QUADRATIC FORMS AND CONGRUENCE SUBGROUPS

GREG MCSHANE

ABSTRACT. The primes p represented by a quadratic fors x^2+y^2, x^2+2y^2, x^2+3y^2, is a subject that was first studied by Fermat. Heath-Brown and Zagier

1. INTRODUCTION

Consider the following pair of well known theorems from elementary number theory:

Theorem 1.1 (Fermat). *Let p be a prime then the equation*

$$x^2 + y^2 = p$$

has a solution in integers iff $p = 2$ or $p - 1$ is a multiple of 4.

These results are intimately linked and often one deduces the second as a corollary of the first, for example, by using unique factorisation in the Gaussian integers. We present a geometric approach to this results using the theory of group actions and in particular an application of Burnsides's Lemma. We also investigate this approach applied to the forms $x^2 + 2y^2$ and $x^2 + 3y^2$.

As in Zagier's remarkable proof [5] both results follow from showing that a certain involution has a fixed point. Amusingly Burnsides's Lemma reduces this to showing that another involution has exactly two fixed points:

- In the proof of Theorem 2.1 this is a consequence of the fact that a quadratic equation over a field has at most two solutions.
- In the proof of Theorem 1.1 this follows from some geometry and the fact that any odd integer can be written as the difference of two squares.

1.1. **Organisation, Remarks.** In Section 2 we recall the statement of Burnsides Lemma and apply it to a Klein four group generated by involutions of $\mathbb{F}_p^*$ yielding a proof of Theorem 2.1. In Section 3 we introduce the congruence subgroups and involutions involved. In Section 4 we study how the involutions act on arcs and give explicit computations. Finally in Sections 5 and 6 we prove the main technical result and show how it is applied to the quadratic forms x^2+y^2, x^2+2y^2 and $x^2 + 3y^2$.

1.1.1. *References.* The reader should not need any other references to understand this paper if they are already familiar with Burnsides Lemma. For background the book by Cox [1] is highly recommended.

1.1.2. *Burnside and signatures.* The astute reader will surely realise that Burnside is not essential to our argument and that one can achieve the same reduction by considering the signature of the permutations associated to the involutions we consider. In fact the first author set this as an undergraduate exam question some years ago.

1.1.3. *Bezouts Theorem.* In previous work [**?**] we were implicitly using Bezout's Theorem when we assert that

- $\mathrm{SL}(2,\mathbb{Z})$ is transitive on $\mathbb{Q}\cup\infty$ (which is equivalent to Bezout's Theorem.)
- $\Gamma(2)$ has exactly three orbits on $\mathbb{Q}\cup\infty$.

We will see in Section 4 how to determine the actions of the various involutions explicit using Bezout's Theorem.

1.2. **Thanks.** The first author thanks Louis Funar and the second author for many useful conversations over the years concerning this subject. He would also like to thank Xu Binbin for reading early drafts of the manuscript.

2. Burnside Lemma

Recall the following result from elementary number theory

Theorem 2.1. *Let p be a prime then the equation*

$$x^2 = -1$$

admits a solution in $\mathbb{F}_p$ iff $p = 2$ or $p-1$ is a multiple of 4.

We give a proof of this using the Burnside Lemma.

Recall that if G is a group acting on a finite set X then the Burnside Lemma says

$$|G||X/G| = \sum_g |X^g| \tag{1}$$

where, as usual, X^g denotes the set of fixed points of the element g and X/G the orbit space.

Let $p \neq 2$, $X = \mathbb{F}_p^*$ and G be the group generated by the two involutions

$$\begin{aligned} x &\mapsto -x \\ x &\mapsto 1/x. \end{aligned}$$

The group G has exactly four elements namely:

- the trivial element which has $p-1$ fixed points
- $x \mapsto -x$ which has no fixed points

- $x \mapsto 1/x$ has exactly two fixed points namely 1 and -1.
- $g : x \mapsto -1/x$ is the remaining element and the theorem is equivalent to the existence of a fixed point for it.

Note that since $\mathbb{F}_p$ is a field $|X^g| = \sharp\{x^2 = -1,\, x \in \mathbb{F}_p^*\}$ is either 0 or 2. Now for our choice of X and G equation (1) yields

$$4|X/G| = (p-1) + 2 + |X^g|. \tag{2}$$

The LHS is always divisible by 4 so the RHS is too and it follows from this that

$$|X^g|. = \begin{cases} 0 & (p-1) = 2 \mod 4 \\ 2 & (p-1) = 0 \mod 4 \end{cases}$$

This proves Theorem 2.1.

Note. As was noted in the introduction one can obtain the same conclusion by calculating the signature of $x \mapsto -1/x$ using the fact that it is the composition of $x \mapsto -x$ and $x \mapsto 1/x$.

3. Congruence subgroups

We denote by Γ the group of invertible matrices with integer coefficients $\mathrm{SL}(2,\mathbb{Z})$ and define $\Gamma(n)$ to be the kernel of the homomorpisms $\mathrm{SL}(2,\mathbb{Z}) \mapsto \mathrm{SL}(2,\mathbb{Z}/n\mathbb{Z})$. The latter group is a subgroup of $\Gamma_0(n)$

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}/n\mathbb{Z}) \,, c = 0 \mod p \right\}.$$

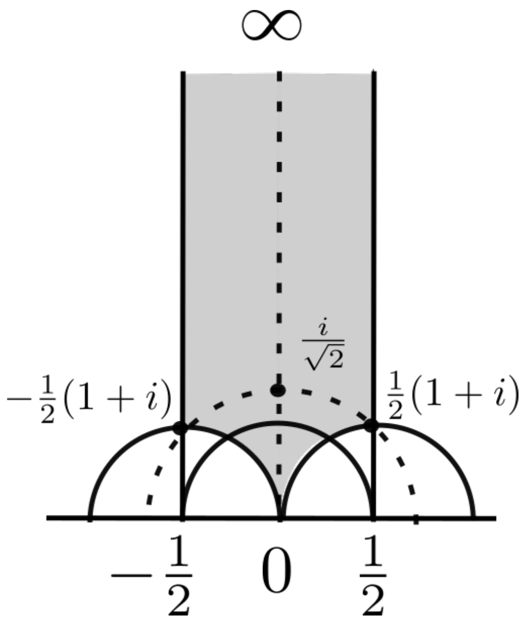

Figure 1. Standard fundamental domain for $\Gamma_0(2)$

It is well known that the involution

$$\tau_n = \begin{pmatrix} 0 & -1/\sqrt{n} \\ \sqrt{n} & 0 \end{pmatrix}$$

normalises $\Gamma_0(n)$.

For $n = 2, 3$ or 4 it is easy to find a fundamental domain for the group Γ^* generated by τ_n and $\Gamma_0(n)$. Let U denote the strip between the Poincare geodesics (vertical lines) joining ∞ to $\pm\frac{1}{2}$ and V_n the image of this strip under τ_n this is easy to compute and it is the region between the pair of Poincare geodesics joining 0 to $\pm\frac{2}{n}$. The region U is the standard fundamental domain

The intersection $U \cap V_n$ contains a fundamental domain for Γ^*. For $n = 2, 3, 4$ this intersection is finite area and is a fundamental domain for $\Gamma_0(n)$ with side pairings P and Q. The Poincare geodesic μ_n joining $-2/n$ to $2/n$ divides it into two regions each of which is a fundamental domain for Γ^*.

From covering theory, an isometry of $\mathbb{H}$ induces an automorphism of the surface $\mathbb{H}/G$ iff it normalises the covering group i.e. G. The fundamental domain for $\Gamma_0(n)$ is invariant under the pair of orientation reversing involutions

$$z \mapsto -\bar{z},\; z \mapsto \frac{1}{n\bar{z}}$$

and these maps preserve side pairings so normalise $\Gamma_0(n)$. One checks that their composition coincides with τ_n. Thus these three involutions together with the identity form a Klein four group that normalises $\Gamma_0(n)$.

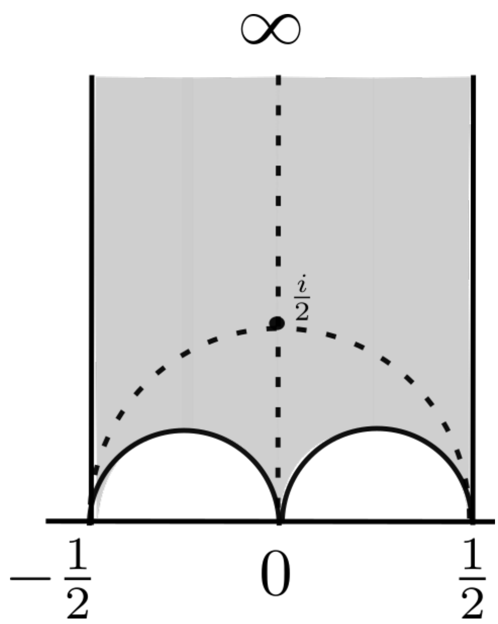

FIGURE 2. Standard fundamental domain for $\Gamma_0(4)$

4. Arcs and involutions

There is a natural map from the set of coprime integers into the rationals namely:

$$(p, q) \mapsto \frac{p}{q}.$$

This map can be extended to be a surjection onto the extended rationals $\mathbb{Q} \cup \infty$ by

$$(1, 0) \mapsto \infty.$$

The modular group Γ acts transitively on $\mathbb{Q} \cup \infty$ and the stabiliser of ∞ is generated by the parabolic $z \mapsto z + 1$. Since $\Gamma_0(n)$ is finite index in Γ the extended rationals is exactly the set of fixed points of the parabolic elements of $\Gamma_0(n)$.

We define an *arc* to be a Poincare geodesic joining distinct points of p/q, $p'/q' \in \mathbb{Q} \cup \infty$. To each such arc we associate an 2x2 matrix with integer coefficients:

$$\begin{pmatrix} p & p' \\ q & q' \end{pmatrix}.$$

We will be concerned with the action of our Klein four group on the set set of geodesics joining ∞ to a point k/q where q is some prime. The corresponding matrices have the form

$$\begin{pmatrix} 1 & k \\ 0 & q \end{pmatrix}.$$

We can compute the action of τ_n on 'the endpoints of an arc:

$$\infty \mapsto 0,\ k/q \mapsto -q/nk$$

so the correspondng arc has matrix

$$\begin{pmatrix} 0 & -q \\ 1 & nk \end{pmatrix}.$$

Since q and nk are coprime there exist integers a and b such that $bnk + aq = 1$. And so the matrix

$$\begin{pmatrix} a & -b \\ nk & q \end{pmatrix}$$

is in $\Gamma_0(n)$.

Now we have:

$$\begin{pmatrix} a & -b \\ nk & q \end{pmatrix} \begin{pmatrix} 0 & -q \\ -1 & nk \end{pmatrix} = \begin{pmatrix} -b & -aq - bnk \\ q & 0 \end{pmatrix} = \begin{pmatrix} -b & -1 \\ q & 0 \end{pmatrix}.$$

So the action of τ_n on the set of arcs can be computed using the euclidean algorithm to find b.

4.1. **Explicit computations for $n = 2$ and $p = 7, 11$.** We will compute the action of τ_2 on the set of arcs joining ∞ to a point $k/7$ or $k/11$. For $k = 1, 2, 3 \ldots p-1$ we denote k'/p the image of k/p under τ_2

So for $p = 7$ we have:

k	1	2	3	4	5	6
k'	-4	-2	1	-1	2	-3

So for $p = 11$ we have:

k	1	2	3	4	5	6	7	8	9	10
k'	-6	-3	-2	4	1	-1	-4	2	3	-5

5. Invariant arcs and solutions to $x^2 \pm ny^2 = p$

Theorem 5.1. *Let n be a positive integer and p a prime:*

(1)
$$x^2 + ny^2 = p$$
has a solution iff τ_q leaves one of the arcs joining ∞ to k/p invariant;

(2)
$$x^2 - ny^2 = p$$
has a solution iff the involution $z \mapsto 1/(n\bar{z})$ leaves one of the arcs ∞ to k/p invariant;

Proof. Suppose there is an arc joining a/b to a'/b' that is invariant under τ_n. Then $a' = -b$ and $b' = na$ since
$$\tau_n(a/b) = \frac{-1}{n(a/b)} = \frac{-b}{na}.$$
So the matrix corresponding to the arc is (up to permuting the columns)
$$\begin{pmatrix} a & -b \\ b & na \end{pmatrix}.$$
If this arc is equivalent by an element of $\Gamma_0(n)$ to an arc of the form $\infty \to k/p$ ie whose matrix is
$$\begin{pmatrix} 1 & k \\ 0 & p \end{pmatrix}.$$
for some k then the determinant of the associated matrix, that is $b^2 + na^2$, is p.

The enuation $x^2 + ny^2 = p$ is proven in a similar way. $\square$

6. Existence of solutions to $x^2 \pm ny^2 = p$

Let p be an odd prime.

6.1. **The case $x^2+4y^2=p$.** This is Fermat's result from the introduction. We will prove that the equation has a solution iff $p \equiv 1 \pmod 4$.

If $p = 4k+1$ then there is an (essentially unique) solution to $x^2 + 4y^2 = p$. The involution $z \mapsto 1/4\bar{z}$ actually leaves a pair of arcs invariant One can see this from the theorem and the fact that

$$x^2 - 4y^2 = (x-2y)(x+2y).$$

Since p if $p = x^2-4y^2$ then $x-2y=1$ and $x+2y=p$ so $x=(p+1)/2$ and $y=(p-1)/4$.

Now applying our argument with the Burnside Lemma we see that τ_4 leaves a pair of arcs invariant so there are solutions to $x^2+4y^2=p$.

6.2. **The case $x^2+2y^2=p$.** If p can be written as $8k+1$ or $8k+3$ then there is a solutions to $x^2+2y^2=p$.

The case $8k+3$ follows from the argument with Burnside again. The involution $z \mapsto 1/2\bar{z}$ leaves no arc invariant. By the theorem this is equivalent to saying that there is p cannot be written as x^2-2y^2 and this is easy to prove using congruences mod 8.

Reciprocally $p = 8k+7$ can be written as x^2-2y^2 as it cannot be written as x^2+2y^2.

Unfortunately the case $p=8k+1$ is not so easy. There seems to be no way to deduce this by using a parity argument although Generalov [2] does claim to have a proof using involutions.

References

[1] David A. Cox Primes of the Form x^2+ny^2: Fermat, Class Field Theory, and Complex Multiplication, Second Edition 2013 John Wiley Sons

[2] 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 . J Math Sci 140, 690–691 (2007).

[3] G. McShane, *Simple geodesics and a series constant over Teichmuller space* Invent. Math. (1998)

[4] R. C. Penner, *The decorated Teichmueller space of punctured surfaces*, Communications in Mathematical Physics 113 (1987), 299–339.

[5] D. Zagier, *A one-sentence proof that every prime p = 1 (mod 4) is a sum of two squares*, American Mathematical Monthly, 97 (2): 144

Institut Fourier 100 rue des maths, BP 74, 38402 St Martin d'Hères cedex, France
Email address: `mcshane at univ-grenoble-alpes.fr`