

QUADRATIC RINGS

PETE L. CLARK

1. QUADRATIC FIELDS AND QUADRATIC RINGS

Let D be a squarefree integer not equal to 0 or 1. Then \sqrt{D} is irrational, and $\mathbb{Q}[\sqrt{D}]$, the subring of \mathbb{C} obtained by adjoining \sqrt{D} to \mathbb{Q} , is a field.

From an abstract algebraic perspective, an explanation for this can be given as follows: since \sqrt{D} is irrational, the polynomial $t^2 - D$ is irreducible over \mathbb{Q} . Since the ring $\mathbb{Q}[t]$ is a PID, the irreducible element $t^2 - D$ generates a maximal ideal $(t^2 - D)$, so that the quotient $\mathbb{Q}[t]/(t^2 - D)$ is a field. Moreover, the map $\mathbb{Q}[\sqrt{D}] \rightarrow \mathbb{Q}[t]/(t^2 - D)$ which is the identity on \mathbb{Q} and sends $\sqrt{D} \mapsto t$ is an isomorphism of rings, so $\mathbb{Q}[\sqrt{D}]$ is also a field. We may write $\mathbb{Q}[\sqrt{D}] = \{a + b\sqrt{D} \mid a, b \in \mathbb{Q}\}$, so that a basis for $\mathbb{Q}[\sqrt{D}]$ as a \mathbb{Q} -vector space is $1, \sqrt{D}$. In particular $\mathbb{Q}[\sqrt{D}]$ is two-dimensional as a \mathbb{Q} -vector space: we accordingly say it is a **quadratic field**.

It is also easy to check by hand that the ring $\mathbb{Q}[\sqrt{D}]$ is a field. For this and for many other things to come, the key identity is

$$(a + b\sqrt{D})(a - b\sqrt{D}) = a^2 - Db^2.$$

For rational numbers a and b which are not both zero, the rational number $a^2 - Db^2$ is also nonzero: equivalently there are no solutions to $D = \frac{a^2}{b^2}$, because \sqrt{D} is irrational. It follows that – again, for a, b not both 0 – we have

$$(a + b\sqrt{D}) \cdot \left(\frac{a}{a^2 - Db^2} - \frac{b}{a^2 - Db^2} \sqrt{D} \right) = 1,$$

which gives a multiplicative inverse for $a + b\sqrt{D}$ in $\mathbb{Q}[\sqrt{D}]$.

We wish also to consider **quadratic rings**, certain integral domains whose fraction field is a quadratic field $\mathbb{Q}(\sqrt{D})$. Eventually we will want a more precise and inclusive definition, but for now we consider $\mathbb{Z}[\sqrt{D}] = \{a + b\sqrt{D} \mid a, b \in \mathbb{Z}\}$.¹

2. FERMAT'S TWO SQUARES THEOREM

The rings $\mathbb{Z}[\sqrt{D}]$ occur naturally when we study Diophantine equations. E.g:

¹This equality is a fact which is not difficult to check; it is *not* the definition of $\mathbb{Z}[\sqrt{D}]$. By way of comparison, we recommend that the reader check that the ring $\mathbb{Z}[\frac{\sqrt{D}}{2}]$ is *not* of the form $\mathbb{Z}\alpha + \mathbb{Z}\beta$ for any two fixed elements α, β of $\mathbb{Z}[\frac{\sqrt{D}}{2}]$. In fact its additive group is not finitely generated as an abelian group.

Question 1. Which prime numbers p can be expressed as a sum of two squares? More precisely, for for which prime numbers p are there integers x and y such that

$$(1) \quad x^2 + y^2 = p?$$

Remark: Evidently 2 is a sum of squares: $1^2 + 1^2 = 2$. Henceforth we assume $p > 2$.

At the moment we have exactly one general technique² for studying Diophantine equations: congruences. So let's try to apply it here: if we reduce the equation $x^2 + y^2 = p$ modulo p , we get $x^2 + y^2 \equiv 0 \pmod{p}$. Suppose $(x, y) \in \mathbb{Z}^2$ is a solution to (1). It cannot be that $p \mid x$, because then we would also have $p \mid y$, and then $p^2 \mid x^2 + y^2 = p$, a contradiction. So both x and y are nonzero modulo p (hence invertible in $\mathbb{Z}/p\mathbb{Z}$) and the congruence can be rewritten as

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

That is, a necessary condition for (1) to have solutions is that -1 be a square modulo p . From [Handout: Pythagorean Triples, Lemma 7] we know that this occurs if and only if $p \equiv 1 \pmod{4}$. Thus we have shown that (1) has no \mathbb{Z} solutions unless $p \equiv 1 \pmod{4}$.

What about the converse: if $p \equiv 1 \pmod{4}$, is p necessarily a sum of two squares?

By Fermat's Lemma, there exists $x \in \mathbb{Z}$ such that $x^2 \equiv -1 \pmod{p}$, i.e., there exists $n \in \mathbb{Z}$ such that $pn = x^2 + 1$. Now factor the right hand side over $\mathbb{Z}[\sqrt{-1}]$:

$$pn = (x + \sqrt{-1})(x - \sqrt{-1}).$$

Suppose that p is prime as an element of $\mathbb{Z}[\sqrt{-1}]$. Then it satisfies Euclid's Lemma: if $p \mid \alpha\beta$, then $p \mid \alpha$ or $p \mid \beta$. Here, if p is prime in $\mathbb{Z}[\sqrt{-1}]$, then we get $p \mid x \pm i$. But this is absurd: what this means is that the quotient $\frac{x \pm i}{p} = \frac{x}{p} \pm \frac{1}{p}i$ is an element of $\mathbb{Z}[\sqrt{-1}]$, i.e., that both $\frac{x}{p}$ and $\frac{1}{p}$ are integers. But obviously $\frac{1}{p}$ is not an integer. Therefore p is not prime, so³ there exists a nontrivial factorization

$$(2) \quad p = \alpha\beta,$$

where $\alpha = a + b\sqrt{-1}, \beta = c + d\sqrt{-1} \in \mathbb{Z}[\sqrt{-1}]$ are nonunit elements. Taking complex conjugates of the above equation, we get

$$(3) \quad \bar{p} = p = \overline{\alpha\beta} = \overline{\alpha}\bar{\beta}.$$

Multiplying (2) and (3) we get

$$(4) \quad p^2 = (\alpha\bar{\alpha})(\beta\bar{\beta}) = (a^2 + b^2)(c^2 + d^2).$$

Now, since α and β are evidently nonzero, we have $a^2 + b^2, c^2 + d^2 > 0$. We claim that indeed $a^2 + b^2 \neq 1$ and $c^2 + d^2 \neq 1$. Indeed, if $a + b\sqrt{-1} \in \mathbb{Z}[\sqrt{-1}]$ with $a^2 + b^2 = 1$, then its multiplicative inverse in $\mathbb{Q}[\sqrt{-1}]$ is $\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\sqrt{-1} = a - b\sqrt{-1}$ which again lies in $\mathbb{Z}[\sqrt{-1}]$. In other words, $a^2 + b^2 = 1$ implies that

²By a "general technique", we mean a technique that can always be applied, not one that is guaranteed always to succeed. In that stronger sense there are provably no general technique for Diophantine equations!

³A gap occurs in the argument here. It has been deliberately inserted for pedagogical reasons. Please keep reading at least until the beginning of the next section!

$a + b\sqrt{-1}$ is a unit in $\mathbb{Z}[\sqrt{-1}]$, contrary to our assumption. Having ruled out that either $a^2 + b^2 = 1$ or $c^2 + d^2 = 1$, (4) now immediately implies

$$a^2 + b^2 = c^2 + d^2 = p.$$

But that is what we wanted: p is a sum of two squares! Thus we have (apparently...please read on!) proved the following theorem.

Theorem 1. (*Fermat's Two Squares Theorem*) *A prime number p can be expressed as the sum of two integer squares if and only if $p = 2$ or $p \equiv 1 \pmod{4}$.*

3. FERMAT'S TWO SQUARES THEOREM LOST

The above proof of Theorem 1 was surprisingly quick and easy, especially compared to Fermat's original one: not having the notion of factorization in domains other than the integers, Fermat's uses (as is typical of him) a more intricate argument by descent. In fact, the way we have presented the above argument, it is **too easy**: there is a gap in the proof. The gap is rather subtle, and is analogous to a notorious mistake made by the early 19th century mathematician Lamé. Rather than expose it directly, let us try to squeeze more out of it and see what goes wrong.

Namely, instead of just working with $x^2 + y^2 = p$ and the corresponding quadratic ring $\mathbb{Z}[\sqrt{-1}]$, let consider the equation

$$(5) \quad x^2 + Dy^2 = p,$$

where p is still a prime and D is a squarefree integer different from 0 or 1. We can mimic the above argument in two steps as follows:

Step 1: By reducing modulo p , we get exactly as before that the existence of an integral solution to (5) implies that $-D$ is a square modulo p .

Step 2: Conversely, assume that $-D$ is a square modulo p , i.e., there exists $x \in \mathbb{Z}$ such that $-D \equiv x^2 \pmod{p}$. Again this leads means there exists $n \in \mathbb{Z}$ such that

$$pn = x^2 + D,$$

and thus leads to a factorization in $\mathbb{Z}[\sqrt{-D}]$, namely

$$pn = (x + \sqrt{-D})(x - \sqrt{-D}).$$

Now if p were a prime element in $\mathbb{Z}[\sqrt{-D}]$ it would satisfy Euclid's Lemma, and therefore since it divides the product $(x + \sqrt{-D})(x - \sqrt{-D})$, it must divide one of the factors: $p \mid x \pm \sqrt{-D}$. But since $\frac{x}{p} \pm \frac{1}{p}\sqrt{-D}$ is still not in $\mathbb{Z}[\sqrt{-D}]$, this is absurd: p is not a prime element in $\mathbb{Z}[\sqrt{-D}]$. So it factors nontrivially: $p = \alpha\beta$, for α, β nonunits in $\mathbb{Z}[\sqrt{-D}]$. Let us now define, for any $\alpha = a + b\sqrt{-D} \in \mathbb{Q}(\sqrt{-D})$, $\bar{\alpha} = a - b\sqrt{-D}$. When $-D < 0$ this is the usual complex conjugation. When $-D > 0$ it is the conjugate in the sense of high school algebra (and also in Galois theory). It is entirely straightforward to verify the identity

$$\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$$

in either case, so that again by taking conjugates we get also

$$p = \bar{p} = \bar{\alpha}\bar{\beta},$$

and multiplying the two equations we get

$$p^2 = (\alpha\bar{\alpha})(\beta\bar{\beta}) = (a^2 - Db^2)(c^2 - Db^2).$$

As above, if $a^2 - Db^2 = \pm 1$, then the inverse of α lies in $\mathbb{Z}[\sqrt{-D}]$, so α is a unit in $\mathbb{Z}[\sqrt{-D}]$, which we assumed it not to be. This time, the conclusion we get is

$$p = \pm(a^2 - Db^2).$$

In particular, if $D < 0$, the conclusion we get is that p is of the form $a^2 + |D|b^2$ if and only if $-D$ is a square mod p .

Unfortunately we can easily see that this conclusion is very often wrong. Namely, suppose $D \leq -3$ and take $p = 2$.⁴ Then the condition that $-D$ is a square modulo 2 is quite vacuous: the only elements of $\mathbb{Z}/2\mathbb{Z}$ are $0 = 0^2$ and $1 = 1^2$, so every integer is congruent to a square modulo 2. Thus the above argument implies there are integers x and y such that

$$2 = x^2 + |D|y^2.$$

But this is absurd: if $y = 0$ it tells us that 2 is a square; whereas if $y \geq 1$, $x^2 + |D|y^2 \geq |D| \geq 3$. In other words, 2 is certainly *not* of the form $x^2 + |D|y^2$.

So what went wrong?!?

4. FERMAT'S TWO SQUARES THEOREM (AND MORE!) REGAINED

It is time to come clean. We have been equivocating over the definition of a prime element in an integral domain. Recall that we did not actually define such a thing. Rather, we defined an **irreducible** element in R to be a nonzero nonunit f such that $f = xy$ implies x or y is a unit. Then in the integers we proved Euclid's Lemma: if an irreducible element f of \mathbb{Z} divides ab , then either f divides a or f divides b . Of course this was not obvious: rather, it was all but equivalent to the fundamental theorem of arithmetic.

Let us now review how this is the case. By definition, a domain R is a unique factorization domain if it satisfies two properties: first, that every nonzero nonunit factor into irreducible elements, and second that this factorization be unique, up to ordering of factors and associate elements.

Suppose R is a UFD. We claim that if f is an irreducible element of R , and $f \mid ab$, then $f \mid a$ or $f \mid b$. We already proved this for $R = \mathbb{Z}$ and the general argument is the same: we leave to the reader the very important exercise of looking back at the argument for $R = \mathbb{Z}$ and adapting it to the context of R a UFD.

We define a **prime element** in a domain R to be a nonzero nonunit p which satisfies the conclusion of Euclid's Lemma: if $p \mid ab$, then $p \mid a$ or $p \mid b$.

Proposition 2. *Let R be an integral domain.*

a) Any prime element is irreducible.

⁴Taking $p = 2$ should always raise alarm bells in your head: it is often said that "2 is the oddest prime." In this case, please do look back the above argument to see that $p = 2$ was not – and did not need to be – excluded from consideration.

b) R is a UFD if and only if it is a factorization domain in which every irreducible element is prime.

Because this is pure algebra, we prefer to not discuss the proof here. It is a good exercise for the reader. See also

<http://www.math.uga.edu/~pete/factorization.pdf>

for a careful treatment of the theory of factorization in integral domains.

We can now recast the results of the previous two sections as follows. First, the “proof” which we gave assumed that either p is a prime element of $\mathbb{Z}[\sqrt{D}]$ or that p is not an irreducible element: i.e., it factors as $p = \alpha\beta$ with α, β nonunit elements of $\mathbb{Z}[\sqrt{D}]$. What was missing was the third possibility: p is an irreducible element of $\mathbb{Z}[\sqrt{D}]$ which is not prime. Because of Proposition 2, this third possibility cannot occur if $\mathbb{Z}[\sqrt{D}]$ is a UFD, so what we actually proved was:

Theorem 3. *Let D be a squarefree integer different from 0 and 1. We **assume** that the ring $\mathbb{Z}[\sqrt{D}]$ is a UFD. Then, for a prime number p , TFAE:*

(i) *There exist $x, y \in \mathbb{Z}$ such that $p = |x^2 - Dy^2|$.*

(ii) *There exists $x \in \mathbb{Z}$ such that $D \equiv x^2 \pmod{p}$.*

Moreover, simply by noticing that for $D < -2$ we had that D is a square mod 2 but 2 is not of the form $|x^2 - Dy^2|$, we also deduce:

Corollary 4. *For no $D < -2$ is the ring $\mathbb{Z}[\sqrt{D}]$ a UFD.*

To complete the proof of Theorem 1 it suffices to show that at least $\mathbb{Z}[\sqrt{-1}]$ is a UFD. Fortunately for us, it is. We show this in a way which is again a generalization of one of our proofs of the fundamental theorem of arithmetic: namely, by showing that we can perform, in a certain appropriate sense, division with remainder in $\mathbb{Z}[\sqrt{-1}]$. The formalism for this is developed in the next section.

4.1. Euclidean norms.

A norm $N : R \rightarrow \mathbb{N}$ is called **Euclidean** if it satisfies the following property: for all $a \in R, b \in R \setminus \{0\}$, there exist $q, r \in R$ such that $a = qb + r$ and $N(r) < N(b)$.

First let us see that this suffices: we claim that any domain R endowed with a Euclidean norm function is a principal ideal domain. Indeed, let I be any nonzero ideal of such a domain R , and let a be any element of minimal nonzero norm. We claim that in fact $I = (a) = \{ra \mid r \in R\}$. The proof is the same as for integers: suppose that $b \in I$ and apply the Euclidean property: there exist $q, r \in R$ such that $b = qa + r$ with $N(r) < N(a)$. But $r = b - qa$ and $a, b \in I$, so $r \in I$. If $r \neq 0$ then $N(r) \neq 0$ and we have found an element with nonzero norm smaller than $N(a)$, contradiction. So we must have $r = 0$, i.e., $b = qa \in (a)$.

Side remark: Note that in our terminology a “norm” $N : R \rightarrow \mathbb{N}$ is multiplicative, and indeed in our present application we are working with such norms. However, the multiplicativity property was not used in the proof. If R is a domain, let us define a **generalized Euclidean norm** on R to be a function $N : R \rightarrow \mathbb{N}$ such that $N(r) = 0 \iff r = 0$ and such that for all $a \in R, b \in R \setminus \{0\}$, there exist

$q, r \in R$ with $a = qb + r$ and $N(r) < N(b)$. Then what we have actually shown is that any domain which admits a generalized Euclidean norm is a PID.⁵

4.2. PIDs and UFDs.

One also knows that any PID is a UFD. This is true in general, but in the general case it is somewhat tricky to establish the existence of a factorization into irreducibles. In the presence of a multiplicative norm function $N : R \rightarrow \mathbb{N}$ – i.e., a function such that $N(x) = 0 \iff x = 0$, $N(x) = 1 \iff x \in R^\times$, $N(xy) = N(x)N(y) \forall x, y \in R$ – this part of the argument becomes much easier to establish, since for any nontrivial factorization $x = yz$ we have $N(y), N(z) < N(x)$. Complete details are available in *loc. cit.*

4.3. Some Euclidean quadratic rings.

Finally, we will show that our norm function on $\mathbb{Z}[\sqrt{-1}]$ is Euclidean. At this point it costs nothing extra, and indeed is rather enlightening, to consider the more general case of $\mathbb{Z}[\sqrt{D}]$ endowed with the norm function $N(a + b\sqrt{D}) = |a^2 - Db^2|$. According to the characterization of (multiplicative) Euclidean norms in the previous subsection, what we must show is: for all $\alpha \in \mathbb{Q}(\sqrt{D})$, there exists $\beta \in \mathbb{Z}[\sqrt{D}]$ with $N(\alpha - \beta) < 1$. A general element of α is of the form $r + s\sqrt{D}$ with $r, s \in \mathbb{Q}$, and we are trying to approximate it by an element $x + y\sqrt{D}$ with $x, y \in \mathbb{Z}$.

Let us try something easy: take x (resp. y) to be an integer nearest to r (resp. s). If z is any real number, there exists an integer n with $|z - n| \leq \frac{1}{2}$, and this bound is sharp, attained for all real numbers with fractional part $\frac{1}{2}$.⁶ So let $x, y \in \mathbb{Z}$ be such that $|r - x|, |s - y| \leq \frac{1}{2}$. Is then $\beta = x + y\sqrt{D}$ a good enough approximation to $\alpha = r + s\sqrt{D}$? Consider the following quick and dirty estimate:

$$(6) \quad N(\alpha - \beta) = |(r - x)^2 - D(s - y)^2| \leq |r - x|^2 + |D||s - y|^2 \leq \frac{|D| + 1}{4}.$$

Evidently $\frac{|D| + 1}{4} < 1$ iff $|D| < 3$.

So $D = -1$ works – i.e., the norm N on $\mathbb{Z}[\sqrt{-1}]$ is Euclidean, so $\mathbb{Z}[\sqrt{-1}]$ is a UFD, which fills in the gap in our proof of Theorem 1.

Also $D = 2$ and $D = -2$ work: the rings $\mathbb{Z}[\sqrt{-2}]$ and $\mathbb{Z}[\sqrt{2}]$ are UFDs.

It is natural to wonder whether the quick and dirty estimate can be improved. We have already seen that it *cannot* be for $D < -2$, since we already know that for such D , $\mathbb{Z}[\sqrt{D}]$ is not a UFD. This can be confirmed as follows: when $D < 0$, $N(a + b\sqrt{D}) = a^2 + Dy^2 = ||a + \sqrt{|D|}i||^2$; that is, our norm function is simply the square of the usual Euclidean length function evaluated on the complex number

⁵In fact further generalization is possible: in order for this simple argument to go through it is not necessary for the codomain of the norm function to be the natural numbers, but only a well-ordered set!

⁶Moreover when $|z - n| < \frac{1}{2}$, the nearest integer is unique, whereas for half-integral real numbers there are evidently two nearest integers. This does not matter to us: in this case take either nearest integer, i.e., either $z - \frac{1}{2}$ or $z + \frac{1}{2}$.

$a + \sqrt{|D|}i$. Moreover, in this case the ring $\mathbb{Z}[\sqrt{-D}]$ lives naturally inside \mathbb{C} as a *lattice*, whose fundamental parallelogram is simply a rectangle with sides 1 and \sqrt{D} . The problem now is to find, for a given point $z = a + bi \in \mathbb{C}$ with $a, b \in \mathbb{Q}$, the closest lattice point. But it is geometrically clear that the points which are furthest away from lattice points are precisely those which lie at the center of the corresponding rectangles, e.g. $\frac{1}{2} + \frac{1}{2}|D|i$. This shows that nothing was lost in our quick and dirty estimate.

The situation $D < 0$ is quite different, most of all because the geometric picture is different: $\mathbb{Z}[\sqrt{D}]$ now lives inside \mathbb{R} , but unlike in the previous case, it is not discrete. Rather, (Exercise X.X) it is *dense*: any interval (a, b) in \mathbb{R} with $a < b$ contains some point $\alpha \in \mathbb{Z}[\sqrt{D}]$. So it is not clear that the above “coordinatewise nearest integer approximation” is the best possible approximation.

But even keeping the same approximating element β as above, we lose ground in the estimate $|(\frac{1}{2})^2 - D(\frac{1}{2})^2| = |\frac{1}{4} - \frac{D}{4}| \leq \frac{D+1}{4}$. Rather, we want $|\frac{1}{4} - \frac{D}{4}| < 1$, or $|D - 1| < 4$, so $D = 3$ also works. Thus, $\mathbb{Z}[\sqrt{3}]$ has a Euclidean norm, hence is a UFD. In summary, we get the following “bonus theorem”:

Theorem 5. *a) A prime p is of the form $x^2 + 2y^2$ iff -2 is a square modulo p .
 b) A prime p is of the form $|x^2 - 2y^2|$ iff 2 is a square modulo p .
 c) A prime p is of the form $|x^2 - 3y^2|$ iff 3 is a square modulo p .*

Of course this brings attention to the fact that for an integer D , we do not know how to characterize the set of primes p such that D is a square mod p , except in the (easiest) case $D = -1$. The desire to answer this question is an excellent motivation for the **quadratic reciprocity law**, coming up shortly.

5. COMPOSITES OF THE FORM $x^2 - Dy^2$

Now that we have determined which primes are of the form $x^2 + y^2$, it is natural to attempt to determine all nonzero integers which are sums of two squares.

An honest approach to this problem would begin by accumulating data and considering various special cases. Here we must unfortunately be somewhat more succinct.

Somewhat more generally, fix D a squarefree integer as before, and put

$$\mathcal{S}_D = \{n \in \mathbb{Z} \setminus \{0\} \mid \exists x, y \in \mathbb{Z}, n = x^2 - Dy^2\},$$

the set of all nonzero integers of the form $x^2 - Dy^2$. Because of the multiplicativity of the norm function – or more precisely, the function $x + y\sqrt{D} \mapsto x^2 - Dy^2$, which takes on negative values when $D > 0$ – the subset \mathcal{S}_D is closed under multiplication.

Remark: Certainly $1 \in \mathcal{S}_D$: $1 = 1^2 - D \cdot 0^2$. Therefore the multiplicative property can be rephrased by saying that \mathcal{S}_D is a **submonoid** of the monoid $(\mathbb{Z} \setminus \{0\}, \cdot)$.

Now we know that the following positive integers are all sums of two squares: 1, 2, and prime $p \equiv 1 \pmod{4}$, and n^2 for any integer n : $n^2 = (n)^2 + 0^2$. Now let n be any positive integer, and write p_1, \dots, p_r for the distinct prime divisors of n

which are congruent to 1 modulo 4, and q_1, \dots, q_s for the distinct prime divisors of n which are congruent to -1 modulo 4, so that

$$n = 2^a p_1^{m_1} \cdots p_r^{m_r} q_1^{n_1} \cdots q_s^{n_s},$$

for $a, m_1, \dots, m_r, n_1, \dots, n_s \in \mathbb{N}$. It follows that so long as n_1, \dots, n_s are all even, n is a product of sums of two squares and therefore itself a sum of two squares.

Finally, we wish to show that we have found all positive integers which are a sum of two squares.⁷ Specifically, what we wish to show is that if $n \in \mathbb{Z}^+$ is a sum of two squares, then for any prime number $p \equiv -1 \pmod{4}$, then $\text{ord}_p(n)$ is even. For this it suffices to show the following

Lemma 6. *Let $p \equiv -1 \pmod{4}$ be a prime number, and suppose that there exist $x, y \in \mathbb{Z}$ such that $p \mid x^2 + y^2$, then $p \mid x$ and $p \mid y$.*

Before proving Lemma 6 let us show how it helps us. Indeed, suppose that a positive integer n is a sum of two squares: $n = x^2 + y^2$. Let p be any prime congruent to $-1 \pmod{4}$, and assume that $p \mid n$ (otherwise $\text{ord}_p(n) = 0$, which is even). Then by Lemma 6 $p \mid x$ and $p \mid y$, so that $\frac{n}{p^2} = (\frac{x}{p})^2 + (\frac{y}{p})^2$ expresses $\frac{n}{p^2}$ as a sum of two integral squares. But now we can repeat this process of repeated division by p^2 until we get an integer $\frac{n}{p^{2k}}$ which is not divisible by p . Thus $\text{ord}_p(n) = 2k$ is even.

Proof of Lemma 6: It follows from the proof of the Two Squares Theorem that if $p \equiv -1 \pmod{4}$ is a prime number, then it remains irreducible in $\mathbb{Z}[\sqrt{-1}]$. Let us recall why: otherwise $p = \alpha\beta$ with $N(\alpha), N(\beta) > 1$, and then taking norms gives

$$p^2 = N(p) = N(\alpha\beta) = N(\alpha)N(\beta),$$

and thus $N(\alpha) = N(\beta) = p$. Writing $\alpha = a + b\sqrt{-1}$, we get $p = N(\alpha) = a^2 + b^2$, so p is a sum of two squares, contrary to Fermat's Lemma.

Since $\mathbb{Z}[\sqrt{-1}]$ is a UFD, the irreducible element p is a prime element, hence Euclid's Lemma applies. We have $p \mid x^2 + y^2 = (x + y\sqrt{-1})(x - y\sqrt{-1})$, so that $p \mid x + y\sqrt{-1}$ or $p \mid x - y\sqrt{-1}$. This implies that $\frac{x}{p}, \frac{y}{p} \in \mathbb{Z}$, i.e., $p \mid x$ and $p \mid y$.

In summary, we have shown:

Theorem 7. *(Full Two Squares Theorem) A positive integer n is a sum of two squares iff $\text{ord}_p(n)$ is even for all primes $p \equiv -1 \pmod{4}$.*

In that Lemma 6 uses (only) that $\mathbb{Z}[\sqrt{-1}]$ is a UFD, similar reasoning applies to other Euclidean quadratic rings $\mathbb{Z}[\sqrt{D}]$. In particular, there is a direct analogue of Theorem 7 for $x^2 + 2y^2$, which we will postpone until we determine for which primes $p \equiv -2$ is a square modulo p . When D is positive, the distinction between $\alpha\bar{\alpha} = a^2 - Db^2$ and $N(\alpha) = |a^2 - Db^2|$ becomes important: for instance as above we obviously have $1 \in \mathcal{S}_D$, but whether $-1 \in \mathcal{S}_D$ is a surprisingly difficult problem: to this day there is not a completely satisfactory solution.

In contrast, if $\mathbb{Z}[\sqrt{D}]$ is not a UFD, then even if we know which primes are of the form $x^2 - Dy^2$, we cannot use the above considerations to determine the set \mathcal{S}_D . For example take $D = -5$. Certainly $2 \neq x^2 + 5y^2$ and $3 \neq x^2 + 5y^2$, but $2 \cdot 3 = 1^2 + 5 \cdot 1^2$.

⁷Why would we think this? Again, trial and error experimentation is the honest answer.