SUMS OF TWO SQUARES

PETE L. CLARK

1. Sums of two squares

We would now like to use the rather algebraic theory we have developed to solve a purely Diophantine problem: which integers n are the sum of two squares?

The relevance of the Gaussian integers is not hard to see: $N(a + bi) = a^2 + b^2$, so we are asking which integers are norms of Gaussian integers.

It turns out to be best to consider first the case of primes. Why would we think to do this? Let us consider the situation for small values of n:

$n = 1$: $1 = 0^2 + 1$;
$n = 2$ (prime): $2 = 1^2 + 1^2$;
$n = 3$ (prime) is not a sum of two squares.
$n = 4$: $4 = 2^2 + 0^2$.
$n = 5$ (prime): $5 = 2^2 + 1^2$.
$n = 6$ is not a sum of two squares.
$n = 7$ (prime) is not a sum of two squares.
$n = 8$: $8 = 2^2 + 2^2$.
$n = 9$: $9 = 3^2 + 0^2$.
$n = 10$: $10 = 3^2 + 1^2$.
$n = 11$ (prime) is not a sum of two squares.
$n = 12$ is not a sum of two squares.
$n = 13$ (prime): $13 = 3^2 + 2^2$.
$n = 14$ is not a sum of two squares.
$n = 15$ is not a sum of two squares.
$n = 16$: $16 = 4^2 + 0^2$.
$n = 17$ (prime): $17 = 4^2 + 1^2$.
$n = 18$: $18 = 3^2 + 3^2$.
$n = 19$ (prime) is not a sum of two squares.
$n = 20$: $20 = 4^2 + 2^2$.
$n = 21$ is not a sum of two squares.
$n = 22$ is not a sum of two squares.
$n = 23$ (prime) is not a sum of two squares.
$n = 24$ is not a sum of two squares.
$n = 25$: $25 = 5^2 + 0^2$.
$n = 26$: $26 = 5^2 + 1$.

We can continue our table, however the apparent pattern in the case of prime values of n is already easy to guess: 2 is a sum of two squares, any prime $p = 4k + 1$

is a sum of two squares, and any prime $p = 4k + 3$ is not a sum of two squares. The general situation seems more complicated, so let us first try to establish these facts.

Proposition 1. *If $n \equiv 3 \pmod 4$, then n is not a sum of two squares.*

Proof: Suppose $n = a^2 + b^2$; then reducing modulo 4 we would have $3 = a^2 + b^2$ in $\mathbb{Z}/4\mathbb{Z}$. In fact this is not possible: the squares in $\mathbb{Z}/4\mathbb{Z}$ are $0 = 0^2 = 2^2$ and $1 = 1^2 = 3^2$. Adding any pair of elements from $\{0, 1\}$, we cannot get 3.

The other part, that every prime $p = 4k + 1$ is a sum of two squares lies much deeper. To get at it, let us try to recast our previous statement in terms of the arithmetic of $\mathbb{Z}[i]$. Namely, suppose $p \equiv 3 \pmod 4$ is a prime number. We claim that p is also irreducible (and hence prime, since $\mathbb{Z}[i]$ is a PID) in $\mathbb{Z}[i]$: indeed, since $N(p) = p\overline{p} = p^2$, any nontrivial factorization would involve an element $a + bi$ with $N(a + bi) = a^2 + b^2 = p$, and we just showed that no such elements exist when p is 3 (mod 4).

Consider however that $5 = N(2 + i) = (2 + i)(2 - i)$. Thus 5 is not prime in $\mathbb{Z}[i]$; it **splits** into the nonassociate primes $2 + i$ and $2 - i$. This is the behavior we would like to generalize: that is, to show that if $p \equiv 1 \pmod 4$, then p splits as a product of two primes in $\mathbb{Z}[i]$.

Unfortunately it is still not clear how to show that a prime p of the form $4k + 1$ splits in this way. To our rescue comes Pierre de Fermat, who long ago discovered the following

Lemma 2. *(Fermat's Lemma) For a prime $p \equiv 1 \pmod 4$, there exists $x \in \mathbb{Z}$ such that $p \mid (x^2 + 1)$.*

In other words, Fermat asserts that at least p divides a sum of two squares of a very special form. Let us for the moment push aside the proof of this claim and see first what it does for us.

Well, what does $x^2 + 1$ have to do with the Gaussian integers? Although this is in general irreducible over $\mathbb{Z}$, it is always reducible over $\mathbb{Z}[i]$: $x^2 + 1 = (x + i)(x - i)$. Aha! Suppose p were irreducible as an element of $\mathbb{Z}[i]$. Then – since $\mathbb{Z}[i]$ is a PID – we may apply Euclid's Lemma to conclude that $p|(x + i)$ or $p|(x - i)$. But in fact this is absurd: in the quotient field $\mathbb{Q}[i]$ the element $(x + i)/p = \frac{x}{p} + \frac{1}{p}i$, and $\frac{1}{p}$ is certainly not an integer! Thus p is not irreducible, meaning that it factors nontrivially:

$$p = (a + bi)(c + di),$$

with $1 < N(a + bi), N(c + di)$. Since $N(p) = p^2$, this means that

$$p = N(a + bi) = (a + bi)(a - bi) = a^2 + b^2,$$

which is what we wanted to show. (It also follows that $c + di = a - bi$; a prime p which is irreducible must split into a product of two conjugate irreducible elements of $\mathbb{Z}[i]$.)

It remains to prove Fermat's Lemma. There are in fact many different proofs, although none is easy. From a modern algebraic perspective, it can be shown to

follow easily from the fact that the group of units $(\mathbb{Z}/p\mathbb{Z})^\times$ of the finite field $\mathbb{Z}/p\mathbb{Z}$ is cyclic. However, we have alluded to but not yet proven this result (at least, not in an understandable way!), so we take a different approach here. First:

Proposition 3. *(Wilson's Theorem) If p is a prime, then $p \mid (p-1)! + 1$.*

Proof: Equivalently, $(p-1)! \equiv -1 \pmod{p}$, or: if we multiply all the nonzero elements of the finite field $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$, we get -1. The statement is certainly true for $p = 2$ (indeed we have equality: $2 = (2-1)!+1$), so we may assume that p is an odd prime. Now consider the map $\iota : \mathbb{F}_p^\times \to \mathbb{F}_p^\times$ given by $x \mapsto x^{-1}$ (inversion). First note that (as for any group), this map is a bijection: every element has a unique inverse. Moreover, the "fixed points" of this map are those elements $x = \iota(x)$ which are equal to their own inverse: $x = x^{-1}$; this is equivalent to the equation $x^2 = 1$. There are two obvious solutions: namely $x = \pm 1$ (which are different since $p \neq 2$). On the other hand, these are all the solutions: the polynomial $x^2 - 1$ can have at most two roots in the field $\mathbb{F}_p$. This means that every element different from ± 1 in $\mathbb{F}_p^\times$ is distinct from its inverse, so when we multiply out the nonzero elements of $\mathbb{F}_p^\times$ we get $1 \cdot -1 \cdot (x_1 \cdot x_1^{-1}) \cdots (x_{\frac{p-3}{2}} \cdot x_{\frac{p-3}{2}}^{-1})$ – i.e., all the elements apart from ± 1 can be paired with their distinct and unique multiplicative inverse. That the product is -1 is now clear!

Remark: It turns out that, like Euclid's Lemma, prime numbers are characterized by satisfying Wilson's Theorem. In other words, if $n > 1$ is an integer such that $n \mid (n-1)! + 1$, then n is prime. (We will leave the proof as a harder – and optional – exercise.) This is worth remarking upon but not very useful as a primality test: computing $(n-1)!$ takes a long time!

Finally, we use Wilson's Theorem to prove Fermat's Lemma. We showed that for a prime p, the product of the first $p-1$ integers is -1 modulo p. But this can be immediately generalized as follows: suppose that S is any **reduced residue system** modulo p, i.e., a set of $p-1$ integers such that

$$\overline{S} = \{\overline{s} \mid s \in S\} = \mathbb{F}_p^\times,$$

i.e., a set that represents every nonzero residue class modulo p exactly once. Then the conclusion of Wilson's theorem must be true for S just as it is for the particular choice $S_0 = \{1, \ldots, p-1\}$: namely, $\prod_{i \in S} i \equiv -1 \pmod{p}$. Consider the set

$$S = \{\frac{-(p-1)}{2}, \frac{-(p-1)}{2} + 1, \ldots, -1, 1, \ldots \frac{(p-1)}{2};$$

in other words, S is the set of nonzero integers of absolute value at most $\frac{p-1}{2}$: note that this set has $p-1$ elements. Moreover, the difference between the least element and the greatest element is precisely $p-1 < p$, so that no two of them can differ by a multiple of p: they are too close together for that. Thus S is a reduced residue system modulo p, so that Wilson's theorem says that

$$-1 \equiv \prod_{i \in S} i \pmod{p}.$$

On the other hand, by "folding in half" – i.e., by multiplying each element in S with its partner of opposite sign – we directly compute that

$$\prod_{i \in S} i = (-1)^{\frac{p-1}{2}} \cdot x^2,$$

where x is (for the record) $(\frac{p-1}{2})!$. If now $p \equiv 1 \pmod 4$, then $\frac{p-1}{2}$ is even, so that $(-1)^{\frac{p-1}{2}} = 1$. Putting everything together we get that in this case $-1 \equiv x^2 \pmod p$, i.e., $p | x^2 + 1$. This proves Fermat's Lemma. Summing up, we have proved:

Theorem 4. *(Fermat's Two Squares Theorem) A prime p is a sum of two integer squares iff $p = 2$ or $p = 4k + 1$.*

2. A characterization of integers $n = a^2 + b^2$

But let us not forget that we treated the case of prime n only for simplicity. What about more general values of n? There is some further "mopping up" to do. The most important observation is the following:

Proposition 5. *For any integers a, b, c, d, we have:*

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

Proof: Of course all we really need to do to prove this result – or any polynomial identity – is just to check it directly by multiplying out both sides. However, this is clearly a suboptimal approach as it does not explain where the identity *comes from*. Having studied the Gaussian integers, the "right" explanation is easy to find: namely, since the norm map is multiplicative, we have

$$N(a + bi)N(c + di) = N((a + bi)(c + di)).$$

The use of this proposition is not in the particular form of the identity (although that it is interesting and it is natural to wonder about the existence of similar-looking identities in more variables) but rather in its immediate consequence that the set of positive integers n which are sums of two squares is closed under multiplication. (If you enjoy such terminology, it is a submonoid of $\mathbb{Z}^+$.) From this we conclude immediately that any integer of the form $2^{a_0} p_1^{a_1} \cdots p_r^{a_r}$, in which the p_i's are all primes of the form $4k + 1$, is a sum of two squares. On the other hand, trivially any square n^2 is a sum of two squares – we allow (for now!) the representation $n^2 = n^2 + 0$ – so this means that we can have prime factors of the form $q_j^{b_j}$ with $q_j \equiv 3 \pmod 4$ as long as all the exponents b_j are even. In other words, any integer n for which $\mathrm{ord}_p(n)$ is even for all primes $p \equiv 3 \pmod 4$ can be represented as a sum of two squares.

It is now time to look more extensively at the data to see whether any other integers are sums of squares. We can easily program a computer program to verify that for n up to, say, $10,000$, it is exactly the integers of the above form which are sums of two squares, which leads us to the statement of the following result.

Theorem 6. *A positive integer n is a sum of two squares iff $\mathrm{ord}_p(n)$ is even for all primes $p \equiv 3 \pmod 4$.*

Remark: An equivalent statement of the theorem, which we will use in the proof, is: n is a sum of two squares iff it factors as $n = ab^2$, where a has no prime factor $p \equiv 3 \pmod 4$.

For the proof we will need the following:

Lemma 7. *If a prime number $p \equiv 3 \pmod 4$ divides $a^2 + b^2$, then $p|a$ and $p|b$.*

Proof: By the previous section we know that such a prime number p remains irreducible in $\mathbb{Z}[i]$, so by Euclid's Lemma, since $p \mid a^2 + b^2 = (a + bi)(a - bi)$, we must have $p \mid a + bi$ or $p \mid a - bi$. In either case we get the desired conclusion (so in fact, p must divide both factors).

Let us now complete the proof of Theorem 6. What remains to be shown is that if n is a sum of two squares, then is of the form ab^2, where a has no prime factor $p \equiv 3 \pmod 4$. We will argue by induction: assume that every smaller sum of squares is of this form, and suppose $n = x^2 + y^2$. Since n is certainly of the required form if it has no prime divisors which are $3 \mod 4$, let us assume that it is divisible by such a p. Then by the previous lemma, $p|x$ and $p|y$; say $x = pX$, $y = pY$, so $n = p^2(X^2 + Y^2)$. Since $X^2 + Y^2$ is a smaller sum of squares, we may by induction write it in the form ab^2, with a not divisible by any prime congruent to $3 \mod 4$, so $n = a(pb)^2$ is itself of this form. This completes the proof of the theorem.

3. SOME SUPPLEMENTARY REMARKS

As mentioned in class, the discovery of which primes (and probably even which integers) are sums of two squares is due to Fermat. It is generally believed that he had a valid proof, although this is to a degree speculative since he almost never wrote out proofs of any of his theorems.

Almost certainly he did not have in mind the proof we have just given, which is much more algebraically sophisticated: indeed, as far as I know the Gaussian integers are correctly named (i.e., were first studied by Gauss, born in 1777), whereas Fermat died in 1665. There are however many other proofs; let me briefly comment on some of them.

Apparently the first published proof is due to Euler, in 1749. He used a form of "infinite descent", a technique developed by Fermat in his one completely written down proof (which we will be seeing soon). The key step in the proof is certainly to see that a prime $p \equiv 1 \pmod 4$ is a sum of two squares, and what we showed first was that at least there exists x such that $p \mid x^2 + 1$; in particular, p divides a sum of two squares, or equivalently, some multiple of p is a sum of two squares. Then a natural line of attack – indeed, probably more natural than a week-long digression on factorization in $\mathbb{Z}[i]$ – would be to try to show that if $mp = a^2 + b^2$ and $m > 1$, then also $np = c^2 + d^2$ for some $0 < n < m$. If so, then we work our way down to $1p = y^2 + z^2$, which is what we want.

It is possible to give a quite elementary and reasonably short proof (shorter than a week-long...) of this argument. This has its merits: on the one hand, it shows that one does not need abstract algebra to prove the two squares theorem, and on the other hand a very similar descent argument works to show that *any* prime is a sum of four squares, given the (easier!) lemma that every prime divides a sum of

four squares. At a moment of more leisure, we will outline this descent argument in some exercises.

There are many other proofs. Perhaps the most interesting argument was given by the leading analytic number theorist Roger Heath-Brown in 1971 (but only published in 1984): it is described in Aigner and Ziegler's *Proofs from THE BOOK*, a book-length compendium of "the most beautiful proofs."[1] Moreover there are many minor variants of the proof we have given: we have allowed ourselves to use relatively sophisticated algebraic concepts but not more esoteric, but elementary, statements about congruences, some of which we will develop later in the course. If we allowed ourselves the luxury of using that the multiplicative group of the finite field $\mathbb{Z}/p\mathbb{Z}$ is cyclic, it follows very rapidly that $-1 \equiv x^2$ can be solved in $\mathbb{Z}/p\mathbb{Z}$ iff $p = 2$ or $p \equiv 1 \pmod 4$. I feel though that most of the proofs of the Two Squares Theorem that do not use Gaussian integers have a somewhat *ad hoc* feeling, and moreover, that developing the theory of unique factorization in PID's and using division to show that $\mathbb{Z}[i]$ forms a PID are very important ideas, probably more important than the two squares theorem itself.

4. A Variant

This section was not discussed in class.
Suppose instead we are considering which integers are of the form $x^2 + 2y^2$. It is interesting to see how much of the previous discussion carries over. We give an outline only, asking the interested reader to fill in the details.

Step 1: An integer n is of the form $x^2 + 2y^2$ iff it is of the form $N(x + y\sqrt{-2})$; i.e., the ring to be looking at here is $\mathbb{Z}[\sqrt{-2}]$.

Step 2: The ring $\mathbb{Z}[\sqrt{-2}]$ (which, alas, does not have a spiffy name like "the Gaussian integers") is a PID, and hence satisfies Euclid's Lemma. (Hint: run through the proof for $\mathbb{Z}[i]$, especially the key division lemma; it works in this case too, albeit with less room to spare.)

Step 3: Experiment with primes p to see if you can find a pattern as to which are of the form $x^2 + 2y^2$ and which are not. (Hint: look again for a congruence condition, although not necessarily modulo 4.)

Step 4: Guess an analogue of Fermat's Lemma: namely that for all primes p in certain "good" congruence classes, $p \mid (x^2 + 2)$.

Step 5: **Assuming** your analogue of Fermat's Lemma is correct, use it and Step 2 to show that odd primes in your "good" congruence classes cannot be irreducible in $\mathbb{Z}[\sqrt{-2}]$, and conclude that all primes in your "good" classes are sums of squares.

Step 6: Can you show that primes in your "bad" congruence classes are not of the form $x^2 + 2y^2$?

---

[1]This was the proof given in a 2006 Math Club talk by Patrick Corn.

5. A NON-VARIANT

This section was not discussed in class.
Let us return to the infamous case of $\mathbb{Z}[\sqrt{-5}]$ and the associated question of which primes are of the form $x^2 + 5y^2$. Note that for instance $p = 3$ is a prime such that $p|(x^2+5)$ for some x, e.g. $x = 1$. If Euclid's Lemma held in $\mathbb{Z}[\sqrt{-5}]$ then we would have $3|((1 + \sqrt{-5})(1 - \sqrt{-5}))$, so that 3 would divide at least one of the factors. Clearly 3 does not divide $1 \pm \sqrt{-5}$, so the above pattern of argument would lead us to try to factor 3 into a product of two primes of $\mathbb{Z}[\sqrt{-5}]$, which would allow us to conclude that $3 = x^2 + 5y^2$. But as previously noted, 3 is clearly not of this form!

The problem, of course, is that Euclid's Lemma does not hold in $\mathbb{Z}[\sqrt{-5}]$, as we have noted already. The point I want to make here is that there is something quite anomalous going on: namely, 3 is a prime that "the congruence conditions suggest" should be of the form $x^2 + 5y^2$, but it isn't: congruence conditions can give information about solutions to Diophantine equations, and in very simple cases this information may be enough to decide the matter, but in only slightly less simple cases they do not tell the whole story.