Gaussian integers

1 Units in $\mathbb{Z}[i]$

An element $x = a + bi \in \mathbb{Z}[i]$, $a, b \in \mathbb{Z}$ is a unit if there exists $y = c + di \in \mathbb{Z}[i]$ such that $xy = 1$. This implies

$$1 = |x|^2|y|^2 = (a^2 + b^2)(c^2 + d^2)$$

But a^2, b^2, c^2, d^2 are non-negative integers, so we must have

$$1 = a^2 + b^2 = c^2 + d^2.$$

This can happen only if $a^2 = 1$ and $b^2 = 0$ or $a^2 = 0$ and $b^2 = 1$. In the first case we obtain $a = \pm 1, b = 0$; thus $x = \pm 1$. In the second case, we have $a = 0, b = \pm 1$; this yields $x = \pm i$. Since all these four elements are indeed invertible we have proved that

$$U(\mathbb{Z}[i]) = \{\pm 1, \pm i\}.$$

2 Primes in $\mathbb{Z}[i]$

An element $x \in \mathbb{Z}[i]$ is *prime* if it generates a prime ideal, or equivalently, if whenever we can write it as a product $x = yz$ of elements $y, z \in \mathbb{Z}[i]$, one of them has to be a unit, i.e. $y \in U(\mathbb{Z}[i])$ or $z \in U(\mathbb{Z}[i])$.

2.1 Rational primes p in $\mathbb{Z}[i]$

If we want to identify which elements of $\mathbb{Z}[i]$ are prime, it is natural to start looking at primes $p \in \mathbb{Z}$ and ask if they remain prime when we view them as elements of $\mathbb{Z}[i]$. If $p = xy$ with $x = a + bi, y = c + di \in \mathbb{Z}[i]$ then

$$p^2 = |x|^2|y|^2 = (a^2 + b^2)(c^2 + d^2).$$

Like before, $a^2 + b^2$ and $c^2 + d^2$ are non-negative integers. Since p is prime, the integers that divide p^2 are $1, p, p^2$. Thus there are three possibilities for $|x|^2$ and $|y|^2$:

1. $a^2 + b^2 = 1$ and $c^2 + d^2 = p^2$;

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

3. $a^2 + b^2 = p^2$ and $c^2 + d^2 = 1$.

In the first case, $a^2 + b^2 = 1 \implies x \in U(\mathbb{Z}[i])$. Similarly, in the third case $c^2 + d^2 = 1 \implies y \in U(\mathbb{Z}[i])$.

Therefore we have the following result.

Proposition 1. *A prime number $p \in \mathbb{Z}$ fails to be a prime element of $\mathbb{Z}[i]$ if and only if p can be written as the sum of two squares, i.e. $p = a^2 + b^2$ for some $a, b \in \mathbb{Z}, a, b > 0$.*

We also have the following observation.

Lemma 2. *If a prime number p can be written as the sum of two squares, then $p = 2$ or $p \equiv 1 \pmod 4$.*

Proof. Assume $p = a^2 + b^2$ with $a, b \in \mathbb{Z}$. We know that $a^2, b^2 \equiv 0, 1 \pmod 4$. Thus $a^2 + b^2 \equiv 0, 1, 2 \pmod 4$. Since p is prime, it cannot be divisible by 4. So we have either $p \equiv 2 \pmod 4$ (and in this case $p = 2$) or $p \equiv 1 \pmod 4$.

□

We would like to prove the converse of this statement. Namely, our goal is to prove the following.

Theorem 3. *A prime $p \in \mathbb{Z}$ can be written as the sum of two squares if and only if $p = 2$ or $p \equiv 1 \pmod 4$.*

Corollary 4. *A prime integer p is a prime element of $\mathbb{Z}[i]$ if and only if $p \equiv 3 \pmod 4$.*

We have already proved one direction of Theorem 3. For the other direction, it is clear that $p = 2 = 1^2 + 1^2$. So for the rest of the Section we assume that p is an *odd* prime. Now we have to show that if $p \equiv 1 \pmod 4$, then it can be written as the sum of two squares. To this end, we will follow Euler's proof. It might not be the shortest proof one can write down, but it has the advantage that it illustrates the concept of descent (which was the idea Fermat used in his sketch of the proof) and reciprocity.

Reciprocity step: A prime $p \equiv 1 \pmod 4$, then it divides $N = a^2 + b^2$ with a and b relatively prime integers.

Descent step: If a prime p divides a number N of the form $N = a^2 + b^2$, where $(a, b) = 1$, then p itself can be written as $p = x^2 + y^2$ for some $(x, y) = 1$.

Clearly the these two claims imply our result.

The reciprocity step is encapsulated (and generalized) in the following Proposition.

Proposition 5. *If $p > 2$ is a prime and d is an integer not divisible by p, then*

$$p | a^2 + db^2 \text{ for some } a, b \in \mathbb{Z} \text{ with } (a, b) = 1 \iff \left(\frac{-d}{p}\right) = 1.$$

Proof. ($\implies$) Since $p|a^2+db^2$ and $(a,b)=1$, it follows that $p \nmid a$ and $p \nmid b$. (Because $p \mid a \implies p \mid db^2 \implies p \mid b$, and this would contradict the fact that a, b are relatively prime.) Since $p \nmid b$, there exists $c \in \mathbb{Z}$, not divisible by p such that $bc \equiv 1 \pmod{p}$. Then

$$a^2+db^2 \equiv 0 \pmod{p} \implies a^2 \equiv -db^2 \pmod{p} \implies (ac)^2 \equiv -d \pmod{p}.$$

Thus $-d$ is a quadratic residue modulo p, hence

$$\left(\frac{-d}{p}\right) = 1.$$

($\Longleftarrow$) $\left(\frac{-d}{p}\right) = 1 \implies -d \equiv a^2 \pmod{p}$ for some $a \not\equiv 0 \pmod{p}$. Then $p \mid a^2+d = a^2+d(1)^2$ and $(a,1)=1$.

□

Proof of the reciprocity step. We take $d=1$ in the previous Proposition. Hence

$$p|a^2+b^2 \ (a,b)=1 \iff \left(\frac{-1}{p}\right) = 1 \iff p \equiv 1 \pmod{4}.$$

□

For the descent step, we need a preliminary result.

Lemma 6. *If N is an integer of the form $N=a^2+b^2$ for some $(a,b)=1$ and $q=x^2+y^2$ is a* prime *divisor of N, then there exist relatively prime integers c and d such that $N/q = c^2+d^2$.*

Proof. First note that since q has no trivial divisors, x and y are forced to be relatively prime. We have

$$x^2N - a^2q = x^2(a^2+b^2) - a^2(x^2+y^2) = x^2b^2 - a^2y^2 = (xb-ay)(xb+ay).$$

Since $q \mid N$, it follows that $x^2N - a^2q \equiv 0 \pmod{q}$, and so

$$(xb-ay)(xb+ay) \equiv 0 \pmod{q}.$$

Since q is a prime, this can happen only if one of the factors is divisible by q. Since we can change the sign of a without affecting our theorem, we can assume that $q \mid xb-ay$, that is $xb-ay = dq$ for some integer d.

We would like to show that $x \mid a+dy$. Since $(x,y)=1$, this is equivalent to showing that $x \mid y(a+dy)$. But

$$y(a+dy) = ay+dy^2 = xb-dq+dy^2 = xb-d(x^2+y^2)+dy^2 = xb-dx^2$$

which is divisible by x. Thus $x \mid a+dy$, so there exist an integer c such that $a+dy = cx$. Therefore

$$cxy = (a + dy)y = xb - dx^2 = x(b - dx)$$

and so

$$cy + dx = b.$$

Next we see that

$$N = a^2 + b^2 = (cx - dy)^2 + (cy + dx)^2 = (x^2 + y^2)(c^2 + d^2) = q(c^2 + d^2).$$

Since $(a, b) = 1$ it follows that $(c, d) = 1$ and the proof is complete.

□

Proof of the descent step. Recall that Fermat's idea (which he used on a number of other occasions), formalized in this case by Euler in this case, is to show that if we have a solution to a diophantine equation, then we can find a "smaller" (in some sense) solution. Iterating this process means that we can find smaller and smaller positive integers. Hence the process needs to terminate at some point, or we reach a contradiction.

Assume that we have an odd prime p (and thus $p > 2$) that divides a number N of the form $N = a^2+b^2$ with $(a, b) = 1$. We want to show that p can be written as a sum of squares.

First, note that we can add or subtract any multiple of p from a or b without changing the problem. That is, we can find integers a_1, b_1 with $|a_1|, |b_1| < p/2$ such that $p|N_1 = a_1^2+b_1^2$. In particular, $N_1 < p^2/2$. Denote $d = (a_1, b_1)$ Then $d < p/2$, so $p \nmid d$. We also know that $a_1 = da_2, b_1 = db_2$ and $(a_2, b_2) = 1$. Note that $|a_2| \leq |a_1| < p/2$ and likewise $|b_2| < p/2$. Therefore $N_2 = a_2^2 + b_2^2 < p^2/2$.

We have

$$p \mid a_1^2 + b_1^2 = d^2(a_2^2 + b_2^2).$$

Since p is a prime that does not divide d, it follows that $p|N_2 = a_2^2 + b_2^2$.

So we showed that our prime p has to divide a number $M = u^2+v^2 < p^2/2$ with $(u, v) = 1$ and $|u|, |v| < p/2$. The positive integer $m = M/p$ will have to be $m < p/2$.

Let q be a *prime* divisor of m. Clearly $q \neq p$ since $q \leq m < p/2$. In particular $q < p$ and $p \mid \frac{M}{q}$.

Assume that q can be written as the sum of two squares. By Lemma 6, we have $M/q = x^2 + y^2$ for some integers $(x, y) = 1$. But then $p \mid x^2 + y^2 < u^2 + v^2 = M$.

So if all the prime factors of M different from p can be written as sums of two squares, then so can p. Since we assumed that this is not the case, it follows that M has some prime divisor $p_1 < p$ that cannot be written as the sum of two squares. By repeating the argument for p_1 it follows that there must exist another prime $p_2 < p_1$ that cannot be written as the sum of two squares. This argument cannot continue indefinitely, so at some point we are bound to hit the prime number $5 = 2^2 + 1^2$ which **can** obviously be written as the sum of two squares. The descent step is now proven and this completes the proof of Theorem 3.

Note that we implicitly used the fact that if $(x, y) = 1$ then $3 \nmid x^2 + y^2$. To see this, recall that for any integer x we have $x \equiv 0, 1$ or $-1 \pmod 3$, so $x^2 \equiv 0$ or $1 \pmod 3$. Since $(x, y) = 1$ we cannot have $x^2 \equiv y^2 \equiv 0 \pmod 3$, so $x^2 + y^2 \not\equiv 0 \pmod 3$.

□

3 Arithmetic

The field is equipped with a norm map $N : \mathbb{Q}(i) \to \mathbb{Q}_{\geq 0}$ given by $N(x + iy) = x^2 + y^2 = |x + iy|^2 = (x + iy)(x - iy) = (x + iy)\overline{(x + iy)}$ where the absolute value is the one on $\mathbb{C}$ and the bar denotes the complex conjugate.

The norm map is clearly multiplicative, i.e. $N(\alpha\beta) = N(\alpha)N(\beta)$. Moreover, $N(\alpha) = 0 \iff \alpha = 0$.

Definition. *There is a notion of* divisibility *in $\mathbb{Z}[i]$ that mirrors the notion from $\mathbb{Z}$. Namely, if $\alpha, \beta \in \mathbb{Z}[i]$ we say that $\alpha \mid \beta$ if $\dfrac{\beta}{\alpha} \in \mathbb{Z}[i]$.*

Example. Does $3 - 4i$ divide $2 + i$? We can do the division by taking the ratio and rationalizing the denominator (i.e. multiply both top and bottom of the fraction by the complex conjugate of the denominator).

$$\frac{2+i}{3-4i} = \frac{(2+i)(3+4i)}{(3-4i)(3+4i)} = \frac{2+11i}{25} = \frac{2}{25} + i\frac{11}{25} \notin \mathbb{Z}[i] \implies 3 - 4i \nmid 2 + i.$$

Example. Does 18 divide $5 + 17i$?

$$\frac{5+17i}{18} = \frac{5}{18} + \frac{17}{18}i \notin \mathbb{Z}[i] \implies 18 \nmid 5 + 17i.$$

Lemma 7. *An integer $c \in \mathbb{Z}$ divides a gaussian integer $a + bi$ if and only if $c|a$ and $c|b$ in $\mathbb{Z}$.*

Proof.

$$c \mid a + bi \iff \frac{a+bi}{c} \in \mathbb{Z}[i] \iff \frac{a}{c} + \frac{b}{c}i \in \mathbb{Z}[i] \iff \frac{a}{c}, \frac{b}{c} \in \mathbb{Z} \iff c|a \text{ and } c|b.$$

□

Proposition 8. *If $\alpha, \beta \in \mathbb{Z}[i]$ and $\alpha \mid \beta$ in $\mathbb{Z}[i]$, then $N(\alpha) \mid N(\beta)$ as integers.*

Proof.

$$\alpha \mid \beta \implies \beta = \alpha\gamma \text{ for some } \gamma \in \mathbb{Z}[i] \implies N(\beta) = N(\alpha)N(\gamma) \implies N(\alpha) \mid N(\beta).$$

□

The converse is not true in general. For instance, $N(5)|N(3-4i)$ (both norms are equal to 25), but $5 \nmid 3-4i$.

There is however an exception. The element $1+i$ has norm $N(1+i)=2$. We have the following result.

Proposition 9. *Let $\alpha = a+bi \in \mathbb{Z}[i]$. Then $N(1+i) \mid N(\alpha)$ if and only if $(1+i) \mid \alpha$.*

Proof. Clearly $(1+i) \mid \alpha$ implies $N(1+i) \mid N(\alpha)$. Conversely, we know that $2 \mid a^2+b^2$. We want to show that there exist $m, n \in \mathbb{Z}$ such that

$$a+bi = (1+i)(m+ni).$$

Expanding the right hand side we see that

$$a+bi = (1+i)(m+ni) \iff a = m-n, b = m+n.$$

We solve this 2×2 linear system in the unknowns m, n and find that $m = \frac{a+b}{2}$ and $n = \frac{b-a}{2}$. In order for these two numbers to be integers we need $a+b$ and $b-a$ to be even. But we know that $2 \mid a^2+b^2$, so a and b are either both even or both odd, i.e. $a \equiv b \pmod 2$. Hence $2 \mid a+b, b-a$ and the result is proved. □

The norm N gives $\mathbb{Z}[i]$ a euclidean ring structure.

Theorem 10. *For any $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$ there exist $\gamma, \rho \in \mathbb{Z}[i]$ such that $\alpha = \beta\gamma + \rho$ and $0 \leq N(\rho) < N(\beta)$.*

Proof. Done in 104A or abstract algebra. □

Note that, as opposed to the division algorithm on $\mathbb{Z}$, we **do not** have uniqueness for γ and ρ. For instance, take $\alpha = -9, \beta = -5$. We have

$$-9 = (-5) \times 1 + (-4)$$

and

$$-9 = (-5) \times 2 + 1.$$

Both 1 and -4 have norm strictly smaller than the norm of -5. Indeed, $N(1) = 1 < 25 = N(-5)$ and $N(-4) = 16 < 25 = N(5)$.

Since $\mathbb{Z}[i]$ is an euclidean ring, it is also a unique factorization domain. But note that the factorization is unique up to multiplication by units. This was already the case over $\mathbb{Z}$. Namely, we know that "any nonzero integer n can be written as a product of primes". But what we really mean by this is that

$$n = (\pm 1)p_1 \dots p_r$$

with $p_1, \dots, p_r$ prime integers (and here is understood that they are positive). But in $\mathbb{Z}[i]$ there is no positivity, so for instance $1+i$ and $1-i$ really represent the same prime. That is, we see that they only differ by a unit, i.e.

$$\frac{1+i}{1-i} = \frac{(1+i)^2}{(1+i)(1-i)} = \frac{2i}{2} = i, \text{ so } (1+i) = i(1-i).$$

In particular, $1+i$ and $1-i$ generate the same ideal of $\mathbb{Z}[i]$. We will write $\pi_1 \sim \pi_2$ if they are two primes in $\mathbb{Z}[i]$ and there is a unit $\mu \in \mathbb{Z}[i]$ such that $\pi_1 = \mu\pi_2$. The unique factorization in $\mathbb{Z}[i]$ tells us that any nonzero gaussian integer α can be written as

$$\alpha = \mu\pi_1 \dots \pi_r$$

where $\mu \in U(\mathbb{Z}[i])$ and $\pi_1, \dots, \pi_r$ are prime gaussian integers (not necessarily distinct. Moreover, if we have another factorization

$$\alpha = \nu\sigma_1 \dots \sigma_s$$

then $r = s$ and for each $1 \leq j \leq r$ there exits $1 \leq k_j \leq r$ such that $\pi_j \sim \sigma_{k_j}$.

This fact allows us to prove the following result about prime elements of $\mathbb{Z}[i]$.

Proposition 11. *If $\pi \in \mathbb{Z}[i]$ is a prime element and $\pi \mid \alpha\beta$ for some gaussian integers α, β then $\pi \mid \alpha$ or $\pi \mid \beta$.*

Proof. Take the prime factorizations of α and β,

$$\alpha = \mu_1\pi_1 \dots \pi_n, \quad \beta = \mu_2\pi_{n+1} \dots \pi_{n+m}.$$

Here the primes π_j are not necessarily distinct. Since $\pi \mid \alpha\beta$ it follows that there exists $\gamma \in \mathbb{Z}[i]$ such that $\alpha\beta = \pi\gamma$. The gaussian integer γ also has a prime factorization

$$\gamma = \nu\delta_1 \dots \delta_r.$$

Thus

$$\mu_1\mu_2\pi_1 \dots \pi_{n+m} = \alpha\beta = \pi\gamma = \nu\pi\delta_1 \dots \delta_r.$$

But the factorization into primes of $\alpha\beta$ is unique (up to multiplication by units), so there exists $1 \leq j \leq$ such that $\pi = u\pi_j$ for some $u \in U(\mathbb{Z}[i])$. If $1 \leq j \leq n$, then $\pi \mid \alpha$. Otherwise, $\pi \mid \beta$. □

Lemma 12. *A gaussian integer α is a unit in $\mathbb{Z}[i]$ if and only if $N(\alpha) = 1$.*

Proof. A unit is an invertible element of the ring. Assume α is a unit. Then there exists $\beta \in \mathbb{Z}[i]$ such that $\alpha\beta = 1$. This implies that $N(\alpha)N(\beta) = 1$. Since the two norms are non-negative integers, we must have $N(\alpha) = N(\beta) = 1$.
Conversely, if $N(\alpha) = 1$ it means that $\alpha\overline{\alpha} = 1$, and since the complex conjugate of a gaussian integer is also a gaussian integer, this makes α a unit. □

Proposition 13. *If $\alpha \in \mathbb{Z}[i]$ has norm $N(\alpha) = p$ a prime integer, then α is a prime element of $\mathbb{Z}[i]$.*

Proof. Assume that $\alpha = \beta\gamma$ with $\beta, \gamma \in \mathbb{Z}[i]$. Then $N(\beta)N(\gamma) = N(\alpha) = p$ is a prime number. Since $N(\beta), N(\gamma) \in \mathbb{Z}$, it follows that either $N(\beta) = 1$ or $N(\gamma) = 1$. By the previous lemma, this means that either β is a unit or γ is a unit. Hence α has no nontrivial divisors, and is therefore a prime gaussian integer. □

Definition. *Just like for regular integers, we say that two gaussian integers α, β are* relatively prime *if they have only unit factors in common.*

Definition. *For two elements $\alpha, \beta \in \mathbb{Z}[i]$,* a greatest common divisor *is a divisor of maximal norm.*

Note that this definition does not define a unique gaussian integer. If you have found a greatest common divisor δ of α, β then $\pm\delta, \pm i\delta$ (that is, δ multiplied by the units) are also divisors with maximal norm. But this is all the indeterminacy, since a greatest common divisor δ of two numbers with prime factorizations

$$\alpha = \mu_1 \pi_1^{m_1} \dots \pi_r^{m_r} \quad \beta = \mu_2 \pi_1^{n_1} \dots \pi_r^{n_r}$$

with $\pi_1, \dots, \pi_r$ prime elements of $\mathbb{Z}[i]$, μ_1, μ_2 units, $m_j, n_j \geq 0$ is of the form

$$\delta = \mu \pi_1^{\min\{m_1, n_1\}} \dots \pi_r^{\min\{m_r, n_r\}}$$

for some $\mu \in U(\mathbb{Z}[i])$.

Lemma 14. *(i) Assume that $\alpha \mid \beta\gamma$ are gaussian integers and that α, β are relatively prime. Then $\alpha \mid \gamma$.*

(ii) Assume that $\alpha, \beta, \gamma \in \mathbb{Z}[i]$ such that $\alpha \mid \gamma$ and $\beta \mid \gamma$. If α and β are relatively prime, then $\alpha\beta \mid \gamma$.

Proof. Follows from unique factorization. □

Proposition 15. *If $\alpha, \beta \in \mathbb{Z}[i]$ and there exist $\gamma, \delta \in \mathbb{Z}[i]$ such that $\alpha\gamma + \beta\delta$ is a unit in $\mathbb{Z}[i]$, then α, β are relatively prime. In particular, if $a, b \in \mathbb{Z}$ are relatively prime integers, then they are also relatively prime in $\mathbb{Z}[i]$.*

Proof. If $\tau \in \mathbb{Z}[i]$ divides both α and β, then $\tau \mid (\alpha\gamma + \beta\delta)$, which is a unit. Therefore $N(\tau) \mid 1$. Thus $N(\tau) = 1$ and τ has to be a unit.

Since $a, b \in \mathbb{Z}$ are relatively prime, it follows that there exist $m, n \in \mathbb{Z}$ such that

$$am + bn = 1.$$

By the first part, they are relatively prime in $\mathbb{Z}[i]$. □

4 Applications to the arithmetic of $\mathbb{Z}$

4.1 Primality testing: Fermat primes

Fermat conjectured that the numbers of the form $2^{2^n}+1$ are prime. Indeed,

- for $n=0: 2^1+1=3$ is prime;
- for $n=1: 2^2+1=5$ is prime;
- for $n=2: 2^3+1=17$ is prime.

The others get a bit too big for us to be able to tell at a glance that they are prime. But $2^8+1=257$ can be checked by hand without too much trouble that it is prime. Same for $2^{2^4}+1=4097$.

However

$$2^{2^5}=2^{32}+1=4294967297$$

is too big to check by hand easily. Note that $2^{32}+1=(2^{16})^2+1^2$ is the sum of two squares. Euler found that it has another very different representation as sum of two squares, namely

$$(2^{16})^2+1^2=4294967297=62264^2+20449^2.$$

The following theorem implies that the fifth Fermat number is in fact **not** a prime.

Theorem 16. *If p is a prime that can be written as sum of two squares, then it can be written like that in essentially one way. That is, if $p=a^2+b^2=c^2+d^2$, with $a,b,c,d,\in\mathbb{Z}$, then either $a=\pm c, b=\pm d$ or $a=\pm d, b=\pm c$.*

Proof. We have

$$(a+bi)(a-bi)=a^2+b^2=c^2+d^2=(c-di)(c+di)$$

and

$$N(a+bi)=N(a-bi)=a^2+b^2=p$$
$$N(c+di)=N(c-di)=c^2+d^2=p.$$

Since p is a prime, it follows from Proposition 13, that $a+bi, a-bi, c+di, c-di$ are all prime elements in $\mathbb{Z}[i]$. By the unique factorization, it follows that either $a+bi=\mu(c+di)$ for some $\mu\in U(\mathbb{Z}[i])$ or $a+bi=\mu(c-di)$ for some $\mu\in U(\mathbb{Z}[i])$.

If $a+bi=\mu(c+di)$ we have four possibilities.

- $\mu=1 \implies a=c, b=d$
- $\mu=-1 \implies a=-c, b=-d$
- $\mu=i \implies a=-d, b=c$

- $\mu = -i \implies a = d, b = -c$

The other case is similar. □

Note that the Theorem does not mention $\mathbb{Z}[i]$, it is a statement purely about integers. The proof however hinges on the arithmetic of the gaussian integers.

4.2 Pythagorean triples revisited

We start with the diophantine equation

$$a^2 + b^2 = c^2. \tag{1}$$

As before we reduce to the case where a, b, c are positive integers with $(a, b) = (b, c) = (c, a) = 1$, a odd and b even. Then c must also be odd.

Our equation (1) can be rewritten as

$$(a + bi)(a - bi) = c^2. \tag{2}$$

Claim 1 $a + bi$ and $a - bi$ are relatively prime.

Proof. Assume $\delta \mid a + bi$ and $\delta \mid a - bi$. Then $\delta \mid 2a$ and $\delta \mid 2bi \implies \delta \mid 2b$. If δ and $2 = -i(1 + i)^2$ were not relatively prime, then $1 + i \mid \delta \implies 2 \mid N(\delta)$. On the other hand, $\delta \mid c^2$, so $N(\delta) \mid c^4$ which is odd. This is a contradiction, so δ and 2 are relatively prime in $\mathbb{Z}[i]$. Then, Lemma 14 implies that

$$\delta \mid a, \delta \mid b.$$

But a, b are relatively prime, so δ must be a unit. □

Claim 2 There exist $\alpha, \beta \in \mathbb{Z}[i]$ such that either $a + bi = \alpha^2$ and $a - bi = \beta^2$ or $a + bi = i\alpha^2$ and $a - bi = -i\beta^2$.

Proof. Exercise. □

Since $a + bi$ and $a - bi$ are relatively prime, α, β are also relatively prime.

Thus $a + bi = \alpha^2 = (m + ni)^2$ or $a + bi = i\alpha^2 = i(m + ni)^2$ for some $m, n \in \mathbb{Z}$. Expanding the square leads to

$$a + bi = m^2 - n^2 + 2mni \quad \text{or} \quad a + bi = -2mn + i(m^2 - n^2).$$

However, we want a to be odd, so the second case cannot occur. We are therefore in the first case and $a + bi$ is after all a perfect square in the $\mathbb{Z}[i]$, i.e.

$$a + bi = (m + ni)^2 \text{ with } m, n \in \mathbb{Z}. \tag{3}$$

The derivation of (3) from unique factorization in $\mathbb{Z}[i]$ is the key step in this proof. The rest is a matter of (careful) book-keeping. Identifying the real and imaginary parts above gives

$$a = m^2 - n^2, b = 2mn$$

and therefore

$$c^2 = a^2+b^2 = (m^2-n^2)^2+(2mn)^2 = m^4+n^4-2m^2n^2+4m^2n^2 = m^4+n^4+2m^2n^2 = (m^2+n^2)^2.$$

Since $c > 0$ it follows that

$$c = m^2 + n^2.$$

We also have $b > 0$ so both m, n have to have the same sign, and by changing that sign we can assume without changing the values of a, b, c that $m, n > 0$. Since $a > 0$ we must have $m > n$. They also have to be relatively prime, since a, b are relatively prime. Lastly, since a is odd, $m \not\equiv n \pmod 2$.

We need to check that our solution $(a, b, c) = (m^2 - n^2, 2mn, m^2 + n^2)$ with $m > n > 0, (m, n) = 1, m \not\equiv n \pmod 2$, satisfies (1) and that a, b, c are positive integers with $(a, b) = (b, c) = (c, a) = 1$, a odd and b even. Indeed,

$$a^2 + b^2 = (m^2 - n^2)^2 + (2mn)^2 = (m^2 + n^2)^2 = c^2.$$

On the other hand, $m \not\equiv n \pmod 2 \implies a$ odd, and $b = 2mn$ is clearly even. Since $m > n > 0$ we have $a, b, c > 0$. If a prime p divides two of $m^2 - n^2, 2mn, m^2 + n^2$ then $p \neq 2$ since the first and third numbers are odd. Assume $p|m^2 - n^2$ and $p|2mn$. Then $p|mn$ so $p|m$ or $p|n$. Assume $p|m$. Then $p|n^2$ and so $p|n$. This cannot happen since $(m, n) = 1$. The other case is similar.

To recap: we found a way of producing Pythagorean triples on demand. That is, take any gaussian integer α with both the real and imaginary parts non-zero, square it and get $\alpha^2 = a + bi$. Then $(|a|, |b|, N(\alpha))$ is a Pythagorean triple.

Example. $(23+10i)^2 = 529-100+460i = 429+460i$ and $23^2+10^2 = 629$. So $(429, 460, 629)$ is a Pythagorean triple (grab a calculator and check!) And since $(23, 10) = 1$ the triple is relatively prime.

4.3 Other diophantine equations

To better appreciate this approach to Pythagorean triples, let's apply it to the diophantine equation

$$a^2 + b^2 = c^3. \tag{4}$$

Theorem 17. *The integral solutions to*

$$a^2 + b^2 = c^3$$

with $(a, b) = 1$ *are given by the parametric formulas*

$$a = m^3 - 3mn^2, b = 3m^2n - n^3, c = m^2 + n^2, \text{ where } (m, n) = 1, m \not\equiv n \pmod 2.$$

Different choices of m, n *give different solutions* (a, b, c).

Proof. Note that a and b cannot both be even since they are relatively prime. On the other hand, if they were both odd we would have $a^2, b^2 \equiv 1 \pmod 8$ and therefore $c^3 \equiv 2 \pmod 8$ which is impossible. (In this case, c would have to be even, which would make c^3 a multiple of 8.) This $a \not\equiv b \pmod 2$ and c is odd. We can rewrite (4) as

$$c^3 = (a + bi)(a - bi). \tag{5}$$

We first show that $a + bi$ and $a - bi$ are relatively prime gaussian integers. Let $\delta \in \mathbb{Z}[i]$ such that $\delta \mid a + bi$ and $\delta \mid a - bi$. As we have seen in the previous section, this implies that

$$\delta \mid 2a, \delta \mid 2b.$$

On the other hand $N(\delta) \mid a^2 + b^2$ which is odd, so $N(\delta)$ is odd and thus δ and 2 are relatively prime. It follows that

$$\delta \mid a, \delta \mid b.$$

Since $(a, b) = 1$ it follows that δ is a unit in $\mathbb{Z}[i]$.
Hence $a + bi$ and $a - bi$ are relatively prime. From (5) follows that $a + bi = \mu\alpha^3$ for some $\mu \in U(\mathbb{Z}[i]), \alpha \in \mathbb{Z}[i]$. On the other hand, every unit in $\mathbb{Z}[i]$ is itself a cube:

$$1 = 1^3, -1 = (-1)^3, i = (-i)^3, -i - i^3.$$

Therefore we can write $\mu = \nu^3$ and so $a + bi = (\nu\alpha)^3$ and $\nu\alpha = \beta \in \mathbb{Z}[i]$. Thus

$$a + bi = (m + ni)^3 \text{ for some } m, n \in \mathbb{Z}.$$

Every prime $p \in \mathbb{Z}$ that divides both m and n would have to divide a and b. Therefore $(m, n) = 1$. We expand the cube and see that

$$a = m^3 - 3mn^2 \quad b = 3m^2n - n^3.$$

If $m \equiv n \pmod 2$, then we would get

$$a \equiv m^3 - 3m^3 \pmod 2 \equiv 0 \pmod 2$$

and

$$b \equiv 3m^3 - m^3 \pmod 2 \equiv 0 \pmod 2$$

which contradicts the fact that $a \not\equiv b \pmod 2$. Hence $m \not\equiv n \pmod 2$. Plugging the value of $a + bi = (m + ni)^3$ into (5) we see that

$$c^3 = (m + ni)^3(m - ni)^3 = (m^2 + n^2)^3 \implies c = m^2 + n^2.$$

We have to check two things. First, that our parametric equations give a solution (a, b, c) to (4) with $(a, b) = 1$ and $a + bi = (m + ni)^3$. Second, that changing the m, n changes the (a, b, c). The first is a direct calculation. The second uses the fact that the only cube root of unity in $\mathbb{Z}[i]$ is 1 itself. We know this since the norm of a cube root of 1 would have to be 1 and so the cube roots of unity are in $U(\mathbb{Z}[i])$. But we have seen what the cubes of the elements in $U(\mathbb{Z}[i])$ are from above. This means that if $(m + ni)^3 = (m' + n'i)^3$ we must have $m + ni = m + n'i$ hence $m = m'$ and $n = n'$.

□

Here's a table with a few solutions (a, b, c) to (4) for various choices of m and n.

m	n	$a = m^3 - 3mn^2$	$b = 3m^2n - n^3$	$c = m^2 + n^2$
1	0	1	0	1
2	1	2	11	5
3	2	-9	46	13
4	1	52	47	17
4	3	-44	9	25
7	2	259	286	53

Another application is to show that a perfect square in $\mathbb{Z}$ cannot come just before a perfect cube.

Theorem 18. *The only integers a, b satisfying $a^2 = b^3 - 1$ are $a = 0, b = 1$.*

On the face of it, this is not at all obvious. Besides, there are plenty of perfect cubes that come just before a perfect square: -1 and 0; 0 and 1; 8 and 9.

Proof. Clearly $a = 0, b = 1$ satisfy $a^2 = b^3 - 1$. We want to show that there are no other solutions. Assume $a, b \in \mathbb{Z}$ do satisfy

$$a^2 = b^3 - 1.$$

We can rewrite this equation as

$$b^3 = (a - i)(a + i)$$

and follow the blueprint from the previous theorem. If we know that $a + i$ and $a - i$ are relatively prime, then, recalling that the only cube root of unity is 1 itself, we see that $a + i$ and $a - i$ have to be perfect cubes in $\mathbb{Z}[i]$ and we would get

$$a = m^3 - 3mn^2, 1 = 3m^2n - n^3$$

for some $m, n \in \mathbb{Z}$. The second relation shows that $n \mid 1$, so $n = \pm 1$. If $n = 1$, we have $1 = 3m^2 - 1$ so $3m^2 = 2$ which is impossible. Thus $n = -1$ and therefore $1 = 1 - 3m^2$ so $m = 0$. This leads us to $a = 0$ and $b^3 = 1$, so $b = 1$.

It remains to show that $a + i$ and $a - i$ are relatively prime gaussian integers. Assume δ is a common divisor. Then $\delta | 2a$ and $\delta \mid 2i = (1 + i)^2$. Thus, up to units, δ is either 1 or $1 + i$ or $(1 + i)^2$. Assume δ is not a unit. Then $(1 + i) \mid \delta$ and therefore $(1 + i) \mid b^3$. Since $1 + i$ is a prime gaussian integer, we then have $(1 + i) \mid b$, hence $b^2 = N(b)$ is even. Thus b must be even, and therefore $a \equiv -1 \pmod 4$ which is impossible. □

Remark. In 1850, Lebesgue used $\mathbb{Z}[i]$ to show that, for $d \geq 2$, the only integral solution to

$$y^2 = x^d - 1$$

is $x = 1, y = 0$.

5 Back to the prime elements of $\mathbb{Z}[i]$

We want to find all elements $\pi = a + bi \in \mathbb{Z}[i]$ that are prime gaussian integers. We have seen in Proposition 13 that if the norm of a gaussian integer α is prime, then α is prime in $\mathbb{Z}[i]$. On the other hand, 3 is a prime in $\mathbb{Z}[i]$ but its norm $N(3) = 9$ is not a prime.

Lemma 19. *If $\pi \in \mathbb{Z}[i]$ is a prime element, then there exist a prime number $p \in \mathbb{Z}$ such that $\pi \mid p$ in $\mathbb{Z}[i]$.*

Proof. The norm $N(\pi)$ is a positive integer, and therefore it factors as a product of primes in $\mathbb{Z}$,

$$N(\pi) = p_1 \dots p_r.$$

On the other hand, $N(\pi) = \pi\bar{\pi}$, so $\pi \mid N(\pi)$. Thus

$$\pi \mid p_1 \dots p_r \implies \pi \mid p_j \text{ for some } 1 \leq j \leq r.$$

□

Theorem 20. *Let $p \in \mathbb{Z}$ be a (positive) prime. Its factorization in $\mathbb{Z}[i]$ is determined by its residue class modulo 4 as follows.*

(i) $2 = (1+i)(1-i) = -i(1+i)^2 = i(1-i)^2$ *and* $1+i = i(1-i)$ *represent the same prime ideal in* $\mathbb{Z}[i]$.

(ii) If $p \equiv 1 \pmod 4$ *then* $p = \pi\bar{\pi}$ *where* $\pi, \bar{\pi}$ *are two prime gaussian integers that are complex conjugates, but not unit multiples. In particular, they generate different prime ideals in* $\mathbb{Z}[i]$.

(iii) If $p \equiv 3 \pmod 4$ *then p is prime in* $\mathbb{Z}[i]$.

Proof. *(i)* Direct calculation.

(ii) If $p \equiv 1 \pmod 4$, Theorem 3 says that p can be written as a sum of two relatively prime squares $p = a^2 + b^2 = (a + bi)(a - bi)$. Set $\pi = a + bi$. Then $\bar{\pi} = a - bi$ and clearly $p = \pi\bar{\pi}$. Moreover $N(\pi) = N(\bar{\pi}) = a^2 + b^2 = p$ is prime, so π and $\bar{\pi}$ are prime gaussian integers by Proposition 13. Lastly,

$$\frac{a+bi}{a-bi} = \frac{(a+bi)^2}{a^2+b^2} = \frac{a^2-b^2}{a^2+b^2} + i\frac{2ab}{a^2+b^2} \notin \mathbb{Z}[i]$$

since both fractions have absolute value smaller than 1.

(iii) If $p \equiv 3 \pmod 4$, this is Corollary 4.

□

Lemma 19 tells us that any gaussian prime is a factor of a prime $p \in \mathbb{Z}$. Theorem 20 tells us how integral primes factor in $\mathbb{Z}[i]$. We put these two results together and get a complete characterization of the prime elements in $\mathbb{Z}[i]$.

Theorem 21. *Every prime gaussian integer is a unit multiple of one of the following primes:*

(i) $1+i$;

(ii) π *or* $\bar{\pi}$ *where* $N(\pi)=p$ *is a prime integer* $p \equiv 1 \pmod 4$;

(iii) a prime p *in* $\mathbb{Z}$ *with* $p \equiv 3 \pmod 4$. *In this case* $N(p)=p^2$.

Note that in the first two cases the prime gaussian integers have nonzero real and imaginary parts, while in the third case we get $\pm p, \pm ip$ which have either the real or the imaginary part equal to 0. Moreover, the only prime gaussian integer of even norm is $1+i$ (up to unit multiples). Therefore we see that if we have $\alpha \in \mathbb{Z}[i]$ and its prime factorization

$$\alpha = \mu\pi_1 \dots \pi_r$$

does not contain a prime multiple of $1+i$, then $N(\alpha)=N(\pi_1)\dots N(\pi_r)$ is an odd integer. We find this way another proof of the fact that $N(\alpha)$ is even $\iff 1+i \mid \alpha$.

6 Representing integers as sums of squares

We saw that a prime can be written as a sum of two squares essentially only one way, if at all (Theorem 16). We have seen that other integers though can be written as sums of two squares in multiple ways. For instance, $50=5^2+5^2=7^2+1^2$. We can use the arithmetic in $\mathbb{Z}[i]$ to *systematically* construct integers that are sums of two squares in more than one way. Take the factorizations of 5 and 10 in $\mathbb{Z}[i]$. We have

$$5=(1+2i)(1-2i) \quad 10=(1+3i)(1-3i).$$

Thus 50 factors in two ways

$$50=5\cdot 10=\big((1+2i)(1+3i)\big)\cdot\big((1-2i)(1-3i)\big)=\big((1+2i)(1-3i)\big)\cdot\big((1-2i)(1+3i)\big)$$

This becomes

$$50=(-5+5i)(-5-5i)=(7-i)(7+i),$$

which gives

$$50=5^2+5^2=7^2+1^2.$$

Different representations of an integer as a sum of two squares in $\mathbb{Z}$ correspond to rearranging prime factors in $\mathbb{Z}[i]$. Here's another example. Consider the factorizations of 5 and 13. We have

$$5=(1+2i)(1-2i) \quad 13=(2+3i)(2-3i).$$

Therefore

$$65=5\cdot 13=\big((1+2i)(2+3i)\big)\cdot\big((1-2i)(2-3i)\big)=\big((1+2i)(2-3i)\big)\cdot\big((1-2i)(2+3i)\big)$$

which becomes

$$65 = (-4 + 7i)(-4 - 7i) = (8 + i)(8 - i).$$

The two factorizations yield two ways of writing 65 as a sum of squares:

$$65 = 4^2 + 7^2 = 8^2 + 1^2.$$

Let us find an integer which can be written as the sum of two squares in *three* different ways. Start with

$$5 = (1 + 2i)(1 - 2i) \quad 13 = (2 + 3i)(2 - 3i) \quad 17 = (1 + 4i)(1 - 4i).$$

Consider the following products

$$\alpha = (1 + 2i)(2 + 3i)(1 + 4i) \quad \beta = (1 - 2i)(2 + 3i)(1 + 4i) \quad \gamma = (1 + 2i)(2 - 3i)(1 + 4i).$$

Then

$$\alpha = -32 - 9i \quad \beta = 12 + 31i \quad \gamma = 4 + 33i$$

are gaussian integers with

$$N(\alpha) = N(\beta) = N(\gamma) = 5 \cdot 13 \cdot 17 = 1105.$$

Therefore

$$1105 = 32^2 + 9^2 = 12^2 + 31^2 = 4^2 + 33^2.$$

Using this method you can construct systematically and without having to guess integers that can be represented as sums of two squares in four, five, ..., twenty, ... ways.

Moreover, the arithmetic of $\mathbb{Z}[i]$ allows us to classify the integers that can be represented as sums of two squares.

Lemma 22. *(a) An integer n can be written as the sum of two squares if and only if n is the norm of some gaussian integer.*

(b) If $m, n \in \mathbb{Z}$ can be written as sums of two squares, then mn can also be written as the sum of two squares.

Proof. (a) $n = a^2 + b^2 \iff n = N(a + bi)$

(b) By part (a), we know that there exist $\alpha, \beta \in \mathbb{Z}[i]$ such that $n = N(\alpha)$ and $m = N(\beta)$. Then $mn = N(\alpha\beta)$ is also the sum of two squares.

□

Theorem 23. *An integer $n > 1$ is a sum of two squares exactly when any prime factor of n which is $\equiv 3 \pmod 4$ occurs with even multiplicity.*

Proof. First we show any integer $n > 1$ having even multiplicity at its prime factors which are $\equiv 3 \pmod 4$ can be written as a sum of two squares. The prime $2 = 1^2 + 1^2$ and any prime $p \equiv 1 \pmod 4$ can be written as the sum of two squares, by Theorem 3. On the other hand, $q^2 = q^2 + 0^2$ is trivially the sum of two squares, for any prime $q \equiv 3 \pmod 4$. Since sums of two squares are closed under multiplication (part (b) of the Lemma), it follows that n can be written as the sum of two squares.

Now we treat the converse direction: any $n > 1$ which is a sum of two squares has even multiplicity at any prime factor which is $\equiv 3 \pmod 4$. We argue by induction on n.
The fact is obviously true for $n = 2 = 1^2 + 1^2$.
Let $n \geq 3$ be a sum of two squares. We assume that in any sum of two squares $< n$ the prime factors that are $\equiv 3 \pmod 4$ occur with even powers. (This is the induction hypothesis.)
If n has no prime factors congruent to 3 modulo 4, then we have nothing to prove and the result is obviously true.
In case $n = a^2 + b^2$ has some prime factor $p \equiv 3 \pmod 4$, we get that

$$p \mid (a + bi)(a - bi).$$

On the other hand p is prime in $\mathbb{Z}[i]$, so $p \mid a + bi$ or $p \mid a - bi$. But in either case we can take complex conjugates and obtain that $p = \bar{p} \mid a - bi$ *and* $p = \bar{p} \mid a + bi$. Thus

$$p^2 \mid (a + bi)(a - bi) = n.$$

On the other hand, $p \mid a + bi$ implies that $p \mid a$ and $p \mid b$ (Lemma 7). Then we can write $a = pa_1, b = pb_1$ for some $a_1, b_1 \in \mathbb{Z}$. Thus

$$n = p^2(a_1^2 + b_1^2)$$

and $n_1 = a_1^2 + b_1^2 < n$. By the induction hypothesis, any prime $\equiv 3 \pmod 4$ that appears in the factorization of n_1 appears with an even exponent. Therefore the same holds for n and our proof is complete.

□

Example. For primes we have seen that $p \equiv 1 \pmod 4 \implies p$ can be written as the sum of two squares. But the number $21 = 3 \cdot 7$ cannot be written as the sum of two squares, even though $21 \equiv 1 \pmod 4$. In general, we need to factor an integer $n > 1$ in order to decide if it can be written as the sum of two squares.