Introduction to Number Theory
Supplement on Gaussian Integers

Spring 2016

Last Updated: April 10, 2016

This is a brief supplemental note on the Gaussian integers, written for my Spring 2016 Elementary Number Class at Brown University. With respect to the book, the nearest material is the material in Chapters 35 and 36, but we take a very different approach.

In this note, we cover the following topics.

(1) What are the Gaussian integers?
(2) Unique factorization within the Gaussian integers.
(3) An application of the Gaussian integers to the Diophantine equation $y^2 = x^3 - 1$.
(4) Other integer-like sets: general rings.
(5) Specific examples within $\mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[\sqrt{-5}]$.

1. What are the Gaussian Integers?

The Gaussian Integers are the set of numbers of the form $a + bi$, where a and b are normal integers and i is a number satisfying $i^2 = -1$. As a collection, the Gaussian Integers are represented by the symbol $\mathbb{Z}[i]$, or sometimes $\mathbb{Z}[\sqrt{-1}]$. These might be pronounced either as *The Gaussian Integers* or as *Z append i*.

In many ways, the Gaussian integers behave very much like the regular integers. We've been studying the qualities of the integers, but we should ask — which properties are really properties of the integers, and which properties hold in greater generality? Is it the integers themselves that are special, or is there something bigger and deeper going on?

These are the main questions that we ask and make some progress towards in these notes. But first, we need to describe some properties of Gaussian integers.

We will usually use the symbols $z = a+bi$ to represent our typical Gaussian integer. One adds and multiples two Gaussian integers just as you would add and multiply two complex numbers. Informally, you treat i like a polynomial indeterminate X, except that it satisfies the relation $X^2 = -1$.

Definition 1. *For each complex number $z = a+bi$, we define the **conjugate of** z, written as $\overline{z}$, by*

$$\overline{z} = a - bi.$$

*We also define the **norm of** z, written as $N(z)$, by*

$$N(z) = a^2 + b^2.$$

You can check that $N(z) = z\overline{z}$ (and in fact this is one of your assigned problems). You can also chack that $N(zw) = N(z)N(w)$, or rather that the norm is multiplicative (this is also one of your assigned problems).

Even from our notation, it's intuitive that $z = a + bi$ has two parts, the part corresponding to a and the part corresponding to b. We call a the **real part of** z, written as $\operatorname{Re} z = a$, and we call b the **imaginary part of** z, written as $\operatorname{Im} z = b$. I should add that the name "imaginary number" is a poor name that reflects historical reluctance to view complex numbers as acceptable. For that matter, the name "complex number" is also a poor name.

As a brief example, consider the Gaussian integer $z = 2 + 5i$. Then $N(z) = 4 + 25 = 29$, $\operatorname{Re} z = 2$, $\operatorname{Im} z = 5$, and $\overline{z} = 2 - 5i$.

We can ask similar questions to those we asked about the regular integers. What does it mean for $z \mid w$ in the complex case?

Definition 2. *We say that a Gaussian integer z **divides** another Gaussian integer w if there is some Gaussian integer k so that $zk = w$. In this case, we write $z \mid w$, just as we write for regular integers.*

For the integers, we immediately began to study the properties of the primes, which in many ways were the building blocks of the integers. Recall that for the regular integers, we said p was a prime if its only divisors were ± 1 and $\pm p$. In the Gaussian integers, the four numbers $\pm 1, \pm i$ play the same role as ± 1 in the usual integers. These four numbers are distinguished as being the only four Gaussian integers with norm equal to 1.

That is, the only solutions to $N(z) = 1$ where z is a Gaussian integer are $z = \pm 1, \pm i$. We call these four numbers the **Gaussian units**.

With this in mind, we are ready to define the notion of a prime for the Gaussian integers.

Definition 3. *We say that a Gaussian integer z with $N(z) > 1$ is a **Gaussian prime** if the only divisors of z are u and uz, where $u = \pm 1, \pm i$ is a Gaussian unit.*

Remark. *When we look at other integer-like sets, we will actually use a different definition of a prime.*

It's natural to ask whether the normal primes in $\mathbb{Z}$ are also primes in $\mathbb{Z}[i]$. And the answer is no. For instance, 5 is a prime in $\mathbb{Z}$, but

$$5 = (1 + 4i)(1 - 4i)$$

in the Gaussian integers. However, the two Gaussian integers $1 + 4i$ and $1 - 4i$ are prime. It also happens to be that 3 is a Gaussian prime. We will continue to investigate which numbers are Gaussian primes over the next few lectures.

With a concept of a prime, it's also natural to ask whether or not the primes form the building blocks for the Gaussian integers like they form the building blocks for the regular integers. We take up this in our next topic.

2. Unique Factorization in the Gaussian Integers

Let us review the steps that we followed to prove unique factorization for $\mathbb{Z}$.

(1) We proved that for a, b in $\mathbb{Z}$ with $b \neq 0$, there exist unique q and r such that $a = bq + r$ with $0 \leq r < b$. This is called the Division Algorithm.
(2) By repeatedly applying the Division Algorithm, we proved the Euclidean Algorithm. In particular, we showed that the last nonzero remainder was the GCD of our initial numbers.
(3) By performing reverse substition on the steps of the Euclidean Algorithm, we showed that there are integer solutions in x, y to the Diophantine equation $ax + by = \gcd(a, b)$. This is often called Bezout's Theorem or Bezout's Lemma, although we never called it by that name in class.
(4) With Bezout's Theorem, we showed that if a prime p divides ab, then $p \mid a$ or $p \mid b$. This is the crucial step towards proving Unique Factorization.
(5) We then proved Unique Factorization.

Each step of this process can be repeatred for the Gaussian integers, with a few notable differences. Remarkably, once we have the division algorithm, each proof is almost identical for $\mathbb{Z}[i]$ as it is for $\mathbb{Z}$. So we will prove the division algorithm, and then give sketches of the remaining ideas, highlighting the differences that come up along the way.

In the division algorithm, we require the remainder r to "be less than what we are dividing by." A big problem in translating this to the Gaussian integers is that the Gaussian integers *are not ordered*. That is, we don't have a concept of being *greater than* or *less than* for $\mathbb{Z}[i]$.

When this sort of problem emerges, we will get around this by taking norms. Since the norm of a Gaussian integer is a typical integer, we will be able to use the ordering of the integers to order our norms.

Theorem 4. *For z, w in $\mathbb{Z}[i]$ with $w \neq 0$, there exist q and r in $\mathbb{Z}[i]$ such that $z = qw + r$ with $N(r) < N(w)$.*

Proof. Here, we will cheat a little bit and use properties about general complex numbers and the rationals to perform this proof. One can give an entirely intrinsic proof, but I like the approach I give as it also informs how to actually compute the q and r.

The entire proof boils down to the idea of writing z/w as a fraction and approximating the real and imaginary parts by the nearest integers.

Let us now transcribe that idea. We will need to introduce some additional symbols. Let $z = a_1 + b_1 i$ and $w = a_2 + b_2 i$.

Then

$$\begin{aligned}\frac{z}{w} &= \frac{a_1+b_1i}{a_2+b_2i} = \frac{a_1+b_1i}{a_2+b_2i}\frac{a_2-b_2i}{a_2-b_2i} \\ &= \frac{a_1a_2+b_1b_2}{a_2^2+b_2^2} + i\frac{b_1a_2-a_1b_2}{a_2^2+b_2^2} \\ &= u+iv.\end{aligned}$$

By rationalizing the denominator by multiplying by $\overline{w}/\overline{w}$, we are able to separate out the real and imaginary parts. In this final expression, we have named u to be the real part and v to be the imaginary part. Notice that u and v are normal rational numbers.

We know that for any rational number u, there is an integer u' such that $|u-u'| \leq \frac{1}{2}$. Let u' and v' be integers within $1/2$ of u and v above, respectively.

Then we claim that we can choose $q = u'+iv'$ to be the q in the theorem statement, and let r be the resulting remainder, $r = z - qw$. We need to check that $N(r) < N(w)$. We will check that explicitly.

We compute

$$N(r) = N(z-qw) = N\left(w\left(\frac{z}{w}-q\right)\right) = N(w)N\left(\frac{z}{w}-q\right).$$

Note that we have used that $N(ab) = N(a)N(b)$. In this final expression, we have already come across $\frac{z}{w}$ before — it's exactly what we called $u+iv$. And we called $q = u'+iv'$. So our final expression is the same as

$$N(r) = N(w)N(u+iv-u'-iv') = N(w)N\left((u-u')+i(v-v')\right).$$

How large can the real and imaginary parts of $(u-u')+i(v-v')$ be? By our choice of u' and v', they can be at most $1/2$.

So we have that

$$N(r) \leq N(w)N\left((\tfrac{1}{2})^2+(\tfrac{1}{2})^2\right) = \frac{1}{2}N(w).$$

And so in particular, we have that $N(r) < N(w)$ as we needed. □

Note that in this proof, we did not actually show that q or r are unique. In fact, unlike the case in the regular integers, it is not true that q and r are unique.

Example. *Consider* $3+5i, 1+2i$. *Then we compute*

$$\frac{3+5i}{1+2i} = \frac{3+5i}{1+2i}\frac{1-2i}{1-2i} = \frac{13}{5} + i\frac{-1}{5}.$$

The closest integer to $13/5$ *is* 3, *and the closest integer to* $-1/5$ *is* 0. *So we take* $q = 3$. *Then* $r = (3+5i) - (1+2i)3 = -i$, *and we see in total that*

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

Note that $N(-i) = 1$ *and* $N(1+2i) = 5$, *so this choice of* q *and* r *works.*

As $13/5$ *is sort of close to* 2*, what if we chose* $q = 2$ *instead? Then* $r = (3 + 5i) - (1 + 2i)2 = 1 + i$*, leading to the overall expression*

$$3_5 i = (1 + 2i)2 + (1 + i).$$

Note that $N(1 + i) = 2 < N(1 + 2i) = 5$*, so that this choice of* q *and* r also works*.*

This is an example of how the choice of q *and* r *is not well-defined for the Gaussian integers. In fact, even if one decides to choose* q *to that* $N(r)$ *is minimal, the resulting choices are still not necessarily unique.*

This may come as a surprise. The letters q and r come from our tendency to call those numbers the *quotient* and *remainder* after division. We have shown that the quotient and remainder *are not well-defined*, so it does not make sense to talk about "the remainder" or "the quotient." This is a bit strange!

Are we able to prove unique factorization when the process of division itself seems to lead to ambiguities? Let us proceed forwards and try to see.

Our next goal is to prove the Euclidean Algorithm. By this, we mean that by repeatedly performing the division algorithm starting with two Gaussian integers z and w, we hope to get a sequence of remainders with the last nonzero remainder giving a greatest common divisor of z and w.

Before we can do that, we need to ask a much more basic question. What do we mean by a greatest common divisor? In particular, the Gaussian integers are not ordered, so it does not make sense to say whether one Gaussian integer is bigger than another.

For instance, is it true that $i > 1$? If so, then certainly i is positive. We know that multiplying both sides of an inequality by a positive number doesn't change that inequality. So multiplying $i > 1$ by i leads to $-1 > i$, which is absurd if i was supposed to be positive!

To remedy this problem, we will choose a common divisor of z and w with the greatest norm (which makes sense, as the norm is a regular integer and thus is well-ordered). But the problem here, just as with the division algorithm, is that there may or may not be multiple such numbers. So we cannot talk about "the greatest common divisor" and instead talk about "a greatest common divisor." To paraphrase Lewis Carroll's[1] Alice, things are getting curiouser and curiouser!

Definition 5. *For nonzero* z, w *in* $\mathbb{Z}[i]$*, a* ***greatest common divisor*** *of* z *and* w*, denoted by* $\gcd(z, w)$*, is a common divisor with largest norm. That is, if* c *is another common divisor of* z *and* w*, then* $N(c) \leq N(\gcd(z, w))$*.*

If $N(\gcd(z, w)) = 1$*, then we say that* z *and* w *are relatively prime. Said differently, if* 1 *is a greatest common divisor of* z *and* w*, then we say that* z *and* w *are relatively prime.*

[1] Carroll was also a mathematician, and hid some nice mathematics inside some of his works.

Remark. *Note that* $\gcd(z, w)$ *as we're writing it is not actually well-defined, and may stand for any greatest common divisor of* z *and* w*.*

With this definition in mind, the proof of the Euclidean Algorithm is almost identical to the proof of the Euclidean Algorithm for the regular integers. As with the regular integers, we need the following result, which we will use over and over again.

Lemma 6. *Suppose that* $z \mid w_1$ *and* $z \mid w_2$*. Then for any* x, y *in* $\mathbb{Z}[i]$*, we have that* $z \mid (xw_1 + yw_2)$*.*

Proof. As $z \mid w_1$, there is some Gaussian integer k_1 such that $zk_1 = w_1$. Similarly, there is some Gaussian integer k_2 such that $zk_2 = w_2$.

Then $xw_1 + yw_2 = zxk_1 + zyk_2 = z(xk_1 + yk_2)$, which is divisible by z as this is the definition of divisibility. □

Notice that this proof is identical to the analogous statement in the integers, except with differently chosen symbols. That is how the proof of the Euclidean Algorithm goes as well.

Theorem 7. *let* z, w *be nonzero Gaussian integers. Recursively apply the division algorithm, starting with the pair* z, w *and then choosing the quotient and remainder in one equation the new pair for the next. The last nonzero remainder is divisible by all common divisors of* z, w*, is itself a common divisor, and so the last nonzero remainder is a greatest common divisor of* z *and* w*.*

Symbolically, this looks like

$$\begin{aligned} z &= q_1 w + r_1, \quad N(r_1) < N(w) \\ w &= q_2 r_1 + r_2, \quad N(r_2) < N(r_1) \\ r_1 &= q_3 r_2 + r_3, \quad N(r_3) < N(r_2) \\ \cdots &= \cdots \\ r_k &= q_{k+2} r_{k+1} + r_{k+2}, \quad N(r_{k+2}) < N(r_{k+1}) \\ r_{k+1} &= q_{k+3} r_{k+2} + 0, \end{aligned}$$

where r_{k+2} *is the last nonzero remainder, which we claim is a greatest common divisor of* z *and* w*.*

Proof. We are claiming several thing. Firstly, we should prove our implicit claim that this algorithm terminates at all. Is it obvious that we should eventually reach a zero remainder?

In order to see this, we look at the norms of the remainders. After each step in the algorithm, the norm of the remainder is smaller than the previous step. As the norms are always nonnegative integers, and we know there does not exist an infinite list of decreasing positive integers, we see that the list of nonzero remainders is finite. So the algorithm terminates.

We now want to prove that the last nonzero remainder is a common divisor and is in fact a greatest common divisor. The proof is actually identical to the proof in the integer case, merely with a different choice of symbols.

Here, we only sketch the argument. Then the rest of the argument can be found by comparing with the proof of the Euclidean Algorithm for $\mathbb{Z}$ as found in the course textbook.

For ease of exposition, suppose that the algorithm terminated in exatly 3 steps, so that we have

$$\begin{aligned} z &= q_1 w + r_1, \\ w &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + 0. \end{aligned}$$

On the one hand, suppose that d is a common divisor of z and w. Then by our previous lemma, $d \mid z - q_1 w = r_1$, so that we see that d is a divisor of r_1 as well. Applying to the next line, we have that $d \mid w$ and $d \mid r_1$, so that $d \mid w - q_2 r_1 = r_2$. So every common divisor of z and w is a divisor of the last nonzero remainder r_2.

On the other hand, $r_2 \mid r_1$ by the last line of the algorithm. Then as $r_2 \mid r_1$ and $r_2 \mid r_1$, we know that $r_2 \mid q_2 r_1 + r_2 = w$. Applying this to the first line, as $r_2 \mid r_1$ and $r_2 \mid w$, we know that $r_2 \mid q_1 w + r_1 = z$. So r_2 is a common divisor.

We have shown that r_2 is a common divisor of z and w, and that every common divisor of z and w divides r_2. How do we show that r_2 is a greatest common divisor?

Suppose that d is a common divisor of z and w, so that we know that $d \mid r_2$. In particular, this means that there is some nonzero k so that $dk = r_2$. Taking norms, this means that $N(dk) = N(d)N(k) = N(r_2)$. As $N(d)$ and $N(k)$ are both at least 1, this means that $N(d) \leq N(r_2)$.

This is true for every common divisor d, and so $N(r_2)$ is at least as large as the norm of any common divisor of z and w. Thus r_2 is a greatest common divisor.

The argument carries on in the same way for when there are more steps in the algorithm. □

Theorem 8. *The greatest common divisor of z and w is well-defined, up to multiplication by $\pm 1, \pm i$. In other words, if $\gcd(z, w)$ is a greatest common divisor of z and w, then all greatest common divisors of z and w are given by $\pm \gcd(z, w), \pm i \gcd(z, w)$.*

Proof. Suppose d is a greatest common divisor, and let $\gcd(z, w)$ denote a greatest common divisor resulting from an application of the Euclidean Algorithm. Then we know that $d \mid \gcd(z, w)$, so that there is some k so that $dk = \gcd(z, w)$. Taking norms, we see that $N(d)N(k) = N(\gcd(z, w)$.

But as both d and $\gcd(z, w)$ are greatest common divisors, we must have that $N(d) = N(\gcd(z, w))$. So $N(k) = 1$. The only Gaussian integers with norm one are $\pm 1, \pm i$, so we have that $du = \gcd(z, w)$ where u is one of the four Gaussian units, $\pm 1, \pm i$.

Conversely, it's clear that the four numbers $\pm \gcd(z, w), \pm i \gcd(z, w)$ are all greatest common divisors. □

Now that we have the Euclidean Algorithm, we can go towards unique factorization in $\mathbb{Z}[i]$. Let g denote a greatest common divisor of z and w. Reverse substitution in the Euclidean Algorithm shows that we can find Gaussian integer solutions x, y to the (complex) linear Diophantine equation

$$zx + wy = g.$$

Let's see an example.

Example. *Consider $32 + 9i$ and $4 + 11i$. The Euclidean Algorithm looks like*

$$\begin{aligned} 32 + 9i &= (4 + 11i)(2 - 2i) + 2 - 5i, \\ 4 + 11i &= (2 - 5i)(-2 + i) + 3 - i, \\ 2 - 5i &= (3 - i)(1 - i) - i, \\ 3 - i &= -i(1 + 3i) + 0. \end{aligned}$$

So we know that $-i$ is a greatest common divisor of $32+9i$ and $4+11i$, and so we know that $32 + 9i$ and $4 + 11i$ are relatively prime. Let us try to find a solution to the Diophantine equation

$$x(32 + 9i) + y(4 + 11i) = 1.$$

Performing reverse substition, we see that

$$\begin{aligned} -i &= (2 - 5i) - (3 - i)(1 - i) \\ &= (2 - 5i) - (4 + 11i - (2 - 5i)(-2 + i))(1 - i) \\ &= (2 - 5i) - (4 + 11i)(1 - i) + (2 - 5i)(-2 + 1)(1 - i) \\ &= (2 - 5i)(3i) - (4 + 11i)(1 - i) \\ &= (32 + 9i - (4 + 11i)(2 - 2i))(3i) - (4 + 11i)(1 - i) \\ &= (32 + 9i)3i - (4 + 11i)(2 - 2i)(3i) - (4 + 11i)(1 - i) \\ &= (32 + 9i)3i - (4 + 11i)(7 + 5i). \end{aligned}$$

Multiplying this through by i, we have that

$$1 = (32 + 9i)(-3) + (4 + 11i)(5 - 7i).$$

So one solution is $(x, y) = (-3, 5 - 7i)$.

Although this looks more complicated, the process is the same as in the case over the regular integers. The apparent higher difficulty comes mostly from our lack of familiarity with basic arithmetic in $\mathbb{Z}[i]$.

The rest of the argument is now exactly as in the integers.

Theorem 9. *Suppose that z, w are relatively prime, and that $z \mid wv$. Then $z \mid v$.*

Proof. This is left as an exercise (and will appear on the next midterm in some form — cheers to you if you've read this far in these notes). But it's now the almost the same as in the regular integers. □

Theorem 10. *Let z be a Gaussian integer with $N(z) > 1$. Then z can be written uniquely as a product of Gaussian primes, up to multiplication by one of the Gaussian units $\pm 1, \pm i$.*

Proof. We only sketch part of the proof. There are multiple ways of doing this, but we present the one most similar to what we've done for the integers. If there are Gaussian integers without unique factorization, then there are some (maybe they tie) with minimal norm. So let z be a Gaussian integer of minimal norm without unique factorization. Then we can write

$$p_1 p_2 \cdots p_k = z = q_1 q_2 \cdots q_\ell,$$

where the p and q are all primes. As $p_1 \mid z = q_1 q_2 \cdots q_\ell$, we know that p_1 divides one of the q (by Theorem 9), and so (up to units) we can say that p_1 is one of the q primes. We can divide each side by p_1 and we get two supposedly different factorizations of a Gaussian integer of norm $N(z)/N(p_1) < N(z)$, which is less than the least norm of an integer without unique factorization (by what we supposed). This is a contradiction, and we can conclude that there are no Gaussian integers without unique factorization. □

If this seems unclear, I recommend reviewing this proof and the proof of unique factroziation for the regular integers. I should also mention that one can modify the proof of unique factorization for $\mathbb{Z}$ as given in the course textbook as well (since it is a bit different than what we have done). Further, the course textbook does proof of unique factorization for $\mathbb{Z}[i]$ in Chapter 36, which is very similar to the proof sketched above (although the proof of Theorem 9 is very different.)

3. An application to $y^2 = x^3 - 1$.

We now consider the nonlinear Diophantine equation $y^2 = x^3 - 1$, where x, y are in $\mathbb{Z}$. This is hard to solve over the integers, but by going up to $\mathbb{Z}[i]$, we can determine all solutions.

In $\mathbb{Z}[i]$, we can rewrite

$$y^2 + 1 = (y + i)(y - i) = x^3. \tag{1}$$

We claim that $y + i$ and $y - i$ are relatively prime. To see this, suppose that d is a common divisor of $y + i$ and $y - i$. Then $d \mid (y + i) - (y - i) = 2i$. It happens to be that $2i = (1 + i)^2$, and that $(1 + i)$ is prime. To see this, we show the following.

Lemma 11. *Suppose z is a Gaussian integer, and $N(z) = p$ is a regular prime. Then z is a Gaussian prime.*

Proof. Suppose that z factors nontrivially as $z = ab$. Then taking norms, $N(z) = N(a)N(b)$, and so we get a nontrivial factorization of $N(z)$. When $N(z)$ is a prime, then there are no nontrivial factorizations of $N(z)$, and so z must have no nontrivial factorization. □

As $N(1+i) = 2$, which is a prime, we see that $(1+i)$ is a Gaussian prime. So $d \mid (1+i)^2$, which means that d is either $1, (1+i)$, or $(1+i)^2$ (up to multiplication by a Gaussian unit).

Suppose we are in the case of the latter two, so that $(1+i) \mid d$. Then as $d \mid (y+i)$, we know that $(1+i) \mid x^3$. Taking norms, we have that $2 \mid x^6$.

By unique factorization in $\mathbb{Z}$, we know that $2 \mid x$. This means that $4 \mid x^2$, which allows us to conclude that $x^3 \equiv 0 \pmod 4$. Going back to the original equation $y^2 + 1 = x^3$, we see that $y^2 + 1 \equiv 0 \pmod 4$, which means that $y^2 \equiv 3 \pmod 4$. A quick check shows that $y^2 \equiv 3 \pmod 4$ has no solutions y in $\mathbb{Z}/4\mathbb{Z}$.

So we rule out the case then $(1+i) \mid d$, and we are left with d being a unit. This es exactly the case that $y+i$ and $y-i$ are relatively prime.

Recall that $(y+i)(y-i) = x^3$. As $y+i$ and $y-i$ are relatively prime and their product is a cube, by unique factorization in $\mathbb{Z}[i]$ we know that $y+i$ and $y-i$ much each be Gaussian cubes. Then we can write $y+i = (m+ni)^3$ for some Gaussian integer $m+ni$. Expanding, we see that

$$y + i = m^3 - 3mn^2 + i(3m^2n - n^3).$$

Equating real and imaginary parts, we have that

$$y = m(m^2 - 3n^2)$$
$$1 = n(3m^2 - n^2).$$

This second line shows that $n \mid 1$. As n is a regular integer, we see that $n = 1$ or -1.

If $n = 1$, then that line becomes $1 = (3m^2 - 1)$, or after rearranging $2 = 3m^2$. This has no solutions.

If $n = -1$, then that line becomes $1 = -(3m^2 - 1)$, or after rearranging $0 = 3m^2$. This has the solution $m = 0$, so that $y + i = (-i)^3 = i$, which means that $y = 0$. Then from $y^2 + 1 = x^3$, we see that $x = 1$.

And so the only solution is $(x, y) = (1, 0)$, and there are no other solutions.

4. Other Rings

The Gaussian integers have many of the same properties as the regular integers, even though there are some differences. We could go further. For example, we might consider the following integer-like sets,

$$\mathbb{Z}(\sqrt{d}) = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}.$$

One can add, subtract, and multiply these together in similar ways to how we can add, subtract, and multiply together integers, or Gaussian integers.

We might ask what properties these other integer-like sets have. For instance, do they have unique factorization?

More generally, there is a better name than "integer-like set" for this sort of construction.

Suppose R is a collection of elements, and it makes sense to add, subtract, and multiply these elements together. Further, we want addition and multiplication to behave similarly to how they behave for the regular integers. In particular, if r and s are elements in R, then we want $r+s = s+r$ to be in R; we want something that behaves like 0 in the sense that $r+0 = r$; for each r, want another element $-r$ so that $r+(-r) = 0$; we want $r \cdot s = s \cdot r$; we want something that behaves like 1 in the sense that $r \cdot 1 = r$ for all $r \neq 0$; and we want $r(s_1 + s_2) = rs_1 + rs_2$. Such a collection is called a **ring**. (More completely, this is called a commutative unital ring, but that's not important.)

It is not important that you explicitly remember exactly what the definition of a ring is. The idea is that there is a name for things that are "integer-like" and that we might wonder what properties we have been thinking of as properties of the integers are *actually* properties of rings.

As a total aside: there are very many more rings too, things that look much more different than the integers. This is one of the fundamental questions that leads to the area of mathematics called *Abstract Algebra*. With an understanding of abstract algebra, one could then focus on these general number theoretic problems in an area of math called *Algebraic Number Theory*.

5. The rings $\mathbb{Z}[\sqrt{d}]$

We can describe some of the specific properties of $\mathbb{Z}[\sqrt{d}]$, and suggest how some of the ideas we've been considering do (or don't) generalize. For a general element $n = a + b\sqrt{d}$, we can define the conjugate $\overline{n} = a - b\sqrt{d}$ and the norm $N(n) = n \cdot \overline{n} = a^2 - db^2$. We call those elements u with $N(u) = 1$ the *units* in $\mathbb{Z}[\sqrt{d}]$.

Some of the definitions we've been using turn out to not generalize so easily, or in quite the ways we expect. If n doesn't have a nontrivial factoriation (meaning that we cannot write $n = ab$ with $N(a), N(b) \neq 1$), then we call n an **irreducible**. In the cases of $\mathbb{Z}$ and $\mathbb{Z}[i]$, we would have called these elements prime.

In general, we call a number p in $\mathbb{Z}\sqrt{d}$ a **prime** if p has the property that $p \mid ab$ means that $p \mid a$ or $p \mid b$. Of course, in the cases of $\mathbb{Z}$ and $\mathbb{Z}[i]$, we showed that irreducibles are primes. But it turns out that this is not usually the case.

Let us look at $\mathbb{Z}\sqrt{-5}$ for a moment. In particular, we can write 6 in two ways as

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

Although it's a bit challenging to show, these are the only two fundamentally different factorizations of 6 in $\mathbb{Z}[\sqrt{-5}]$. One can show (it's not very hard, but it's not particularly illuminating to do here) that neither 2 or 3 divides $(1 + \sqrt{-5})$ or $(1 - \sqrt{-5})$ (and vice versa), which means that none of these

four numbers are primes in our more general definition. One can also show that all four numbers are irreducible.

What does this mean? This means that 6 can be factored into irreducibles in fundamentally different ways, and that $\mathbb{Z}[\sqrt{-5}]$ does not have unique factorization.

It's a good thought exercise to think about what is really different between $\mathbb{Z}[\sqrt{-5}]$ and $\mathbb{Z}$. At the beginning of this course, it seemed extremely obvious that $\mathbb{Z}$ had unique factorization. But in hindsight, is it really so obvious?

Understanding when there is and is not unique factorization in $\mathbb{Z}[\sqrt{d}]$ is something that people are still trying to understand today. The fact is that we don't know! In particular, we really don't know very much when d is positive.

One reason why can be seen in $\mathbb{Z}[\sqrt{2}]$. If $n = a + b\sqrt{2}$, then $N(n) = a^2 - 2b^2$. A very basic question that we can ask is *what are the units?* That is, which n have $N(n) = 1$?

Here, that means trying to solve the equation

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

We have seen this equation a few times before. On the second homework assignment, I asked you to show that there were infinitely many solutions to this equation by finding lines and intersecting them with hyperbolas. We began to investigate this Diophantine equation because each solution leads to another square-triangular number.

So there are infinitely many units in $\mathbb{Z}[\sqrt{2}]$. This is strange! For instance, $3 + 2\sqrt{2}$ is a unit, which means that it behaves just like ± 1 in $\mathbb{Z}$, or like $\pm 1, \pm i$ in $\mathbb{Z}[i]$. Very often, the statements we've been looking at and proving are true "up to multiplication by units." Since there are infinitely many in $\mathbb{Z}[\sqrt{2}]$, it can mean that it's annoying to determine even if two numbers are actually the same up to multiplication by units.

As you look further, there are many more strange and interesting behaviours. It is really interesting to see what properties are very general, and what properties vary a lot. It is also interesting to see the different ways in which properties we're used to, like unique factorization, can fail.

For instance, we have seen that $\mathbb{Z}[\sqrt{-5}]$ does not have unique factorization. We showed this by seeing that 6 factors in two fundamentally different ways. In fact, some numbers in $\mathbb{Z}[\sqrt{-5}]$ do factor uniquely, and others do not. But if one does not, then it factors in at most two fundamentally different ways.

In other rings, you can have numbers which factor in more fundamentally different ways. The actual behaviour here is also really poorly understood, and there are mathematicians who are actively pursuing these topics.

It's a very large playground out there.