A CLASSIFICATION OF GAUSSIAN PRIMES

LEE A. BUTLER

Abstract. This essay gives an elementary classification of the prime elements of $\mathbb{Z}[i]$, the Gaussian integers.

The history of our number system is a well documented one; there are a plethora of books that describe that slow advancement we made as a species from using those most intuitive and yet amazingly abstract objects, the natural numbers, up to the 19th century where complex numbers were finally given a rigorous footing and a level of acceptance. The story has a neat pattern to it, with each new set of numbers, be they rationals, irrationals, or negative numbers, being accepted slowly and reluctantly. Their use was always introduced due to necessity, specifically for the solution to certain equations. Perhaps the rational numbers were not explicitly introduced because a Greek merchant wished to know the solution to $3x + 4 = 6$, but from a mathematician's point of view most real world questions could be boiled down to such a problem.

So with the rationals fell equations such as those above, negative numbers toppled those such as $x + 1 = 0$ which defied natural numbers. Irrationals, the bane of Pythagoras, dealt with those troublesome quadratics that refused to fit into rational shoes. And finally, with the invention/discovery of the complex numbers, came the payoff for two thousand years of development. Now all quadratics could be solved, and even better, the Fundamental Theorem of Algebra assures us that now any polynomial can be solved with any coefficients from the complex numbers. Having gained this much, apparently everything we could ever need when it comes to numbers, why move further? Introducing new numbers seems pointless since they would only solve the very equations they themselves created. So moving forward would be counterproductive, and surely moving back would be nothing more than a history lesson. We moved slowly from natural numbers to rationals to negatives to irrationals to complex numbers... no stages were skipped so nothing new is to be gleaned from a step backwards. Is it?

German wunderkind Johann Carl Friedrich Gauss suggested the resolute answer: yes. What he did was not strictly speaking a step backward, nor was it totally a step forward. What Gauss did was to consider "integers" to be all complex numbers $a + bi$ where both a and b are real integers. The reasons for him doing this are not hard to see.

Definition. The Gaussian integers are the elements of the set

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

Both $\mathbb{Z}$ and $\mathbb{Z}[i]$[1] are commutative rings with identity (respectively 1 and $1+0i$), which a quick check of the ring axioms will confirm. Neither are division rings though, owing to the lack of a multiplicative inverse for most of the elements of each set. I say 'most' because both sets do have a finite set of units within them, that is elements u for which there is an element v with the property that $uv = vu = 1$. In $\mathbb{Z}$ the units are 1 and -1, since $1 \cdot 1 = 1$ and $(-1) \cdot (-1) = 1$. In $\mathbb{Z}[i]$ there are four units, again we have 1 and -1, but also we have i and $-i$, noting that $i \cdot (-i) = 1$ and $(-i) \cdot i = 1$. So in $\mathbb{Z}[i]$, just like in $\mathbb{Z}$, most elements have no inverse. If we take $57 - 11i$ then there is no Gaussian integer z such that $(57 - 11i) \cdot z = 1$. But how can we be sure that this is true? The Gaussian Integers are a big set, after all. In fact they are countably infinite, a fact derived directly from the fact that $|\mathbb{Z} \times \mathbb{Z}| = \aleph_0$. So how can we be sure that only those four members mentioned above have a multiplicative inverse?

[1] Denoted $k(i)$ by some authorities[1].

The proof that there are indeed just four units in the Gaussian Integers uses a simple proof by contradiction. The first step is to define the norm of a Gaussian Integer, although it is the same as for complex numbers in general.

Definition. The norm of $a + bi \in \mathbb{Z}[i]$ is $N(a + bi) = a^2 + b^2$.

So, for example, $N(57 - 11i) = 3370$. Some authorities prefer to define the norm as the value $\sqrt{a^2 + b^2}$, that is the square root of our norm, and is just the modulus of the number. While doing so provides the direct link between the norm of the complex number $a + bi$ and the length of the vector (a, b), it also requires us to think of the norm-squared for most purposes. However, both definitions have the simple property laid out in Lemma 1.

Lemma 1. For all Gaussian integers s and t, $N(s)N(t) = N(st)$.

Proof. Let $s = a + bi$ and $t = c + di$. First note that

$$st = (a + bi)(c + di) = (ac - bd) + i(ad + bc).$$

Then simple algebra gives the result:

$$\begin{aligned}
N(st) &= (ac - bd)^2 + (ad + bc)^2 \\
&= a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2 \\
&= a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2 \\
&= a^2(c^2 + d^2) + b^2(c^2 + d^2) \\
&= (a^2 + b^2)(c^2 + d^2) \\
&= N(s)N(t).
\end{aligned}$$

□

Now we can clear up the issue of units in $\mathbb{Z}[i]$. First we assume that we have a unit of the set, say $u + vi$. That means there is another element, call it $x + yi$, with the property that

$$(u + vi)(x + yi) = 1.$$

Lemma 1 then assures us of the following:

$$N(u + vi)N(x + yi) = N((u + vi)(x + yi)) = N(1) = 1.$$

And so

$$(u^2 + v^2)(x^2 + y^2) = 1.$$

Recall that u,v,x, and y are all integers, and since the integers form a ring the interior of each set of parentheses must also be an integer. As previously mentioned the only units in $\mathbb{Z}$ are 1 and -1. But each bracketed term is positive so we must have that

$$u^2 + v^2 = 1.$$

The only integer solutions to this equation are $(u, v) = (1, 0), (0, 1), (-1, 0)$, and $(0, -1)$. Putting these back into the original number we get that the only units of $\mathbb{Z}[i]$ are $1, i, -1$, and $-i$.

A second result that will come in useful later is Lemma 2 below, the proof of which is more algebraic manipulation similar to Lemma 1, and so the proof is not included here.

Lemma 2. For all Gaussian integers s and t with $t \neq 0$,

$$N\left(\frac{s}{t}\right) = \frac{N(s)}{N(t)}.$$

The above lemma may look innocuous enough but it side-steps an important issue. I mentioned earlier that $\mathbb{Z}[i]$ is not a division ring, so can we divide at all? That is, given Gaussian Integers s and t, is s/t a Gaussian Integer too? The answer is a straightforward: sometimes. Not all such divisions yield a Gaussian Integer, and as always division by zero is forbidden. Taking our Gaussian Integer $57 - 11i$ we can easily find a second number that doesn't divide to give a third element of the set. For instance, $\frac{57-11i}{14+3i} = 3.731 - 1.585i$ to three decimal places, which is clearly not in $\mathbb{Z}[i]$. Rather more challenging, but apparently sometimes possible, is finding a Gaussian Integer that does divide $57 - 11i$. "Divides" here is the same definition as in the integers, that is:

Definition. We say that the Gaussian integer $a + bi$ divides the Gaussian integer $c + di$ if and only if we can find a Gaussian integer $e + fi$ such that

$$c + di = (a + bi)(e + fi).$$

We write this as $a + bi \mid c + di$.

So $14 + 3i \nmid 57 - 11i$ since $3.731 - 1.585i \notin \mathbb{Z}[i]$. But if we try some other random Gaussian Integer, say $7 - 25i$, then we find that $\frac{57-11i}{7-25i} = 1 + 2i$. And so $7 - 25i \mid 57 - 11i$, and clearly we also have that $1 + 2i \mid 57 - 11i$.

The next obvious step after finding that these numbers can sometimes be divided by one another is to ask about factorisations. The regular integers have those remarkable numbers the primes as their building blocks, and the Fundamental Theorem of Arithmetic (with a few tweaks to incorporate the negative integers as well as the positive naturals) tells us that every integer can be factorised uniquely into a product of primes. "Unique" here only takes into account the primes used. So if someone wanted to disprove the FTA by pointing out that -28 factorises as both $(-1) \cdot 2^2 \cdot 7$ and $2 \cdot 1 \cdot 7 \cdot (-1) \cdot 2$ then they would be wrong. For these two are clearly the same, and the FTA ignores the order that the multiplication is performed in as well as repeated multiplication by the units. And quite rightly too.

By drawing parallels with the integers we should expect the units of $\mathbb{Z}[i]$ to trivially divide any given Gaussian Integer, and indeed they all do. Below is the case for the unit $-i$.

$$\frac{a + bi}{-i} = \frac{(a + bi)i}{(-i)i} = -b + ai.$$

So in our attempts to factorise Gaussian Integers we should avoid repeated factors of any of the four units. We should also try to draw another parallel between Gaussian and normal integers. Normal (natural) primes have several definitions, the most concise probably being p is prime $\Leftrightarrow \phi(p) = p - 11^2$. However in the integers the most useful definition for us is: p is a prime if and only if the only divisors of p are $1, -1, p$, and $-p$. This leads us directly to a definition for primes in the ring of Gaussian Integers:

[2] $\phi(p)$ being Euler's phi-function, defined not quite concisely as "the number of positive integers not exceeding p which are relatively prime to p."[6]

Definition. A Gaussian integer γ is called a Gaussian prime if and only if the only Gaussian integers that divide γ are:

$$1, -1, i, -i, \gamma, -\gamma, \gamma i, \text{ and } -\gamma i.$$

That seems like a lot initially. Primes, after all, are supposed to be those numbers with the least possible number of factors, and yet here they are with no less than eight of them. And yet this is right, for aside from the units, every Gaussian Integer has *at least* eight divisors, and so the ones with *only* eight deserve to be called our primes.

Now that we have identified the requirements a Gaussian prime needs to fulfil, the questions suddenly pour forth. Are there any Gaussian Primes? Are there infinitely many? Does every Gaussian Integer factor uniquely into a product of Gaussian primes? Are all the primes from $\mathbb{Z}$ also primes in $\mathbb{Z}[i]$? And so on and so forth.

I will address some of these questions in this essay, starting of course with the last one. The simple and possibly unsurprising answer is no. Take, for the easiest example, the smallest prime, 2. This factors as $(1+i)(1-i)$, and so is not prime. Why not kill two hedgehogs with one stone, you may say, and answer another of the questions while we're here. For surely 2 also factorises as $(-1+i)(-1-i)$, so prime factorisations are not unique! But hold on. First are we sure that all these $\pm 1 \pm i$'s are primes? Well let's see. They are all rather small numbers, each having a norm of two, so any number that divides them must have a norm of either one or two, a fact provided by lemma 1 or 2, take your pick. But now we're in troublesome territory, for if we factor one of them, say σ, as $\sigma = \alpha \cdot \beta$ then one of the two factors must have a norm of one, and hence it will be a unit. But this means the only divisors of our number are $1, -1, i,$ and $-i$, because they always are, and $\sigma, -\sigma, i\sigma,$ and $-i\sigma$. Hence σ will be a prime, and so we have found our first primes. This also means that 2 does factor into Gaussian primes and so is not one itself. Moreover we can note that $(1+i)(1-i) = (-i)(-1+i)i(-1-i)$, and so we can consider the factorisation unique, for each of our two representations is the same, just with multiples of units included. Two factorisations of a Gaussian Integer that are intrinsically the same, that is the only difference is the multiples of units, will be called equivalent.

Back to factorising our numbers. We know now that 2 is not a Gaussian prime, but it is a special case being even, so let us look at an odd prime, 3. There aren't many Gaussian integers small enough to multiply up to three, so it is possible to check all possible combinations and see whether any give 3 as a result. But none do as can be shown a bit more succinctly.

By Lemma 1 we know that if α and β are two non-unit factors of 3 then they will satisfy

$$N(\alpha)N(\beta) = N(\alpha\beta) = N(3).$$

Taking $\alpha = a + bi$ and $\beta = c + di$ this converts to

$$(a^2 + b^2)(c^2 + d^2) = 9.$$

Recall that neither of our factors are units, so neither expression in the brackets can be equal to 1. Since $9 = 3^2$ this gives us

$$a^2 + b^2 = 3 \qquad \text{and} \qquad c^2 + d^2 = 3.$$

A moment's thought will assure us that these have no integer solutions, and hence 3 has no factors in Gaussian integers except for the units and the units multiplied by 3 itself. So 3 is a Gaussian prime. What does 3 have that 2 doesn't to allow it to retain its status as prime? The clue is in the above working. If we try the same approach to factor 2 we get the two equations below

$$a^2 + b^2 = 2 \qquad \text{and} \qquad c^2 + d^2 = 2.$$

These do have solutions, namely any combination of $(a, b) = (\pm 1, \pm 1)$, and the corresponding values for (c, d) to ensure that $(a+bi)(c+di)$ does indeed equal 2. So if we wish to factorise an integer n into Gaussian integers then we first factorise its norm as $N(n) = XY$ with neither X nor Y equal to 1. We can always do this since $N(n) = n^2$ so it will have at least two (not necessarily distinct) non-unit factors. Then we just have to solve the equations

$$a^2 + b^2 = X \qquad \text{and} \qquad c^2 + d^2 = Y$$

while bearing in mind that we also need $(a + bi)(c + di) = n$. For integers we can use the fact that $\Im(n) = 0$ to get a nice formula for the factors, but rather than do that we shall instead extend the idea to all Gaussian integers.

The same reasoning still holds true for Gaussian integers. We first factor the norm of the number, say $N(\gamma) = UV$, then solve the equations

$$a^2 + b^2 = U \qquad \text{and} \qquad c^2 + d^2 = V.$$

So, for instance, $N(57 - 11i) = 3370 = 10 \cdot 337$, so we want to solve $a^2 + b^2 = 10$ and $c^2 + d^2 = 337$. These have many solutions, among them $(a, b) = (3, 1)$ which then gives $(c, d) = (16, -9)$; and hence $57 - 11i = (3 + i)(16 - 9i)$.

This approach gives us our first fact about Gaussian Primes: if the norm of a Gaussian Integer is prime, then the Gaussian Integer is prime. This is simply because if $N(\gamma) = p = 1 \cdot p$ then we will always have one factor being a unit.

The problem seems to have boiled down to finding out which integers can be written as the sum of two squares[3]. This is not really an easy question either, but it is approachable with the Number Theory technique of experimentation.

Numbers that can be written as the sum of two squares	$1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26, 29, 32, 34, 36, 37, 40, 41, 45, 49, 50$
Numbers that can't be written as the sum of two squares	$3, 6, 7, 11, 12, 14, 15, 19, 21, 22, 23, 24, 27, 28, 30, 31, 33, 35, 38, 39, 42, 43, 44, 46, 47, 48$

This table is not particularly telling unless one really knows one's numbers. The one below may be more instructive.

Numbers that can be written as the sum of two squares (modulo 4)	$1, 2, 0, 1, 0, 1, 2, 1, 0, 1, 2, 0, 1, 2, 1, 0, 2, 0, 1, 0, 1, 1, 1, 2$
Numbers that can't be written as the sum of two squares (modulo 4)	$3, 2, 3, 3, 0, 2, 3, 3, 1, 2, 3, 0, 3, 0, 2, 3, 1, 3, 2, 3, 2, 3, 0, 2, 3, 0$

Now a pattern emerges, although unfortunately not a very encompassing one. It does appear that if $n \equiv 3 \pmod 4$ then it cannot be written as a sum of two squares. But aside from that the table reveals little. The one below looks at just the primes.

[3] A problem that dates back at least two thousand years to the time of Diophantus[2].

Primes that can be written as the sum of two squares (modulo 4)	$2, 1, 1, 1, 1, 1, 1$
Primes that can't be written as the sum of two squares (modulo 4)	$3, 3, 3, 3, 3, 3, 3, 3$

Admittedly the table is looking somewhat low on entries now to be convincing, but that is no reason not to blindly surmise that if $p \equiv 1 \pmod 4$ then it can be written as the sum of two squares, and if $p \equiv 3 \pmod 4$ then it cannot. Longer tables would support this claim, but enough tables, now for proofs.

What we would like to show is that for all odd primes p, p can be written as the sum of two squares if and only if $p \equiv 1 \pmod 4$. We can ignore 2 as an anomaly. The contrapositive of this proposition then gives us the corresponding result for when $p \equiv 3 \pmod 4$.

We can prove the result in one direction immediately.

Theorem 3(i). For all odd primes p, if p can be written as the sum of two squares then $p \equiv 1 \pmod 4$.

Proof. Suppose that $p = a^2 + b^2$. Since p is odd exactly one of a and b must be odd, otherwise the sum of their squares would be even. So wlog take a as the odd one, and set $a = 2m + 1$ and $b = 2n$. Then we have:

$$\begin{aligned} p &= a^2 + b^2 \\ &= (2m+1)^2 + (2n)^2 \\ &= 4m^2 + 4m + 1 + 4n^2 \\ &\equiv 1 \pmod 4. \end{aligned}$$

And so we are done. □

The proof in the opposite direction is unfortunately much more challenging. Girard is tenuously given credit by some authorities[4] for coming up with the first proof, though some prefer Fermat who claimed to have a proof several years later in his letters to Huygens, and who claimed to know a proof based on his Method of Descent, but who of course never wrote it down. His claim is offered support by Euler who wins the prize for the first published proof which does indeed use Fermat's Method of Descent[5].

This is the approach I will use here[6]. Given our prime $p \equiv 1 \pmod 4$ we try to write some multiple of p as a sum of two squares. Legendre's and Euler's Law of Quadratic Reciprocity, which I won't prove here, assures us that we can do this and find a multiple less than the prime itself. Thus we have integers A_1, B_1, and M_1 such that $A_1^2 + B_1^2 = M_1 p$ and $M_1 < p$. If $M_1 = 1$ then we are done, so assume also that $M_1 \geq 2$. Fermat's Method of Descent requires us to then use this fact to find three new integers, A_2, B_2, and M_2, such that $A_2^2 + B_2^2 = M_2 p$, and $M_2 < M_1$. We can then apply this procedure repeatedly to find A_i, B_i, and M_i for $i = 3, 4, \ldots$

[4] Including the renowned H. Davenport[2].

[5] Despite all this the result is known as Thue's Lemma after the Norwegian mathematician Axel Thue.

[6] Parts of proof taken from Silverman, pp. 176-181 [3]

until $M_i = 1$. Then we will have proved the second part of our sum of two squares theorem. So let us do just that.

Theorem 3(ii). For all odd primes p, if $p \equiv 1 \pmod 4$ then p can be written as the sum of two squares.

Proof. As just mentioned, the Legendre-Euler-Gauss Law of Quadratic Reciprocity assures us that we can find integers A_1 and B_1 such that $A_1^2 + B_1^2$ divides p, specifically we will have

$$A_1^2 + B_1^2 = M_1 p$$

with $M_1 < p$. Now choose two more integers u and v with $u \equiv A_1 \pmod{M_1}$ and $v \equiv B_1 \pmod{M_1}$, both satisfying $-\frac{1}{2}M_1 \leq u, v \leq \frac{1}{2}M_1$. By the elementary properties of congruences we will have

$$\begin{aligned} u^2 + v^2 &\equiv A_1^2 + B_1^2 \pmod{M_1} \\ &\equiv 0 \pmod{M_1}. \end{aligned}$$

Which implies that

$$u^2 + v^2 = M_1 r.$$

Here I assert that the number r satisfies $1 \leq r < M_1$. Clearly r must be non-negative since it is the sum of real squares. So to prove that $r \geq 1$ we only need to show that $r \neq 0$. We can show this by contradiction, so assume that $r = 0$. From the previous line this means that $u^2 + v^2 = 0$, and so $u = v = 0$. But remember that $u \equiv A_1 \pmod{M_1}$ and $v \equiv B_1 \pmod{M_1}$, so if $r = 0$ then $A_1 \equiv B_1 \equiv 0 \pmod{M_1}$, that is A_1 and B_1 are divisible by M_1. It follows that $A_1^2 + B_1^2$ is divisible by M_1^2, but we already know that $A_1^2 + B_1^2 = M_1 p$, so this implies that M_1 divides p. But we also know that $M_1 < p$, so by the fact that p is prime we must have $M_1 = 1$. But then we have that $A_1^2 + B_1^2 = p$, and so we will be done. We are assuming it does not work straight away, and so we have arrived at our contradiction, thus $r \neq 0$, and so $r \geq 1$.

I also asserted that $r < M_1$, which we can see directly using the fact that u and v are between $-\frac{1}{2}M_1$ and $\frac{1}{2}M_1$. So:

$$r = \frac{u^2 + v^2}{M_1} \leq \frac{(M_1/2)^2 + (M_1/2)^2}{M_1} = \frac{M_1}{2} < M_1$$

as required.

So we now have that $u^2 + v^2 = M_1 r$ and $A_1^2 + B_1^2 = M_1 p$, we can multiply these expressions together to get

$$(u^2 + v^2)(A_1^2 + B_1^2) = M_1^2 r p.$$

At this point we need a useful identity which apparently stems from nowhere, but does in fact come from the very world we are investigating, the complex numbers. The identity is the following.

Lemma 4. $(u^2 + v^2)(A_1^2 + B_1^2) = (uA_1 + vB_1)^2 + (vA_1 - uB_1)^2$.

The obvious proof is to just multiply out the brackets and show equality. How does this identity[7] relate to the complex numbers, you may wonder. Well consider

[7] Discovered by Fibonacci, known to some as Leonardo de Pisa, and published in 1202 [4].

the two complex numbers $u + vi$ and $A_1 + B_1 i$. Now look back at the proof of Lemma 1.

You should notice that the above is exactly the result that $N(u + vi)N(A_1 + B_1 i) = N((u + vi)(A_1 + B_1 i))$. So we have already proved this lemma, as well as shown where it comes from.

Back to our proof of Theorem 3(ii), and we had that $(u^2+v^2)(A_1^2+B_1^2) = M_1^2 rp$. Using the identity of Lemma 4 gives us

$$(uA_1 + vB_1)^2 + (vA_1 - uB_1)^2 = M_1^2 rp.$$

If we could divide by M_1^2 at this point then we would gain an expression of the form $A_2^2 + B_2^2 = M_2 p$, with $A_2 = \frac{uA_1+vB_1}{M_1}, B_2 = \frac{vA_1-uB_1}{M_1}$, and $M_2 = r < M_1$. In this case we would have achieved what we set out to do, and by repeating all of this enough times M_i would shrink to one and as we saw above we would have $A_i^2 + B_i^2 = p$. All that we need to do, then is to show that we can indeed divide $(uA_1+vB_1)^2+(vA_1-uB_1)^2$ by M_1^2, or equivalently that $uA_1+vB_1 \equiv vA_1-uB_1 \equiv 0 \pmod{M_1}$. This is reassuringly straightforward.

First we have

$$uA_1 + vB_1 \equiv A_1A_1 + B_1B_1 \equiv M_1 p \equiv 0 \pmod{M_1}.$$

And next

$$vA_1 - uB_1 \equiv B_1A_1 - A_1B_1 \equiv M_1 p \equiv 0 \pmod{M_1}.$$

And thus we are done, and Theorem 3 is complete. □

We now know which primes can be written as the sum of two squares, those congruent to 1 modulo 4, and the special case of 2. But our original desire was to find which natural numbers could be written as the sum of two squares. As is often the case, we can use the fact that prime numbers are the building blocks of the natural numbers in order to build up a proof for all natural numbers using the fact we know the result for the primes. The key to our attack is Lemma 4.

Suppose we have a number $n = p_1 p_2 \cdots p_r$, with the p_i being distinct primes. Also suppose that each $p_i \equiv 1 \pmod 4$ or is equal to 2. Then by Theorem 3 we can write each p_i as the sum of two squares, say $p_i = \alpha_i^2 + \beta_i^2$, so we now have

$$n = (\alpha_1^2 + \beta_1^2)(\alpha_2^2 + \beta_2^2) \cdots (\alpha_r^2 + \beta_r^2).$$

But by Lemma 4 the product of the sums of two squares is equal to a single sum of two squares, so that $(\alpha_1^2 + \beta_1^2)(\alpha_2^2 + \beta_2^2) = (\alpha_1\alpha_2 + \beta_1\beta_2)^2 + (\beta_1\alpha_2 - \alpha_1\beta_2)^2 = A_1^2+A_2^2$, where the A_i are chosen for equality. But we can then apply this procedure repeatedly to transform our expression for n into a single sum of two squares. Formally, a simple proof by induction will show this. So we now know that if all the prime factors of a given number n are either congruent to 1 modulo 4 or equal to 2 then n can be written as a sum of two squares.

That is not quite the whole story, though, since this implication is one way, the fact that n can be written as the sum of two squares does not imply that all its prime factors are 1 modulo 4 or 2. Take, for example, $98 = 7^2 + 7^2$ which factors as $2 \cdot 7 \cdot 7$. Not all of its prime factors are congruent to 1 modulo 4 nor equal to 2, yet 98 has just been written as the sum of two squares. Or how about $490 = 7^2 + 21^2$, that factors as $2 \cdot 5 \cdot 7 \cdot 7$. Notice that in both cases the two numbers being squared are divisible by seven, and (clearly) the result of their sum is divisible by seven squared. Dividing through by 7^2 we get that these two results are simply $2 = 1^2 + 1^2$ and $10 = 1^2 + 3^2$ masquerading as something more poignant.

More generally, for any number n which can be written as the sum of two squares, say

$$n = a^2 + b^2,$$

we can simply multiply through by any square number s^2 to get the new case

$$s^2 n = (sa)^2 + (sb)^2.$$

Thus $s^2 n$ can be written as the sum of two squares. And conversely if s^2 divides the number n then it must divide the two squares, thus we can divide through to get

$$\frac{n}{s^2} = \left(\frac{a}{s}\right)^2 + \left(\frac{b}{s}\right)^2,$$

so that n/s^2 can also be written as the sum of two squares. What this tells us is that if n can be written as $s^2\tilde{n}$ then we can ignore the s^2 since it will only cloud the issue. What we need to really look at is the prime factorisation of $\tilde{n}$ where $\mu(\tilde{n}) \neq 0$ [8]. Then the same result applies as we noticed earlier in the page.

Theorem 5. If n is a positive integer of the form $n = p_1 p_2 \cdots p_r S^2$ where the p_i are distinct primes, then n can be written as the sum of two squares if and only if every p_i satisfies one of $p_i = 2$ and $p_i \equiv 1 \pmod 4$.

This theorem has given us almost everything we need in order to classify the Gaussian primes, which is done in the following theorem.

Theorem 6(i) (Classification of Gaussian primes Part 1)**.** Let u be a unit of $\mathbb{Z}[i]$, then the following are all Gaussian primes.

(i) $u(1+i)$,
(ii) $u(a+bi)$ where $a^2 + b^2 = p$ for some prime number $p \equiv 1 \pmod 4$,
(iii) uq where q is a prime in $\mathbb{Z}$ satisfying $q \equiv 3 \pmod 4$.

Proof. We observed just before embarking on our look at which numbers can be written as the sum of two squares that if $N(\alpha)$ is an ordinary prime then α is a Gaussian prime. Since $N(1+i) = 2$ is prime and by hypothesis $N(a+bi)$ is prime then we have that (i) and (ii) are indeed Gaussian primes. Finally for (iii) we take an ordinary prime p such that $p \equiv 3 \pmod 4$ and assume it can be factorised as $(a+bi)(c+di)$. We can then use Lemma 1 to see that $N(p) = N(a+bi)N(c+di)$, or that

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

Since p is prime the only ways to factor p^2 are as $p \cdot p$ or as $1 \cdot p^2$. But we don't want either of the factors to be a unit so we require

$$a^2 + b^2 = p \qquad \text{and} \qquad c^2 + d^2 = p.$$

But $p \equiv 3 \pmod 4$ so by Theorem 3 the above two equations have no solutions, and so p cannot be factorised, thus it is a Gaussian prime. □

[8] Möbius' mu-function is defined using the prime factorisation of the number in question. If $n = p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$ then $\mu(n) = \begin{cases} t \text{ if } e_1 = e_2 = \ldots = e_t = 1 \\ 0 \text{ if } e_i > 1 \text{ for some } i. \end{cases}$

This is not the whole classification. We have shown that the above Gaussian Integers (hereon referred to as Category (i), (ii), and (iii) primes) are indeed primes, but we have not shown that all Gaussian Primes are one of the above Gaussian Integers. We only need one more result to do that.

Lemma 7. For any Gaussian integer $\gamma = a + bi$:

(1) If 2 divides $N(\gamma)$ then $1 + i$ divides γ;
(2) If q is a category (iii) prime and q divides $N(\gamma)$ in $\mathbb{Z}$ then q divides γ in $\mathbb{Z}[i]$;
(3) If $\pi = s + ti$ is a category (ii) prime and $N(\pi) = p$ divides $N(\gamma)$ in $\mathbb{Z}$ then one or both of π and $\overline{\pi}$ divide γ in $\mathbb{Z}[i]$.

Proof. (1) By hypothesis, 2 divides $N(\gamma) = a^2 + b^2$, so clearly a and b are either both odd or both even. Either way 2 must divide $a + b$ and $-a + b$, so $\frac{a+bi}{1+i} = \frac{(a+b)+(-a+b)i}{2}$ is itself a Gaussian integer, and so by definition, $1 + i$ divides $a + bi$.

(2) We know that $q \equiv 3 \pmod 4$ and that q divides $a^2 + b^2$. That means that $a^2 + b^2 \equiv 0 \pmod q)$, or equivalently that $a^2 \equiv -b^2 \pmod q$. Here we need to use Legendre's symbols to complete the result. For an odd prime p and an integer $m(\not\equiv 0 \pmod p)$ the Legendre symbol $(m \mid p)$ is defined in terms of "quadratic residues". Essentially, if there exists an integer x such that $m \equiv x^2 \pmod p$ then $(m \mid p) = 1$. If no such x exists then $(m \mid p) = -1$. One of the many useful facts about Legendre symbols is that they multiply quite nicely, that is: $(m \mid p) \cdot (n \mid p) = (mn \mid p)$. Using this fact[9] we can see that $a^2 \equiv -b^2 \pmod q$ implies that $(a^2 \mid q) = (-b^2 \mid q)$. But also that $(a^2 \mid q) = (a \mid q)^2$ and that $(-b^2 \mid q) = (-1|q)(b^2 \mid q) = (-1|q)(b|q)^2$. We now once again use the Law of Quadratic Reciprocity (see Theorem 3(ii)) which tells us that since $p \equiv 3 \pmod 4$, $(-1 \mid q) = -1$. Thus we now have:

$$(a \mid q)^2 = -(b \mid q)^2.$$

But this is surely a contradiction since Legendre symbols equal ±1, and so this seems to say that $1 = -1$. The contradiction arises because of the proviso that we defined $(m \mid p)$ for $m \not\equiv 0 \pmod p$. The only way out it seems is to suppose that $a \equiv b \equiv 0 \pmod q$. Thus $a = q\alpha$ and $b = q\beta$, and $\gamma = a + bi = q\alpha + q\beta i = q(\alpha + \beta i)$, thus γ is divisible by q as required.

(3) By hypothesis p divides $N(\gamma)$, or equivalently $N(\gamma) = a^2 + b^2 = pR$ for some natural number R. We just need to show that one of

$$\frac{\gamma}{\pi} = \frac{(as + bt) + (-at + bs)i}{p} \quad \text{and} \quad \frac{\gamma}{\overline{\pi}} = \frac{(as - bt) + (at + bs)i}{p}$$

is a Gaussian integer.

Remember that $p = N(\pi) = s^2 + t^2$. Note first the following,

$$\begin{aligned}(as + bt)(as - bt) &= a^2s^2 - b^2t^2\\ &= a^2s^2 - b^2(p^2 - s^2)\\ &= (a^2 + b^2)s^2 - pb^2\\ &= pRs^2 - pb^2.\end{aligned}$$

[9] Proved by Gauss, among others, in Theorem 98 of Disquisitiones Arithmeticae[5].

Since p divides $pRs^2 - pb^2$ it must also divide $(as+bt)(as-bt)$. But one of the most important facts in number theory is that if $p|xy$ then $p|x$ or $p|y$, so we know that p divides at least one of $(as+bt)$ and $(as-bt)$. Using exactly the same techniques shows that p divides at least one of $(-at+bs)$ and $(at+bs)$. So p definitely divides at least one real part and one imaginary part in the above two numerators. The worst case scenario is that p divides just one real part and one imaginary part, and that these parts are not from the same number. If p divides $(as+bt)$ *and* $(-at+bs)$ then we are done and γ/π is a Gaussian Integer. Similarly if p divides $(as-bt)$ *and* $(at+bs)$ then we are done and γ/π is a Gaussian Integer. So suppose that neither of these are true, and that p divides $(-at+bs)$ and $(as-bt)$. So p also divides any linear combination of these two, specifically, p divides

$$(-at+bs)a-(as-bt)b = -a^2t+abs-abs+b^2t = (b^2-a^2)t.$$

Recall that $p = s^2+t^2$ so p cannot divide t, thus p must divide b^2-a^2, as well as a^2+b^2 by hypothesis. So we can combine these to see that p divides both $2a^2$ and $2b^2$ since

$$2a^2 = (a^2+b^2)-(b^2-a^2) \qquad \text{and } 2b^2 = (a^2+b^2)+(b^2-a^2).$$

Since $p \neq 2$ we must have p dividing a and b, so let $a = pa'$ and $b = pb'$. Then

$$\gamma = a+bi = p(a'+b'i) = (s^2+t^2)(a'+b'i) = \pi\overline{\pi}(a'+b'i).$$

So in fact in this case γ is divisible by π and $\overline{\pi}$.

Exactly the same argument shows that in the case of the other bad scenario exactly the same thing happens. Thus (3) is proved and the Lemma complete. We are now ready to complete the classification.

□

Theorem 6(ii) (Classification of Gaussian Primes Part 2)**.** Every Gaussian Prime is of one of the forms (i), (ii), and (iii) from Theorem 6(i).

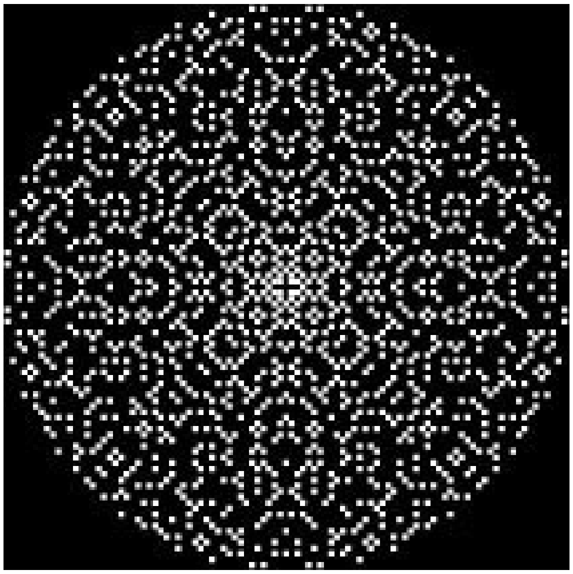
Proof. Suppose that $\rho = a+bi$ is a Gaussian prime. By definition $N(\rho) \geq 2$, so there is at least one prime q (in the integers) that divides $N(\rho)$.

First suppose that $q = 2$. Part (1) of Lemma 7 tells us that $1+i$ divides ρ, but ρ is a prime so it must be of the form $u(1+i)$ for some unit u, i.e. in category (i).

Now suppose that $q \equiv 3 \pmod 4$. Part (2) of Lemma 7 tells us that q divides ρ, but again, ρ is a prime so this can only happen if $\rho = uq$ for some unit u, that is if ρ is of category (iii).

Finally assume that $q \equiv 1 \pmod 4$. Theorem 3 tells us that q can be written as the sum of two squares, say $q = \alpha^2+\beta^2$, while part (3) of Lemma 7 tells us that $\rho$is divisible either by $\alpha+\beta i$ or $\alpha-\beta i$. Thus ρ is equal to either $u(\alpha+\beta i)$ or $u(\alpha-\beta i)$. In particular $N(\rho) = a^2+b^2 = \alpha^2+\beta^2 = q$, so ρ is from category (ii). Since all primes are either 2 or odd they are all either 2, congruent to 1 mod 4, or congruent to 3 mod 4. Thus we have covered all primes and completed our classification of the Gaussian Primes. □

When plotted on the complex plane the Gaussian Primes appear as in the diagram below. The rotational symmetry is a side effect of the fact that if $a + bi$ is a Gaussian prime then so is $b + ai$, and so are $u(a + bi)$ and $u(b + ai)$ for any unit u, which acts as a rotation. Just as with the normal integers, the Gaussian Primes form a foundation for the Gaussian Integers, and every Gaussian integer can be expressed uniquely (up to multiplication by units and the order of the factors) as a product of Gaussian Primes. This can be proved using results from Algebra II, by showing that $\mathbb{Z}[i]$ forms a Euclidean domain, and that every Euclidean Domain is a Unique Factorisation Domain. The results in this essay are more number theoretical and less algebraic, and hopefully highlight some of the underlying beauty of the numbers Gauss considered to be the true integers.



References

[1] G.H. Hardy & E.M. Wright An Introduction to the Theory of Numbers 5th Edition (Oxford Science Publications)

[2] H. Davenport The Higher Arithmetic 6th Edition (Cambridge University Press)

[3] J.H. Silverman A Friendly Introduction to Number Theory 2nd Edition (Prentice Hall)

[4] Paul Erdös & János Surányi Topics in the Theory of Numbers (Springer)

[5] Carl Friedrich Gauss Disquisitiones Arithmeticae; translated by Arthur A. Clarke (Yale University Press)

[6] Leonard Eugene Dickson History of the Theory of Numbers Volume I (Chelsea Publishing Company)