# 6

# The Gaussian integers

## PREVIEW

The Gaussian integers $\mathbb{Z}[i]$ are the simplest generalization of the ordinary integers $\mathbb{Z}$ and they behave in much the same way. In particular, $\mathbb{Z}[i]$ enjoys *unique prime factorization*, and this allows us to reason about $\mathbb{Z}[i]$ the same way we do about $\mathbb{Z}$. We do this because $\mathbb{Z}[i]$ is the natural place to study certain properties of $\mathbb{Z}$. In particular, it is the best place to examine *sums of two squares*, because in $\mathbb{Z}[i]$ we can factorize a sum of two integer squares into linear factors: $x^2+y^2=(x-yi)(x+yi)$.

In the present chapter we use this idea to prove a famous theorem of Fermat: *if $p>2$ is prime then $p=a^2+b^2$, for some natural numbers $a$ and $b$, if and only if $p=4n+1$ for some natural number $n$*. The Fermat two square theorem turns out to be related, not only to unique prime factorization in $\mathbb{Z}[i]$, but also to the actual "primes" of $\mathbb{Z}[i]$, the so-called *Gaussian primes*.

The Gaussian primes are easily shown to include the ordinary primes that are not sums of two squares, and the factors $a-bi$ and $a+bi$ of each ordinary prime of the form $a^2+b^2$. Unique prime factorization in $\mathbb{Z}[i]$ establishes that these are the only Gaussian primes, up to multiples by $\pm 1$ and $\pm i$.

An easy congruence argument shows that ordinary primes of the form $4n+3$ are not sums of two squares. The two square theorem then shows that the primes that *are* sums of two squares are 2 and all the remaining odd primes, namely, those of the form $4n+1$.

The proof of the two square theorem involves an important lemma proved with the help of Wilson's theorem: each prime $p=4n+1$

divides a number of the form $m^2+1$. Since $m^2+1$ factorizes in $\mathbb{Z}[i]$, it follows from unique prime factorization that $p$ does also. The factorization of $p$ turns out to be of the form $(a-bi)(a+bi)$, hence $p=(a-bi)(a+bi)=a^2+b^2$, as claimed.

## 6.1 $\mathbb{Z}[i]$ and its norm

In the last chapter we saw that certain questions about $\mathbb{Z}$ are clarified by working with generalized integers, in particular, working in $\mathbb{Z}[\sqrt{n}]$ to solve $x^2-ny^2=1$ in $\mathbb{Z}$. The role of $\mathbb{Z}[\sqrt{n}]$ in this case is to allow the factorization

$$x^2-ny^2=(x-y\sqrt{n})(x+y\sqrt{n}).$$

Similarly, when studying $x^2+y^2$, it helps to use the *Gaussian integers*

$$\mathbb{Z}[i]=\{a+bi : a,b\in\mathbb{Z}\}$$

because $x^2+y^2=(x-yi)(x+yi)$.

Sums of two squares, $x^2+y^2$, are the oldest known topic in number theory. We have already seen results about them found by the Babylonians, Euclid, and Diophantus. In fact, it could be said that some properties of $\mathbb{Z}[i]$ itself go back this far; at least, as far as Diophantus.

Diophantus apparently knew the two square identity (Section 1.8)

$$(a_1^2+b_1^2)(a_2^2+b_2^2)=(a_1a_2-b_1b_2)^2+(a_1b_2+b_1a_2)^2$$

because he knew that the product of sums of two squares is itself the sum of two squares. Today we recognize this formula as equivalent to the *multiplicative property of absolute value*,

$$|z_1||z_2|=|z_1z_2|,$$

where $z_1=a_1+b_1i$ and $z_2=a_2+b_2i$. And Diophantus' identity is exactly the formula

$$\text{norm}(a_1+b_1i)\text{norm}(a_2+b_2i)=\text{norm}((a_1+b_1i)(a_2+b_2i)), \qquad (*)$$

where "norm" denotes the norm of $\mathbb{Z}[i]$,

$$\text{norm}(a+bi)=|a+bi|^2=a^2+b^2.$$

### Exercises

When discussing factorization there are always trivial factors, called *units*, that we prefer to ignore. For example, in $\mathbb{N}$ the only unit is 1, in $\mathbb{Z}$ the units are 1 and $-1$, and in $\mathbb{Z}[i]$ the units are the elements of norm 1.

**6.1.1** Show that the units of $\mathbb{Z}[i]$ are $\pm 1$, $\pm i$.

Likewise, the units of $\mathbb{Z}[\sqrt{n}]$ are its elements of norm 1, that is, the numbers $a+b\sqrt{n}$ with $a^2-nb^2=1$.

**6.1.2** Describe the units of $\mathbb{Z}[\sqrt{2}]$.

**6.1.3** Show that $\mathbb{Z}[\sqrt{n}]$ has infinitely many units for any nonsquare natural number $n$.

## 6.2 Divisibility and primes in $\mathbb{Z}[i]$ and $\mathbb{Z}$

The $\mathbb{Z}[i]$ norm

$$\text{norm}(a+bi) = |a+bi|^2 = a^2+b^2$$

is more useful in number theory than the absolute value because the norm is always an ordinary integer. The *multiplicative property of the norm* (*) implies that, if a Gaussian integer $\alpha$ divides a Gaussian integer $\gamma$, that is, if

$$\gamma = \alpha\beta \quad \text{for some } \beta \in \mathbb{Z}[i],$$

then

$$\text{norm}(\gamma) = \text{norm}(\alpha)\text{norm}(\beta),$$

that is, $\text{norm}(\alpha)$ *divides* $\text{norm}(\gamma)$.

Because of this, questions about divisibility in $\mathbb{Z}[i]$ often reduce to questions about divisibility in $\mathbb{Z}$. In particular, it is natural to define a *Gaussian prime* to be a Gaussian integer that is not the product of Gaussian integers of smaller norm. Then we can answer various questions about Gaussian primes by looking at norms.

**Examples.**

1. $4+i$ is Gaussian prime.
   Because $\text{norm}(4+i) = 16+1 = 17$, which is a prime in $\mathbb{Z}$. Hence $4+i$ is not the product of Gaussian integers of smaller norm, because no such norms divide 17.

2. 2 is not a Gaussian prime.
   Because $2 = (1-i)(1+i)$ and both $1-i$ and $1+i$ have norm 2, which is smaller than $\text{norm}(2) = 4$.

3. $1-i$, $1+i$ are Gaussian prime factors of 2.
   Because $\text{norm}(1-i) = \text{norm}(1+i) = 2$ is a prime in $\mathbb{Z}$, hence $1-i$ and $1+i$ are not products of Gaussian integers of smaller norm.

**Prime factorization in $\mathbb{Z}[i]$.** *Any Gaussian integer factorizes into Gaussian primes.* The proof is similar to the proof in $\mathbb{Z}$.

**Proof.** Consider any Gaussian integer $\gamma$. If $\gamma$ itself is a Gaussian prime, then we are done. If not, then $\gamma = \alpha\beta$ for some $\alpha, \beta \in \mathbb{Z}[i]$ with smaller norm. If $\alpha$, $\beta$ are not both Gaussian primes, we factorize into Gaussian integers of still smaller norm, and so on. This process must terminate since norms, being natural numbers, cannot decrease forever. Hence we eventually get a Gaussian prime factorization of $\gamma$. □

As in $\mathbb{Z}$, it is not immediately clear that the prime factorization is unique. However, we see in Section 6.4 that unique prime factorization holds in $\mathbb{Z}[i]$ for much the same reasons as in $\mathbb{Z}$.

## Exercises

An equivalent way to define Gaussian primes, in line with a common way of defining ordinary primes, is to say that $\varpi$ is a Gaussian prime if $\varpi$ is divisible only by units and units times $\varpi$. (It is conventional to use the Greek letter pi to denote primes in $\mathbb{Z}[i]$ and other generalizations of $\mathbb{Z}$, the way $p$ is used to denote ordinary primes. However, to avoid confusion with $\pi = 3.14159\ldots$ I prefer to use $\varpi$, the variant form of pi.)

**6.2.1** Explain why this definition is equivalent to the one above.

**6.2.2** Prove that 3 is a Gaussian prime by considering the divisors of norm(3).

Ordinary primes are not always Gaussian primes, as we have already seen in the case of 2. In fact, 2 is "almost a square" in $\mathbb{Z}[i]$.

**6.2.3** Show that a unit times 2 is a square in $\mathbb{Z}[i]$.

**6.2.4** Factorize 17 and 53 in $\mathbb{Z}[i]$.

## 6.3 Conjugates

The *conjugate* of $z = a + bi$ is $\overline{z} = a - bi$. The basic properties of conjugation (not only in $\mathbb{Z}[i]$ but for all complex numbers $z$) are

$$\begin{aligned} z\overline{z} &= |z|^2, \\ \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2}, \\ \overline{z_1 - z_2} &= \overline{z_1} - \overline{z_2}, \\ \overline{z_1 \times z_2} &= \overline{z_1} \times \overline{z_2}. \end{aligned}$$

These can be checked by writing $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$ and working out both sides of each identity. We use these properties of conjugation to take the first step towards a classification of Gaussian primes.

**Real Gaussian primes.** *An ordinary prime $p \in \mathbb{N}$ is a Gaussian prime $\Leftrightarrow$ $p$ is not the sum of two squares. (And obviously $p < 0$ is a Gaussian prime $\Leftrightarrow -p \in \mathbb{N}$ is a Gaussian prime.)*

**Proof.** ($\Leftarrow$) Suppose that we have an ordinary prime $p$ that is not a Gaussian prime, so it factorizes in $\mathbb{Z}[i]$:

$$p = (a + bi)\gamma,$$

where $a + bi$ and $\gamma$ are Gaussian integers with norm $<$ the norm $p^2$ of $p$ (and hence also of norm $> 1$). Taking conjugates of both sides we get

$$p = (a - bi)\overline{\gamma},$$

since $p$ is real and hence $p = \overline{p}$. Multiplying these two expressions for $p$ gives

$$\begin{aligned} p^2 &= (a - bi)(a + bi)\gamma\overline{\gamma} \\ &= (a^2 + b^2)|\gamma|^2, \end{aligned}$$

where both $a^2 + b^2, |\gamma|^2 > 1$. But the only such factorization of $p^2$ is $pp$, hence $p = a^2 + b^2$.

($\Rightarrow$) Conversely, *if an ordinary prime $p$ equals $a^2 + b^2$ with $a, b \in \mathbb{Z}$ then $p$ is not a Gaussian prime* because it has the Gaussian prime factorization

$$p = (a - bi)(a + bi)$$

into factors of norm $a^2 + b^2 = p < \text{norm}(p) = p^2$. □

Notice also that *the factors $a-bi$ and $a+bi$ of $p$ are Gaussian primes* because their norm is the prime number $a^2+b^2=p$. Moreover, all Gaussian primes $a+bi$, where $a,b\neq 0$, come in conjugate pairs like this. This is so because if one member of the pair factorizes into $\alpha\beta$ then its conjugate factorizes into $\overline{\alpha}\overline{\beta}$.

What is not yet clear is whether *all* Gaussian primes $a+bi$ with $a,b$ nonzero are factors of ordinary primes $p=a^2+b^2$. It is conceivable that $a+bi$ could be a Gaussian prime while $a^2+b^2$ is a product of two or more ordinary primes. In Section 6.4 we rule this out with the help of unique prime factorization in $\mathbb{Z}[i]$.

At any rate, we can see that further clarification of the nature of Gaussian primes depends on finding another way to describe the ordinary primes that are sums of two squares. We saw in Section 3.7 (Example 1) that ordinary primes that are *not* sums of two squares are of the form $4n+3$. The complement to this result—that any prime of the form $4n+1$ *is* a sum of two squares—is a famous theorem discovered by Fermat. It is proved in Section 6.5.

## Exercises

**6.3.1** Verify the basic properties of conjugation mentioned above.

The proof of the classification of real Gaussian primes has the following interesting consequences.

**6.3.2** Show that each ordinary prime has a distinct Gaussian prime associated with it.

**6.3.3** Deduce that there are infinitely many Gaussian primes.

Since the real positive Gaussian primes are those of the form $4n+3$, another way to prove that there are infinitely many Gaussian primes is to show that there are infinitely many ordinary primes of the form $4n+3$. The proof is along lines similar to Euclid's proof in Section 1.1.

**6.3.4** Show that the product of numbers of the form $4n+1$ is of the same form. Deduce that any number of the form $4n+3$ has a prime divisor of the form $4n+3$.

**6.3.5** If $p_1,p_2,\ldots,p_k$ are primes of the form $4n+3$, show that $2p_1p_2\cdots p_k+1$ is also of the form $4n+3$.

**6.3.6** Deduce from Exercises 6.3.4 and 6.3.5 that there are infinitely many primes of the form $4n+3$.

## 6.4 Division in $\mathbb{Z}[i]$

Unique prime factorization in $\mathbb{Z}[i]$, as in $\mathbb{Z}$, relies on the Euclidean algorithm, which depends in turn on:

**Division property of $\mathbb{Z}[i]$.** *If $\alpha, \beta \neq 0$ are in $\mathbb{Z}[i]$ then there is a* quotient $\mu$ *and a* remainder $\rho$ *such that*

$$\alpha = \mu\beta + \rho \quad \textit{with} \quad |\rho| < |\beta|.$$

**Proof.** This property becomes obvious once one sees that the Gaussian integer multiples $\mu\beta$ of any Gaussian integer $\beta \neq 0$ form a square grid in the complex plane.

This is because multiplication of $\beta$ by $i$ rotates the vector from 0 to $\beta$ through 90°, hence 0, $\beta$, and $i\beta$ are three corners of a square. All other multiples of $\beta$ are sums (or differences) of $\beta$ and $i\beta$, hence they lie at the corners of a square grid. (Figure 6.1.)

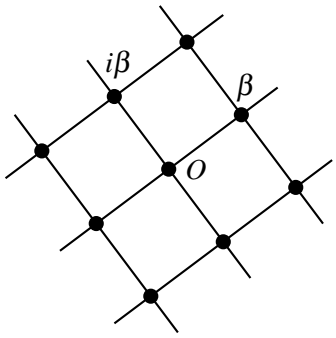


Figure 6.1: Multiples of a Gaussian integer

Any Gaussian integer $\alpha$ lies in one of these squares, and there is a nearest corner $\mu\beta$ (not necessarily unique, but no matter). Then

$$\alpha = \mu\beta + \rho, \quad \text{where} \quad |\rho| = \text{distance to nearest corner},$$

so $|\rho|$ is less than the side of a square, namely $|\beta|$. □

Thanks to the division property we have

1. A Euclidean algorithm for $\mathbb{Z}[i]$
2. $\gcd(\alpha, \beta) = \mu\alpha + \nu\beta$ for some $\mu, \nu \in \mathbb{Z}[i]$.

3. The *prime divisor property*: if a prime $\varpi$ divides $\alpha\beta$ then $\varpi$ divides $\alpha$ or $\varpi$ divides $\beta$.

4. Unique prime factorization up to order and factors of norm 1, namely $\pm 1$ and $\pm i$. Elements of norm 1 are called *units* and unique prime factorization usually comes with the qualification "up to unit factors". This is true even in $\mathbb{Z}$, where the units are $\pm 1$ and hence primes may vary up to sign.

As a first application of unique prime factorization in $\mathbb{Z}[i]$ we complete the description of Gaussian primes begun in Section 6.3. There we found that the real Gaussian primes are ordinary primes that are not sums of two squares, and their negatives. It is also clear that the *pure imaginary* Gaussian primes are of the form $\pm ip$, where $p$ is a real Gaussian prime. Thus it remains to describe the Gaussian primes $a+bi$ with $a$, $b$ nonzero.

**Imaginary Gaussian primes.** *The Gaussian primes $a+bi$ with $a$, $b$ nonzero are factors of ordinary primes $p$ of the form $a^2+b^2$.*

**Proof.** First, as noted in Section 6.3, if $a+bi$ is a Gaussian prime then so is $a-bi$ (because if $a-bi=\alpha\beta$ is not prime, neither is $a+bi=\overline{\alpha}\overline{\beta}$).

Next, $(a-bi)(a+bi)$ is a (necessarily unique) Gaussian prime factorization of

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

But $p$ must then be an ordinary prime. Indeed, if

$$p=rs \quad \text{with} \quad 1<r,s<p \quad \text{and} \quad r,s\in\mathbb{Z},$$

then the Gaussian prime factors of $r$ and $s$ give a Gaussian prime factorization of $p$ different from $(a-bi)(a+bi)$ (either two real factors $r$ and $s$, or $\geq$ four complex factors). □

## Exercises

Using unique prime factorization we can prove results on squares and cubes in $\mathbb{Z}[i]$, similar to those on squares and cubes in $\mathbb{N}$ proved in Section 2.5. The only difference is that we have to take account of units, as indeed we already do in $\mathbb{Z}$.

**6.4.1** Is it true in $\mathbb{Z}$ that relatively prime factors of a square are themselves squares? If not, how should the statement be modified to make it correct?

**6.4.2** Show that relatively prime factors of a cube in $\mathbb{Z}$ are themselves cubes.

**6.4.3** Formulate a theorem about relatively prime factors of a square in $\mathbb{Z}[i]$.

**6.4.4** Show that relatively prime factors of a cube in $\mathbb{Z}[i]$ are themselves cubes.

## 6.5 Fermat's two square theorem

In Section 3.7 we used congruence mod 4 to show that primes of the form $4n+3$ are not sums of two squares. Fermat's two square theorem says that the remaining odd primes—those of the form $4n+1$—are all sums of two squares.

We apply the theory of $\mathbb{Z}[i]$ to a prime $p = 4n+1$ with the help of an $m \in \mathbb{Z}$ such that $p$ divides $m^2+1$. Such an $m$ always exists by a result of Lagrange (1773) that follows from *Wilson's theorem* in Section 3.5: for any prime $p$

$$1 \times 2 \times 3 \times \cdots \times (p-1) \equiv -1 \pmod{p}.$$

**Lagrange's lemma.** *A prime $p = 4n+1$ divides $m^2+1$ for some $m \in \mathbb{Z}$.*

**Proof.** If we apply Wilson's theorem to the prime $p = 4n+1$ we get

$$\begin{aligned}
-1 &\equiv 1 \times 2 \times 3 \times \cdots \times 4n \pmod{p} \\
&\equiv (1 \times 2 \times \cdots \times 2n) \times \\
&\quad ((2n+1) \times \cdots \times (4n-1) \times (4n)) \pmod{p} \\
&\equiv (1 \times 2 \times \cdots \times 2n) \times \\
&\quad ((-2n) \times \cdots \times (-2)(-1)) \pmod{p} \qquad \text{since } p-k \equiv -k \pmod{p} \\
&\equiv (1 \times 2 \times \cdots \times 2n)^2 (-1)^{2n} \pmod{p} \\
&\equiv (1 \times 2 \times \cdots \times 2n)^2 \pmod{p}
\end{aligned}$$

Taking $m = (2n)!$ we get $m^2 \equiv -1 \pmod{p}$. That is, $p$ divides $m^2+1$. □

**Fermat's two square theorem.** *If $p = 4n+1$ is prime, then $p = a^2+b^2$ for some $a, b \in \mathbb{Z}$.*

**Proof.** Given $p$, let $m \in \mathbb{Z}$ be such that $p$ divides $m^2+1$, as in the lemma. In $\mathbb{Z}[i]$, $m^2+1$ has the factorization

$$m^2+1 = (m-i)(m+i).$$

And, even though $p$ divides $m^2+1$, $p$ does *not* divide $m-i$ or $m+i$ because $\frac{m}{p} - \frac{i}{p}$ and $\frac{m}{p} + \frac{i}{p}$ are not Gaussian integers.

By the Gaussian prime divisor property of Section 6.4, it follows that *$p$ is not a Gaussian prime*. But then $p = a^2+b^2$, as proved in Section 6.3. □

It also follows that

$$p = (a - bi)(a + bi)$$

is a factorization into Gaussian primes, and we now know that any such factorization is unique. So in fact we have a stronger form of Fermat's two square theorem: *each prime* $p = 4n + 1$ *is a sum* $a^2 + b^2$ *of two squares for a unique pair of natural numbers a, b.*

### Exercises

Here is another way in which $\mathbb{Z}[i]$ throws light on sums of two squares. The following exercises develop a proof of a theorem of Euler (1747): *if* $\gcd(a,b) = 1$ *then any divisor of* $a^2 + b^2$ *is of the form* $c^2 + d^2$ *where* $\gcd(c,d) = 1$. The main steps depend on unique prime factorization in $\mathbb{Z}[i]$.

**6.5.1** Give an example that shows why the condition $\gcd(a,b) = 1$ is necessary.

**6.5.2** Show that each integer divisor $e > 1$ of $a^2 + b^2$ is a product of Gaussian prime divisors $q + ir$ of $a^2 + b^2$, unique up to unit factors.

**6.5.3** Show that each of the Gaussian primes $q + ir$ divides either $a - ib$ or $a + ib$. Deduce that none of them is an ordinary prime $p$.

**6.5.4** Show that, along with each Gaussian prime factor $q + ir$ of $e$, its conjugate $q - ir$ is also a factor.

**6.5.5** Deduce from Exercise 6.5.4 that $e$ is of the form $c^2 + d^2$ where $c + di$ divides $a + bi$.

**6.5.6** Deduce from Exercise 6.5.5 that $\gcd(c,d) = 1$.

## 6.6 Pythagorean triples

Now is a good time to revisit the primitive Pythagorean triples, whose relationship with $\mathbb{Z}[i]$ was suggested in Section 1.8. Since odd squares are congruent to 1 (mod 4) and even squares are congruent to 0 (mod 4), a sum of two odd squares is not a square. Hence in a primitive triple $(x,y,z)$ one of $x$, $y$ is even and $z$ is odd. The argument in Section 1.8 was that if

$$x^2 + y^2 = z^2$$

then

$$(x - yi)(x + yi) = z^2,$$

so $x - yi$ and $x + yi$ are Gaussian prime factors of an odd square, $z^2$. Then we wanted to say that:

1. If $x$ and $y$ are relatively prime (in $\mathbb{Z}$) then so are $x-yi$ and $x+yi$ (in $\mathbb{Z}[i]$).

2. In $\mathbb{Z}[i]$, relatively prime factors of a square are squares.

The first statement is correct. If $\gcd(x,y)=1$ in $\mathbb{Z}$ then $\gcd(x,y)=1$ in $\mathbb{Z}[i]$. This is so since a common Gaussian prime divisor is accompanied by its conjugate, and their product is a common divisor $>1$ in $\mathbb{Z}$. A common divisor of $x-yi$ and $x+yi$ also divides their sum $2x$ and difference $2iy$. Therefore, since $\gcd(x,y)=1$, any common prime divisors of $x-iy$ and $x+iy$ are primes $\pm 1 \pm i$ dividing 2. No such divisors are present, since they imply that $(x-iy)(x+iy)=z^2$ is even.

The second statement is not quite correct, but the following amendment of it is: *relatively prime factors of a square are squares, up to unit factors*. This follows from unique prime factorization in $\mathbb{Z}[i]$.

Since $x-yi$ and $x+yi$ have no common Gaussian prime factor, while each prime factor of $z^2$ occurs to an even power, each prime factor of $x-yi$, and each prime factor of $x+yi$, must also occur to an even power. A product of primes, each occurring to an even power, is obviously a square (compare with the same argument for natural numbers in Section 2.5). Hence each of $x-yi$ and $x+yi$ is a unit times a square, since their only possible nonprime factors are units. □

The amended second statement is good enough to give us the conclusion we expect. We have shown that $x-yi$ is a unit times a square, hence it is one of

$$(s-ti)^2, \quad -(s-ti)^2, \quad i(s-ti)^2, \quad -i(s-ti)^2, \quad \text{for some } s,t \in \mathbb{Z}.$$

That is, it is one of

$$(s^2-t^2)-2sti, \quad t^2-s^2+2sti, \quad 2st+(s^2-t^2)i, \quad -2st+(t^2-s^2)i.$$

In each case, equating real and imaginary parts gives one of $x$ and $y$ in the form $u^2-v^2$ and the other in the form $2uv$ for some natural numbers $u$ and $v$. Thus the result is essentially the same as that obtained by the loose argument in Section 1.8, but better, because it does not force the even member of the pair, $2uv$, to be first.

Moreover, we necessarily have $\gcd(u,v)=1$ because any common prime divisor of $u$ and $v$ is a common divisor of $u^2-v^2$ and $2uv$, hence of $x$ and $y$. Thus the correct outcome of the speculation in Section 1.8 is:

**Primitive Pythagorean triples.** *If $x^2+y^2=z^2$ for some relatively prime natural numbers $x$ and $y$, then one of $x$ and $y$ is of the form $u^2-v^2$ and the other of the form $2uv$, for relatively prime natural numbers $u$ and $v$.* □

We also find in each case that $z=u^2+v^2$, because

$$(u^2-v^2)^2+(2uv)^2=u^4+2u^2v^2+v^4=(u^2+v^2)^2.$$

Thus $z$ is a sum of two squares. Since $u$ and $v$ are any relatively prime numbers, and a prime $u^2+v^2$ necessarily has $\gcd(u,v)=1$, $z$ can be any *prime* sum of two squares. Thus we get a geometric characterization of the primes that are sums of two squares.

**Prime hypotenuses.** *The primes that are sums of two squares are those that occur as hypotenuses of right-angled triangles with integer sides.* □

## Exercises

The last result, together with Fermat's two square theorem, shows that the primes of the form $4n+1$ are precisely those occurring as hypotenuses of integer right-angled triangles.

**6.6.1** Find integer right-angled triangles with hypotenuses 5, 13, 17 (you should know these), and 29, 37, and 41.

**6.6.2** Given a prime $p=4n+1$, is the integer right-angled triangle with hypotenuse $p$ unique?

The argument above shows that, if $(x,y,z)$ is a primitive Pythagorean triple, then $x+yi$ is a unit times a square in $\mathbb{Z}[i]$. But once we know that $x=u^2-v^2$, $y=2uv$ we can say more.

**6.6.3** If $(x,y,z)$ is a primitive Pythagorean triple with $x$ odd, show that $x+yi$ is a square in $\mathbb{Z}[i]$.

**6.6.4** Verify directly that $3+4i$ is a square in $\mathbb{Z}[i]$.

It should be clear from your answer to Question 6.6.3 that finding the parameters $u$ and $v$ for a given primitive Pythagorean triple $(x,y,z)$, with $x$ odd, is equivalent to finding the square root(s) of a complex number.

**6.6.5** Find the square root of $5+12i$.

**6.6.6** If you have some software for computing square roots of complex numbers, verify that each entry $(x,y,z)$ in Plimpton 322 (Section 1.6), except the triple $(60,45,75)$, yields a $y+xi$ that is a square in $\mathbb{Z}[i]$. (Note: this includes the last triple $(90,56,106)$, which is clearly not primitive.)

**6.6.7** Explain how to compute the square root of a complex number by hand, using quadratic equations.

## 6.7 *Primes of the form $4n+1$

Lagrange's lemma, proved in the Section 6.5, is actually half of an important result concerning the so-called "quadratic character of $-1$" that we study further in Chapter 9. Here we use it to prove that there are infinitely many primes of the form $4n+1$, complementing the corresponding easy result about primes of the form $4n+3$ proved in Exercises 6.3.4–6.3.6.

**Quadratic character of $-1$.** *The congruence $x^2 \equiv -1 \pmod{p}$, where $p$ is an odd prime, has a solution precisely when $p = 4n+1$.*

**Proof.** When $p = 4n+1$, Lagrange's lemma gives an $x$ with $x^2 \equiv -1 \pmod{p}$. To show that $x^2 \equiv -1 \pmod{p}$ has no solution when $p = 4n+3$ we suppose, on the contrary, that it does.

If

$$x^2 \equiv -1 \pmod{p = 4n+3}$$

then raising both sides to the power $2n+1$ gives

$$(x^2)^{2n+1} \equiv (-1)^{2n+1} \equiv -1 \pmod{p = 4n+3}.$$

Since $2(2n+1) = 4n+2 = p-1$, this says that

$$x^{p-1} \equiv -1 \pmod{p},$$

contrary to Fermat's little theorem. Hence $x^2 \equiv -1 \pmod{p}$ has no solution when $p = 4n+3$. □

Thus solutions of $x^2 \equiv -1 \pmod{p}$ occur precisely when the odd prime $p$ is of the form $4n+1$. To put it another way: *the odd primes $p$ that divide values of $x^2+1$, for $x \in \mathbb{Z}$, are precisely the primes $p = 4n+1$.*

**Infinitude of primes $4n+1$.** *There are infinitely many primes of the form $p = 4n+1$.*

**Proof.** From what we have just proved, it suffices to show that infinitely many primes divide values of $x^2+1$ for $x \in \mathbb{Z}$. Suppose on the contrary that only finitely many primes $p_1, p_2, \ldots, p_k$ divide values of $x^2+1$.

Now consider the polynomial

$$(p_1 p_2 \cdots p_k y)^2 + 1 = g(y).$$

Clearly, any prime $p$ that divides a value of $g(y)$, for $y \in \mathbb{Z}$, also divides a value of $x^2+1$ (namely, for $x = p_1 p_2 \cdots p_k y$). But none of $p_1, p_2, \ldots, p_k$ divides $g(y)$, because each leaves remainder 1.

Therefore, *no* prime divides $g(y)$, for any $y \in \mathbb{Z}$, and hence the only possible values of the integers $g(y)$ are $\pm 1$. In other words,

$$(p_1 p_2 \cdots p_k y)^2 + 1 = \pm 1 \quad \text{for all } y \in \mathbb{Z}.$$

But this is absurd, because each of the quadratic equations

$$(p_1 p_2 \cdots p_k y)^2 + 1 = 1 \quad \text{and} \quad (p_1 p_2 \cdots p_k y)^2 + 1 = -1$$

has at most two solutions $y$. This contradiction shows that $x^2 + 1$ is divisible by infinitely many primes, as required. □

It now follows, by Fermat's two square theorem, that infinitely many primes are sums $a^2 + b^2$ of two squares. Hence there are infinitely many Gaussian primes $a + ib$ that are neither real nor pure imaginary.

## Exercises

The argument just used to prove that $x^2 + 1$ is divisible by infinitely many primes can be generalized to any nonconstant polynomial $f(x)$ with integer coefficients. We suppose that

$$f(x) = a_m x^m + \cdots + a_1 x + a_0, \quad \text{where } a_0, a_1, \ldots, a_m \in \mathbb{Z} \text{ and } a_0, a_m \neq 0,$$

has values divisible only by the primes $p_1, p_2, \ldots, p_k$, and consider the polynomial

$$f(a_0 p_1 p_2 \cdots p_k y) = a_0 g(y),$$

where $g(y)$ is a polynomial of degree $m$.

**6.7.1** Show that $g(y)$ has integer coefficients, constant term 1, and that any prime dividing a value of $g(y)$, for $y \in \mathbb{Z}$, also divides a value of $f(x)$, for $x \in \mathbb{Z}$.

**6.7.2** Show, however, that none of $p_1, p_2 \ldots, p_k$ divide $g(y)$ when $y \in \mathbb{Z}$.

**6.7.3** Deduce from Exercise 6.7.2 that $g(y) = \pm 1$ for any $y \in \mathbb{Z}$.

**6.7.4** Show that the equations $g(y) = 1$ and $g(y) = -1$ have only finitely many solutions, which contradicts Exercise 6.7.3. (Where have you assumed that $f(x)$ is nonconstant?)

This contradiction shows that $f(x)$ is divisible by infinitely many primes. But now notice: we did *not* assume that infinitely many primes exist, hence this is a self-contained proof of Euclid's theorem that there are infinitely many primes.

**6.7.5** Is this argument essentially different from Euclid's?

## 6.8 Discussion

The two square theorem was stated without proof by Fermat in 1640, though he claimed to have a proof by descent: assuming there is a prime that is of the form $4n+1$ but *not* a sum of two squares he could show that there is a smaller prime with the same property. The first known proof of the theorem was in fact by descent, and published by Euler (1755). It cost him several years of effort.

Today it is possible to give quite simple proofs with the help of the result we called Lagrange's lemma in Section 6.5. Lagrange himself proved this lemma by means of Fermat's little theorem and his own theorem on the number of solutions of congruences mod $p$ (Section 3.5).

Lagrange (1773) used his lemma together with his theory of equivalence of quadratic forms (Section 5.6) to give a new proof of the two square theorem. The part of the proof involving quadratic forms was simplified by Gauss (1801), long before his creation of the Gaussian integers. It seems that Gauss had the main results about $\mathbb{Z}[i]$, including unique prime factorization, around 1815, but they were first published in 1832. The proof used in this chapter, combining unique prime factorization in $\mathbb{Z}[i]$ with Lagrange's lemma, is due to Dedekind (1894).

Yet another popular proof uses the *geometry of numbers*, developed in the 1890s by Minkowski. It may be found in Scharlau and Opolka (1985) together with a historical introduction to Minkowski's results.

Parallel to all the popular proofs of the two square theorem there are analogous proofs of the *four square theorem* of Lagrange (1770): *every natural number is the sum of* (at most) four natural number squares. Most use the following counterpart of Lagrange's lemma: any prime $p$ divides a number of the form $l^2+m^2+1$. The counterpart turns out to be easier. What is harder is the *four square identity* discovered by Euler (1748b). It is analogous to the two square identity of Section 6.1 but is much more complicated (see Section 8.3). It can be mechanically checked by multiplying out both sides, but what does it mean?

The Gaussian integer proof is favored in this book because $\mathbb{Z}[i]$ is a natural structure and the two square identity is a natural part of it—the multiplicative property of the norm—rather than an accidental identity of formal expressions. In Chapter 8 we give a similar "structural" proof of the four square theorem that uses the *quaternion integers*. These are a remarkable four-dimensional structure from which the four square identity

emerges naturally as the multiplicative property of the *quaternion norm*. Again, the key to the proof is unique prime factorization (or rather the prime divisor property, which happens to be somewhat easier than unique prime factorization in the quaternion case).

Fermat's two square theorem was generalized in another direction by Fermat himself. In 1654 Fermat announced similar theorems on primes of the forms $x^2+2y^2$ and $x^2+3y^2$:

$$p = x^2+2y^2 \Leftrightarrow p = 8n+1 \text{ or } p = 8n+3,$$
$$p = x^2+3y^2 \Leftrightarrow p = 3n+1.$$

Our proof of the two square theorem in Section 6.5 can be adapted to Fermat's $x^2+2y^2$ and $x^2+3y^2$ theorems with the help of unique prime factorization theorems for numbers of the forms $a+b\sqrt{-2}$ and $a+b\sqrt{-3}$ respectively. Such theorems will be proved in the next chapter.

The other thing we need to adapt is Lagrange's lemma: if $p = 4n+1$ then $p$ divides a number of the form $m^2+1$ for some $m \in \mathbb{Z}$. In Section 6.7 we described this lemma (together with its converse) as the *quadratic character of* $-1$ because it says that $-1$ is congruent to a square mod $p$ precisely when $p = 4n+1$.

To prove Fermat's theorems on primes of the form $x^2+2y^2$ and $x^2+3y^2$ we similarly need the quadratic characters of $-2$ and $-3$. They are:

$$-2 \equiv \text{square } (\text{mod } p) \Leftrightarrow p = 8n+1 \text{ or } 8n+3,$$
$$-3 \equiv \text{square } (\text{mod } p) \Leftrightarrow p = 3n+1.$$

Instead of finding quadratic characters one by one, in Chapter 9 we prove the sweeping *law of quadratic reciprocity*, which allows us to tell when any integer is congruent to a square mod $p$. Quadratic reciprocity was first observed by Euler and proved by him in special cases, such as those involved in Fermat's theorems. The first general proof is due to Gauss (1801), and quadratic reciprocity has since been proved in many different ways. In fact, it has been proved more often than any other theorem except its distant ancestor, the Pythagorean theorem.

Springer

http://www.springer.com/978-0-387-95587-2

Elements of Number Theory

Stillwell, J.
2003, XII, 256 p., Hardcover
ISBN: 978-0-387-95587-2