6 The Gaussian integers

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

In the last chapter we looked at *real* quadratic ring $\mathbb{Z}[\sqrt{n}]$ (meaning $\mathbb{Z}[\sqrt{n}] \subseteq \mathbb{R}$) to solve Pell's equation

$$x^2 - ny^2 = 1$$

because, over $\mathbb{Z}[\sqrt{n}]$, $x^2 - ny^2$ factors as

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

Precisely, we defined the conjugate of an element $\alpha = x + y\sqrt{n}$ to be $\overline{\alpha} = x - y\sqrt{n}$, and then the norm to bet

$$N : \mathbb{Z}[\sqrt{n}] \to \mathbb{Z}$$

by

$$N(\alpha) = \alpha\overline{\alpha} = x^2 - ny^2.$$

So the solutions to Pell's equation are precisely the elements of norm 1 in $\mathbb{Z}[\sqrt{n}]$, which we showed form a group under multiplication that is generated by two elements ϵ and -1.

Similarly if one wants to study the equation

$$x^2 + ny^2 = k$$

it makes sense to look at the imaginary quadratic ring $\mathbb{Z}[\sqrt{-n}]$ (meaning $\mathbb{Z}[\sqrt{-n}] \subseteq \mathbb{C}$ and $\mathbb{Z}[\sqrt{-n}] \nsubseteq \mathbb{R}$). The imaginary quadratic rings can be treated in the same basic way as the real quadratic rings theoretically, however their flavor is quite different. For example $x^2 + ny^2 = 1$ has only finitely many solutions, but $x^2 - ny^2 = 1$ has infinitely many (for n nonsquare).

For now, we will just treat the simplest case, the *Gaussian integers*, which were first studied in detail by Gauss. This ring is related to questions about Pythagorean triples, and more generally, which numbers are sums of two squares.

Definition 6.1. *The* Gaussian integers *are the ring*[1]

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

For $\alpha = a + bi$, the conjugate of α *is $\overline{\alpha} = a - bi$, and the* norm *is*

$$N(\alpha) = |\alpha|^2 = \alpha\overline{\alpha} = a^2 + b^2.$$

Note that if we draw α as a vector in the complex plane, $c = |\alpha|$ denotes the length of this vector, so the norm of α is just the square of the length of the vector α, i.e., $N(\alpha) = c^2$. Hence the formula for the norm is precisely the Pythagorean theorem:

$$\alpha\overline{\alpha} = a^2 + b^2 = c^2 = N(\alpha).$$

[1] As before, I will sometime call a set of numbers a "ring" to stress that it is like $\mathbb{Z}$ in some sense. It contains 0, 1 and -1, and you can add, subtract, and multiply as in $\mathbb{Z}$. The formal definition comes in Chapter 10.

The reason we want the norm to be the square of the length, instead of just the length is because c^2 is always an integer, but c rarely is, e.g., $1^2 + 1^2 = \sqrt{2}^2$. With this definition, the norm is a map from the ring $\mathbb{Z}[i]$ into the ring of integers $\mathbb{Z}$,

$$N : \mathbb{Z}[i] \to \mathbb{Z}.$$

(In fact the norm will map into $\mathbb{N} \cup \{0\}$, unlike in the case of $\mathbb{Z}[\sqrt{-n}]$.) The fact that it maps into the integers is very important because of the following.

Proposition 6.2. *$N : \mathbb{Z}[i] \to \mathbb{Z}$ is multiplicative, i.e., $N(\alpha\beta) = N(\alpha)N(\beta)$ for any $\alpha, \beta \in \mathbb{Z}[i]$.*

Proof. We did this twice already, in Chapters 1 and 5. Check that $\overline{\alpha\beta} = \overline{\alpha} \cdot \overline{\beta}$ by simple multiplication. Then

$$N(\alpha\beta) = \alpha\beta\overline{\alpha\beta} = \alpha\overline{\alpha}\beta\overline{\beta} = N(\alpha)N(\beta).$$

□

Hence, through the norm map $N : \mathbb{Z}[i] \to \mathbb{Z}$, we will be able to relate the multiplicative structure on $\mathbb{Z}$ to the multiplicative structure on $\mathbb{Z}[i]$ (and hence also the theory of divisibility, primes, etc.). In general, the norm is a fundamental tool in algebraic number theory.

Definition 6.3. *Let $\alpha \in \mathbb{Z}[i]$. We say α is a* unit *of $\mathbb{Z}[i]$ if $N(\alpha) = 1$.*

The units of $\mathbb{Z}[i]$ are $\pm 1, \pm i$. (See exercise below.)

Exercise 6.1. *Exercises 6.1.1, 6.1.2 (be precise), 6.1.3.*

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

Recall that the notion of divisibility in $\mathbb{Z}$ is defined as follows: if $a, b \in \mathbb{Z}$, we say $b|a$ if $a = bc$ for some $c \in \mathbb{Z}$. (Formally we only defined it for $\mathbb{N}$, but this definition is the same.) Note that this techincally means (i) any number divides 0, and (ii) if $d|n$, then $(-d)|n$.

Similarly we say $p \in \mathbb{Z}$ is *prime* if (i) $p \neq 0$, (ii) $p \neq \pm 1$ and (iii) p has no factors other than ± 1 and $\pm p$, i.e., up to ± 1, the *units* of $\mathbb{Z}$. In other words, the primes p in $\mathbb{Z}$ are the (non-zero) non-units which only have the trivial factorizations $p = 1 \cdot p$ and $p = (-1)(-p)$. Hence, for us, the primes of $\mathbb{Z}$ are

$$\ldots, -5, -3, -2, 2, 3, 5, \ldots.$$

We define divisibility in $\mathbb{Z}[i]$ the exact same way.

Definition 6.4. *Let $\alpha, \beta \in \mathbb{Z}[i]$. We say α* divides *β, or $\beta|\alpha$, if*

$$\alpha = \beta\gamma$$

for some $\gamma \in \mathbb{Z}[i]$. We say $\alpha \neq 0$ is a Gaussian prime*, or* prime in $\mathbb{Z}[i]$*, if α is a non-unit such that the only divisors of α are $\pm 1, \pm i$, $\pm\alpha$ and $\pm i\alpha$, i.e., if the only divisiors of α, up to units, are* 1 *and α.*

Equivalently, we can say α is prime in $\mathbb{Z}[i]$ if (i) $\alpha \neq 0$ (ii) α is not a unit ($N(\alpha) \neq 1$), and (iii) the only factorzations $\alpha = \beta\gamma$ are the trivial ones, i.e., where one of β and γ is a unit (and therefore the other is a unit times α).

We can relate divisibility in $\mathbb{Z}[i]$ with divisibility in $\mathbb{Z}$ via the norm map.

Lemma 6.5. *If $\beta|\alpha$ in $\mathbb{Z}[i]$, then $N(\beta)|N(\alpha)$ in $\mathbb{Z}$.*

Proof. Suppose $\beta|\alpha$. Then $\alpha = \beta\gamma$, which implies $N(\alpha) = N(\beta)N(\gamma)$ by multiplicativity of the norm. Hence $N(\beta)|N(\alpha)$. □

Example. *$\alpha = 4 + i$ is a Gaussian prime.*

Proof. The reason is $N(\alpha) = 4^2 + 1^2 = 17$ which is prime in $\mathbb{Z}$. Hence if $\beta|\alpha$, we can write $\alpha = \beta\gamma$, and $N(\beta)N(\gamma) = N(\alpha) = 17$. The norm is always non-negative, so the only possibilities are $N(\beta) = 1, N(\gamma) = 17$ or $N(\gamma) = 1, N(\beta) = 17$. By symmetry, we may assume the former. Then β is a unit (± 1 or $\pm i$), so it has an inverse $\beta^{-1} \in \mathbb{Z}[i]$. Hence $\gamma = \beta^{-1}\alpha$, i.e., a unit times α.

In other words, the only divisors of α are the units and the units times α. □

Exercise 6.2. *Let $\alpha \in \mathbb{Z}[i]$. Show α is a unit of $\mathbb{Z}[i]$ if and only if α is invertible in $\mathbb{Z}[i]$, i.e., $\alpha\beta = 1$ for some $\beta \in \mathbb{Z}[i]$. (Hint: use the multiplicative property of the norm.) In fact, this is the general definition of a unit in a ring: something is a unit means its invertible (in the ring). The units always form a multiplicative group. The units of $\mathbb{Z}$ are just ± 1. The role of $\{1, -1, i, -i\}$ in $\mathbb{Z}[i]$ is exactly analogous to the role of ± 1 in $\mathbb{Z}$.*

The text defines α to be a Gaussian prime to be an element of $\mathbb{Z}[i]$ such that α is not a product of two elements of smaller norm. The following exercise shows our definition and the book's are equivalent.

Exercise 6.3. *Using the definition of Gaussian prime we gave in class, show the following is true: Suppose α is a non-zero, non-unit. Then α is a Gaussian prime if and only if $\beta|\alpha$ implies $N(\beta) = 1$ or $N(\beta) = N(\alpha)$. (Cf. Exercise 6.2.1. Hint: look at the example above.) Conclude that if $N(\alpha)$ is prime in $\mathbb{Z}$, α is a Gaussian prime.*

Example. *$\alpha = 2$ is not a Gaussian prime.*

Proof. Note $N(\alpha) = 2^2 = 4$, so the only reason it couldn't be a prime in $\mathbb{Z}[i]$ is if it is the product of two elements of norm 2. What are the elements $a + bi$ of norm 2? Well, $N(a + bi) = a^2 + b^2$, so there are 4 elements of norm 2: $1 + i$, $1 - i$, $-1 + i$, and $-1 - i$. We observe that

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

□

Example. *The elements of norm 2: $1 + i$, $1 - i$, $-1 + i$ and $-1 - i$ are all Gaussian primes. This follows from the previous exercise, because their norm is prime in $\mathbb{Z}$.*

At this point you might ask if any prime in $\mathbb{Z}$ remains prime in $\mathbb{Z}[i]$. The answer, reassuringly, is yes. In fact, in a way that can be made precise, exactly half of the primes are still prime in $\mathbb{Z}[i]$.

Exercise 6.4. *Show there are no elements in $\mathbb{Z}[i]$ whose norm is of the form $4n + 3$. Conclude that if $p = 4n + 3$ is prime in $\mathbb{Z}$, then p is also a Gaussian prime. (Cf. Section 6.3)*

We will soon see that the primes which do not remain prime in $\mathbb{Z}[i]$, are precisely the primes of the form $p = x^2 + y^2$, which will turn out to be 2 and all primes of the form $4n + 1$.

Proposition 6.6. (Existence of prime factorization.) *Let $\alpha \in \mathbb{Z}[i]$ be a non-zero, non-unit. Then α factors into a finite product of Gaussian primes.*

The proof is the same as for the ordinary integers.

Proof. Here it's easier to use the book's definition of Gaussian prime (though our definition is the more natural one). Either $\alpha = \beta\gamma$ where $N(\beta), N(\gamma) < N(\alpha)$ or not. If not, α is a Gaussian prime and we are done. If there is such a factorization, we repeat the process with β and γ. E.g., either $\beta = \beta'\beta''$ with $N(\beta'), N(\beta'') < N(\beta)$ or not. If not, β is a Gaussian prime. If so, we repeat with β', β''. Because the norms must get smaller and are always at least 1, this process terminates at some point, leaving us with a prime factorization of α in $\mathbb{Z}[i]$. □

Exercise 6.5. *Factor 17 and 53 in $\mathbb{Z}[i]$. (Exercise 6.2.4.)*

6.3 Conjugates

Proposition 6.7. *Let $p \in \mathbb{N}$ be prime. Then p is also prime in $\mathbb{Z}[i]$ if and only if p is not the sum of two squares.*

Proof. We will prove the negation of this statement, which is equivalent: p is a sum of two squares if and only if p is not prime in $\mathbb{Z}[i]$.

($\Rightarrow$) Suppose $p = a^2 + b^2$. Then $p = (a + bi)(a - bi)$. Since p is not a square, both a and b must be nonzero here, hence $a \pm bi$ is not a unit. Thus $(a + bi)(a - bi)$ is a non-trivial factorization of p in $\mathbb{Z}[i]$.

($\Leftarrow$) This is a typical use of the norm map. Suppose p is not prime in $\mathbb{Z}[i]$. Then $p = \alpha\beta$ where $\alpha, \beta \in \mathbb{Z}[i]$ are both non-units, i.e., neither has norm 1. But then $N(\alpha)N(\beta) = N(p) = p^2$, so $N(\alpha)|p^2$ but $N(\alpha), N(\beta) \neq 1$ implies $N(\alpha) = N(\beta) = p$ since p is prime in $\mathbb{Z}$. Writing $\alpha = a + bi$, $N(\alpha) = p$ means $p = a^2 + b^2$. (The exercise below says that in fact $\beta = \overline{\alpha}$.) □

Exercise 6.6. *Suppose $p \in \mathbb{N}$ is prime in $\mathbb{Z}$ but factors as $p = \alpha\beta$ where α, β are non-units in $\mathbb{Z}[i]$. Show $\beta = \overline{\alpha}$.*

In fact, a second look at the proof of ($\Leftarrow$) above gives the following.

Corollary 6.8. *Suppose $p \in \mathbb{N}$ is prime and $p = a^2 + b^2$. Then $\alpha = a + bi$ and $\overline{\alpha} = a - bi$ are primes in $\mathbb{Z}[i]$.*

Proof. We know $N(\alpha) = N(\overline{\alpha}) = p$, i.e., their norms are prime in $\mathbb{N}$. This means α and $\overline{\alpha}$ are prime by Exercise 6.3. □

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

Recall that the way we proved unique factorization in $\mathbb{Z}$ was via the prime divisor property ($p|ab \implies p|a$ or $p|b$), which we proved using the Euclidean algorithm. We will show unique factorization in $\mathbb{Z}[i]$ in the same way.

Proposition 6.9. *If $\alpha, \beta \in \mathbb{Z}[i]$ are non-zero, then there is a quotient $\mu \in \mathbb{Z}[i]$ and remainder $\rho \in \mathbb{Z}[i]$ such that*

$$\alpha = \mu\beta + \rho \text{ where } N(\rho) < N(\beta).$$

Proof. In the complex plane, multiplication by i is just rotation by 90 degrees. Hence the vectors β and $i\beta$ are orthogonal. Now the set of $\mu\beta$ where $\mu \in \mathbb{Z}[i]$ is just the set of integer linear combinations of β and $i\beta$, i.e., the square lattice (since β and $i\beta$ have the same length $|\beta|$) in $\mathbb{C}$ generated by β and $i\beta$.

Now α lies in one of the squares in the lattice (possibly on a line in the lattice), but the furthest α can be from all 4 corners of a square with side length $|\beta|$ is $|\beta|/\sqrt{2}$ (when α is at the center). Hence α is within distance $|\beta|/\sqrt{2}$ to some (not necessarily unique) corner $\mu\beta$.

Then $\rho = \mu\beta - \alpha$ has length at most $|\beta|/\sqrt{2}$. Since $N(\rho) = |\rho|^2$, we see $N(\rho) \leq |\beta|^2/2 = N(\beta)/2$. □

As before the remainder ρ is 0 if and only if $\beta|\alpha$, however the μ and ρ are no longer uniquely determined.

Example. *$\alpha = 1 + 3i$, $\beta = 2 + 2i$. Then we can write $\alpha = 1 \cdot \beta + (i - 1) = (i + 1)\beta + (1 - i)$.*

However, what is important for us is the following.

Lemma 6.10. *Suppose $\alpha = \mu\beta + \rho$ where $N(\rho) < N(\beta)$. Then $\rho = 0$ iff $\beta|\alpha$.*

Proof. ($\Rightarrow$) Clearly if $\rho = 0$, $\beta|\alpha$.

($\Leftarrow$) Suppose $\beta|\alpha$ but $\rho \neq 0$, i.e., $\alpha = (\mu + \tau)\beta$ for some non-zero $\tau \in \mathbb{Z}[i]$, so $\rho = \tau\beta$. But then $N(\beta)|N(\rho) \implies N(\rho) \geq N(\beta)$, a contradiction. Hence $\rho = 0$. □

Definition 6.11. *Suppose $\delta|\alpha$ and $\delta|\beta$. We say δ is a* common divisor *of α and β. We say δ is a* greatest common divisor (gcd) *of α and β if it is a common divisor of α and β with maximum possible norm.*

Note that the gcd is not unique here. For example, $\pm 5, \pm 5i$ are all gcd's of 5 and 10. It is not even clear *a priori* if the gcd is always unique up to a unit. This will follow from the prime divisor property or unique factorization.

Lemma 6.12. *Suppose $\alpha = \mu\beta + \rho$ where $N(\rho) < N(\beta)$. If $\rho = 0$, β is a gcd for α and β. If not, a gcd for α and ρ is also a gcd for α and β (and vice versa).*

Proof. The first statement is clear from the above lemma, because no divisor of β can have norm larger than $N(\beta)$.

So suppose $\rho \neq 0$. Any common divisor of α and β must be a common divisor of α and ρ, and vice versa. This implies that any gcd of α and β will be a gcd of α and ρ, and vice versa. □

This lemma shows that the following algorithm to determine a gcd for α and β in $\mathbb{Z}[i]$ always works. Just as in case of $\mathbb{Z}$, it terminates in a finite number of steps by decent on the norm of the remainder.

Euclidean algorithm for $\mathbb{Z}[i]$

1. Suppose $N(\alpha) \geq N(\beta)$. Let $(\alpha_1, \beta_1) = (\alpha, \beta)$. Set $i = 1$.

2. Write $\alpha_i = \mu_i\beta_i + \rho_i$ where $N(\rho_i) < N(\beta_i)$. If $\rho_i = 0$, then β_i is a gcd for α and β. If not, continue.

3. Let $\alpha_{i+1} = \beta_i, \beta_{i+1} = \rho_i$. Note that $N(\alpha_{i+1}) > N(\beta)_{i+1}$ from Step 2. Increase i by 1 and repeat Step 2 (i.e., repeat Step 2 for α_{i+1} and β_{i+1}).

Exercise 6.7. *Use the Euclidean algorithm to determine a gcd for $\alpha = 5$ and $\beta = 3+4i$ in $\mathbb{Z}[i]$.*

Just like in Section 2.3, we can run the Euclidean algorithm in reverse to get that some gcd of α and β must be of the form $\mu\alpha + \nu\beta$ for some $\mu, \nu \in \mathbb{Z}[i]$.

Proposition 6.13. (Prime divisor property) *Let π be a prime in $\mathbb{Z}[i]$. If $\pi | \alpha\beta$ then $\pi|\alpha$ or $\pi|\beta$.*

Proof. Suppose $\pi|\alpha\beta$, but $\pi \nmid \alpha$. We want to show $\pi|\beta$. This is the essentially the same argument we gave for $\mathbb{N}$.

Since $\pi \nmid \alpha$ and π is prime in $\mathbb{Z}[i]$, any gcd of π and α must be a unit. (If not, it would have to be a non-trivial divisor of π.) By the above, some unit u (some gcd) is a linear combination of π and α, i.e., $u = \mu\pi + \nu\alpha$ for some $\mu, \nu \in \mathbb{Z}[i]$. Hence

$$u\beta = \mu\pi\beta + \nu\alpha\beta.$$

But π divides both terms on the right, $\pi|u\beta$, so $\pi|\beta$ by the simple exercise below. □

Exercise 6.8. *Let $\pi, \beta \in \mathbb{Z}[i]$ and u be a unit of $\mathbb{Z}[i]$. Show that $\pi|u\beta \iff \pi|\beta$.*

Another simple exercise is the following, which we will use in the proof of unique factorization.

Exercise 6.9. *Suppose π and π' are primes of $\mathbb{Z}[i]$. Show $\pi|\pi'$ implies $\pi = u\pi'$ where u is a unit of $\mathbb{Z}[i]$.*

Theorem 6.14. (Unique factorization) *Let $\alpha \neq 0$ be a non-unit in $\mathbb{Z}[i]$. Suppose $\alpha = \pi_1 \cdots \pi_m$ and $\alpha = \pi_1' \cdots \pi_n'$ are two factorizations of α into Gaussian primes π_i and π_j'. Then $m = n$, and up to a reordering of the π_j''s, we have*

$$\pi_i = u_i \pi_i'$$

for each i, where u_i is a unit in $\mathbb{Z}[i]$.

Proof. The proof again is almost exactly the same as in the case of $\mathbb{N}$. Suppose the theorem is false. Cancelling any common factors, assume

$$\pi_1 \cdots \pi_m = \pi_1' \cdots \pi_n'$$

where $\pi_i \neq u\pi_j'$ for any i, j and unit u. Clearly $\pi_1 | \pi_1' \cdots \pi_n'$. Since π_1 is prime, by the prime divisor property $\pi_1|\pi_1'$ or $\pi_1|(\pi_2' \cdots \pi_n')$. Now $\pi_1|\pi_1'$ is impossible since $\pi_1 \neq u\pi_1'$ for any unit u. Hence $\pi_1|(\pi_2' \cdots \pi_n')$. Repeating this argument eventually gives $\pi_1|\pi_n'$, a contradiction. □

Example. *Let $\alpha = 6i = 2 \cdot 3i$. In $\mathbb{Z}[i]$, 3 remains prime since it is not a sum of two squares. This means $3i$ is also a prime in $\mathbb{Z}[i]$ since i is just a unit. On the other hand, $2 = (1+i)(1-i)$ and we know $1 \pm i$ are prime in $\mathbb{Z}[i]$ because their norms (both 2) are prime in $\mathbb{N}$. Unique factorization says that*

$$6i = (1+i)(1-i)(3i)$$

is the only factorization of 6 into primes of $\mathbb{Z}[i]$, up to reordering and multiplying each factor by a unit, i.e., we consider

$$6i = (-3)(i-1)^2$$

to be equivalent to the factorzation above since $-3 = i \cdot (3i)$, $(i-1) = i \cdot (1+i)$ and $(i-1) = -(1-i)$.

Exercise 6.10. *Let u be a unit in $\mathbb{Z}[i]$. Show π is prime in $\mathbb{Z}[i]$ if and only if $u\pi$ is prime in $\mathbb{Z}[i]$.*

We say the primes π and $u\pi$ (u a unit) are **associated**. Hence we say primes π and π' are **nonassociated** if they do not differ by units. Just as in the case of $\mathbb{Z}$, unique factorization implies the following.

Corollary 6.15. *Let $\pi_1, \ldots, \pi_n$ be nonassociated primes of $\mathbb{Z}[i]$. Suppose $\alpha = \pi_1^{e_1} \cdots \pi_n^{e_n}$ and $\beta = \pi_1^{f_1} \cdots \pi_n^{f_n}$ (where $e_i, f_i \geq 0$). Then any gcd of α and β equals a unit times $\pi_1^{\min(e_1,f_1)} \cdots \pi_n^{\min(e_n,f_n)}$.*

Hence "the" gcd of two elements is unique up to units. Moreover, it satisfies the property that if μ and ν are relatively prime (i.e., their gcd is 1) common divisors of α and β, then $\mu\nu$ divides "the" gcd of α and β.

Example. *To see what can go wrong with the* gcd *when we don't have unique factorization, consider the example of*

$$6 = 2 \cdot 3 = (1+\sqrt{-5})(1-\sqrt{-5}) \in \mathbb{Z}[\sqrt{-5}].$$

Then what is the gcd *of $\alpha = 6$ and $\beta = 2(1+\sqrt{-5})$? Both 2 and $1+\sqrt{-5}$ (which themselves are relatively prime) are common divisors, but their product is not a common divisor. In this specific example, we could say that $1+\sqrt{-5}$ should be "the"* gcd *because it has larger norm, but philosophically 2 should also be a* gcd*, because no multiple of 2 divides both α and β.*

Proposition 6.16. *The Gaussian primes, up to units, are precisely the following: (i) primes $p \in \mathbb{N}$ not of the form $x^2 + y^2$, (ii) $\alpha = a + bi$ where $N(\alpha)$ is prime in $\mathbb{N}$.*

Proof. We know that elements of type (i) are prime in $\mathbb{Z}[i]$ by Section 6.3, and elements of type (ii) are prime in $\mathbb{Z}[i]$ by Exercise 6.3. So we just have to show any prime α of $\mathbb{Z}[i]$ is, up to a unit, of type (i) or (ii).

Consider $N(\alpha) = \alpha\overline{\alpha} = n$. Since $\overline{\alpha}$ is also prime, $\alpha\overline{\alpha}$ is the unique prime factorization of n in $\mathbb{Z}[i]$. Either n is prime in $\mathbb{N}$ or not. If it is, then α is of type (ii) and we are done. If not, then n has a non-trivial factorization $n = ab$ in $\mathbb{N}$. But since any prime factorization of n in $\mathbb{Z}[i]$ has only two elements, a and b must be prime in $\mathbb{Z}[i]$, and the factorizations $n = \alpha\overline{\alpha} = ab$ are the same. Thus $u\alpha \in$ equals a or b for some unit u. Moreover $u\alpha$ must be prime in $\mathbb{N}$ since it is prime in $\mathbb{Z}[i]$. Hence $u\alpha$ is of type (i) by Section 6.3. □

Exercise 6.11. *Suppose α is prime in $\mathbb{Z}[i]$. Show $N(\alpha) = p$ or $N(\alpha) = p^2$ for some prime p of $\mathbb{N}$.*

6.5 Fermat's two square theorem

Now we can apply the theory of $\mathbb{Z}[i]$ to the classical problem, which number k are sums of two squares, i.e., for which k is

$$Q(x,y) = x^2 + y^2 = k$$

solvable for $x, y \in \mathbb{Z}$. Just by calculation, we can check that the numbers $1 \leq k \leq 50$ which are sums of two squares are

$$1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26, 29, 32, 34, 36, 37, 40, 41, 45, 49, 50.$$

Here there is no apparent pattern, except perhaps that, as we have seen before, this sequence avoids the numbers of the form $4n + 3$. Now, if I ask you to check which primes $1 \leq p \leq 50$ are sums of two squares you will see

$$2, 5, 13, 17, 29, 37, 41$$

so the primes < 50 which are not sums of two squares are precisely

$$3, 7, 11, 19, 23, 31, 43, 47.$$

Here, if we look at the mod 4 congruence classes we see the patter than $p = x^2 + y^2$ if and only if $p \not\equiv 3 \bmod 4$. This is the content of Fermat's two square theorem, and it will answer the more general question about when is $n = x^2 + y^2$.

Lemma 6.17. (Lagrange) *A prime $p = 4n + 1$ divides $m^2 + 1$ for some $m \in \mathbb{Z}$.*

Proof. By Wilson's theorem

$$(4n)! \equiv -1 \bmod p.$$

Note

$$(4n)! \equiv (2n)! \times (2n+1)(2n+2)\cdots 4n \equiv (2n)! \times (-2n)(-2n+1)\cdots(-1) \bmod p.$$

Hence

$$(4n)! \equiv (2n)!(-1)^{2n}(2n)! \equiv ((2n)!)^2 \bmod p.$$

Thus $p|((2n)!)^2 + 1$. □

Theorem 6.18. (Fermat) *Let p be prime. Then $p = x^2 + y^2$ for some $x, y \in \mathbb{Z}$ if and only if $p = 2$ or $p \equiv 1 \bmod 4$.*

Proof. ($\Rightarrow$) Suppose $p = x^2 + y^2$. We saw in Section 3.7 that $p \not\equiv 3 \bmod 4$. However no prime is 0 mod 4, and 2 is the only prime which is 2 mod 4.

($\Leftarrow$) We know $2 = 1^2 + 1^2$ so suppose $p = 4n + 1$ for some n. By Lagrange's lemma, $p|m^2 + 1$ for some $m \in \mathbb{Z}$. If p were a Gaussian prime, then $p|m^2 + 1 = (m+i)(m-i)$ would imply $p|(m+i)$ or $p|(m-i)$ in $\mathbb{Z}[i]$. However this is not the case since $\frac{m}{p} \pm \frac{i}{p} \notin \mathbb{Z}[i]$. So p is not a Gaussian prime, i.e., p is the sum of two squares by Section 6.3. □

Corollary 6.19. *Let p be a prime which is not 3 mod 4. Then $p = x^2 + y^2$ has exactly 1 solution for $x, y \in \mathbb{N}$ with $x \geq y$.*

Proof. There is a solution $x^2 + y^2 = p$ by the theorem. We know that $p = (x + yi)(x - yi)$ is a factorization into primes in $\mathbb{Z}[i]$. (Section 6.3) If we also have $p = x'^2 + y'^2$ then $p = (x' + y'i)(x' - y'i)$ is also a prime factorization in $\mathbb{Z}[i]$. But since prime factorization is unique, we have $x' + y'i = u(x + yi)$ or $u(x - yi)$ for some unit u.

Suppose the former. If $u = \pm 1$ then $x' = \pm x$ and $y' = \pm y$. If $u = \pm i$ then $u = \pm i$, then $x' = \mp y$ and $y' = \pm x$. Hence x' and y' are up to sign, x and y in some order. The case $x' + y'i = u(x - yi)$ is similar. So there is only one solution to $p = x^2 + y^2$ with $x, y > 0$, up to interchanging x and y. □

Corollary 6.20. *The primes of $\mathbb{Z}[i]$ are precisely units times (i) p where $p = 4n + 3$ is a prime in $\mathbb{N}$, (ii) $\alpha = a + bi$ where $N(\alpha)$ is either 2 or a prime of $\mathbb{N}$ congruent to 1 mod 4.*

Proof. Combine Proposition 6.16 with Fermat's two square theorem. □

Theorem 6.21. *A natural number n is a sum of two squares if and only if any prime $p|n$ such that $p \equiv 3 \bmod 4$ occurs to an even power in the prime factorization of n.*

Proof. Clearly $n = x^2 + y^2$ if and only if $n = N(\alpha)$ where $\alpha = x + yi$, i.e., n is a sum of two squares if and only if it is the norm of some element of $\mathbb{Z}[i]$.

Let

$$n = \prod p_i^{e_i} \prod q_j^{f_j}$$

be the prime factorization of n in $\mathbb{N}$ where each $p_i \equiv 3 \bmod 4$ and $q_j \not\equiv 3 \bmod 4$. Each $q_j = \pi_j \overline{\pi}_j$ where π_j is some Gaussian prime. Thus

$$n = \prod p_i^{e_i} \prod \pi_j^{f_j} \overline{\pi}_j^{f_j}$$

is the prime factorization of n in $\mathbb{Z}[i]$, where each p_i is of type (i) and each π_j, $\overline{\pi}_j$ is of type (ii), in the notation of the above corollary.

($\Rightarrow$) Suppose $n = x^2 + y^2$, i.e., $n = N(\alpha)$ for some $\alpha \in \mathbb{Z}[i]$. Let

$$\alpha = \prod r_i^{h_i} \prod \phi_j^{k_j}$$

be the prime factorization of α in $\mathbb{Z}[i]$, again where each r_i is of type (i) and each ϕ_j is of type (ii). Then

$$\overline{\alpha} = \prod \overline{r}_i^{h_i} \prod \overline{\phi}_j^{k_j} = \prod r_i^{h_i} \prod \overline{\phi}_j^{k_j}$$

is the prime factorization of $\overline{\alpha}$ since each $r_i \in \mathbb{N}$. But then

$$\alpha\overline{\alpha} = \prod r_i^{2h_i} \prod \phi_j^{k_j} \overline{\phi}_j^{k_j} = n = \prod p_i^{e_i} \prod \pi_j^{f_j} \overline{\pi}_j^{f_j}$$

Hence up to reordering the primes (assuming all were distinct), we have $e_i = 2h_i$ is even, which is what we wanted to prove.

($\Leftarrow$) Homework. (This direction is easier and you don't need to use the unique factorization in $\mathbb{Z}[i]$.) □

This theorem, which was also known to Fermat, along with Corollary 6.19, shows the power of having unique factorization in $\mathbb{Z}[i]$. Now finish the proof yourself.

Exercise 6.12. *Suppose $n = \prod p_i^{2e_i} \prod q_j^{f_j}$ where each p_i, q_j are primes of $\mathbb{N}$ such that each $p_i \equiv 3 \bmod 4$ and $q_j \equiv 1, 2 \bmod 4$. (i) Show each $p_i^{2e_i}$ and $q_j^{f_j}$ is the norm of an element in $\mathbb{Z}[i]$. (ii) Deduce $n = x^2 + y^2$ for some $x, y \in \mathbb{Z}$.*

In fact, Gauss and Jacobi went further than Fermat's theorem and proved the following (which Fermat it seems also knew, though perhaps did not have a proof).

Theorem 6.22. *Let $r_2(n)$ denote the number of solutions to $x^2 + y^2 = n$. Write $n = 2^f n_1 n_2$ where n_1 is a product of primes $\equiv 1 \bmod 4$ and n_2 is a product of primes $\equiv 3 \bmod 4$. Then $r_2(n) = 0$ if n_2 is not a perfect square, and $r_2(n)$ is 4 times the number of divisors of n_1 otherwise.*

We will not prove this now (although we could using the unique factorization in $\mathbb{Z}[i]$), but one of my goals for next semester is to prove a theorem of Dirichlet which essentially contains this as a special case.

Exercise 6.13. *Find an n such that $n = x^2 + y^2$ in at least two distinct ways (with $x, y > 0$ and $x \geq y$). Write down all solutions (with $x, y > 0$, $x \geq y$). Using this, show there are two elements $\alpha, \beta \in \mathbb{Z}[i]$ such that $N(\alpha) = N(\beta)$ but α and β do not differ by units.*

6.6 Pythagorean triples

Definition 6.23. *Let $\alpha, \beta \in \mathbb{Z}[i]$. If the only common divisors of α and β are units, we say α and β are* **relatively prime.**

Lemma 6.24. *Suppose (x, y, z) is a primitive Pythagorean triple. Then $x + yi$ and $x - yi$ are relatively prime in $\mathbb{Z}[i]$, i.e., they have no common prime divisors in $\mathbb{Z}[i]$.*

Proof. Suppose instead, $x + yi$ and $x - yi$ have a common prime divisor $\pi \in \mathbb{Z}[i]$. Then π divides their sum $2x$ and their difference $2yi$. Since x and y have no common prime factors in $\mathbb{Z}$, they have no common prime factors in $\mathbb{Z}[i]$. Thus π must be a prime dividing 2, i.e., $\pi = \pm 1 + \pm i$. Then

$$N(\pi) = \pi\overline{\pi} = 2|(x + yi)(x - yi) = x^2 + y^2 = z^2.$$

This means z is even, so $x^2 + y^2 \equiv 0 \bmod 4$, which implies x and y are both even, a contradiction. □

Lemma 6.25. *Suppose $\alpha, \beta \in \mathbb{Z}[i]$ are relatively prime. If $\alpha\beta = \gamma^2$ is a square in $\mathbb{Z}[i]$, then $u\alpha$ and $u^{-1}\beta$ are squares for some unit u of $\mathbb{Z}[i]$.*

Proof. Note that this is trivial if γ is a unit (and vacuous if $\gamma = 0$). So assume $\alpha\beta$ is the square of some $\gamma \in \mathbb{Z}[i]$, where γ is a non-zero non-unit. Then γ has a prime factorization in $\mathbb{Z}[i]$:

$$\gamma = \prod \pi_i^{e_i}.$$

Thus the prime factorization of $\alpha\beta$ is

$$\alpha\beta = \prod \pi_i^{2e_i}.$$

Up to a reordering of primes, we have

$$\alpha = u^{-1}\pi_1^{2e_1} \cdots \pi_j^{2e_j}$$

$$\beta = u\pi_{j+2}^{2e_{j+1}} \cdots \pi_k^{2e_k}$$

for some unit u. □

Exercise 6.14. *Give an example of relatively prime α, β in $\mathbb{Z}[i]$ such that $\alpha\beta$ is a square in $\mathbb{Z}[i]$, but α and β are not squares in $\mathbb{Z}[i]$.*

Recall that (x, y, z) is a primitive Pythagorean triple if $x, y, z \in \mathbb{N}$ if $\gcd(x, y) = 1$. If (x, y, z) is a primitive Pythagorean triple, then $(\lambda x, \lambda y, \lambda z)$ is also a Pythagorean triple. It is also clear that all Pythagorean triples are multiples of the primitive ones. Hence to determine all Pythagorean triples, it suffices to determine the primitive ones, which we now see how to do using $\mathbb{Z}[i]$.

Proposition 6.26. *(x, y, z) is a primitive Pythagorean triple if and only if x and y are (in some order) $u^2 - v^2$ and $2uv$ for u, v relatively prime in $\mathbb{N}$ with $u > v$ and u, v not both odd. In this case, $z = u^2 + v^2$.*

Proof. ($\Leftarrow$) This was Exercise 2.1.5.

($\Rightarrow$) Suppose (x, y, z) is a primitive Pythagorean triple, so $x^2 + y^2 = (x + yi)(x - yi) = z^2$. By the first lemma, $x + yi$ and $x - yi$ are relatively prime, and by the second they are units times squares. In particular $x + yi = \pm\alpha^2$ or $x + yi = \pm i\alpha^2$ for some $\alpha \in \mathbb{Z}[i]$. Since -1 is a square in $\mathbb{Z}[i]$, we may absorb the possible minus sign into α and write either $x + yi = \alpha^2$ or $x + yi = i\alpha^2$.

Write $\alpha = u + vi$, and we get in the first case

$$x + yi = (u + vi)^2 = u^2 + v^2 + 2uvi$$

and

$$x + yi = i(u + vi)^2 = -2uv + (u^2 + v^2)i.$$

In the first case we have $x = u^2 + v^2$, $y = 2uv$. In the second, we may replace u by $-u$ to write $x = 2uv$, $y = u^2 + v^2$. Then the conditions $\gcd(u, v) = 1$, $u > v$ and u, v not both odd all follow from the facts that $\gcd(x, y)$ and $x, y > 0$ just as in Exercise 2.1.5.

The last statement is obvious. □

Corollary 6.27. *Let $p \in \mathbb{N}$ be prime. Then p occurs as the hypoteneuse of a right-angle triangle with integer length sides if and only if $p > 2$ is a sum of two squares, i.e., if and only if $p \equiv 1 \bmod 4$.*

Proof. The second equivalence is Fermat's two square theorem, so it suffices to prove the first.

($\Rightarrow$) Suppose p is such a hypotenuse. Clearly $p \neq 2$. Now $x^2 + y^2 = p^2$. This implies $\gcd(x, y) = 1$. Hence by the proposition $p = u^2 + v^2$ for some u, v.

($\Leftarrow$) Suppose $p = u^2 + v^2$ is odd. Then $u \neq v$ and u and v are not both odd. Furthermore, we may assume $u > v$. By the proposition $(u^2 - v^2, 2uv, p)$ is a primitive Pythagorean triple. □

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

I mentioned earlier that precisely half of the primes in $\mathbb{Z}$ remain prime in $\mathbb{Z}[i]$, and half factor in $\mathbb{Z}[i]$. We saw that the primes in the first groups are the primes which are 3 mod 4, and the primes in the second group are 2 and those 1 mod 4. My claim then is that, in a sense which can be made precise, half of the (odd) primes in $\mathbb{N}$ are 1 mod 4 and half are 3 mod 4.

In this section we will prove that there are infinitely many primes of the form $4n+1$ and, in the exercises, infinitely many primes of the form $4n + 3$.

Proposition 6.28. *Let p be an odd prime. Then $x^2 \equiv -1 \bmod p$ has a solution if and only if $p = 4n + 1$.*

Proof. ($\Rightarrow$) Suppose $x^2 \equiv -1 \bmod p$ but $p = 4n + 3$. Then

$$x^{p-1} \equiv x^{4n+2} \equiv (x^2)^{2n+1} \equiv (-1)^{2n+1} \equiv -1 \bmod p,$$

which contradicts Fermat's little theorem. Hence $p \equiv 1 \bmod 4$.

($\Leftarrow$) Suppose $p = 4n+1$. Then Lagrange's lemma (Lemma 6.17) says $p|m^2+1$ for some $m \in \mathbb{Z}$. Hence $m^2 \equiv -1 \bmod p$. □

The above proposition is a special case of Gauss's golden theorem in number theory, *quadratic reciprocity*. We will treat quadratic reciprocity in Chapter 9, but for now, we will use the above to prove the following.

Proposition 6.29. *There are infinitely many primes of the form $4n + 1$.*

Proof. The above proposition says an odd prime p is of the form $4n+1$ if and only if $p|x^2+1$ for some $x \in \mathbb{Z}$. Hence it suffices to show that there are infinitely many primes dividing the values of the polynomial $f(x) = x^2+1$.

Suppose instead only finitely many primes $p_1, \ldots, p_k$ (including 2) divide the values of $f(x)$. Let

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

The values of $g(y)$ are a subset of the values of $f(x)$ so the only primes which divide values of $g(y)$ are $p_1, \ldots, p_k$. However

$$g(y) \equiv 1 \bmod p_i$$

for $i = 1, 2, \ldots, k$. Hence $g(1) = (p_1 \cdots p_k)^2 + 1 > 1$ is not divisible by any prime, a contradiction. □

Exercise 6.15. *Let $f(x)$ be any nonconstant polynomial over $\mathbb{Z}$. Show there are infinitely many primes dividing the values of $f(x)$. (Cf. Exercises 6.7.1—6.7.4.)*

Exercise 6.16. *Show that there are infinitely many primes of the form $4n+3$ (Cf. Exercises 6.3.4—6.3.6. Note that this argument is similar to the $4n+1$ case with the polynomial $f(x) = 2x^2+1$. If you like, you may try to use this idea and apply the previous exercise.)*

The above results are a special case of the following.

Theorem 6.30. (Dirichlet's theorem on arithmetic progressions) *Suppose* $\gcd(a, n) = 1$. *There are infinitely many prime $p \equiv a \bmod n$.*

The proof was surprisingly novel, using an analytic tool that Dirichlet invented called ***L*-functions**, which form a central topic in modern number theory. We will not prove this theorem this semester, but will come back to it next semester when we introduce Dirichlet's L-functions.

Now that we know there are infinitely many primes which are 1 mod 4 and 3 mod 4, you might wonder whether there's something further about my assertion that half of the (odd) primes are 1 mod 4 and half are 3 mod 4. The answer is yes, Dirichlet proved something even more precise than the above theorem.

Theorem 6.31. (Dirichlet) *The odd primes, with the natural ordering, are equally distributed in the two congruence classes $4\mathbb{Z}+1$ and $4\mathbb{Z}+3$, i.e.,*

$$\lim_{n\to\infty} \frac{\#\{p < n : p \equiv 1 \bmod 4 \textit{ prime}\}}{\#\{p < n : p \equiv 3 \bmod 4 \textit{ prime}\}} = 1.$$

Dirichlet in fact proved that mod any n, the primes (not dividing n), are equally distributed (in the above sense) among the $\phi(n)$ congruence classes $n\mathbb{Z}+a$ where $\gcd(a, n) = 1$.

Nevertheless, Chebyshev noticed an interesting phenomenon, more amazing in light of the equal distribution result of Dirichlet: there *appear* to be more primes of the latter form. Precisely, if we actually count the primes in each class less than n, "most of the time" we have

$$\#\{p < n : p \equiv 3 \bmod 4 \text{ prime}\} > \#\{p < n : p \equiv 1 \bmod 4 \text{ prime}\}$$

For example, the first time the right hand side is greater is for $n = 26861$. One might wonder if Chebyshev's observation just happens to be true for small values of n. In fact for infinitely many n the left hand side is greater, and infinitely many n the right hand side is greater. Nevertheless, Chebyshev was right. In 1994, Rubinstein and Sarnak showed, in an appropriate way of quantifying things, the above inequality holds about 99.59% of the time. Very roughly, one reason why there are more primes in $4\mathbb{Z}+3$ is because $4\mathbb{Z}+1$ must contain all the odd squares, leaving less room for primes. If you are interested in learning more about this, see the excellent, very readable exposition *Prime Number Races* by Andrew Granville and Greg Martin (Jan. 2006 *American Math Monthly*, available online).

6.8 Discussion

Again, I think it's worth reading.