2$, then $x \equiv y \equiv 0 \bmod p$.*

Proof of Lemma 2: By Fermat's little theorem: Let p be prime and x an integer, then

$$x^{p-1} \bmod p = \begin{cases} 0 & \text{if } x \equiv 0 \bmod p \\ 1 & \text{if } x \not\equiv 0 \bmod p. \end{cases}$$

If $p \equiv 3 \bmod 4$, then

$$(x^2+y^2)(x^{p-3}-x^{p-5}y^2+x^{p-7}y^4 \mp \cdots + y^{p-3}) = x^{p-1}+y^{p-1}.$$

Since $x^2+y^2 \equiv 0 \bmod p$, one also has $x^{p-1}+y^{p-1} \equiv 0 \bmod p$. As $p > 2$, we must have, by Fermat's observation above, that $x \equiv y \equiv 0 \bmod p$. □

The above lemmata reduce the proof of Theorem 1 to a proof of Theorem 2.

There is a multitude of proofs of Theorem 2. Most of these use quite essentially the fact that for a prime $p \equiv 1 \bmod 4$ there is a solution of $x^2 \equiv -1 \bmod p$. This follows for example from $x = \frac{p-1}{2}!$ or $x = g^{\frac{p-1}{4}}$, where g is a generating element of the group $(\mathbb{Z}/p\mathbb{Z})^\times$ or g is a nonresidue modulo p. However, checking the details in this calculation from first principles is already half of the proof.

The methods involved in these various proofs include e.g. congruence computations, Minkowski's theorem, the pigeon hole principle, properties of Gaussian integers, continued fractions and the like. The book by Hardy and Wright [19] gives several different proofs. For other proofs see also [33], [40], [6].

A very different second type of proof goes back to Liouville. In a series of eighteen papers Liouville describes a quite general method, a special case of which gives Theorem 2. Liouville's work is described in the books by Bachmann [3], Dickson [8], Uspensky and Heaslet [34], Venkov [36], and Nathanson [29].

This special case was considerably simplified by Heath-Brown [20]. Zagier [41] reformulated Heath-Brown's proof to write it in one sentence, however leaving elementary calculations to the reader.

This proof has generated a considerable literature explaining the proof for teaching purposes [35], [39], [12], [31], [5] or extending it to related results: [4], [13], [14], [16], [21], [22], [23], [32]. The collection of beautiful proofs "Proofs from the BOOK" by Aigner and Ziegler [1] explains in its first edition Zagier's version of the proof, but changed to Heath-Brown's version for the 2nd edition.

A key ingredient is an ingenious choice of a set which allows a partition into orbits of length 1 or 2. In this way a simple parity check guarantees the decomposition into two squares. The reader who is familiar with Liouville's method will appreciate the simplifications made by Heath-Brown and Zagier. Still, the proof is quite mysterious. We make an attempt to demystify the proof, i.e. explain how the details can be motivated.

In addition to the study of this second type of proof we apply the idea of orbits of length 1 or 2 to a proof based on lattice points, which is more in the spirit of the first type of proof. After reviewing the history of these, i.e. discuss contributions by Lucas, Grace and others we present in section 1.6 a quite short version of the proof, which admittedly also requires some routine checking, as is the case with the proofs by Zagier and Heath-Brown.

1.2 Zagier's proof

Here is the famous one-sentence-proof for primes $p = 4k+1$, quoting from Zagier [41].

"The involution on the finite set $S = \{(x,y,z) \in \mathbb{N}^3 : x^2+4yz = p\}$ defined by

$$(x,y,z) \mapsto \begin{cases} (x+2z,z,y-x-z) & \text{if } x < y-z \\ (2y-x,y,x-y+z) & \text{if } y-z < x < 2y \\ (x-2y,x-y+z,y) & \text{if } x > 2y \end{cases}$$

has exactly one fixed point, so $|S|$ is odd and the involution defined by $(x,y,z) \to (x,z,y)$ also has a fixed point. $\square$"

Quite a few routine checks are necessary to verify all these implicit claims. For the reader's ease we would like to add that the first map, α (say), defines a partition $S = S_1 \cup S_2 \cup S_3$ with $S_1 = \{(x,y,z) \in S : x < y-z\}, S_2 = \{(x,y,z) \in S : y-z < x < 2y\}, S_3 = \{(x,y,z) \in S : x > 2y\}$. There are no solutions with $y-z = x$ or $x = 2y$, since otherwise x^2+4yz is not a prime. Solutions with $x < y-z$ are mapped to solutions with $x > 2y$, and vice versa. Solutions with $y-z < x < 2y$ are mapped to solutions with the very same property. That is $\alpha(S_1) = S_3, \alpha(S_3) = S_1, \alpha(S_2) = S_2$. Thus fixed points of α must lie in S_2 and therefore satisfy $(x,y,z) = (2y-x,y,x-y+z)$, i.e. $x = y$. Since p is prime, the only fixed point is $(1,1,(p-1)/4)$.

Writing out all details, which we do not do here, makes the proof actually quite a bit longer.

1.3 Heath-Brown's proof

Heath-Brown reformulated Liouville's work in 1971. His version [20] appeared in 1984 in a student magazine, issued by the undergraduate mathematical society at Oxford University. Meanwhile a retyped version is available, see the bibliography. Since Heath-Brown's proof was slightly different, we describe his proof briefly.

Let us define

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & -1 \end{pmatrix}.$$

Define the sets

$$S = \{(x,y,z) \in \mathbb{Z}^3 : p = 4xy+z^2, \quad x,y > 0\},$$

$$T = \{(x,y,z) \in S : z > 0\}, \quad U = \{(x,y,z) \in S : x+z > y\}.$$

One can check that $X_1^2 = X_2^2 = X_3^2 = I$. Moreover, X_1 maps S to itself, X_2 maps T to itself, and X_3 maps U to itself. One also verifies that $|T| = |X_1 T|$ and $|U| = |X_1 U|$. Since S is the disjoint union of T and $X_1 T$ it follows that $|S| = |T| + |X_1 T| = 2|T|$ and similarly $|S| = 2|U|$. This implies $|T| = |U|$. Since the map X_3 acting on U has exactly one orbit of length 1 (for $y = z = 1$), and since all other orbits have length two, we find that $|U|$ must be odd. So, $|T|$ is also odd, and the action of X_2 on T must have an orbit of length 1, i.e. there is a fixed point with $x = y$, giving $p = 4x^2 + z^2$.

This is an impressive example that the right choice of a set, group action and orbit counting can simplify existing proofs. Another example of this principle is McKay's proof [28] of a Theorem of Cauchy in group theory.

1.4 Grace' lattice point proof

In this section we describe a proof based on lattice points, due to Grace [17]. It is one of the proofs in Hardy and Wright's book [19].

The proof starts with the fact that $a^2 \equiv -1 \bmod p$ has a solution. Take those lattice points in $\mathbb{Z} \times \mathbb{Z}$ with $ax \equiv y \bmod p$. Note that if (x,y) and (x',y') belongs to the set, then also $(x \pm x', y \pm y')$ belong to it, so that the set of these points define a discrete lattice. Let $P_1 = (x,y)$ be one of the points with minimal distance to the origin $P_0 = (0,0)$. Since $-ay \equiv x \bmod p$, the point $P_2 = (-y,x)$ also belongs to the lattice. These points together with $P_3 = (x-y, x+y)$ define the fundamental domain. Observe that there are no further lattice points in this fundamental domain, since otherwise the distance from $(0,0)$ to (x,y) was not minimal. Also observe that in this situation the fundamental domain is not only a parallelogram, but even a square.

In a very large circle about the origin, the proportion of points belonging to the lattice is $\frac{1}{p}$ so that the area of the fundamental domain is p. Hence the side lengths of the square satisfy by Pythagoras' theorem: $x^2 + y^2 = p$.

The lattices can also be understood as coming from the problem of regular solutions of placing p non-taking queens on a $p \times p$ chessboard, with reduction modulo p, i.e. a chessboard on a torus. This approach has been studied by Polya [30], Kraitchik [24] and Larson [25]. These proofs also make use of counting the lengths of orbits and are similar in spirit to those discussed below.

1.5 Lucas' work on regular Satins

In 1867 Édouard Lucas [26] had similar ideas on regular "Satin" squares which were thought of in connection with patterns of fabrics. As Decaillot [7] writes, in France

at that time there was a group of mathematicians writing as accessible as possible for a wide audience.

Without assuming that there is a solution of $a^2 \equiv -1 \bmod p$, he considered those integer lattices with slopes $2,3,\ldots,\frac{p-1}{2}$. He paired off those lattices with slopes s_i and s_j where $s_i s_j \equiv \pm 1$. For a given s_i there is a unique s_j in this set. He interpreted this in terms of the geometric pattern. Starting with an odd number of lattices, one lattice remains. This remaining lattice is associated to itself, and has a square unit.

In this paper, Lucas did not actually conclude the two squares theorem, namely that a prime $p \equiv 1 \bmod 4$ is a sum of two squares, but rather the opposite.

The reason for this apparently comes from the historical background. The question, for which moduli regular lattices exist, was asked by Édouard Gand, also in 1867, in connection with fabric patterns, and Gand's question was answered by Lucas.

However, there is some indirect evidence that Lucas later actually proved Theorem 2 using this method. Dickson [8] (Volume 2, page 245) gives [27] (which does not contain that proof) and Aubry [2] as references. Decaillot [7] mentions a comment by Aubry in Fermat's collected works [15] (note 27 of the 4th volume). Here Aubry writes that the two squares theorem is "perhaps the most beautiful of all of Fermat's theorems", and Aubry refers to a graphical proof by Lucas.

Decaillot [7] constructed a proof that possibly was the one given by Lucas. It is very similar to the proof by Grace discussed above.

1.6 A short proof

In this section we aim to modify the two approaches above to assemble a proof which can be formulated in one sentence. However, as is the case with Zagier's proof, several additional words of explanations are appropriate, and several routine calculations required. The author believes that memorizing this proof may be easier than memorizing Zagier's proof.

1.6.1 The long version

- Let $p \equiv 1 \bmod 4$ be a prime and let $S = \{2,3,\ldots,\frac{p-1}{2}\}$. For $z \in S$ let us define the lattices
$$L_z = \{(x \bmod p,\ zx \bmod p) : 0 \le x < p\}$$
as subsets of $\mathbb{Z}_p \times \mathbb{Z}_p$ (which can be thought of as a torus). To see that these are lattices take any two points $(x \bmod p,\ xz \bmod p)$ and $(y \bmod p,\ yz \bmod p)$. The sum $(x+y \bmod p, (x+y)z \bmod p)$ is again in S_z and the same follows for integer multiples $(\lambda x \bmod p, \lambda xz \bmod p)$.
For $p \equiv 1 \bmod 4$, the number $|S| = \frac{p-1}{2} - 1$ of lattices is odd. For a better understanding we draw these for $p = 17$.

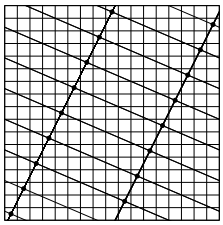

Fig. 1 L_2

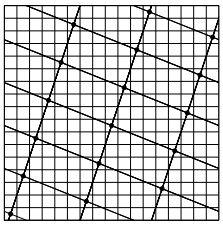

Fig. 2 L_3

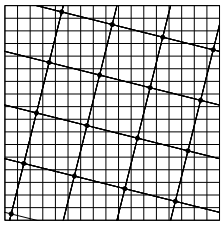

Fig. 3 L_4

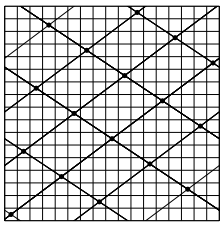

Fig. 4 L_5

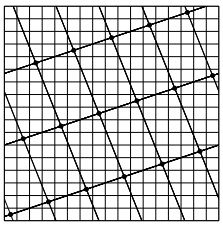

Fig. 5 L_6

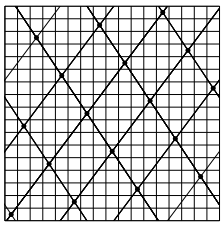

Fig. 6 L_7

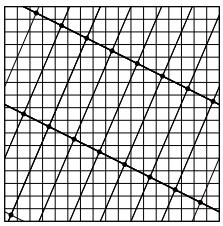

Fig. 7 L_8

In the pictures we include the parallelograms which define fundamental domains of the lattices. A fundamental domain is a parallelogram, spanned by a point and two of its 4 closest neighbours in two linear independent directions. In this sense, each point uniquely corresponds to a fundamental domain, so that there are p fundamental domain, and for a given lattice all of these parallelograms are congruent, understood modulo p.

- But the fundamental domains for different lattices are in general not congruent to each other. In the above example with $p = 17$ the shape of the fundamental domain is the same for L_2 and L_8, for L_3 and L_6, for L_5 and L_7. The lattice L_4 (which turns out to deliver the solution $x^2 \equiv -1 \bmod 17$ and finally the decomposition of $17 = 1^2 + 4^2$) does not have a corresponding partner. Generally this can be described by means of the following map: Let $S = \{2 \le a \le \frac{p-1}{2}\}$. Let $f : S \to S$ with

$$a \mapsto \begin{cases} a^{-1} \bmod p & \text{if } 2 \le (a^{-1} \bmod p) \le \frac{p-1}{2}, \\ -a^{-1} \bmod p & \text{otherwise.} \end{cases}$$

Observe that for $p = 17$ one has that $f(2) = 8, f(8) = 2, f(3) = 6, f(6) = 3, f(5) = 7, f(7) = 5, f(4) = 4$. Here the representatives of the residue classes modulo p are assumed to be in the interval $0 \le b < p$. It can be easily checked that f is an involution. We have to show that for all $a \in S$: $f(f(a)) = a$. If the first alternative holds for the inner argument, then also at the second time so that $f(f(a)) = f(a^{-1}) = (a^{-1})^{-1} = a$ and similarly $f(f(a)) = f(-a^{-1}) = -(-a^{-1})^{-1} = a$. Since $|S|$ is odd, there must be an odd number (i.e. at least one) of elements with $a = f(a)$. Since $-1, 1 \notin S$, it follows that $(a+1)(a-1) \equiv 0 \bmod p$ has no solution in S which implies that $a \equiv a^{-1} \bmod p$ has no solution. But then there must be an element with $a \equiv -a^{-1} \bmod p$. It is this element which satisfies $a^2 \equiv -1 \bmod p$, but we better leave it as $a \equiv -a^{-1} \bmod p$. In this form we see that the slopes a and $-a^{-1}$ of the sides of the parallelogram are orthogonal. The lattice is invariant under the map f which means it is invariant under a rotation by 90^o. This proves why for prime $p \equiv 1 \bmod 4$ there must be a lattice amongst the lattices L_z, of which the fundamental domain is a square.

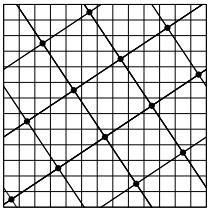

Fig. 8 L_z with $z^2 \equiv -1 \bmod p$ being a fixed point, here $p = 13, z = 5$.

- Since the fundamental domains are defined by a point and its closest neighbours, the fundamental domains do not contain any lattice point in their interior. Thus the fundamental domains cover the $p \times p$ board without overlap. Since for each of the p points there is exactly one fundamental domain, its area is $\frac{p^2}{p} = p$, so that the length of a side is $\sqrt{p}$. An alternative argument here could be the one by Grace [17].
- Finally, an application of Pythagoras' theorem to the grid decomposition of the base side of the square shows that $p = (\sqrt{p})^2 = a^2 + b^2$ holds.

It seems particularly pleasant that we did not explicitly need the solution of $a^2 \equiv -1 \bmod p$, but could rather directly conclude from $a \equiv -a^{-1} \bmod p$ that the parallelogram is a square.

1.6.2 A short version of the proof

Having said all this, the reader can see that the following one sentence version of the proof, written in the spirit of Zagier's proof [41], contains essentially all the necessary information and is perhaps easier to work with, or memorize, than other proofs of this theorem. The amount of hidden routine checking may be comparable with that in Heath-Brown's or Zagier's version.

The involution on the finite set $S = \{2 \le a \le \frac{p-1}{2}\}$ defined by

$$a \mapsto \begin{cases} a^{-1} \bmod p & \text{if } 2 \le (a^{-1} \bmod p) \le \frac{p-1}{2}, \\ -a^{-1} \bmod p & \text{otherwise,} \end{cases}$$

has at least one fixed point z, so the fundamental domain of the lattice defined by

$$L_z = \{(x,\ zx \bmod p), 0 \le x < p\}$$

is a square with area p, so that the two squares theorem follows by an application of Pythagoras' theorem. □

2 How Zagier's involution can be motivated

We will give two explanations, how Zagier's map can be motivated. One was found by the present author, and was described in [12, 13, 14]. We will show that this approach gives a method to search systematically for proofs of related theorems on quadratic forms.

An alternative motivation can be found in lecture notes by E.W. Dijkstra.

2.1 First motivaton

It is possible to *construct* the "complicated" involution by means of some fairly easy assumptions, (see also [12]). These assumptions ensure that the final mapping would be as simple as possible.

If we look for a mapping that

I) can be described by a matrix $B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, with integer entries which are independent of k, (linearity),

II) maps the solutions (in positive integers) of $p = 4k+1 = x^2+4yz$ onto such solutions, (invariance),
III) has the easiest solution, namely $(1,1,k)$ as its only fixed point, (simplicity),

then we are uniquely led to $B = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

This can be seen as follows: Property (III) gives

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ k \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 \\ 1 \\ k \end{pmatrix}.$$

In particular $a+b+ck = 1$. But since the coefficients are supposed to be independent of k, we have that

$$\begin{aligned} a+b+ck = 1 &\Rightarrow c = 0,\ a+b = 1 \\ d+e+fk = 1 &\Rightarrow f = 0,\ d+e = 1 \\ g+h+ik = k &\Rightarrow i = 1,\ \ g+h = 0. \end{aligned}$$

Property (II) gives

$$(x')^2 + 4y'z' = (ax+by+cz)^2 + 4(dx+ey+fz)(gx+hy+iz) \stackrel{!}{=} x^2+4yz.$$

Hence, a comparison of the coefficients shows that

$$\begin{array}{ll} x^2 & a^2+4dg = 1 \\ xy & 2ab+4(dh+eg) = 0 \\ \vdots & \vdots \\ yz & \text{with } c = f = 0, i = 1 : ei = 1 \Rightarrow e = 1, \Rightarrow d = 0. \end{array}$$

Now $a^2+4dg = 1$ is simplified to $a^2 = 1$. Suppose that $a = 1$. Then $b = 0$ and from $2ab+4(dh+eg) = 0$ we get $g = 0$ and finally $h = 0$. Then, the matrix would be the identity matrix I. This is not what we want, since the map shall have only one fixed point.

Thus, $a = -1$, and so $b = 2, g = 1$, and finally $h = -1$. So, we have found the matrix B.

Surprisingly, we did not even need that our map shall be an involution but we can readily check that $B^2 = I$.

This only works for $-x+2y > 0$ and $x-y+z > 0$. For the other cases one apparently needs a different matrix. Let us see how we can manipulate B to yield a corresponding row condition $x-2y > 0$. We look for a matrix X which turns the row conditions of B into $(1,-2,0)$ and $(1,-1,1)$. Let $X = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. The matrices $A = BX$ and $C = XB$ cover all cases. Let $A = BX = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix}$ and

$C = XB = \begin{pmatrix} 1 & -2 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. We see that the row conditions perfectly fit to each other and induce a partition of all solutions.

Alternatively, one can find these matrices A and C by choosing small primes $(p = 13, 17, 29)$ and observing that here the sets of solutions with $-x+2y<0$ or $x-y+z<0$ only have one or two elements. For $p=13$, we find that $(1,3,1)$ must be mapped to $(3,1,1)$ and vice versa. For $p=17$, we find that $(1,4,1)$ must be mapped to $(3,1,2)$ and vice versa. For $p=29$ there are two possibilities. One excludes by the partially known mapping that $(1,7,1)$ is mapped to $(5,1,1)$ and finds that $(1,7,1)$ is mapped to $(3,1,5)$, from which A and C uniquely follow.

Even though we did not know about the partition of S into three sets we have found the map $\alpha : S \to S$ with

$$\alpha = \begin{cases} \alpha_1 & \text{decribed by matrix A, if } -x+y-z>0 \\ \alpha_2 & \text{described by matrix B, if } -x+2y>0 \text{ and } x-y+z>0 \\ \alpha_3 & \text{decribed by matrix C, if } x-2y>0 \text{ and } x-y+z>0. \end{cases}$$

This is precisely the mapping given by Zagier. Of course, α as a whole is not a linear map, so that property (I) is not strictly satisfied. We obtain in this way the easiest involution, α, with the required property, namely that we know the set of fixed points.

Let us remark that the intersection into three subsets was caused since we work with positive x,y,z. In Heath-Brown's version negative values are allowed, and so he did not need this division into three cases.

Zagier's second mapping, β (say), with $\beta : S \to S$ and $(x,y,z) \mapsto (x,z,y)$ corresponds to the matrix

$$Y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

2.2 Making the proof constructive

In his paper, Zagier mentioned the proof only shows the existence of the solution. Combining the two involutions α and β, we can give a constructive proof. Starting with the only fixed point of α, and iterating $\beta, \alpha ...$ we must arrive at a period.

$$(1,1,k) \xrightarrow{\beta} (1,k,1) \xrightarrow{\alpha} (3,1,k-2) \xrightarrow{\beta} \cdots \xrightarrow{\beta} (3,1,k-2) \xrightarrow{\alpha} (1,k,1) \xrightarrow{\beta} (1,1,k).$$

Since the maps are bijective, there is no pre-period. So, we eventually come back to $(1,1,k)$ with β. The number of elements in the period is even. By symmetry, there must be another fixed point in the middle of the cycle. Since there is only one

fixed point of α, this iteration constructs a fixed point of β, that is a solution of $p = x^2 + 4y^2$.

Applying this algorithm to a composite non-square integer $n = 4k+1$ the very same argument shows that any cycle containing $(1,1,k)$ must also contain another fixed point. Since n is no longer prime we may well come to another fixed point of α which corresponds to a factorization of n. To see that this can happen, let us concentrate on products of two distinct primes $n = p_1 p_2$ with $p_1 \equiv p_2 \equiv 3 \bmod 4$. Here β does not have a fixed point, since n cannot be written as a sum of two squares. Hence, in this case the iteration $\beta, \alpha, \beta \ldots$ must eventually come to another fixed point of α which corresponds to $x = y$, i.e. a factorization of n.

This algorithm for finding the decomposition into 2 squares is very slow. For some details see Bagchi [4]. Shiu [32] describes how one can accelerate this algorithm. It turns out to have an interpretation in the theory of continued fractions. A fast algorithm is described by Wagon [37].

2.3 A motivation due to Dijkstra

A different, and very elegant derivation of Zagier's map was also given by Dijkstra [9]. His notes are written in the language of a computer scientist and are extraordinary detailed. I will try keeping the flavour of his exposition, but will have to shorten his account. After some general remarks on involutions Dijkstra concludes that to write p as a sum of two integer squares it is enough to look at

$$(x,y) : x^2 + 4y^2 = p. \qquad (*)$$

In order to establish the desired correspondence between solutions of this equation and the fixed points of an involution "we do something with which every computer scientist is very familiar: replacing in a target relation" $(*)$ "something by a fresh variable". Dijkstra refers to "Leibniz' principle" (informally: substituting equals for equals) to rewrite $(*)$ as

$$(x,y,z) : x^2 + 4yz = p \text{ and } y = z.$$

Let $S = \{(x,y,z) : x,y,z \in \mathbb{N} : x^2 + 4yz = p\}$. Exploiting the symmetry in y and z, Dijkstra chooses a first involution inv_0 by $S \to S : (x,y,z) \mapsto (x,z,y)$. The fixed points of inv_0 satisfy $y = z$. Hence it is enough to show that inv_0 has at least one fixed point. In order to do this one intends to construct a second involution inv_1 on S, which has exactly one fixed point.

Next, Dijkstra gathers some elementary facts:
$x > 0, y > 0, z > 0, x \neq \pm(y-z)$, since p is odd and not a square.

Next, "can we think of operators on (x,y,z) for which $x^2 + 4yz = p$ is an invariant", i.e. an operator which maps solutions of S onto such solutions?

Dijkstra then studies operators of the type

$$(x,y,z) \mapsto (x+\Delta x, y+\Delta y, z+\Delta z).$$

Here Dijkstra implicitly assumes that Δ is an operator, for which

$$\Delta f(x) = f(x+\Delta x) - f(x)$$

so that for example $\Delta(x^2) = (x+\Delta x)^2 - x^2 = 2x\Delta x + (\Delta x)^2$.

Since $\Delta x = 0$ would too easily lead back to inv_0, he assumes $\Delta x \neq 0$. Since for all elements of S, x is odd, Δx is even, so that $\Delta x = 2b$, say.

The invariance assumption $\Delta : S \to S$, i.e. $(x')^2 + 4y'z' = p$ means that

$$\Delta(x^2 + 4yz) = 0.$$

So,

$$\begin{gathered}\Delta(x^2+4yz) = 0\\ \Delta(x^2) = -4\Delta(yz)\\ 2x(\Delta x) + (\Delta x)^2 = -4((y+\Delta y)(z+\Delta z) - yz)\\ b(x+b) = -y\Delta z - z\Delta y - \Delta y\Delta z.\end{gathered}$$

In order to simplify this expression Dijkstra chooses $\Delta y = 0$ and arrives at

$$b(x+b) = -y\Delta z.$$

He remarks that this choice does not restrict the generality, since one could arrive at any "move" with $\Delta y \neq 0, \Delta z \neq 0$ by means of two single moves.

Now, the last equation suggests the following 4 possibilities:

1. $b = -y, x+b = \Delta z$, giving $(x,y,z) \mapsto (x-2y, y, z+x-y)$
2. $b = y, x+b = -\Delta z$, giving $(x,y,z) \mapsto (x+2y, y, z-x-y)$
3. $b = \Delta z, x+b = -y$, giving $(x,y,z) \mapsto (-x-2y, y, z-x-y)$
4. $b = -\Delta z, x+b = y$, giving $(x,y,z) \mapsto (2y-x, y, z+x-y)$

In order to satisfy the invariance of $x > 0, y > 0, z > 0$, one sees that the third case above with $x' = -x-2y$ can be discarded from consideration.

So far, we have not yet used the fact inv_1 is supposed to have exactly one fixed point. Now, for a fixed point $(x,y,z) = (x',y',z')$. Here $x = x'$ and $y > 0$ mean that the only remaining case is the 4th case above. Here $x = 2y - x$ shows that a fixed point can only occur if $x = y$ so that $p = x^2 + 4yz = x(x+4z)$ implies that $z = \frac{p-1}{4}$, giving the unique fixed point $(1, 1, \frac{p-1}{4})$.

Dijkstra then completes the construction of the involution inv_1 for those solutions for which $y > z + x$ or $x > 2y$, respectively.

2.4 Comparison

Comparing both constructions in sections 2.1 and 2.3, it can be observed that the principle to keep the construction as simple as possible, but also as general as nec-

essary is quite successful. While in my motivation in section 2.1 the choice of the fixed point $(1,1,k)$ quickly led to the entries $c = f = 0, i = 1$ of the matrix B, and then the invariance of the quadratic form delivered the additional entries. Dijkstra's choice of $\Delta y = 0$, in the language of section 2.1 quickly led to $d = f = 0, e = 1$, and then the invariance of the form and consideration of the fixed point completed the entries of B.

Let us finally ask: is there any application (other than the two squares theorem itself) of the fact discovered by this combinatorial proof that the number of solutions (x,y,z) of a given type (say for $p = x^2 + 4yz$ with $x < y - z$) equals the number of solutions of another type (say here $x > 2y$)? If so, that could be of interest also for the generalizations considered below.

3 Generalization of the method

One can ask for similar involutions α for related question on $p = sx^2 + tyz$, where s and t are fixed constants. For example it is well known that for a prime p the following holds

$$p \equiv 1,3 \bmod 8 \Leftrightarrow p = x^2 + 2y^2 \text{ in positive integers.}$$

It would be nice to have an easy proof of this theorem by the idea of the Heath-Brown—Zagier proof.

Such generalizations were found by the current author in 1996, see [13], and also by Jackson [21, 22, 23] and Generalov [16].

Here we shall derive the following results:

Theorem 3. *Let p denote a prime.*

a) For $p = 8k+3$ there is a solution of $p = x^2 + 2y^2$ in positive integers.
b) For $p = 8k+7$ there is a solution of $p = x^2 - 2y^2$ in positive integers.
c) For $p = 8k+5$ there is a representation as $p = x^2 + y^2$. (A new proof!)

Theorem 4. *Let p denote a prime.*

a) For $p = 12k+7$ there is a solution of $p = 3x^2 + 4y^2$ in positive integers.
b) For $p = 12k+11$ there is a solution of $p = 3x^2 - 4y^2$ in positive integers.

Generalizing the approach of section 2.1 one can prove that the matrix $B = \begin{pmatrix} -1 & 2\frac{m}{n} & 0 \\ 0 & 1 & 0 \\ 4\frac{sm}{tn} & -4\frac{sm^2}{tn^2} & 1 \end{pmatrix}$ maps solutions of $p = sx^2 + tyz$ to such solutions and has the fixed point (m,n,k'). Here m,n,s, and t are fixed non-negative integers. So $k' = \dfrac{p - sm^2}{tn}$. We note that again $B^2 = I$. Unfortunately, in the general case the boundaries induced by the rows, namely $-x + 2\frac{m}{n} > 0$ and $4\frac{sm}{tn}x - 4\frac{sm^2}{tn^2}y + z > 0$, do not induce such a balanced three-partition of the set of solutions.

However, it is possible to construct mappings for $p = x^2 + 2y^2$ and $p = 3x^2 + 4y^2$ consisting of even more matrices. As before, these matrices are generated by B and X.

Note, even though the occurring matrices will be more complicated, the idea of the proof is still the same. The justification of the properties of the map α can -in principle- be left to an automatic system since it requires elementary calculations only.

As before, we try, if $A = BX$ can be useful. As above we use $X = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

At this point, we do not worry about the boundaries or a partition of the set of all solutions.

Geometrically, we can expect that $|\det A| = 1$, since we should not map bijectively a large region to a small one and vice versa.

Consider the eigenvalues of

$$\begin{aligned} A = BX &= \begin{pmatrix} -1 & 2\frac{m}{n} & 0 \\ 0 & 1 & 0 \\ 4\frac{sm}{tn} & -4\frac{sm^2}{tn^2} & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 2\frac{m}{n} \\ 0 & 0 & 1 \\ -4\frac{sm}{tn} & 1 & -4\frac{sm^2}{tn^2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & a \\ 0 & 0 & 1 \\ -c & 1 & -d \end{pmatrix}, \text{ say.} \end{aligned}$$

Noting that $ac = 2d$ we find

$$\begin{aligned} 0 &= (1-\lambda)(0-\lambda)(-d-\lambda) - (1-\lambda) - (-c)(0-\lambda)a \\ &= (\lambda+1)(\lambda^2 + (d-2)\lambda + 1). \end{aligned}$$

We find that $\lambda_1 = -1$ and $\lambda_{2,3} = -\frac{d-2}{2} \pm \sqrt{\left(\frac{d-2}{2}\right)^2 - 1}$.

For integers $d \geq 5$ or $d \leq -1$, the values of $\lambda_{2,3}$ are real but irrational numbers. So the order of A is infinite, and there is little hope of finding a suitable map consisting of finitely many parts. So $d = 0, 1, 2, 3, 4$ and in these cases $|\lambda_1| = |\lambda_2| = |\lambda_3| = 1$. This justifies our expectation that $\det A = 1$.

Recall that $d = \dfrac{4sm^2}{tn^2}$. Since we want to represent primes with $p = sx^2 + tyz = sm^2 + tnz$ we may assume that $\gcd(sm, tn) = 1$. We shall systematically consider all cases.

3.1 $d = 0$

For $d = 0$ we have $sm = 0$, so that $p = tnz$. This case is of no interest.

3.2 $d = 1$

Here $d = \dfrac{4sm^2}{tn^2} = 1$, and $(s,t) = (s,n) = (t,m) = (m,n) = 1$. Hence there are two possibilities:

- $s = m = n = 1, t = 4$. This is precisely the case of Heath-Brown's and Zagier's proof.
- $s = m = t = 1, n = 2$.
  In $p = x^2 + yz$ the solution with $p = x^2 + y^2 = y^2 + x^2$ is counted twice. In order to make the original argument work we need to break the symmetry. This can be done by assuming y and z to be even.
  The involution α is generated by

$$B = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix},$$

  $A = BX$, and $C = A^{-1}$.
  This gives the following variant of the proof of the two squares theorem:
  The involution on the finite set $S = \{(x,y,z) \in \mathbb{N} \times 2\mathbb{N} \times 2\mathbb{N} : x^2 + yz = p\}$ defined by

$$(x,y,z) \mapsto \begin{cases} (x+z, z, -2x+y-z) & \text{if } 2x+z < y \\ (-x+y, y, 2x-y+z) & \text{if } x < y < 2x+z \\ (x-y, 2x-y+z, y) & \text{if } y < x \end{cases}$$

  has exactly one fixed point, so $|S|$ is odd and the involution defined by $(x,y,z) \to (x,z,y)$ also has a fixed point.

3.3 $d = 2$

3.3.1 The case $p = x^2 + 2yz$

Here we consider the case $d = \dfrac{4sm^2}{tn^2} = 2$. By the coprime condition $(sm^2, tn^2) = 1$ we necessarily have that $s = m = n = 1, t = 2$.

Empirically one observes that the number of fixed points varies with the residue classes modulo 8:
a) primes $p \equiv 3 \bmod 8$ induce 1 fixed point,
b) primes $p \equiv 7 \bmod 8$ induce 2 fixed points,
c) primes $p \equiv 5 \bmod 8$ induce 2 fixed points, and
d) primes $p \equiv 1 \bmod 8$ induce 3 fixed points.

Case a) was also proved by Jackson [21] and Generalov [16]. They also observed d), but did not prove it by elementary methods. We shall prove a), b) and c) which

corresponds to our Theorem 3 a,b,c). Unfortunately we do not see either a convenient way to prove d) without appealing to the theory of quadratic forms.

Let $S = \{(x,y,z) \in \mathbb{N}^3 : x^2+2yz = p\}$. The one sentence proof is as before with the following map $\alpha : S \to S$.

$$\alpha = \begin{cases} A = BX & = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ -2 & 1 & -2 \end{pmatrix} \text{ if } -2x+y-2z > 0 \\ E = -XA^2 & = \begin{pmatrix} -3 & 2 & -2 \\ -2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix} \begin{cases} \text{if } -3x+2y-2z > 0 \\ \text{and } 2x-y+2z > 0 \\ \text{(then } -2x+2y-z > 0 \text{ is implied.)} \end{cases} \\ D = -A^2 & = \begin{pmatrix} 3 & -2 & 2 \\ 2 & -1 & 2 \\ -2 & 2 & -1 \end{pmatrix} \begin{cases} \text{if } 3x-2y+2z > 0 \\ \text{and } -2x+2y-z > 0 \\ \text{(then } 2x-y+2z > 0 \text{ is implied.)} \end{cases} \\ B = XA^3 & = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \\ 2 & -2 & 1 \end{pmatrix} \text{ if } -x+2y > 0 \text{ and } 2x-2y+z > 0 \\ C = A^{-1} = A^3 = XB & = \begin{pmatrix} 1 & -2 & 0 \\ 2 & -2 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{array}{l} \text{if } x-2y > 0, \\ (2x-2y+z > 0 \text{ follows trivially.}) \end{array} \end{cases}$$

Note that this map makes use of all matrices of the form $(-1)^{j+1}A^j, (j = 1,2,3)$, and $(-1)^{j+1}XA^j, (j = 2, j = 3)$. Note that also $A^4 = I$. The matrix XA is of no use, since this contains an impossible row condition $-x-2z > 0$.

The map above is equivalent to that given by Jackson and Generalov, here in Jackson's notation [21]:

$$(x,y,z) \mapsto \begin{cases} (x-2y, z+2x-2y, y) & \text{if } \quad y < \frac{x}{2} \\ (2y-x, y, 2x-2y+z) & \text{if } \frac{x}{2} \quad < y < x+\frac{z}{2} \\ (3x-2y+2z, 2x-y+2z, -2x+2y-z) & \text{if } x+\frac{z}{2} \quad < y < \frac{3}{2}x+z \\ (-3x+2y-2z, -2x+2y-z, 2x-y+2z) & \text{if } \frac{3}{2}x+z \quad < y < 2x+2z \\ (x+2z, z, -2x+y-2z) & \text{if } 2x+2z < y. \end{cases}$$

Let us call the subsets of S that correspond to the matrices A,B,C,D,E by $\mathscr{A},\mathscr{B},\mathscr{C},\mathscr{D},\mathscr{E}$. For a complete proof we have to show that

1. $\alpha : S \to S$, i.e. α maps (x,y,z) with $p = x^2+2yz$ to (x',y',z') with $p = x'^2+2y'z'$,
2. $\alpha^2 = id$,
3. the boundaries ($x-2y = 0, 2x-2y+z = 0$ etc.) are never attained,
4. the sets $\mathscr{A},\mathscr{B},\mathscr{C},\mathscr{D},\mathscr{E}$ induce a partition of the set of all solutions,
5. there is only one fixed point.

3.3.2 Proof of theorem 3

1. Since all parts of the mapping α are generated by $-I, X$ and B, it suffices to prove the first property for $-I, X$ and B. It is obvious for $-I$ and X. For B we have:
$$(x')^2 + 2y'z' = (-x+2y)^2 + 2(y)(2x-2y+z) = x^2 + 2yz = p.$$
2. Note that A maps the region $\mathscr{A}$ to the region $\mathscr{C}$. Because of $C = A^{-1}$ the region $\mathscr{C}$ is mapped to the region $\mathscr{A}$. The first assertion follows from $x' - 2y' = (x+2z) - 2z > 0$ and $2x' - 2y' + z' = 2(x+2z) - 2z + (-2x+y-2z) = y > 0$. For the second assertion we need that $-2x' + y' - 2z' > 0$ with $x' = x - 2y, y = 2x - 2y + z, z' = y$ and so $-2x' + y' - 2z' = z > 0$. Note that $B^2 = D^2 = E^2 = I$. So the matrix B maps the set $\mathscr{B}$ onto $\mathscr{B}$. The same holds for $D : \mathscr{D} \to \mathscr{D}$ and $E : \mathscr{E} \to \mathscr{E}$.
3. Suppose the boundaries are attained. This will lead to a contradiction.
   a. For the boundaries in the first row, namely $x - 2y = 0, -x + 2y = 0, 3x - 2y + 2z = 0, -3x + 2y - 2z = 0$, it would follow that x is even. This contradicts $p = x^2 + 2yz$, since p is odd.
   b. $-2x + y - 2z = 0$: $p = x^2 + 2yz = x^2 + 2(2x+2z)z = (x+2z)^2$. This contradicts the primality of p.
   c. $2x - 2y + z = 0$: $p = x^2 + 2yz = x^2 + 2y(2y - 2x) = (x-2y)^2$, contradicting the primality of p.
   d. $2x - y + 2z = 0$: See (b).
   e. $-2x + 2y - z = 0$: See (c).
4. It follows easily from the boundaries in Jackson's notation (given above) that α induces a partition of S.
5. We now look for the fixed points of α. Here we distinguish between the various cases depending on the residue class modulo 8. We see that A and C cannot have fixed points, since the set $\mathscr{A}$ is mapped onto $\mathscr{C}$ and the other way around.
   Suppose that (x,y,z) is a fixed point of B, then
$$B \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x+2y \\ y \\ 2x-2y+z \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$
   Hence $x = y$. Because of $p = x^2 + 2yz$ this is only possible for $x = y = 1$. Hence B has precisely one fixed point.
   For the matrix D we find that
$$D \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x-2y+2z \\ 2x-y+2z \\ -2x+2y-z \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$
   Hence $y = x + z$, and therefore, $p = x^2 + 2yz = x^2 + 2(x+z)z = (x+z)^2 + z^2$.
   Similarly,

$$E\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}-3x+2y-2z\\-2x+2y-z\\2x-y+2z\end{pmatrix}\stackrel{!}{=}\begin{pmatrix}x\\y\\z\end{pmatrix}.$$

Hence $y=2x+z$, which implies that $p=x^2+2yz=x^2+2(2x+z)z=(x+2z)^2-2z^2$.

If $p\equiv 3 \bmod 8$, there is no fixed point coming from D and E. To see this recall that squares modulo 8 only take the values $0,1,4$. So, the only fixed point is in $\mathscr{B}$, and so $|S|$ is odd. As before, the involution β must have an odd number of fixed points. Hence there is at least one fixed point with $y=z$, leading to the solution of $p=8k+3=x^2+2y^2$.

The same consideration of the values of squares modulo 8 shows:

If $p\equiv 7 \bmod 8$, we have again the trivial fixed point of B. There cannot be a fixed point from D. Since there cannot be a representation $p=x^2+2y^2$, we see that there must be a fixed point coming from E. So, $p\equiv 7 \bmod 8$ can be written as $p=x^2-2y^2$. This proves theorem 3b).

If $p\equiv 5 \bmod 8$, there cannot be a fixed point of E. Since $p=x^2+2y^2$ is impossible, there must be a fixed point of D, hence p has a representation of the form x^2+y^2. This gives a new proof for one half of the two squares theorem, here Theorem 3c).

If $p\equiv 1 \bmod 8$, we have a fixed point of B and (by the two squares theorem) of D. In order to prove the existence of the representation $p=x^2+2y^2$ it is enough to prove that there is (precisely) one fixed point of E. We do not see how to prove this with the methods of this paper. For this reason we did not state a theorem for the case $p\equiv 1 \bmod 8$.

3.4 $d=3$

3.4.1 The case $p=3x^2+4y^2$

Here we deal with the case $d=\dfrac{4sm^2}{tn^2}=3$. We have again two sub-cases.

- $s=3, m=n=1, t=4$.
- $s=3, m=t=1, n=2$, with even y and z.

As above in the case $d=1$, both of these sub-cases are equivalent. We will thus concentrate on the first case.

The form $p=3x^2+4yz$ represents only primes $p\equiv 3 \bmod 4$, hence we consider $p=12k+7$ and $p=12k+11$. We will proceed as in the case $d=2$.

The general form of our matrix B is now

$$B=\begin{pmatrix}-1&2&0\\0&1&0\\3&-3&1\end{pmatrix},A=BX=\begin{pmatrix}1&0&2\\0&0&1\\-3&1&-3\end{pmatrix}.$$

In view of $A^6 = I$ we consider the 9 matrices

$$(-1)^{j+1}A^j, (j = 1,\dots,5) \text{ and } (-1)^{j+1}XA^j, (j = 2,\dots,5).$$

(As before, the matrix XA is of no use, in view of the row condition $-x-2z>0$.)

$$-A^2 = \begin{pmatrix} 5 & -2 & 4 \\ 3 & -1 & 3 \\ -6 & 3 & -4 \end{pmatrix}, A^3 = \begin{pmatrix} 7 & -4 & 4 \\ 6 & -3 & 4 \\ -6 & 4 & -3 \end{pmatrix}, -A^4 = \begin{pmatrix} 5 & -4 & 2 \\ 6 & -4 & 3 \\ -3 & 3 & -1 \end{pmatrix},$$

$$A^5 = A^{-1} = XB = \begin{pmatrix} 1 & -2 & 0 \\ 3 & -3 & 1 \\ 0 & 1 & 0 \end{pmatrix}, D = -XA^2 = \begin{pmatrix} -5 & 2 & -4 \\ -6 & 3 & -4 \\ 3 & -1 & 3 \end{pmatrix},$$

$$E = XA^3 = \begin{pmatrix} -7 & 4 & -4 \\ -6 & 4 & -3 \\ 6 & -3 & 4 \end{pmatrix}, F = -XA^4 = \begin{pmatrix} -5 & 4 & -2 \\ -3 & 3 & -1 \\ 6 & -4 & 3 \end{pmatrix}, B = XA^5.$$

The corresponding boundaries are induced by the matrices themselves: for example the matrix $\begin{pmatrix} -5 & 4 & -2 \\ -3 & 3 & -1 \\ 6 & -4 & 3 \end{pmatrix}$ corresponds to $-5x+4y-2z>0$, $-3x+3y-z>0, 6x-4y+3z>0$.

Hence the map α is:

$$(x,y,z) \to \begin{cases} (x-2y, 3x-3y+z, y) & \text{if } y < \frac{x}{2} \\ (-x+2y, y, 3x-3y+z) & \text{if } \frac{x}{2} < y < x+\frac{z}{3} \\ (5x-4y+2z, 6x-4y+3z, -3x+3y-z) & \text{if } x+\frac{z}{3} < y < \frac{5}{4}x+\frac{z}{2} \\ (-5x+4y-2z, -3x+3y-z, 6x-4y+3z) & \text{if } \frac{5}{4}x+\frac{z}{2} < y < \frac{3}{2}x+\frac{3}{4}z \\ (7x-4y+4z, 6x-3y+4z, -6x+4y-3z) & \text{if } \frac{3}{2}x+\frac{3}{4}z < y < \frac{7}{4}x+z \\ (-7x+4y-4z, -6x+4y-3z, 6x-3y+4z) & \text{if } \frac{7}{4}x+z < y < 2x+\frac{4}{3}z \\ (5x-2y+4z, 3x-y+3z, -6x+3y-4z) & \text{if } 2x+\frac{4}{3}z < y < \frac{5}{2}x+2z \\ (-5x+2y-4z, -6x+3y-4z, 3x-y+3z) & \text{if } \frac{5}{2}x+2z < y < 3x+3z \\ (x+2z, z, -3x+y-3z) & \text{if } 3x+3z < y. \end{cases}$$

In order to prove theorem 4, we shall show: For primes $p \equiv 7 \bmod 12$ there is one fixed point of α. For primes $p \equiv 11 \bmod 12$ there are two fixed points of α.

3.4.2 Proof of theorem 4

Suppose that the boundaries are attained. This will lead to a contradiction. Note that for odd primes $p = 3x^2+4yz$ the value of x is odd. This excludes the boundaries $-x+2y=0$, $5x-2y+4z=0$, $5x-4y+2z=0$ and $7x-4y+4z=0$. Since $p = 3x^2+4yz$ is prime ($p>3$), we can deduce that y and z are not divisible by 3. This excludes the boundaries $3x-3y+z=0$, $3x-y+3z=0$, $6x-4y+3z=0$, and $6x-3y+4z=0$.

Now let us look at the fixed points: The mappings $A, -A^2, -A^4, A^5$ cannot have any fixed points, (since A maps the region A onto the region A^5 etc.). The matrices B, A^3, D, E, F are involutions. So we have to check their fixed points.

- As before, B has precisely one fixed point: $(1,1,k' = \frac{p-3}{4})$.
- For A^3 the fixed point condition $A^3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ simplifies to: $3x-2y+2z=0$. This is a contradiction since x is odd.
- Similarly, for D, we need to look at $3x-y+2z=0$. Consider the equation $p = 3x^2+4yz = 3x^2+12xz+8z^2 = 3(x+2z)^2-4z^2$ modulo 3. With $z^2 \equiv 0,1 \bmod 3$ and for $p \equiv 7 \bmod 12$, we find that $1 = 2z^2 \bmod 3$, a contradiction.
- For E, the fixed point condition is $2x-y+z=0$. We look at $p = 3x^2+4yz = 3x^2+8xz+4z^2 = 4(x+z)^2-x^2 = (3x+2z)(x+2z)$, contradicting the primality of p.
- Finally, for F we have to look at $3x-2y+z=0$, and plug this into our ternary form $p = 3x^2+4yz = 3x^2-12xy+8y^2 = 3(x-2y)^2-4y^2$. Again, we consider this modulo 3: For $p = 12k+7$ and with $2y^2 = 1 \bmod 3$ we see that there cannot be a fixed point.

We find that for $p = 12k+7$ there is only the trivial fixed point of B, namely $(1,1,(p-3)/4)$. By the standard argument p can be written as $p = 3x^2+4y^2$.

Since $p = 12k+11$ cannot be written as $p = 3x^2+4y^2$, there must be a fixed point of D or F. Any such fixed point induces a representation of the type $p = 3x^2-4y^2$, (see the analysis of these cases above).

This proves theorem 4.

3.5 $d = 4$

$d = \dfrac{4sm^2}{tn^2} = 4$. Here necessarily $s = t = m = n = 1$, and therefore

$$B = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \\ 4 & -4 & 1 \end{pmatrix}.$$

This matrix generates an infinite partition. Since in the case $s = t = 1$ we do not expect anything new, we do not pursue this case further.

4 On infinite but incomplete mappings

One can also consider corresponding mappings induced by B and X for other values of d. We cannot expect that the number of required matrices is finite.

Consider $p = 3x^2 + 2yz = 24k + 5$. Generate the matrices with

$$B = \begin{pmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \\ 6 & -6 & 1 \end{pmatrix}.$$

Take

$$A = BX \text{ and } C = A^{-1} = XB.$$

$$A = BX, -A^2, A^3, -A^4, A^5 \text{ etc.}$$

$$C = XB, -C^2, C^3, -C^4, \text{ etc.}$$

$$B, -BC, BC^2, -BC^3, BC^4 \text{ etc.}$$

$$-XA^2, XA^3, -XA^4 \text{ etc.}$$

(Note: $-X$ and XA are again omitted.) The matrix A does not have a finite order. This can easily be seen by looking at the eigenvalues of A, namely $-1, -2-\sqrt{3}, -2+\sqrt{3}$.

Taking infinitely many of these matrices, we see: The "region" of each matrix becomes smaller and smaller.

For the powers of $-A$ the row conditions come arbitrarily close to:

$$(3+\sqrt{3})x - y + (2+\sqrt{3})z > 0$$

and

$$-(3+\sqrt{3})x + y - (2+\sqrt{3})z > 0.$$

There are similar row conditions for the other series of matrices. The series

$$C = XB, -C^2, C^3, -C^4, \text{ etc.}$$

corresponds to

$$B, -BC, BC^2, -BC^3, BC^4 \text{ etc.}$$

in that respect that the row conditions of the first and third row are the same and the condition of the first row is reversed. Similarly, the two series

$$A = BX, -A^2, A^3, -A^4, A^5 \text{ etc.}$$

and

$$-XA^2, XA^3, -XA^4 \text{ etc.}$$

have associated boundaries. This latter series tends to a row condition of

$$(3+\sqrt{3})x - (2+\sqrt{3})y + z > 0$$

and

$$-(3+\sqrt{3})x + (2+\sqrt{3})y - z > 0.$$

Unfortunately the two boundaries

$$(3+\sqrt{3})x-y+(2+\sqrt{3})z>0$$

and

$$(3+\sqrt{3})x-(2+\sqrt{3})y+z>0$$

do not correspond. Hence there is a gap in between the regions of these series.

One would need further matrices to close this gap in order to proceed.

Looking at the conditions $ax-by+cz>0$ and $ax-cy+bz>0$, we see how incidental the above described finite mappings are.

In the case studied by Zagier we have $a=b=c$. So there are no problems at all. In the case $p=x^2+2yz$ we had $2x-y+2z$. Here $a=c$ so we still do not clearly see, what the condition in the general case is.

In the case $p=3x^2+4yz$ we had $6x-3y+4z$ and $6x-4y+3z$. Here, we see the importance of the matrices $A^3=BXBXBX$ and $XA^3=XBXBXBX$ with both rows, $6,-3,4$ and $-6,4,-3$. These matrices are the "turning point", reversing the y and z coordinate. We have a complete cycle: $(-3,1,-3)\Rightarrow(3,-1,3)\Rightarrow(-6,3,-4)\Rightarrow(6,-3,4)\Rightarrow(-6,4,-3)\Rightarrow(6,-4,3)\Rightarrow(-3,3,-1)\Rightarrow(3,-3,1)$. These matrices can be discovered by a sub-matrix (omit the first row and column) of the form $\begin{pmatrix} a & -b \\ -b & a \end{pmatrix}$.

In the incomplete mapping above there are no such "turning points".

5 Acknowledgments

The author is grateful to B. Artmann and D. Spalt for introducing him to Zagier's proof and for the challenge to understand how the proof could have been found. Further thanks goes to A.M. Decaillot for clarifying a question on Lucas' work. Sections 2.1 and 2.2 were found in 1990, section 3 in 1996, and section 1.6 in 2001, see also [11], [12], [13], [14].

References

1. Aigner, M., Ziegler, G.M.: Proofs from THE BOOK, 2nd edition, Springer, Berlin, Heidelberg, 2001.
2. Aubry, A.: Les principes de la géométrie des quinconces, L'Enseignement Mathématique 13 (1911), 187–203.
3. Bachmann, P.: Niedere Zahlentheorie, reprint by Chelsea Publishing Co., New York, 1968, originally published 1902/1910.
4. Bagchi, B.: Fermat's two squares theorem revisited, Resonance 4 (7) (1999), 59–67.
5. Barbeau, E. J.: Polynomials. Problem Books in Mathematics. Springer-Verlag, New York, 1995.
6. Clarke, F.W., Everitt, W.N., Littlejohn, L.L., Vorster, S.J.R.: H. J. S. Smith and the Fermat Two Squares Theorem, The American Mathematical Monthly 106 (7) (1999), 652-665.

7. Décaillot A.-M.: Géométrie des tissus. Mosaïques. Échiquiers. Mathématiques curieuses et utiles. Revue d'histoire des mathématiques 8. Vol. 2. (2002), 145-206.
8. Dickson, L. E.: History of the theory of numbers, reprint by Chelsea Publishing Co., New York, 1966, originally published 1919-1923.
9. Dijkstra, E.W.: A derivation of a proof by D. Zagier, Manuscript EWD 1154, 1993, available at http://www.cs.utexas.edu/users/EWD/welcome.html
10. Edwards, H. M.: A genetic introduction to algebraic number theory. Springer-Verlag, New York, 1996.
11. Elsholtz, C.: Primzahlen der Form $4k+1$ sind Summe zweier Quadrate, contribution to "Bundeswettbewerb Jugend forscht" (German National Contest for Young Scientists), 1990/91, published in [12].
12. Elsholtz, C.: Primzahlen der Form $4k+1$ sind Summe zweier Quadrate, Mathematiklehren, no. 62, February 1994, pp. 58–61.
13. Elsholtz, C.: The Liouville—Heath-Brown—Zagier proof of the two squares theorem (Preprint 2001/10, Institut für Mathematik, TU Clausthal, Germany).
14. Elsholtz, C.: Kombinatorische Beweise des Zweiquadratesatzes, Mathematische Semesterberichte 50 (2003), 77-93.
15. Œuvres de Fermat, ed: Paul Tannery and Charles Henry, Paris: Gauthier-Villars et fils 1891-1912.
16. Generalov, A.I.: A combinatorial proof of Euler-Fermat's theorem on the representation of the primes $p = 8k+3$ by the quadratic form x^2+2y^2, Journal of Mathematical Sciences 140 (2007), 690–691.
17. Grace, J.H.: The four square theorem, J. London Math. Soc. 2 (1927), 3–8.
18. Hardy, G.H.: A Mathematician's Apology, Cambridge University Press, 1940.
19. Hardy, G.H.; Wright, E.M.: An introduction to the theory of numbers. 5th edition. Oxford University Press, New York, 1979.
20. Heath-Brown, D.R.: Fermat's two squares theorem, Invariant, 1984, 3–5. Available at http://eprints.maths.ox.ac.uk/677/
21. Jackson, T.: A Short Proof That Every Prime $p = 3 \pmod 8$ is of the Form x^2+2y^2, Amer. Math. Monthly, 107 (2000) p. 447.
22. Jackson, T.: Automorphs and involutions. Tatra Mt. Math. Publ. 20 (2000), 59–63.
23. Jackson, T.: Direct proofs of some of Euler's results. Number theory (Turku, 1999), 163–166, de Gruyter, Berlin, 2001.
24. Kraitchik, M.: Mathematical Recreations. Allen & Unwin, London, 1943.
25. Larson, L.C.: A theorem about primes proved on a chessboard. Mathematics Magazine 50 (2) (1977), 69–74,
26. Lucas, É.: Application de l'arithmétique à la construction de l'armure des satins réguliers. Paris, 1867. Available online at http://edouardlucas.free.fr/gb/index.html
27. Lucas, É.: Les principes fondamentaux de la géométrie des tissus, Congrès de l'Association française pour l'avancement des sciences 40 (1911), 72-87, (based on an article in L'Ingegnere Civile, Torino, 1880). Available at http://www.biodiversitylibrary.org/item/27373#5
28. McKay, James H.: Another proof of Cauchy's group theorem. Amer. Math. Monthly 66 (1959), 119.
29. Nathanson, M.B.: Elementary methods in number theory. Graduate Texts in Mathematics, 195. Springer-Verlag, New York, 2000.
30. Pólya, G.: Über die "doppelt-periodischen" Lösungen des *n*-Damen Problems, in: Ahrens, W. Mathematische Unterhaltungen und Spiele, Teubner, Leipzig, Volume II, 2nd edition 1918, 364–374.
31. Shirali, S.: On Fermat's Two-Square Theorem. Resonance 2, no.3 (1997), 69-73.
32. Shiu, P.: Involutions associated with sums of two squares. Publ. Inst. Math. (Beograd) (N.S.) 59(73) (1996), 18–30.
33. Tikhomirov, V.: Quantum, May/June 1994, pp. 5–7.
34. Uspensky, J. V.; Heaslet, M. A.: Elementary Number Theory. McGraw-Hill Book Company, New York and London, 1939.

35. Varouchas, I.: Une démonstration élémentaire du théorème des deux carrés, I.R.E.M. Bull. 6 (1984), 31-39.
36. Venkov, B.A.: Elementary Number Theory, Wolters-Noordhoff, Groningen, 1970, (originally published in Russian, 1937).
37. Wagon, S.: The Euclidean algorithm strikes again. Amer. Math. Monthly 97 (1990), 125–129.
38. Wells, D.: Are these the most beautiful? Math. Intelligencer 12 (1990), no. 3, 37–41.
39. Williams, K.S.: Heath-Brown's elementary proof of the Girard-Fermat theorem, Carleton Coordinates, (1985), 4-5.
40. Winter, H.: Der Zwei-Quadrate-Satz von Fermat - eine Studie zur Heuristik des Beweisens. Math. Semesterber. 50 (2003), 191-235.
41. Zagier, D.: A one-sentence proof that every prime $p \equiv 1 \pmod 4$ is a sum of two squares, Amer. Math. Monthly 97(2) (1990), 144.