4000/6000 Day 14 Complex numbers,

We have seen how to construct a simple field $Q(\sqrt{2})$ only slightly larger than the rationals and containing a square root of 2. In the same way we construct a field $C = R(\sqrt{-1})$ only slightly larger than the real numbers, and containing a square root of -1. It is traditional to denote $\sqrt{-1}$ by i, (except for electrical engineers, who use i for "current" maybe, and j for $\sqrt{-1}$). More precisely:

Definition: A complex number is defined to be an ordered pair (a,b) of real numbers. We denote (1,0) = 1, and (0,1) = i. Then we write the complex number (a,b) as a+bi.

Addition of complex numbers: The sum of two complex numbers a+bi and c+di, is (a+c) + (b+d)i. In pairs notation, this is "vector addition", i.e. add the components separately: (a,b) + (c,d) = (a+c,b+d).

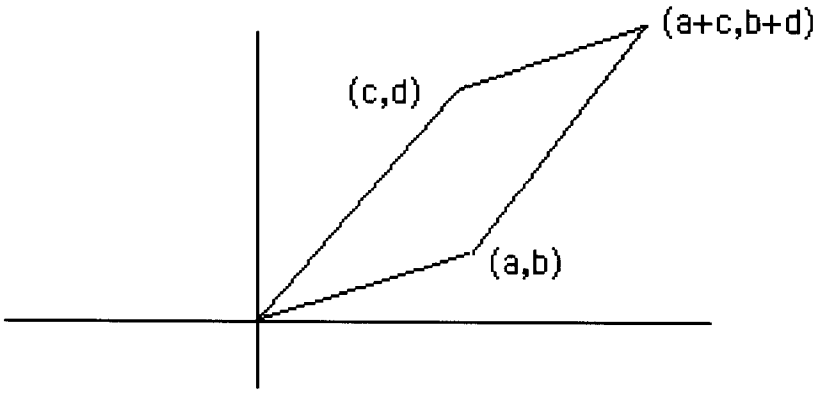
Multiplication of complex numbers: The product of two complex numbers is best remembered in the a+bi notation. I.e.

$(a+bi)(c+di) = ac + bci + adi + bdi^2 = (ac-bd) + (ad+bc)i.$

In the pairs notation it looks like this: (a,b)(c,d) = (ac-bd, ad+bc).

Geometry of complex numbers: Since a complex number is given by an ordered pair of real numbers, it makes sense to graph complex numbers in the plane, i.e. to picture the number a+bi as located at the point (a,b) in the plane. Then we can hope for a geometric interpretation of addition and multiplication.

As we have mentioned, addition is "vector addition", i.e. it satisfies the parallelogram law:

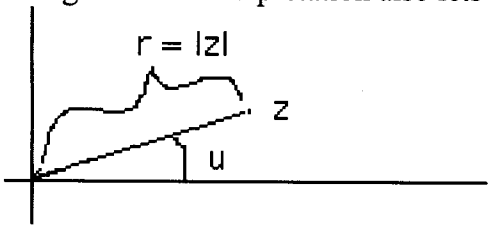


Then it makes sense to define the length or absolute value of a complex number by the distance formula, i.e. by the Pythagorean theorem.

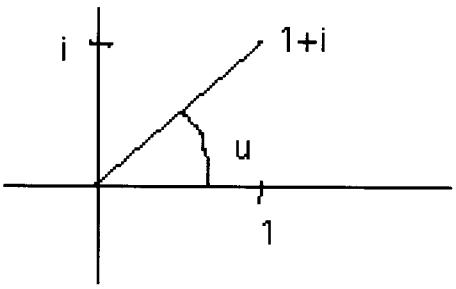
The absolute value of the complex number a+bi is defined to be $(a^2+b^2)^{1/2}$. This is often denoted |a+bi|. If we use the shorthand notation of z = a+bi, then $|z| = (a^2+b^2)^{1/2}$.

Angle (or argument) of a complex number

The geometric interpretation also lets us define the angle of a complex number.



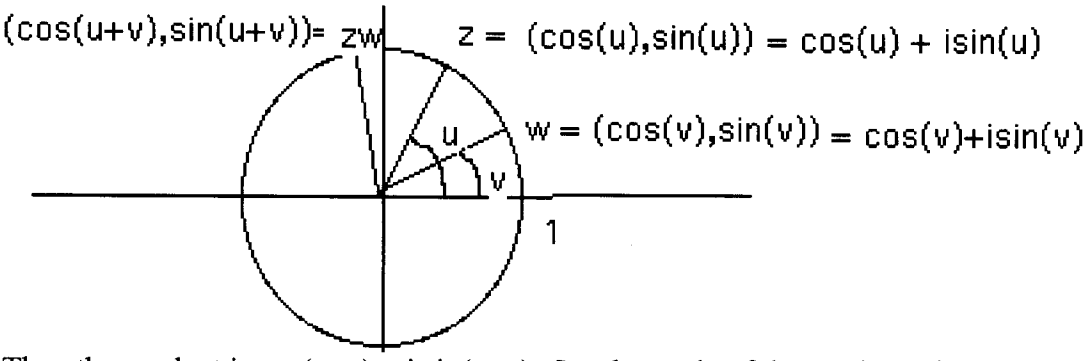
In this case, given the complex number z, draw a line segment from the origin to z. The length of the segment is the absolute value of the complex number r = |z|, and the angle u made by the line segment z with the positive x axis is the angle of z, arg(z).



E.g., $|1+i| = \sqrt{2}$, $u = \arg(1+i) = \pi/4 = 45^0$.

The Geometry of multiplication.
Complex multiplication also has a nice geometric meaning in terms of length and angles. Look back at the formula for multiplication, **(a,b)(c,d) = (ac-bd, ad+bc)**, and think about angles. Ring any bells? Remember the trig formulas for cosine and sine of the sum of two angles?
cos(u+v) = cos(u)cos(v) - sin(u)sin(v),
and **sin(u+v) = cos(u)sin(v) + sin(u)cos(v).**

These formulas are exactly the same as the multiplication rule! I.e. if we have two complex numbers z and w, which happen to have length one, then they define points on the unit circle. Consequently they can be written as cos(u)+i sin(u), and cos(v)+ i sin(v).



Then the product is cos(u+v) + i sin(u+v). I.e. the angle of the product of two complex numbers is the sum of their angles.
for a general complex number, not necessarily of length one, we can write it as a multiple of a complex number of length one as follows. If |z| = r, then z/r has length one, so if u = arg(z), then

$z/r = \cos(u) + i \sin(u)$. Thus $z = r(\cos(u) + i \sin(u))$. Another complex number w, with length $|w| = s$ and angle $\arg(w) = v$, can be written $w = s(\cos(v) + i \sin(v))$.

Thus if $z = r(\cos(u) + i \sin(u))$, and $w = s(\cos(v) + i \sin(v))$,

then

$$zw = r(\cos(u) + i \sin(u))\, s(\cos(v) + i \sin(v))$$

$$= rs([\cos(u)\cos(v) - \sin(u)\sin(v)] + i[\cos(u)\sin(v) + \sin(u)\cos(v)])$$

$$= rs(\cos(u+v) + i \sin(u+v)).$$

I.e. to multiply two complex numbers,
we multiply their lengths and add their angles.

Corollary: I.e. $|zw| = |z||w|$,
and $\arg(zw) = \arg(z) + \arg(w)$ $(\pm 2\pi)$.

Notice that angle 2π = angle 0, so angles are equivalent a bit like modular numbers. I.e. two angles u,v are equivalent if and only if $u - v = 2n\pi$, where n is an integer. Thus arg(z) should really be a point on the unit circle instead of a number. I.e. what is well defined is really the cosine and sine of the angle, i.e. the pair $(\cos(\arg(z)), \sin(\arg(z)))$, rather than the angle itself.

Remark: How to remember the trig addition laws. Since the function arg(z) changes products into sums, i.e. $\arg(zw) = \arg(z) + \arg(w)$, it behaves like a logarithm function, and its inverse $u \dashrightarrow \cos(u) + i \sin(u) = e^{iu}$, is an exponential function. This greatly simplifies the laws of adding angles in cosines and sines. I.e. we know exponentials change addition into multiplication, so $e^{i(u+v)} = e^{iu}e^{iv}$. Since $e^{ix} = \cos(x) + i\sin(x)$, this says that

$\cos(u+v) + i \sin(u+v) = e^{i(u+v)} = e^{iu}e^{iv}$
$= (\cos(u) + i \sin(u))(\cos(v) + i \sin(v))$
$= [\cos(u)\cos(v) - \sin(u)\sin(v)] + i[\cos(u)\sin(v) + \sin(u)\cos(v)]$.

Setting real and imaginary parts equal, gives us back the trig laws for adding angles in cosines and sines. This makes it easier to remember them for most of us, since we only need remember $e^{ix} = \cos(x) + i\sin(x)$.

Solving equations in complex numbers. Now we can easily solve all equations of form $X^n - 1 = 0$. Just subdivide the circle into n equal parts starting at 1 = (1,0). Those are the complex nth roots of 1. Since these points have length one and their angles are evenly distributed around the circle, they are:

$\cos(2\pi/n) + i \sin(2\pi/n)$,

$\cos(4\pi/n) + i \sin(4\pi/n)$,

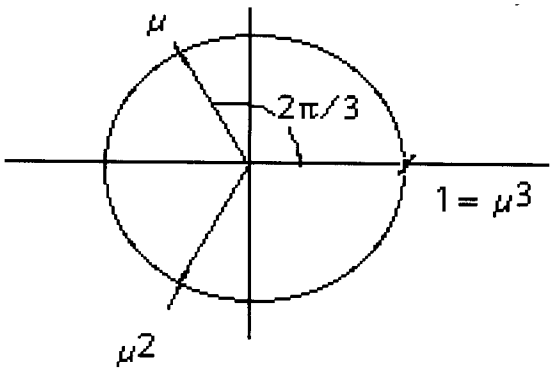
$\cos(6\pi/n) + i \sin(6\pi/n)$

.....

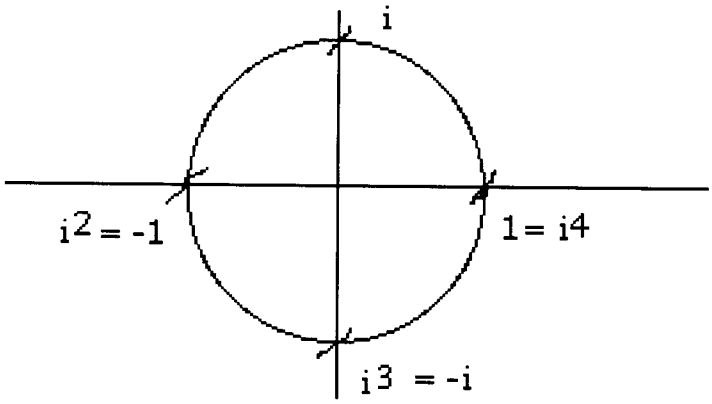
$\cos(2n\pi/n) + i \sin(2n\pi/n) = \cos(2\pi) + i \sin(2\pi) = \cos(0) + i \sin(0) = 1$.

Analog between modular arithmetic and complex nth roots of 1. Notice that if we set the first one, the one with the smallest angle, equal to $\mu = \cos(2\pi/n) + i \sin(2\pi/n)$, then the others are all powers of this one. I.e. Then the nth roots of 1 have form μ, μ^2, μ^3,.....,$\mu^n = 1$. We call the one μ which generates the others, a "primitive nth root of 1". There is a perfect analog here of modular arithmetic, mod n. I.e. the integer k mod n, corresponds to the power μ^k. Then $\mu^k = 1$ if and only if k is congruent to 0 mod n, if and only if n divides k. Recall we often call modular arithmetic "clock arithmetic" because it is modeled on addition on a circle, which is exactly how complex multiplication works for numbers of length one.

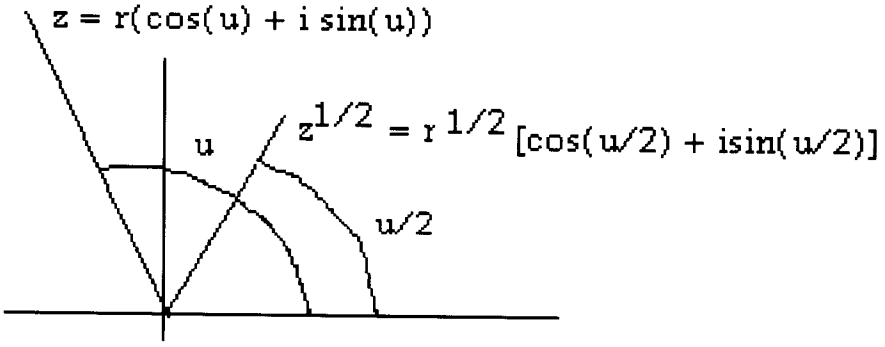
For example, let $u = 2\pi/3 = 120^0$. Then the solutions of $X^3 - 1 = 0$ are $\cos(u) + i \sin(u) = \boldsymbol{\mu}$, $\cos(2u) + i \sin(2u) = \boldsymbol{\mu^2}$, and $1 = \boldsymbol{\mu^3}$.



Another example are the solutions of $X^4 - 1 = 0$. They are at angles of $\pi/2$, π, $3\pi/2$, $4\pi/2 = 2\pi = 0$. I.e. angles 90^0, 180^0, 270^0, and $360^0 = 0^0$. This time we write $i = \cos(\pi/2) + i \sin(\pi/2)$, instead of μ.



The complex number field C is the smallest field containing the reals, in which the equation $X^2 - 1 = 0$ has a solution. It is amazing that then automatically every other polynomial equation with real (or complex) also has a solution in C. Without giving a proof, we can suggest the reason. It turns out all one needs is that every real polynomial of odd degree has a real root, and that every complex number has a square root. Putting these together one can handle equations of all degrees. By the intermediate value theorem we already have solutions of real equations of odd degree. Now I claim we also have solutions of equations of degree two. I.e. the quadratic formula gives a solution of an equation of degree two, provided you can take a square root. But in the complex numbers every number has a complex square root. I.e. just take half the angle, and take the square root of the length.

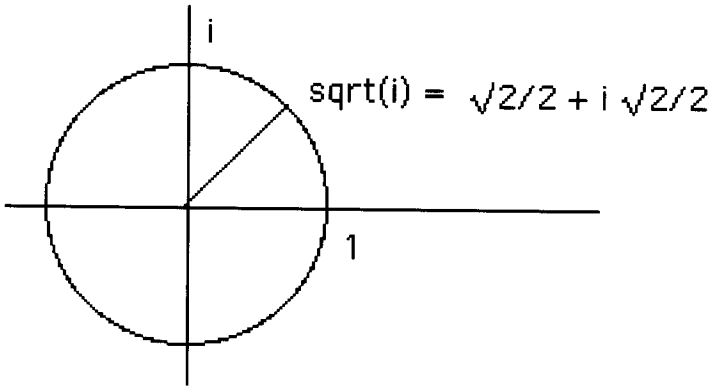


For instance, if z = 1+2i, then |z| = sqrt(1^2+2^2) = sqrt(5), and arg(z) = arctan(2) = u. Hence sqrt(z) = sqrt(5)(cos(u/2) + i sin(u/2)), where u = arctan(2). I didn't say the answer was pretty, I just said it exists.

Here is a prettier one.

Since arg(i) = $\pi/2$, and |i| = 1, arg($\sqrt{i}$) = $\pi/4$, and $|\sqrt{i}|$ = 1, so we have

$\sqrt{i}$ = cos(π/4) + i sin(π/4) = ($\sqrt{2}$/2) + i ($\sqrt{2}$/2).



To see C is a field, since 1 has length one and angle zero, given $z = a+bi \neq 0$, we have to find $w = c+di$ such that the lengths multiply up to one, and the angles add to zero. Just take length $w = 1/|z|$, and $\arg(w) = -\arg(z)$. I.e. if $z = r(\cos(u) + i \sin(u))$ is not zero, then $r \neq 0$, so then

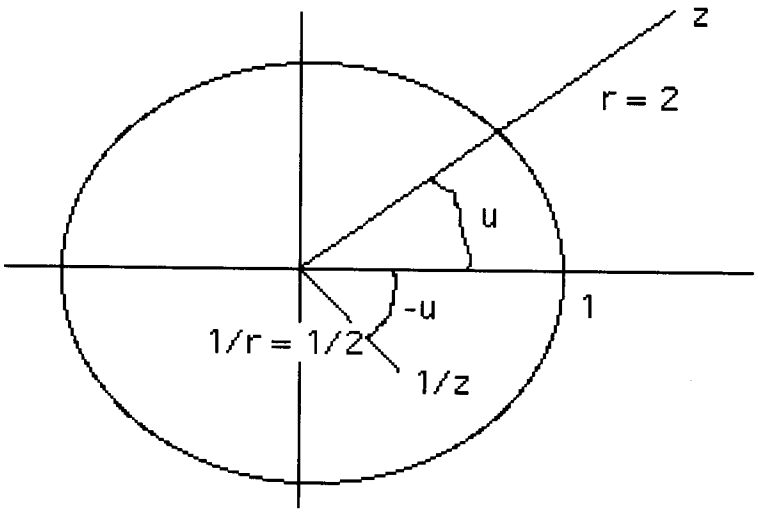
we have $1/z = (1/r)(\cos(-u) + i \sin(-u))$, and since cos is an even function while sin is an odd function, this is $1/z =$

$(1/r)(\cos(u) - i \sin(u))$.

Notice the x coordinate keeps the same sign and the y coordinate changes sign. That means we are reflecting the point about the x axis. Thus to change z into 1/z, we scale the length r to become 1/r, i.e. we "reflect z in the unit circle", and then we reflect about the x axis.

For example, if $z = (x,y)$ is on the unit circle, then $1/z = (x, -y)$, and

if $z = r(x,y)$, where (x,y) is on the unit circle, then $1/z = (1/r)(x, -y)$.



Another way to see the multiplicative inverse is algebraically.

If $z = a+bi \neq 0$, and the "conjugate" $\hat{a}z = a-bi$, then $|z|^2 = z\hat{a}z = a^2+b^2 \neq 0$,

then

$1/z = (1/z)(\hat{a}z/\hat{a}z) = \hat{a}z/z\hat{a}z = \hat{a}z/|z|^2 = (a-bi)/(a^2+b^2)$.

Notice that a-bi is the reflection of a+bi in the x axis. So to get 1/z, we reflect z in the x axis and then divide by the square of the length, i.e. the length of 1/z is the reciprocal of the length of z.

For example, if $z = 2 + 3i$, since $|z|^2 = 2^2 + 3^2 = 13$,

then $1/z = (2-3i)/13 = (2/13) - i\,(3/13)$.

Note since $i = 0+ 1i$, the conjugate is -i, and $|i| = 1$, so we have

$1/i = -i$. That is, $(i)(-i) = 1$.

4000/6000 Day 15, Complex ("Gaussian") integers.
It turns out that we can learn a lot about ordinary integers by studying "complex integers", i.e. complex numbers where both real coefficients are integers. Since Gauss studied these we define the ring Z[i] = {a+bi, where a,b are ordinary integers} to be the "Gaussian integers".

Here is an important connection between them and ordinary integers.
Definition of norm. For each Gaussian integer z = a+bi, define the "norm" of z, Nm(z) = $|z|^2$ = a^2+b^2, to be the square of the length of the integer.

Lemma: If z, w are Gaussian integers, then
(i) Nm(z) is an ordinary non negative integer,
(ii) z = 0 if and only if Nm(z) = 0,
(iii) Nm(zw) = Nm(z)Nm(w).
Proof: Exercise. (Recall we already know that |zw| = |z||w|.)

Just as every ordinary integer can be factored using 1 or -1, every Gaussian integer can be factored using 1, -1, i or -i. I.e. since i(-i) = 1, for any Gaussian integer z we have z = (iz)(-i). E.g. 2 -3i = (3+2i)(-i). We want to ignore such "trivial" factorizations in Z[i], just as we ignored factorizations like 7 = (-7)(-1) in Z. We make the following definition.

Definition of units.
A non zero Gaussian integer z is a "unit" if its multiplicative inverse 1/z is also a Gaussian integer.

Lemma: (i) If z = a+bi is a Gaussian integer then z is a unit if and only if the Nm(z) = $|z|^2$, equals 1.
(ii) The only Gaussian integers which are units are { 1,-1,i,-i }.
Proof: Since the inverse of z = a+bi is 1/z = (a-bi)/(a^2+b^2), it follows that if a^2+b^2 = 1, then 1/z =a-bi is a Gaussian integer, so z is a unit. Conversely if z is a unit, then both a/(a^2+b^2) and b/(a^2+b^2) must be integers. If $z \neq 0$, then at least one of a or b is not zero. If say $a \neq 0$, and a/(a^2+b^2) is an integer, then $|a| \geq (a^2+b^2)$. Since a and b are integers, this is only possible if |a| = 1 and b = 0, i.e. then z = 1 or -1. If on the other hand $b \neq 0$, then since b/(a^2+b^2) is an integer, we must have b = 1 or -1, i.e. z = i or -i. **QED.**

We are interested in factoring Gaussian integers into "primes".
Definition: A Gaussian prime is a Gaussian integer z such that if z = ab, where a and b are Gaussian integers, then at least one of a or b is a unit.
I.e. a Gaussian integer z fails to be prime, if it has a factorization z = ab, where neither a nor b is a unit.

Lemma: If a Gaussian integer z is not prime in Z[i], then its norm Nm(z) = $|z|^2$ is not prime in Z.
Proof: If z is not prime in Z[i], then z = ab, where neither a nor b is a unit in Z[i]. Then if we take norms of both sides we get $|z|^2 = |a|^2\ |b|^2$. Since neither a nor b is a unit, we know neither

number $|a|^2$ nor $|b|^2$ is 1. Since they are both positive, $|z|^2$ is not a prime integer. **QED.**

Corollary: If z is a Gaussian integer, such that Nm(z) is a prime integer, then z is a Gaussian prime.
Proof: This is equivalent to the previous statement. **QED.**

Example: $z = 2+3i$ is a Gaussian prime since $2^2+3^2 = 13$ is a prime integer. Also 5-2i is a Gaussian prime, since 25+4 = 29 is a prime integer. However the test fails for z = 3 and z = 5, since in these cases we have Nm(z) = 9, and Nm(z) = 25, which are not prime integers. This gives us no information. In fact 5 is not a Gauss prime since $5 = (2+i)(2-i) = 2^2+1^2$. It turns out later that 3 is a Gauss prime however. The observation we wish to make is that the difference between these cases is that 5 can be written as a sum of two squares, but 3 cannot!

Next we come to the main point, that connects arithmetic in Z[i] with Fermat's problem of writing a prime as a sum of two squares.
Lemma: Given a prime integer p, p can be written as a sum of two squares in Z, if and only if p has a non trivial factorization in Z[i], i.e. if and only if the prime integer p is not a Gaussian prime.
Proof: Suppose first that $p = a^2+b^2$ with a and b ordinary integers. Then p = (a+bi)(a-bi) is a factorization of p in Z[i]. To see it is non trivial, we must show that a+bi is not a unit. We know a+bi is a unit if and only if $a^2+b^2 = p = 1$. But since p is a prime integer $p \neq 1$.

Next suppose p is not gaussian prime, i.e. that p has a factorization as p = (a+bi)(c+di) where neither factor is a unit. Then taking Norms, gives $p^2 = (a^2+b^2)(c^2+d^2)$, a factorization into natural numbers, where no factor equals 1. By uniqueness of prime factorization in Z, both factors on the right side must equal p. I.e. $p = (a^2+b^2) = (c^2+d^2)$. Thus p is a sum of two squares. **QED.**

Now we can conclude that 3, 7, 11, 19, 23, ... must be Gaussian primes because they cannot be written as sums of two squares.
Corollary: If p is a prime in Z which is congruent to 3 mod 4, then p is also a Gaussian prime.
Proof: If p were a prime in Z which is not a Gaussian prime, then we have just seen that p would be a sum of two squares in Z. But we already know that no number congruent to 3 mod 4, is a sum of two squares in Z. **QED.**

Note: Since $2 = 1^2+1^2$ that 2 = (1+i)(1-i) is not a Gaussian prime. We claim that no primes of form 4k+1 are Gaussian primes either, and that hence they can all be written as sums of two squares. The proof is more subtle than it might at first appear.

Lemma: If a+bi is a Gaussian integer and n an ordinary integer, then n divides a+bi in Z[i], if and only if n divides both a and b in Z.
Proof: If n divides a and b in Z, then a = nc, b = nd, for some integers c,d. Then a+bi = nc+ndi = n(c+di). Conversely if n divides a+bi in Z[i], so that n(c+di) = a+bi, then a = nc and b = nd, by definition of multiplication in Z[i]. **QED.**
In particular, if either a or b is 1 or -1, then no integer ≥ 2 can divide a+bi.

Proposition: If p is a prime in Z of form 4k+1, then p is not a Gaussian prime, and hence $p = X^2 + Y^2$ has a solution in Z.

Attempted Proof: Let p = 4k+1. Then by problem #15, HW #4, there is a square root of -1 mod p. Thus there is an ordinary integer k ,such that $k^2 = -1 \pmod{p}$. I.e. such that p divides k^2+1 in the ordinary integers. I.e. we have an integer m such that $mp = k^2+1$. But this equation still holds in the Gaussian integers, and in the Gaussian integers this says that $mp = k^2+1 = (k+i)(k-i)$. Thus p divides the product (k+i)(k-i) in the Gaussian integers.

What if p were a Gaussian prime? If p were a Gaussian prime, and if Gaussian primes behave the way ordinary primes behave, then we would conclude that p must divide either (k+i) or (k-i). This however is impossible since the coefficients of i in these factors are 1, -1, and since $p \geq 2$, hence p does not divide 1.

The missing link: Thus we would know that p is not a Gaussian prime provided we knew the prime divisibility property were true for Gaussian primes. I.e. this shows p does not have the prime divisibility property for Gaussian integers. But does that say p is not a Gaussian prime? I.e. how do we know that Gaussian primes have the prime divisibility property? This is not obvious, since in $Z[\sqrt{-5}]$, the property fails. I.e. 2 and 3 are prime, but $6 = (2)(3) = (1+\sqrt{-5})(1-\sqrt{-5})$, and 2 does not divide $(1+\sqrt{-5})$.

So we must go back and see what the proof of the prime divisibility property was, and if the proof can be generalized to Gaussian integers.

Review of proof of prime divisibility property.

If p is prime and p divides ab then p divides either a or b.

The proof was by contradiction. I.e. suppose p does not divide say b. Then p is relatively prime to b, so we can write 1 as a linear combination of p and b. I.e. 1 = np + mb. Then multiplying by a, we get a = anp + amb. Then since p divides anp and amb, p divides a.

What do we need for this proof?

To run this proof again we need the part about writing 1 as a linear combination of p and b. I.e. we need to know that if p is a Gaussian prime and if p does not divide the Gaussian integer a+bi, that then we can write 1 as a linear combination of p and a+bi.

How did we write 1 as a linear combination?

It used the concept of "gcd". I.e. we considered the smallest linear combination of p and a+bi and showed this smallest number divides both p and a+bi. Since p was prime, this smallest number must be 1, -1 i or -1. So at least we could write one of those as a linear combination of p and a+bi. Then multiplying through by -1, i or -i, we could also write 1 as a linear combination.

GCD for Gaussian integers.

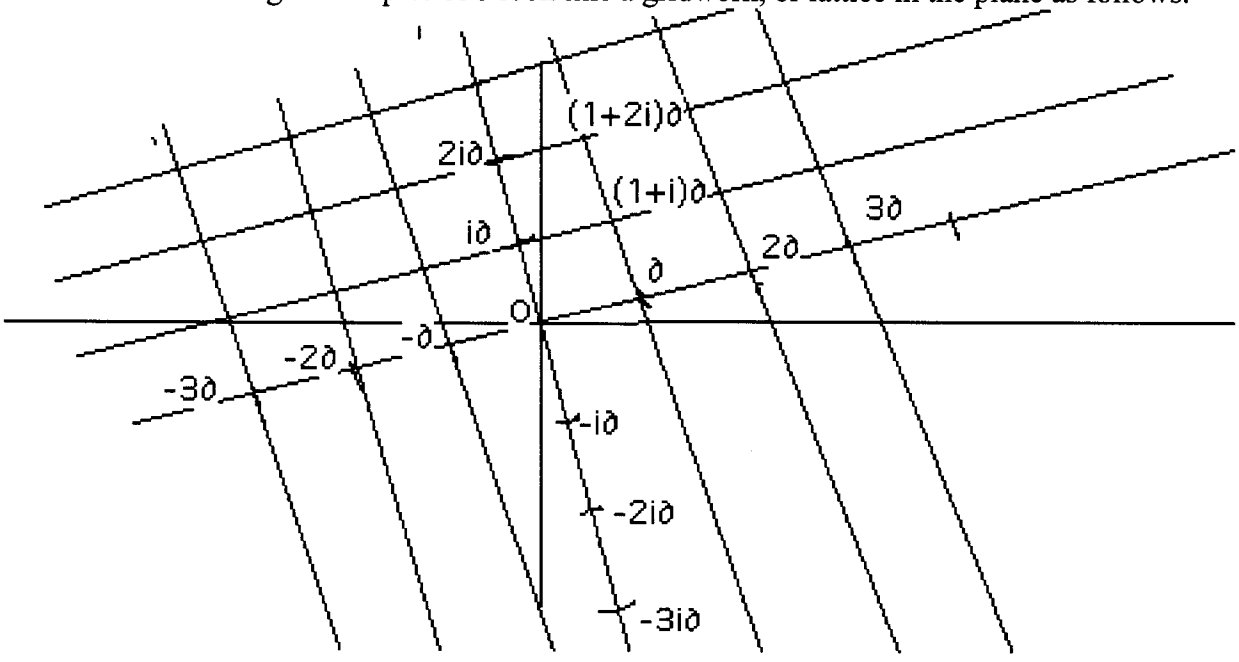
Now what do we mean by smallest or largest for a Gaussian integer? It would be natural to take it to mean the length, or to get an integer, the squared length, or norm. So let ∂ be the Gaussian integer of smallest norm which is a linear combination of p and a+bi. Then we claim that ∂ divides both p and a+bi.

More generally let z, w be Gaussian integers, and let's try to find a "gcd" for them, where

"greatest" means greatest norm.
So given z,w in Z[i], not both zero, let ∂ be a non zero Gaussian integer of least norm such that ∂ is a linear combination of z and w. (Note ∂ is not unique since then $-\partial$, $i\partial$, and $-i\partial$ all have the same norm, and they are also linear combinations of z and w. Just multiply the linear combination equation for ∂ by -1, i or -i, to get a new linear combination.)

We claim that ∂ divides both z and w. How to prove that? We need to divide ∂ into say z, and show that if it does not divide evenly, then the remainder is smaller than ∂, and get a contradiction. I.e. the remainder is also a linear combination of z and w, so cannot be smaller than ∂.
To do that we need a "division algorithm"! We try to imitate the division algorithm for integers, in the new setting of Gaussian integers.

Division algorithm for Gaussian integers.
Given two Gaussian integers z, ∂, where $\partial \neq 0$, there exist Gaussian integers u,v such that
(i) $z = u\partial + v$, and
(ii) $|v| < |\partial|$.
Note: We do not need the uniqueness of u and v, which is a good thing, since they are not going to be unique in this case.
Proof: This says we can find a Gaussian multiple $u\partial$ of ∂, which is closer to z than ∂ is to 0. Just look at the picture of all the multiples of ∂ in the complex plane. Note that integer multiples $n\partial$ of ∂ are just spread out along the line through 0 and ∂, and that $i\partial$ is perpendicular to ∂, so multiples of form $mi\partial$ are spread out along the line perpendicular to ∂.
I.e. all Gaussian integer multiples of ∂ look like a gridwork, or lattice in the plane as follows.



We are asking whether any vertex in this lattice, i.e. any Gaussian integer multiple of ∂, is closer to z than $|\partial|$. But z must lie in one of those squares, and any point of a square is at least as close to some vertex as the center of the square. Now the diagonal of a square of side $|\partial|$ has length $|\partial|(\sqrt{2})$, so the center of the diagonal has distance $|\partial|(\sqrt{2}/2)$ from each vertex. Since $(\sqrt{2}/2) < 1$,

this is less than $|\partial|$. **QED.**

Note what failed with this argument for $Z[\sqrt{-5}]$. Then the grid is composed of rectangles of side $|\partial|$ and $\sqrt{5}|\partial|$, instead of squares. Thus the diagonal has length $|\partial|(\sqrt{6})$, and half of that is $|\partial|(\sqrt{6}/2)$ which is greater than $|\partial|$. So in fact if the number $z = n+m\sqrt{-5}$ is too near the center of one of these rectangles, there might not be a vertex close enough to z to satisfy the division algorithm. But it is not so obvious.

GCD'S for Gaussian integers.
Definition: A Gaussian integer z is "reducible" if and only if it is neither a unit nor prime, i.e. it has a factorization $z = ab$, where neither a nor b is a unit.

Hence a Gaussian integer z is reducible if and only if it can be factored as $z = ab$ where both $Nm(a) > 1$, and $Nm(b) > 1$, equivalently where both $Nm(a) < Nm(z)$ and $Nm(b) < Nm(z)$.

Definition: If z,w are Gaussian integers, not both zero, the gcd(z,w) is a Gaussian integer ∂ such that
(i) ∂ divides both z and w, and
(ii) if μ also divides both z and w, then $Nm(\mu) \leq Nm(\partial)$.

Thus among all common factors of z and w, ∂ is one of largest norm. Since the norm of a common factor of z and w is at most equal to
$\min\{ Nm(z), Nm(w)\}$, there is only a finite set of a possible norms of common factors. Hence some of them have largest possible norm.

We will show that a gcd of z and w is universal in the same way as for ordinary integers. I.e. if ∂ is a gcd of z and w, then any common divisor
of z and w divides ∂. We proceed exactly as before.

Lemma: If z w, are Gaussian integers, not both zero, consider all possible linear combinations az+bw, where a,b are Gaussian integers. If $d = az+bw$ is such a linear combination having smallest non zero norm, then d divides both z and w in Z[i]. In fact d is a gcd of z and w, and any other common divisor of z and w divides d.
Proof: Divide say z by d using the division algorithm. We get $z = ud + r$, where $Nm(r) < Nm(d)$. But by the generalized three term principle, since both z and d are linear combinations of z,w, so is r. Since r is a linear combination of z and w and has smaller norm than d, r must have norm zero. I.e. $r= 0$. Thus d divides z. Similarly d divides w.

Now we know d is a common divisor of z and w. To see any other common divisor divides d, just observe that any common divisor of z and w must divide any linear combination of z and w, by the three term principle. Thus if μ is any other common divisor of z and w, μ divides d, so $Nm(\mu) \leq Nm(d)$. Thus d is a gcd of z and w. **QED.**

Corollary: Conversely if ∂ is any gcd of z and w, then ∂ is also a linear combination of z and w of smallest norm, hence every other common divisor of z and w divides ∂.
Proof: Let d be any linear combination of z and w of smallest norm, as in the lemma. Then d is a common divisor of z and w, and ∂ divides d. Hence $Nm(\partial) \leq Nm(d)$, But ∂ has largest norm among all common divisors of z and w. Hence $Nm(d) = Nm(\partial)$. Thus since $d = \partial\mu$, so that

$Nm(d) = Nm(\partial)Nm(\mu)$, we must have $Nm(\mu) = 1$, so μ must be a unit. If we have $d = az+bw$ and multiply by $1/\mu = \Omega$, we get $\partial = a\Omega z+b\Omega w$, so that ∂ is also a linear combination of z and w. **QED.**

Notice: The gcd of two integers is not unique, since we can multiply any gcd by any unit and get another gcd. Thus each pair of Gaussian integers, not both zero, has exactly four gcd's, obtained from any one of them by multiplying by 1,-1,i,-i. This corresponds to considering either d or -d as the gcd of two ordinary integers. For ordinary integers, even though there were essentially two gcd's, it was possible to single out one of them by choosing only the positive one. For Gaussian integers there is no concept of positive, so we are stuck with considering all 4 of them.

Definition: Two Gaussian integers are called relatively prime, if and only if their gcd is a unit, if and only if their only common divisors are the numbers {1,-1,i,-i}. We could also say their gcd is 1, since that is one of the units, and in this case there is a nice way to pick just one of the four gcd's, since 1 is a distinguished member of the set {1,-1,i,-i}.

Lemma: Two Gaussian integers z,w are relatively prime if and only if we can write 1 as a linear combination of z and w, with coefficients in Z[i].
proof: This is what we have proved above. **QED.**

Prime divisibility property for Gaussian integers.
Proposition: If z is a Gaussian prime, and z divides the product ab of Gaussian integers, then z divides either a or b.
Proof: If z does not divide a then the gcd of z,a is 1, so write $1 = zu+av$. Then multiply by b and get $b = bzu+bav$. Now z divides both terms on the right side, hence z divides b. **QED.**

This completes what we needed for the proof that if p is a prime in Z of form 4k+1, then $p = X^2 + Y^2$ has a solution in Z.

Unique factorization of Gaussian integers.
Recall that once we had the prime divisibility property we could prove the unique factorization theorem. In exactly the same way we get the following theorem in Z[i].
Theorem: If z is any Gaussian integer which is not a unit, i.e. any Gaussian integer except 1,-1, i, -i,
(i) there exist Gaussian primes $w_1,..,w_n$ such that $z = \prod w_j = (w_1)..(w_n)$.
(ii) The prime factors w_j in (i) are unique "up to units and ordering". I.e. if $z = \prod t_j$ is any other factorization of z into Gaussian primes t_j, then there are exactly n of the t_j's, and they may be renumbered so that for every j we have $t_j = w_j$, $t_j = -w_j$, $t_j = iw_j$, or $t_j = -iw_j$. I.e. up to ordering, the t_j are "unit multiples" of the w_j.
Proof: If you want to master this subject, you should go through this proof to see that it works. It will also be a good exercise in reviewing the proof of unique factorization for integers. I.e. go back and take the proof we gave for integers, and rewrite it in the present context. **QED.**

Remarks:(i) The prime divisibility property turns out to be so important, that in algebra we actually change the definition of the word "prime" to mean anything having that property. I.e. we call a number in an integral domain "prime" if it is not a unit, and whenever it divides a product, it must divide at least one of the factors.
(ii) Then we need a new name for the old concept. So we call a non unit "irreducible" it cannot be factored without at least one factor being a unit.
(iii) Since we can easily derive the prime divisibility property from the uniqueness of prime factorization, it follows that the concepts of "prime" and "irreducible" are equivalent in a domain with unique factorization into irreducibles.
More precisely: a prime element, i.e. one with the prime divisibility property, is always irreducible. If we have unique factorization into irreducibles, then conversely every irreducible is prime.

E.g. in $Z[\sqrt{-5}]$, and in $Z[\sqrt{-3}]$, 2 is "irreducible" but not "prime", since 2 divides $6 = (1+\sqrt{-5})(1-\sqrt{-5})$ and $4 = (1+\sqrt{-3})(1-\sqrt{-3})$, but 2 does not divide any of these factors. Hence these rings do not have unique factorization, and thus they do not have a division algorithm either.

Practical division algorithm for Gaussian integers
We have proved if $\partial \neq 0$ and w, are Gaussian integers, there must be a multiple of ∂ which is closer to w than $|\partial|$, but have not shown how to find one. The proof of the division algorithm in the book, and accompanying examples, show how to find one. Namely to divide w by ∂ in Z[i], just divide it in Q(i), and then approximate the quotient by an element of Z[i]. I.e. suppose $w/\partial = u$ in Q(i). We know this problem can be solved in Q(i) since Q(i) is a field. In fact if $\partial = a+bi$, then $1/\partial = (a-bi)/(a^2+b^2)$, so $w/\partial = w(a-bi)/(a^2+b^2)$. Then if $w/\partial = u = x+iy$ where x,y are rational numbers, then we can choose integers c,d as near as possible to x and y. I.e choose integers c,d such that $|c-x| \leq 1/2$, and $|d-y| \leq 1/2$. Consider $\mu = c+di$. Then $u-\mu = (x-c) + i(d-y)$ has coefficients each at most 1/2 in length. Hence $|\mu-u)|^2 = (x-c)^2 + (y-d)^2 \leq 1/4 + 1/4 = 1/2$, so $|\mu-u| \leq \sqrt{2}/2 < 1$. Consequently, since $w/\partial = u$, and $u = \mu + |u-\mu|$, we have $w = \partial(w/\partial) = \partial u = \partial(\mu + [u-\mu]) = \partial\mu + \partial(u-\mu)$, where $|\partial(u-\mu)| = |\partial||u-\mu| < |\partial|$. This does what we want.
E.g. Given $w = 3-4i$, and $\partial = 1+i$, we have $1/\partial = (1-i)/2 =$
$(1/2) - i\,(1/2)$. Thus $w/\partial = (3-4i)(1/2 - i/2) = 3/2 -2 -2i -3i/2 =$
$-1/2 -7i/2$. We may take $\mu = -3i$. Then $w = \partial\mu + r$ where $r = w-\mu\partial =$
$3-4i -(3-3i) = i$. Note $|i| = 1 < |1+i| = |\partial|$. Since -1/2 is equally close to 0 and -1, and -7/;2 is equally close to -3 and -4, w e could also have chosen $\mu = -1-3i$, or -1-4i, or -4i.
Please check me on the arithmetic here.

4000/6000

Factorization of Gaussian integers into Gaussian primes

Definition: A complex or "Gaussian" integer z = a+bi is a Gaussian prime if and only if
(i) z is not a Gaussian unit, i.e. not equal to 1,-1,i,or ,-i, and
(ii) the only way to factor it as z = uw, is when either u or v is a unit.

"Trivial factorizations." Any Gaussian integer z can be factored using a unit, since z = (1)z = (-1)(-z) = (i)(-iz) = (-i)(iz). These are so called "trivial" factorizations. A prime is a non unit that has no factorizations other than these trivial ones. Of course a unit, such as i, also has only trivial factorizations, but we do not choose to call a unit a prime in Z[i], just as we did not call the ordinary integers 1, or -1 primes in Z.

Recognizing primes in Z[i]

It is no easier to recognize primes in Z[i] than it is to recognize primes in Z. So we need some tests that can help. One test we had for primes in Z was the Fermat little theorem. I.e. if p is a prime in Z, then for every integer n, n^p is congruent to n mod p. That might not seem very useful on small numbers p, but it is easy to program a computer to perform this check, and then it is very fast even on really large numbers p. Note this test does not guarantee a number is prime, as it is only a way to check that a number is not prime. I.e. if p fails the test that n^p should be congruent to n for any one integer n, then p is not prime. (Actually Professor Rumely here at UGA showed how to modify this test a few years ago, into a slightly different test that does recognize primes, but I do not know what other conditions his test requires.)

For Gauss integers we have a test that works in the other direction. I.e. the following test guarantees certain Gaussian integers to be Gaussian primes. Recall that for a Gaussian integer z = a+bi, we have $|z| = \sqrt{a^2+b^2}$, and $Nm(z) = |z|^2 = (a^2+b^2)$. Moreover |zw| = |z||w|, and hence also Nm(zw) = Nm(z)Nm(w).

A test that recognizes some Gaussian primes

Lemma: **(i)** If z,w are Gaussian integers such that w divides z, then Nm(w) divides Nm(z).
(ii) If z = a+bi is a Gauss integer such that $Nm(z) = a^2+b^2$ is a prime in Z, then z is a prime in Z[i].

Proof: **(i)** This is because the norm of a product is the product of the norms. I.e. if z = vw, then Nm(z) = Nm(v)(Nm(w).
(ii) Also the norm of a non unit is larger than 1. Thus if z is a non prime in Z[i], and z can be factored as z = vw where neither v nor w is a unit in Z[i], then Nm(z) = Nm(v)Nm(w), whether neither Nm(v) nor Nm(w) is 1. Hence if z is not prime in Z[i], then Nm(z) is not prime in Z. Equivalently, if Nm(z) is prime in Z, then z is prime in Z[i]. **QED.**

Examples of Gauss primes: The previous test lets us construct lots of Gauss primes. E.g. 1-2i is a Gauss prime since its norm is 5, a prime integer. 1+4i is a Gauss prime since it has norm 17, which is prime in Z. 2+5i is a Gauss prime since it has norm 29, a prime integer. 9+4i is a Gauss prime since its norm 97 is prime in Z. Notice however that 3 is a Gauss prime even though this test does not recognize it, since Nm(3) = 9 is not prime. We discuss this case next.

When is an ordinary prime integer also a Gauss prime?

For ordinary integers n, the test above is of no use, since the norm n^2 is never prime, so we need a different test to tell whether an ordinary integer is a Gauss prime. Of course if n is not prime in Z, then n is still not prime in Z[i][, since a non trivial factorization of n in Z, is also a non trivial factorization in Z[i]. So, we only need a test for which prime integers p remain prime in Z[i].

You might think that a prime integer in Z will remain prime in Z[i], but there are more numbers in Z[i] as possible factors, so maybe it can be factored non trivially there. In fact $5 = 1^2+2^2 = (1+2i)(1-2i)$ shows that 5 is not prime in Z[i]. Also $17 = 1^2+4^2 = (1+4i)(1-4i)$ is no longer prime in Z[i]. Again $29 = 5^2+2^2 = (5+2i)(5-2i)$ is not prime in Z[i]. The key point is whether or not the prime p is a sum of two squares. In fact this gives a test that works for all ordinary integers.

Lemma:
(i) An integer n in Z, which is not prime in Z is also not prime in Z[i].
(ii) An integer p which is prime in Z, remains prime in Z[i] if and only if p cannot be written as a sum of two squares in Z.
Proof: (i) If n = ab, where a,b are integers other than 1 or -1, then the equation n = ab is still true in Z[i], and a,b, are still not units in Z[i]. So n is also not prime in Z[i].
(ii) If p is prime in Z and $p = a^2+b^2$ in Z, then $p = (a+bi)(a-bi)$ in Z[i]. Since neither a nor b can be zero, neither a+bi nor a-bi is a unit in Z[i], so we get a non trivial factorization in Z[i].

Conversely, if p is not prime in Z[i], then $p =(a+bi)(c+di)$ where neither factor is a unit in Z[i]. Then taking norms of both sides gives $p^2 = (a^2+b^2)(c^2+d^2)$ in Z, where neither factor is 1. Hence the right side is a non trivial factorization of p^2. Thus the unique prime factorization of p^2 must be obtained by factoring the two factors (a^2+b^2) and (c^2+d^2) into primes. But there are only two prime factors of p^2. Thus both (a^2+b^2) and (c^2+d^2) must equal p or -p. Since those factors are both positive, in fact we must have $(a^2+b^2) = p = (c^2+d^2)$. I.e. we have written p as a sum of two squares. **QED.**

Corollary: Since a prime of form 4k+3 cannot be a sum of two squares it always remains prime in Z[i]. E.g. 19 is prime in Z[i].

Example of prime factorization in Z[i].
(i) Given the integer 390, we factor it in Z as 390 = 13(30) = (13)(2)(3)(5). Then in Z[i], 3 remains prime, but $13 = 2^2+3^2 = (2+3i)(2-3i)$, and $2 = 1^2+1^2 = (1+i)(1-i)$, and $5 = 1^2+2^2 = (1+2i)(1-2i)$. Thus in Z[i], we have $390 = (3)(2+3i)(2-3i)(1+i)(1-i)(1+2i)(1-2i)$. Moreover each of the factors except 3 is also Gaussian prime, since each has prime norm.

(ii) To factor a more interesting Gaussian integer like 7-3i, take the norm getting 58 = 2(29), so it can only be factored into two factors with norms 2 and 29. If we try 1+i we get $(7-3i)/(1+i) = (7-3i)(1-i)/2 = (4-10i)/2 = 2-5i$. So $7-3i = (1+i)(2-5i)$. Now both 1+i and 2-5i have prime norms, so they are Gaussian primes. Of course since (i)-i) = 1, we could also factor $7-3i = i(1+i)(-i)(2-5i) = (-1+i)(-5-2i)$, but that is not really different.

Sums of two squares.
Now we are interested in Fermat's problem of writing primes as sums of two squares, and we see

it is equivalent to the problem of determining when a prime integer remains a Gauss prime. Thus we need tools for determining when prime integers remain Gauss primes. We assert the following theorem of Fermat.

Theorem (Fermat): if p is an odd prime integer, the following properties are all equivalent:
(i) $p = 4k+1$ for some natural number k.
(ii) There is an integer n such that n^2 is congruent to -1 (mod p).
(iii) p is not prime in Z[i].
(iv) $p = a^2 + b^2$ for some integers a,b.
Proof: We have already proved (by working mod 4) that (iv) implies (i), and (in homework #4 solutions) that (i) implies (ii). We proved just above that (iii) is equivalent to (iv). Thus to complete the cycle all we need to do is prove that (ii) implies (iii). We sketched this in class Monday, and it is proved in the previous notes. Recall the approach here.

(ii) implies (iii). If there is an integer n such that n^2 is congruent to -1 (mod p), then n^2+1 is congruent to zero mod p. I.e. p divides n^2+1 in Z. Thus there is an integer k such that $pk = n^2+1$. Now look at this equation in Z[i]. It says that p divides $n^2+1 = (n+i)(n-i)$. But we claim p cannot divide (n+i) because p does not divide 1. I.e. if p(a+bi) = pa + pbi, so if p divides x + yi, then p divides both x and y.

Thus if there is an integer n with n^2+1 congruent to zero mod p, then p does not have the "prime divisibility property" in Z[i]. But does that mean p is not prime in Z[i]? I.e. does every prime in Z[i] have the prime divisibility property? We have not proved this. Thus we need to see if we can prove that primes in Z[i] do have the prime divisibility property.

Let's review the proof from the ring Z, and try to reproduce it here. recall the proof we used before. Suppose p is a prime in Z and that p divides ab, but p does not divide b. Then we want to show that p divides a. Our first step was to show that we could write 1 as a linear combination of p and b. I.e. if we could write 1 = np + mb, then we could multiply by a and get a = anp + amb. Since p divides ab, then p divides both terms on the right, hence p also divides the term on the left. I.e. then p divides a.

Thus we need to prove if z is a Gaussian prime, and z does not divide a Gaussian integer w, then we can write 1 as a linear combination of z and w. Recall how we proved this. If z is prime and z does not divide w, then the only common divisors of z and w must be units. So we needed to show that some linear combination of z and w does divide both z and w. I.e. we need the following result.

Smallest linear combinations.
If z and w are Gaussian integers, and d is the smallest linear combination of z and w, i.e. the one having smallest norm, then d divides both z and w.
Recall that to prove that result, we used the division algorithm.

Division algorithm for Gauss integers.
If ∂,w are Gaussian integers and if $\partial \neq 0$, then there are Gaussian integers μ and r such that:
(i) $w = \partial\mu + r$, and
(ii) $0 \leq Nm(r) < Nm(\partial)$.
proof: (as in book and in Day 15 notes.) To divide w by ∂ in Z[i], just divide it in Q(i), and then

approximate the quotient by an element of Z[i]. I.e. suppose $w/\partial = u$ in Q(i).

We know this problem can be solved in Q(i) since Q(i) is a field. In fact if $\partial = a+bi$, then $1/\partial = (a-bi)/(a^2+b^2)$, so $w/\partial = w(a-bi)/(a^2+b^2)$. Then if $w/\partial = u = x+iy$ where x,y are rational numbers, we can choose integers c,d as near as possible to x and y.

I.e choose integers c,d such that $|c-x| \leq 1/2$, and $|d-y| \leq 1/2$. Consider $\mu = c+di$. Then $u-\mu = (x-c) + i(d-y)$ has coefficients each at most 1/2 in length. Hence $Nm(\mu-u) = |\mu-u|^2 = (x-c)^2 + (y-d)^2 \leq 1/4 + 1/4 = 1/2 < 1$.

Since $w/\partial = u$, and $u = \mu + |u-\mu|$, we have $w = \partial(w/\partial) = \partial u = \partial(\mu + [u-\mu]) = \partial\mu + \partial(u-\mu)$. Then if we define $r = w - \partial\mu$, this implies r is a Gaussian integer such that $w = \partial\mu + r$. By the expression for w just above, we have $r = w - \partial\mu = \partial(u-\mu)$, so $Nm(r) = Nm(\partial)Nm(u-\mu) \leq (1/2)Nm(\partial) < Nm(\partial)$. **QED.**

Corollary: Given Gaussian integers z,w, not both zero, let d be the Gaussian integer of smallest positive norm which is a linear combination of z and w. Then d divides both z and w.
Proof: Give $d = zu + wv$, we claim d divides say z. By the division algorithm, at least there is an equation of form $z = da + r$, where $Nm(r) < Nm(d)$. But then since both z and d are linear combinations of z,w, so is r. Since r has smaller norm than d, Nm(r) must be zero. Thus $r = 0$, and d divides z. Similarly d divides w. **QED.**

Corollary: If z is a Gaussian prime and z does not divide a Gaussian integer w, than we can write 1 as a linear combination of z and w.
Proof: If $d = uz+vw$ is the linear combination of z and w having smallest norm, then d divides both z and w. But z does not divide w, so d is not a unit times z. Hence d is a unit, i.e. $d^{-1} = e$ is also a Gaussian integer. If we multiply through the equation $d = uz+vw$, by e, we get $de = 1 = (eu)z+(ev)w$. Thus 1 is a linear combination of z and w, as desired. **QED.**

Note: Just as in the case of ordinary integers, this proof shows if z,w are two Gaussian integers which are relatively prime, i.e. whose only common factors are units, then we can write 1 as a linear combination of z and w.

Corollary: (Prime divisibility property in Z[i]) If z is a Gaussian prime, and z divides a product uv, then z divides either u or v.
Proof: If z divides uv but does not divide v then we can write $1 = az + bv$. Then multiplying by u gives $u = uaz + ubv$. Then z divides both terms on the right hand side, hence z divides u.
QED.

NOTE: This completes the proof of (ii) implies (iii) in Fermat's theorem above. I.e. if there is an integer n such that n^2 is congruent to -1 mod p, then p does not have the prime divisibility property in Z[i], so p cannot be prime in Z[i]. **QED.**

Finding solutions of Fermat's problem in practice.
Given a prime p of form 4k+1, how do we actually find integers a,b such that $p = a^2+b^2$? I do not know if there is a good answer to this. I.e. we know that $n = (2k)!$ gives a solution to n^2 congruent to -1 (mod p). Also if we have $p = a^2+b^2$, and k is congruent to 1/b mod p, then $(ak) = n$ also solves n^2+1 congruent to -1 (mod p). But can we go backwards? I.e. given n such that

n^2+1 congruent to -1 (mod p), can we find a and b such that $p = a^2+b^2$?

How is it possible for a proof to guarantee that a solution to a problem must exist but give no clue for finding it? This is one of the mysteries of modern mathematics. We have to accept these halfway measures, until something better is found, since at least it is better to know a solution exists than to wonder whether we are wasting our time looking for one.

4000/6000 Day 17 SUMMARY OF FERMAT THEOREM

We have argued precisely that the Fermat theorem follows from the prime divisibility property for Gaussian integers. Thus the key point was to prove that Gaussian integers do indeed have the prime divisibility property. To do this one should review and understand the proof of that property for ordinary integers and then imitate the proof in the case of Gaussian integers. Next we recall the steps in the proof. Ultimately,

The Division algorithm implies prime divisibility property.

Although you should know the proof in complete precision, we will state the parts of the proof here in a less precise way, which makes them hopefully easier to remember, and also makes it easier to see how they are generalized to the case of Gaussian integers.

Rough description of the proof:

1. Basic division theorem:

Any non zero element d can be divided into any other element a with remainder smaller than the divisor d.

(For ordinary integers the notion of "smaller" is just the usual absolute value. I.e. the remiander r satisfies $|r| < |d|$. The version of this theorem for Gaussian integers uses as a notion of size, either the absolute value of the complex numbers, or the square of the absolute value, i.e. the "norm".)

This implies:

2. Basic linear combination property: the smallest non zero linear combination of two elements (which are not both zero) divides both of them.

proof:

Given a,b, if d is their smallest linear combination, we have $d = ax+by$, and we claim d divides both a and b. Check it for a. By division property, at least d divides into a with remainder smaller than a. Thus we have

$a = dq+r$ where r is smaller than a. But now by the generalized three term principle, since a and d are linear combinations of a and b, so is r. Thus r is a smaller linear combination of a and b than d is. Since r is the smallest non zero linear combination of a and b, this is impossible unless r is zero, i.e. unless d divides a. **QED.**

(Again the concept of "smallest" used here can be the real or complex absolute value, or the norm. You should understand and be able to prove the "generalized three term principle" used here.)

This implies:

Relatively prime linear combinations: Given any two relatively prime ring elements a,b, (i.e. the only common factors of a and b are units), then 1 can be written as a linear combination of the two elements.

Proof: Let d be the least non zero linear combination of a and b. Then d divides both a and b, so since they are relatively prime, d is a unit. Multiplying through the linear combination $d = ax+by$ by the inverse e of d, gives 1 as a linear combination of a and b. I.e. then $1 = ed = a(ex) + b(ey)$. **QED.**

This implies the
prime divisibility property: if a prime divides a product of two elements it divides one of them.
Proof: If p is prime and p divides ab, but p does not divide a, then p and a are relatively prime, so we can write $1 = np + ma$. Then $b = bnp + bma$. Then p divides both terms on the right hence p divides the left. **QED.**

By now many people can repeat this last short proof. That is essential. But do not be satisfied with knowing just this one tiny step in the arguments. Learn them all. And how they fit together. And know the precise versions in the other notes, not just these abbreviated, and colloquial versions.

Remarks: The prime divisibility property in turn implies the unique factorization property for integers as before. I.e. the prime divisibility property for ordinary integers implies the unique factorization of ordinary integers, and the prime divisibility property for Gaussian integers implies the unique factorization of Gaussian integers. HOWEVER, we have not even proved the existence of a prime factorization property for Gaussian integers.
QUESTION: How would you prove
1) Every Gaussian integer which is not a unit, has at least one prime factor.
2) Every Gaussian integer which is not a unit, factors completely into Gaussian primes.
After proving that, then the prime divisibility property will prove uniqeness of the prime factorization. Assuming existence of prime factorization, you should be able to prove uniqueness both ordinary and Gaussian integers (uniqueness up to multiplying by units, and reordering the factors of course).

Since appropriate versions of all this hold in Z[i], we obtain:
Every Gaussian prime has the prime divisibility property.

Hence we deduce: If p is an ordinary prime such that $X^2-1 = 0$ has a solution mod p, then p is not a Gaussian prime.
Proof: If n^2-1 is congruent to zero mod p, then p divides n^2-1, i.e. $kp = n^2-1$, for some k in Z. Then in Z[i], p divides $n^2-1 = (n-i)(n+i)$. If p is a Gaussian prime then p divides either n-i or n+i, but this is impossible. (An ordinary integer cannot divide a+bi unless it divides both a and b, and here b = 1 or -1.) Hence p is not a Gaussian prime, so p factors as $p = (a+bi)(c+di)$, where neither factor is a Gaussian unit.

Then we deduce that an ordinary prime p which is not a Gaussian prime must be a sum of two squares in Z.
I.e. Since p is not a Gaussian prime, p factors as $p = (a+bi)(c+di)$, where neither factor is a Gaussian unit, i.e. neither a^2+b^2, nor c^2+d^2 equals 1.
Taking norms of both sides in $p = (a+bi)(c+di)$, gives $p2 = (a^2+b^2)(c^2+d^2)$. By hypothesis, neither norm on the right is 1, i.e. both factors (a^2+b^2) and (c^2+d^2) are ≥ 2. Hence the prime factorization of the right hand side consists of the prime factors of (a^2+b^2) and the prime factors of (c^2+d^2). But it must be the same as the factorization of the left side p^2. Since there are only two primes on the left, namely p and p, there can be only two primes on the right, and both are p. Thus $(a^2+b^2) = p = (c^2+d^2)$. We have written p as a sum of two squares. **QED.**

Since we know (solutions for hw#4, read them!) for every p of form 4k+1, there is a solution of $X^2-1 = 0 \pmod p$, it follows that every prime of form 4k+1 is a sum of two squares a^2+b^2. (Finding a and b is harder.)

Remark: Jan has observed that the two complex cube roots of 1, are remarkably similar to each other. In fact they are "identical" twins. There is no way to tell one from the other arithmetically. The same is true of i and -i. Each is the cube of the other. This is different from the properties of 1 and -1. I.e. $1^2 = 1$, but $(-1)^2 = 1$ also. So 1 and -1 are not interchangeable, as i and -i are.

This symmetry gives rise to the so called Galois group of an equation. I.e. the group of the equation X^n-1 is the symmetries of the nth roots of 1. Not all nth roots of 1 are identical, but the "primitive" ones are. The symmetries of the roots of this equation are just the group of generators of the group Z_n. This concept is a primary tool in studying solutions of equations, and which equations have no solutions formulas. It is studied more in the second semester of this course, i.e. in math 4010/6010.

day 18 monday March 3, review for test

Some hw problems to go over:

proof that every gaussian integer has a prime factor.

proof that every gaussian integer has unique prime factorization.

problems on complex roots of unity such as 13, 19, 21 (HW #6).

irrationality of e, using geometric series.

proof the product of all non zero numbers mod p, for p prime, is -1 mod p.

also the proof that there is a square root of -1 mod p, if $p = 4k+1$.

mention types of proof techniques, by contradiction, induction.

basic tools for arguments involving positive integers: well ordering,

properties of some rings: domains, primes, units, prime divisibility property, a notion of "size" allowing a division algortihm, linear combination property involving size, relatively prime linear combinations.

real numbers. geometric series. decimals.

Subject: outline of real numbers

1. Know that a real number is an infinite decimal, and know when two different infinite decimals give the same real number.

2. Know the reals form an ordered field, and that they satisfy the least upper bound property. In particular be able to state that property.

Be able to state the density property of rationals (prop. 2.4, p. 53). Do NOT learn the proof of 2.4 using l.u.b.'s.

3. Be able to recognize the l.u.b. of some infinite sequences of real or rational numbers. E.g. what is the l.u.b. of the sequence .9, .99, .999, .9999, ...?

what about the l.u.b. of the sequence .23, .2323, .232323, ...?
(see #5 below)

4. Be able to deduce, as in the notes, the fact that the natural numbers are not bounded above, using the l.u.b. property.

5. Be able to convert a fraction into a repeating decimal, and vice versa.

6. Be able to state the intermediate value theorem, and use it to prove certain polynomials have real roots, like X^3 + X +1.

7. Know the formula 1/(1-r) for the sum of the infinite geometric series

1 + r + r^2 + r^3 +....., of ratio r, when $0 < r < 1$.

This next one is harder, but hopefully not too hard.

8. Be able to prove the intermediate value theorem as in the notes using the concepts of continuity and l.u.b.'s.

DO NOT LEARN the PROOF THAT r^k approaches zero as k approaches infinity,

when $0 < r < 1$. This is MUCH harder than anything I expect.

Review of 4000/6000.
(Commutative) Rings and Fields
Know the axioms. We say "ring" to mean commutative ring.

Divisibility and three term principle:
If a,b are elements of a ring R, we say a divides b (in R) if there exists c in R such that ac = b.

Linear combinations: Given two elements a,b of a ring R, any element of form ax+by, with x,y, in R is called a linear combination of a and b.

Three term principles:
(i) If a + b + c = 0, and d divides two of the elements a,b,c, then it also divides the third.
(ii) If a + b + c = 0, and two of the elements a,b,c, are linear combinations of x,y, then so is the third.
(exercise)

"Units" A "unit in a ring R is an element a such that there is an element b in R with ab = ba = 1.

"Fields" A field is a ring in which <u>every non zero element</u> is a unit.

"Zero divisors" A zero divisor in a ring R is a <u>non zero</u> element a such that there is a <u>non zero</u> element b in R with ab = 0.

Note that technically every element "divides" zero since for every a we have a(0) = 0, the "trivial "case. Thus we restrict use of the term "zero divisor" to non trivial cases.

"Domain" A domain or integral domian, is a ring with no (non trivial) zero divisors.

For example the ordinary integers Z are a domain, in which the units are exactly {1, -1}.

Lemma: A unit is never a zero divisor.
proof: If ab = 1, and bc = 0, then c = 1(c) = (ab)c = a(bc) = a(0) = 0. Thus if b is a unit, we cannot have bc = 0 for a non zero c. **QED.**

Corollary: Every field is a domain.

Thus the rationals Q, the reals R, and the complexes C, are all domains.

Corollary: Every ring contained in a field is a domain.
proof: Since the field contains no zero divisors, and the ring is contained in the field, the ring contains no zero divisors either. **QED.**

E.g. The ring of Gaussian integers Z[i] is a domain since it is contained in the complex field. The units in Z[i] are exactly {1,-1,i,-i}. Also the ring Z[sqrt(-3)] of complex numbers of form {a +bsqrt(-3), a,b integers}, is a domain, in which the units appear to be just {1,-1}. Z[sqrt(2)] is a domain in which units are all numbers a+bsqrt(2) such that $a^2-2b^2 = \pm 1$. There are infinitely many of these units, e.g. $\pm 7 \pm 5\text{sqrt}(2)$, $\pm 17 \pm 12\text{sqrt}(2)$. This is one reason we prefer to work with Z[sqrt(negative number)].

Lemma: Conversely, every domain is contained in a unique smallest field, its field of fractions.
proof (sketch of existence): The construction is exactly like the construction of the rationals from the integers. If R is a domain, define the field F to consist of all pairs (a,b) of elements of R with $b \neq 0$, and with equivalence relation (a,b) = (c,d) if and only if ad = bc. Addition and multiplication are as usual. (a,b)+(c,d) = (ad+bc,bd), and (a,b)(c,d) = (ac,bd). Notice the domain property guarantees that $bd \neq 0$ if both $b \neq 0$ and $d \neq 0$. If a and b are $\neq 0$, the inverse of (a,b) is (b,a). Everything is exactly like the rationals. **QED.**

Primes and Factorization

"Primes" A prime element of a ring is an element a such that a is <u>not</u> a unit, and if a = bc, then <u>either b or c is a unit</u>.

Important question: Particularly in a domain, it is of interest to determine <u>what are the primes</u>, and <u>how they can be used to represent other elements</u>. This works especially well in rings with a notion of "size", which can be measured by a non negative integer, and where there is a "division theorem", such that one can always divide any element by any non zero element in such a way that the remainder is smaller than the divisor. The ordinary integers provide the basic example of this.

Fundamental "Well ordering" property of integers:
Every non empty subset of the non negative integers contains a smallest element.

This allows one to prove several important facts.

Existence of prime factors: If n is any integer with $|n| \geq 2$, then n has a prime factor.
proof: (Brian and Tom) Let k be the smallest factor of n with $k \geq 2$. (Since $n \neq 0$, and n, -n are both factors of n, thus n has some factors ≥ 2.) Then k is prime, for if not, then $k = rs$ where neither r nor s is a unit, and since k is positive, we may assume both $r,s \geq 2$. Then $r,s < k$, so both r,s are factors of n smaller than k and ≥ 2, a contradiction. **QED.**

Division theorem for integers: Given integers a,b with $b \neq 0$, there are unique integers q,r with
(i) $a = bq + r$,
(ii) $0 \leq r < |b|$.
proof idea: Let r be the smallest non negative integer of form a-bq, where q is an integer. Then prove that $r < |b|$. (Start by assuming a,b are positive.)

Then one deduces the
Smallest linear combination property: Given two integers a,b, not both zero, the smallest positive linear combination d of a and b, divides both a and b.

Greatest common divisors.
Given two integers a,b, not both zero, **gcd(a,b) is the largest integer dividing both a and b.** If gcd(a,b) = 1, call a and b **relatively prime.**

Corollary: Relatively prime linear combinations If a,b are relatively prime integers, then 1 is a linear combination of a and b.

Corollary: Relatively prime divisibility property. If a,b,c are integers, and gcd(a,b) = 1, and a divides bc, then a divides c.

Corollary: Prime divisibility property: If a,b,p are integers, p is a prime integer, and p divides ab, then p divides either a or b.

(Corollary):
Theorem: Unique factorization of (non unit) integers. If n is any integer except 0, 1, or -1, then
(i) n is prime or can be written as a product of prime integers.
(ii) If $n = (p_1)(....)(p_r) = (q_1)(....)(q_s)$, where all p's and q's are prime, then r = s, and after renumbering the q's we have $p_1 = \pm q_1, ..., p_r = \pm q_r$.
Proof: It suffices to prove the statements for positive integers n.
(i) If n has a prime factorization then so does -n, so let n be the smallest integer ≥ 2 for which the statement is false. Then $n \geq 2$ is not prime but has a prime factor p. Then $n = pm$, where $1 < m < n$, since $n > p \geq 2$. Then m is either prime or a product of primes $m = (q_1)(....)(q_s)$. But then also $n = pm = p(q_1)(....)(q_s)$, a contradiction. **QED existence.**
(ii) Let n be the smallest integer ≥ 2 for which the statement is false. Let $n = (p_1)(p_2)(....)(p_r) = (q_1)(q_2)(....)(q_s)$, be any two prime factorizations of n. Then by the prime divisibility property, p_1 divides both sides, so it divides some q_j, which by renumbering we may call q_1. Then since

q_1 is prime we have $q_1 = \pm p_1$. Then by cancelling, we get $(p_2)(....)(p_r) = (\pm q_2)(....)(q_s) = m < n$. if $m = 1$ we are done, if $m \geq 2$, since m is smaller than n the uniqueness statement is true for m, so we have $r = s$, and after renumbering, we have $p_2 = \pm q_2, ..., p_r = \pm q_r$. Combining that with the fact that $q_1 = \pm p_1$, we are done. **QED.**

Using unique factorization of integers

We can now prove many square roots are irrational.

lemma: If p is any prime integer, then sqrt(p) is irrational, i.e. there is no rational number r with $r^2 = p$.

proof: By contradiction. if $r = n/m$ and $(n/m)^2 = p$, then $n^2 = pm^2$. If we factor n and m each into primes, then take two of each prime, we get prime factorizations of n^2 and m^2. In particular, in the factorizations of n^2 and also of m^2, every prime occurs an even number of times, in particualr p occurs an even number of times (possibly zero, which is an even number). But then in the prime factorization of pm^2, p occurs an odd number of times, (once in front of m^2, and an even number of times "within" m^2). Thus we cannot have $n^2 = pm^2$, since p occurs an even number of times in the prime factorization of the left side but an odd number of times in the prime factorization of the right side. **QED.**

Remark: Since every integer greater than 1 has a prime factor, any two integers with a common factor > 1 also have a common <u>prime</u> factor. Thus two integers are relatively prime if and only if they have no common prime factors. In particular, if a and b are relatively prime, so are a^r, and b^s for any natural numbers r,s, since the same primes occur in powers of a, as occur in a, and the same for b.

"Fractions in lowest terms"

Claim: A non zero rational number r can be represented as a fraction n/m with n,m relatively prime and $m > 0$, in exactly one way.

proof: Assume $r > 0$. Then $r = a/b$ where a,b are positive integers. Assume this is done so that b is the smallest positive integer which can occur in the denominator (possible by well ordering). Then we claim gcd(a,b) = 1. If not, then a,b have a common prime factor say p. Thus $a = pn$, $b = pm$, for some n,m. Dividing both a,b by p, gives a new representation $r = n/m$ where $1 \leq m < b$, a contradiction to choice of b. As for uniqueness assume $a/b = c/d$ with $gcd(a,b) = 1 = gcd(c,d)$, and both $b,d > 0$. Then $ad = bc$ so by relatively prime divisibility property, b divides d and d divides b, and since both are positive they are equal. Then we may cancel b,d from $ad = bc$, getting $a = c$. **QED.**

Rational roots theorem:

If $a_nX^n + ...+ a_1X + a_0$ is a polynomial with integer coefficients $a_0, ..., a_n$, and if $r = c/d$ is a rational root in lowest terms, then c divides a_0 and d divides a_n.

proof: Substituting $X = c/d$, gives zero, and then multiplying through by d^n gives $a_nc^n + a_{n-1}c^{n-1}d + ...+ a_1cd^{n-1} + a_0d^n = 0$. Now by the "n+1 term" principle, since d divides all terms except the first, it also divides the first, i.e. d divides a_nc^n. Since c,d are relatively prime so are c^n and d, hence d divides a_n. Similarly, c divides a_0. **QED.**

Modular arithmetic "new rings from old"
Starting from the integers Z, and a natural number $n \geq 2$, we define a new ring Z_n, by setting all multiples of n equal to 0, and equating two integers which differ by a multiple of n. I.e. elements of Z_n are represented by integers k, where $k = s \pmod n$ if and only if n divides (k-s).

Remark: We do not consider Z_1, since all numbers would be equivalent and our ring would be $\{0\}$ in which case $1 = 0$, an uninteresting ring.

It is of interest to ask what are the **units in the ring Z_n.**
Lemma: In Z_n, the integer k represents a unit if and only if $\gcd(k,n) = 1$.
proof: If $\gcd(k,n) = 1$, then there is a linear combination of form $1 = ak+bn$. Then mod n, we have $1 = ak$, i.e. a and k are units.
Conversely, if k is a unit mod n, then there is some integer a such that $ak = 1 \pmod n$. I.e. $ak-1 = nm$, for some m. Then $1 = ak - nm$ is a linear combination of k,n giving 1. Thus any common divisor of k,n also divides 1. Since the only common factors of k,n are factors of 1, $\gcd(k,n) = 1$. **QED.**

Corollary: If n is prime, then Z_n is a field, since then every non zero element is a unit.

Curiously, it follows also that there are no domains among the Z_n that are not already fields. I.e.

Corollary: Z_n is a domain if and only if n is prime if and only if Z_n is a field.
proof: If n is prime we know Z_n is a field, and if Z_n is a field we know it is a domain. Thus it remains to prove only that if Z_n is a domain then n is prime. It suffices to prove the contrapositive: that if n is not prime then Z_n is not a domain. But n not prime implies $n = ab$ where $2 \leq a,b < n$. Then $ab = 0 \pmod n$, but neither a nor b is 0 mod n. **QED.**

Using modular arithmetic to prove things about integers.
Basic principle: If a polynomial equation with coefficients in Z has a solution in Z, then this solution is also a solution in every Z_n. thus if there is even one n such that the equation has no solution in Z_n, (which means that none of the numbers 0,1,2,.....,n-1, is a solution mod n), then there is no integer solution to the original equation.

Corollary: The equation $X^2 + Y^2 = n$, never has an integer solution if n is an integer of form $4k + 3$. (E.g. $X^2 + Y^2 = 1003$ has no integer solution.)
proof: try $X = 0,1,2,3$, $Y = 0,1,2,3$, mod 4. **QED.**

Corollary: The equation $X^2 + Y^2 + Z^2 = n$, never has an integer solution if n is an integer of form $8k + 7$. (E.g. $X^2 + Y^2 + Z^2 = 1007$, has no integer solution.)
proof: Try all integers 0,1,2,3,.....,7, as values of X,Y,Z. (First find all possibilities for X^2, then for $X^2 + Y^2$, then...) **QED.**

[Review reals and complexes.]

Complex ("Gaussian") integers
Let Z[i] denote the ring of complex numbers of form a+bi where a,b, are integers, and define the "norm" of a+bi as $Nm(a+bi) = a^2+b^2$, and the conjugate to be a-bi. Then the units in this ring are 1,-1, i, -i, and it is a domain. Then a+bi is a unit if and only if Nm(a+bi) = 1, and Nm(zw) = Nm(z)Nm(w), for all elements z,w, of Z[i].

Using the norm as a measure of size, we can prove, in a similar way as for integers, the following:

Existence of prime factors
Every Gaussian integer which is not a unit, has a prime factor.
(Take a factor of smallest norm ≥ 2, and prove it must be prime.)

Division theorem for Gaussian integers: Given Gaussian integers z,w with $z \neq 0$, there are Gaussian integers (not unique) q,r with
(i) $w = zq + r$,
(ii) $0 \leq Nm(r) < Nm(z)$.
proof idea: Let u = w/z as an element of the field Q(i), and take q as a closest possible approximation to u within Z[i]. Then show $Nm(w - zq) < Nm(z)$. I.e. we know we can choose a Gaussian integer q within 1/sqrt(2) of u. I.e. such that u = q+a with $|a| \leq 1/\sqrt{2}$, hence $|a|^2 \leq 1/2$. Then we have w = zu = z(q+a) = zq + za. Then r = za = w - qz is a Gaussian integer and $Nm(r) = Nm(z)|a|^2 \leq (1/2)Nm(z) < Nm(z)$, as desired. **QED.**

Smallest linear combination property: Given two Gaussian integers z,w, not both zero, any linear combination d of z and w, of smallest possible norm, divides both z and w.

Greatest common divisors.
Given two Gaussian integers z,w, not both zero, we could define **gcd(z,w) as a Gaussian integer of largest norm dividing both z and w.** If the only common divisors of z,w are units, i.e. if gcd(z,w) is a unit, call z,w **relatively prime.**

Corollary: Relatively prime linear combinations If z,w are relatively prime Gaussian integers, then 1 is a linear combination of z,w.

Corollary: Relatively prime divisibility property. If z,w,u are Gaussian integers, z,w are relatively prime, and z divides uw, then z divides u.

Corollary: Prime divisibility property: If z,w,u are Gaussian integers, z is prime, and z divides uw, then either z divides u, or z divides w.

(Corollary):
Theorem: Unique factorization of (non unit) Gaussian integers. If z is a non zero, non unit, Gaussian integer, then

(i) z is prime or can be written as a product of prime Gaussian integers.
(ii) If $z = (w_1)(....)(w_r) = (u_1)(....)(u_s)$, where all w's and u's are prime, then r = s, and after renumbering the u's we have $w_1 = (\text{unit})u_1, ..., w_r = (\text{unit})u_r$.
Proof: (i) Let z be a Gaussian integer of smallest norm for which the statement is false. Then z is not prime but has a prime factor w. Then z = wq, where $1 < Nm(q) < Nm(z)$, since $Nm(z) > Nm(w) \geq 2$. Then q is either prime or a product of primes $q = (u_1)(....)(u_s)$. But then also $z = wq = w(u_1)(....)(u_s)$, is a product of primes, a contradiction. **QED existence.**
(ii) Let z be a Gaussian integer of smallest norm for which the statement is false. Let $z = (w_1)(w_2)(....)(w_r) = (u_1)(u_2)(....)(u_s)$, be any two prime factorizations of z. Then by the prime divisibility property, w_1 divides both sides, so it divides some u_j, which by renumbering we may call u_1. Then since u_1 is prime we have $u_1 = (\text{unit})w_1$. Then by cancelling u_1, we get $(w_2)(....)(w_r) = (\text{unit})(q_2)(....)(q_s) = y$ = a number with smaller norm than z. If y is a unit we are done. If y is not a unit, the uniqueness statement is true for y, so we have r = s, and after renumbering we have $w_2 = (\text{unit})u_2, ..., w_r = (\text{unit})\, u_r$. Combining that with the fact $w_1 = (\text{unit})u_1$, we are done. **QED.**

Applying modular arithmetic and unique factorization to prove Fermat's theorem
Theorem: if p is an odd prime, the following are equivalent:
(i) p = 4k+1 for some integer k.
(ii) $X^2 + 1 = 0 \pmod p$ has a solution.
(iii) p is not a prime in Z[i].
(iv) p is a sum of two squares in Z.
proof: That (i) implies (ii) was proved in hw #4, solutions, prob. 15.
To see (ii) implies (iii), note that if k in Z solves $X^2 + 1 = 0 \pmod p$, then $k^2+1 = pn$, for some n in Z. Hence p divides k^2+1 in Z. Thus p also divides $k^2+1 = (k+i)(k-i)$ in Z[i], but p does not divide either factor (why?). Hence p is nor prime in Z[i]. To deduce (iv) from (iii) asume $p = (a+bi)(c+di)$ in Z[i], where neither factor on the right is a unit, and take norms of both sides, getting $p^2 = (a^2+b^2)(c^2+d^2)$ in Z, where neither factor on the right is 1. Then by unique prime factorization in Z, we must have both factors (a^2+b^2) and (c^2+b^2) being prime and hence equal to p. (If they were not prime they would factor into primes and we would get too many prime factors on the right. Since they are prime they must equal the only prime factor on the left, namely p.) That (iv) implies (i) we have proved earlier. **QED.**

Remarks on real numbers:
Density of rationals, and the unboundedness of the natural numbers (Archimedean property) are essentially equivalent properties. I.e. Archimedean proerty says that natural numbers get arbitrarily large. Then taking reciprocals, their inverse 1/n get arbitrarily small. Then taking multiples of form k/n, these number are arbitrarily close together, and that is the density property. It is a little tricky to write dwon, but it is easy to understand.

I.e. given real numbers $0 < x < y$, we want to find a rational in between them. Jus make sure the denominator is big enough. I.e. take n so big that $1/n < (y-x)$, i.e. take $n > 1/(y-x)$. Then take m

the smallest positive number with $(m/n) > x$, i.e. with $m > nx$. Then we claim that $x < m/n < y$. If not, then we would have $m/n \geq y$, and since $1/n < y-x$, we get $-1/n > x-y$. Then adding that to our previous inequality we get $m/n - 1/n > y+(x-y) = x$. Thus $(m-1)/n > x$, contradicting choice of m. **QED.**

4000/6000 Day 20
Division, linear combinations, and unique factorization

The three properties we are studying are so important, names have been given to the rings that satisfy them, euclidean domains, principal ideal domains, and unique factorization domains. Since the properties themselves are more important than the names however we will not stop to discuss those, but continue to analyze the relation between these properties. We are interested at present in domains only, i.e. rings where $ab \neq 0$ whenever both a and b are $\neq 0$..

Size functions:
The first main property is the existence of a size function, defined on non zero elements, such that products of non zero elements have larger size, or at least no smaller size, than their factors, and the smallest elements are the units. It is most convenient if this size is a non negative integer, since that enables us to use the well ordering property to produce elements of smallest possible size in any non empty collection.

If the size of an integer n is $|n|$, then $|ab| \geq |b|$, for $a,b, \neq 0$.
In the case of the ordinary integers, we take the size to be simply the absolute value. Then we have for all non zero integers n,m, if n divides m then $|n| \leq |m|$. I.e. if $m = an$, then $|a| \geq 1$, and $|n| \geq 1$, so multiplying by $|n|$ gives $|a||n| \geq |n|$. Since $|m| = |a||n|$, we are done. Note that the only integers of size 1, are the units 1, -1. Consequently, if $m = an$, where a is not a unit then $|n| < |m|$.

Prime integers.
Recall an integer p is "prime" if and only if it is not zero, not a unit, and whenever $p = ab$, for integers a,b, then either a or b must be a unit.

This lets us conclude that every non zero, non unit integer can be factored into prime integers as follows.

Lemma: If an integer n is not zero, and not a unit, then n can be written as a product of (one or more) prime integers.
Proof: If there are integers which cannot be so written, by well ordering principle there is one of smallest absolute value, say n. If n is prime we have a contradiction, so $n = ab$, where neither a nor b is zero or a unit. Then both a and b have strictly smaller absolute values than does n, so both a and b can be written as products of prime integers, say $a = (p_1)(p_2)(.....)(p_r)$, $b = (q_1)(q_2)(.....)(q_s)$. Then $n = ab = (p_1)(p_2)(.....)(p_r)(q_1)(q_2)(.....)(q_s)$ is also a product of primes, a contradiction. Thus no non zero, non unit integer exists which cannot be written as a product of primes. **QED.**

Exactly the same argument works on Gaussian integers if we define the size of a Gaussian integer to be its norm, the square of the absolute value. (We do this simply to make the norm an ordinary integer, so we can more easily use the well ordering principle.)

Size of a Gaussian integer a+bi is the norm, $Nm(a+bi) = a^2+b^2$
Again we have for all non zero Gaussian integers, a+bi, that $Nm(a+bi) = a^2+b^2 \geq 1$, and $Nm(a+bi) = 1$ if and only if a+bi is a unit, i.e. one of the Gaussian integers 1,-1, i, or -i. Since

$Nm(a+bi) = |a+bi|^2$, we again have $Nm(zw) = Nm(z)Nm(w)$, so $Nm(zw) \geq Nm(w)$, and if z is not a unit, then $Nm(zw) > Nm(w)$.

Prime Gaussian integers
A Gaussian integer $z = a+bi$ is prime if and only if whenever $z = uw$, where both u and w are Gaussian integers, then at least one of u or w is a Gaussian unit.

We can repeat the proof above substituting norm for absolute value.
Lemma: If a Gaussian integer z is not zero, and not a Gaussian unit, then z can be written as a product of (one or more) Gaussian primes.
Proof: If there are Gaussian integers which cannot be so written, by well ordering principle there is one of smallest norm, say z. If z is prime we have a contradiction, so $z = uw$, where neither u nor w is zero or a unit. Then both u and w have strictly smaller norms than does z, so both u and w can be written as products of Gaussian primes, say $u = (p_1)(p_2)(.....)(p_r)$, $w = (q_1)(q_2)(.....)(q_s)$. Then $z = uw = (p_1)(p_2)(.....)(p_r)(q_1)(q_2)(.....)(q_s)$ is also a product of Gaussian primes, a contradiction. Thus no non zero, non unit Gaussian integer exists which cannot be written as a product of Gaussian primes. **QED.**

Define: A non constant polynomial f(X) with coefficients in the field F is called "irreducible" if and only if whenever $f(X) = g(X)h(X)$, then either g or h is a non zero constant.

Exercise: For polynomials over a field, such as the rationals Q, let the size of a non zero polynomial be its degree, i.e. the power of X in its leading term. Then prove that the degree is zero if and only if the (non zero) polynomial is a unit, and that every non constant polynomial can be written as a product of (one or more) irreducible polynomials.

Division algorithm
After well ordering, the most basic property that we use in making proofs about the integers is the division algorithm. it says that we can always divide by non zero integers, so as to obtain a remainder smaller than the divisor. More precisely:

Lemma: Given integers a,b, if $a \neq 0$, then there exist (not necessarily unique) integers q,r such that
(i) $b = aq+r$, and
(ii) either $r = 0$, or $|r| < |a|$.
Ugly Proof: Of all non negative integers r of form $b - ax$, let $r = b - aq$ be one with smallest possible absolute value. We can show that there exist <u>some</u> non negative integers of form b-ax, since if $b \geq 0$, then $b - a(0)$ is ≥ 0, and if $b < 0$, then setting $x = 2ab$, gives the linear combination $b - a(2ab) = b(1 - 2a^2) > 0$. Hence a <u>smallest</u> non negative integer $r = b - aq$ of this form exists by the well ordering principle. We claim that $r < |a|$. If not, and $r \geq |a|$, then in case $a > 0$, we have $r = a+s$ where $r > s \geq 0$. Hence we have $b = aq + r = aq + a+s = a(q+1)+s$, where $0 \leq s < r$. Since then $s = b - a(q+1)$ is a non negative linear combination of form $b - ax$, which is smaller than r, this is a contradiction to choice of r.
If $a < 0$, and $r \geq |a| = -a$, then $r = -a + s$, where $0 \leq s < r$. Then $b = aq + r = aq - a + s = a(q-1)+s$, with $0 \leq s < r$, again a contradiction to choice of r. **QED.**

(I dislike this proof. It is too ugly to memorize, so don't. But at least I hope I got it right, so we can use the result. Sometimes they just do not simplify down as much as we would like. The trouble with ugly proofs is they are also harder to be sure of, so please help me out here and report any mistakes you notice. This is one reason for assuming all numbers are positive in the earlier version of this proof, which is not so bad. I.e. you should be able to do this proof when all numbers in it are positive.)

The first use we make of this division algorithm is to find good linear combinations.

Linear combinations.

Given two elements a,b of a ring R, a linear combination of a,b, (with coefficients in R) is an element of form ax+by where x and y are also in R. A linear combination of three elements a,b,c is an element of form ax+by+cz where x,y,z, are in R. A basic fact about the integers is that we never need to use more than one element to form linear combinations. I.e. given any two elements a,b there is always some one other element d such that linear combinations of a,b, are the same as multiples of d. By induction the same is true of 3 elements, 4 elements, etc.

linear combination property:

Using the division algorithm, it follows that for any two integers a,b, not both zero, there is always a linear combination d = ax+by, such that d divides both a and b. Consequently, linear combinations of a and b are the same as multiples of d. In fact it suffices to take for d the smallest positive linear combination of a and b.

Smallest linear combination property:

Given any two integers a,b, not both zero, let d = ax+by be a non zero linear combination of a and b having smallest absolute value. Then d divides both a and b.

Proof: Since a and b are not both zero there are some non zero linear combinations. If $a \neq 0$ for example, take 1(a) + 0(b). Then by well ordering, there is a linear combination d = ax+by of smallest absolute value. To show d divides a, we divide it and show the remainder is zero. To show it is zero we show the remainder is a linear combination of a and b having smaller absolute value than d, hence by choice of d it cannot be non zero. I.e. we can write a = qd + r where $0 \leq |r| < |d|$. Then substituting d = ax+by into a = dq + r gives a = (ax+by)q + r. Solving for r gives r = a - (ax+by)q = a-axq - byq = a(1-xq) + b(-yq), hence r is a linear combination of a and b. Since d is a linear combination of a and b having smallest absolute value among all non zero linear combinations, and r has smaller absolute value than d, r must be zero. I.e. d divides a. Similarly d divides b. **QED.**

We can make this same argument for Gaussian integers, once we have a division algorithm for them. The following is true for Gaussian integers, where we substitute the concept of norm in place of absolute value.

Lemma: Given two Gaussian integers z,w, with $z \neq 0$, there exist Gaussian integers q,r, not necessarily unique, such that w = zq + r, and
Nm(r) < Nm(z).

Proof: Let u = w/z = a+bi, be the complex rational quotient of w by z, where a,b are rational numbers. Then choose integers x,y such that $|x-a| \leq 1/2$, and $|y-b| \leq 1/2$. We claim q = x+iy works. (It had better work, since it is as close to the actual quotient as we can get with a Gaussian integer.) Well, a+bi = (x+iy) + (c+di), where c,d are rational numbers each of length $\leq$

1/2. Hence $Nm(c+di) = c^2+d^2 \leq 1/4 + 1/4 = 1/2$.
Thus $w = z(w/z) = z(x+iy) + z(c+di) = zq + r$, where $q = x+iy$ and $r = z(c+di)$. Note that $r = z(c+di)$ is a Gaussian integer since it equals $w - qz$, the difference of two Gaussian integers. Then it only remains to show that $Nm(r) < Nm(z)$. But $Nm(r) = Nm(z(c+di)) = Nm(z)Nm(c+di) \leq (1/2)Nm(z) < Nm(z)$, since $Nm(z) > 0$. **QED.**

Say why didn't we give this less ugly proof for integers?
Easy proof of division theorem for integers.
I.e. assuming Q is a field and every rational number can be approximated within 1/2 by an integer we get $b = a(b/a) = a(q) + r$, where q is within 1/2 of (b/a) and $r = b - aq$. Then since $(b/a) = q+s$ where $|s| \leq 1/2$, we get $b = a(b/a) = a(q+s) = aq + as$, where $|as| = |a||s| \leq (1/2)|a| < |a|$, since $|a| \neq 0$. **Done.**

(I guess we did not give this proof for integers because at the time we did not yet have the rationals constructed.)

Now we can repeat the arguments for linear combinations of Gaussian integers, substituting norms for absolute values.

Lemma: Given any two Gaussian integers z,w not both zero, there exists a linear combination of them $\partial = zx+wy$, such that ∂ divides both z and w.
Proof: Since there are some non zero linear combinations of z and w, let $\partial = zx+wy$ be one of smallest possible positive norm. We will prove ∂ divides z. By the division theorem we can write $z = \partial q + r$ where q,r, are Gaussian integers and $Nm(r) < Nm(\partial)$. By choice of ∂, thus r cannot be a non zero linear combination of z and w. But substituting $\partial = zx+wy$ into $z = \partial q + r$ gives $z = (zx+wy)q + r$. Solving for r gives $r = z - (zx+wy)q =$
$z-zxq -wyq = z(1-wq) + w(-yq)$. Since thus r is a linear combination of z,w with smaller norm than ∂, r cannot be non zero. So $r = 0$ and ∂ divides z. Similarly ∂ divides w. **QED.**

A special case of the linear combination property is that of two relatively prime integers.
Corollary: If a,b are relatively prime, then 1 can be written as a linear combination of a and b.
Proof: There is a linear combination $d = ax+by$ which divides both a and b. But a and b are relatively prime, so d must be a unit. If $d = 1$ we are done; if $d = -1$, multiply through by -1, getting $1 = -ax -by$.

This also holds for Gaussian integers.
Corollary: If a,b are relatively prime Gaussian integers, then 1 can be written as a linear combination of a and b.
Proof: There is a linear combination $d = ax+by$ which divides both a and b. But a and b are relatively prime, so d must be a unit. If $e = d^{-1}$, then $ed = 1$, so $1 = a(ex)+b(ey)$. **QED.**

The last key property that follows from the division algorithm is the (relatively) prime divisibility property.

Lemma: If a,b,c are integers, where a,b are relatively prime, and a divides bc, then a divides c.

Proof: Since a,b, are relatively prime, write 1 = ax+by. Then multiplying by c gives c = acx + bcy. Then a divides both terms on the right side, hence a divides the left side, i.e. a divides c. **QED.**

Not surprisingly,
Lemma: If a,b,c are Gaussian integers, where a,b are relatively prime, and a divides bc, then a divides c.
Proof: Since a,b, are relatively prime, write 1 = ax+by. Then multiplying by c gives c = acx + bcy. Then a divides both terms on the right side, hence a divides the left side, i.e. a divides c. **QED.**

A key special case is where a is prime.

Lemma: If a,b,c are either integers or Gaussian integers, a is prime, and a divides bc, then a divides either b or c (or both).
proof: If a divides b we are done. If not then a and b are relatively prime. Then since a divides bc, a divides c. **QED.**

Finally, prime divisibility allows us to prove, in both the case of integers and Gaussian integers, that the factorization of non zero, non units into primes, is unique, except for unit multiples, and reordering the factors.

Theorem: Unique factorization of (non zero, non unit) integers. If n is any integer except 0, 1, or -1, and if $n = (p_1)(....)(p_r) = (q_1)(....)(q_s)$, where all p's and q's are prime, then r = s, and after possibly renumbering the q's we have $p_1 = \pm q_1$, $p_2 = \pm q_2$, ..., $p_r = \pm q_r$.
Proof: If the statement is false for some integers, there is an integer n of smallest absolute value for which it is false. Then n has two factorizations which are not equivalent in the sense of the theorem.
Let $n = (p_1)(p_2)(....)(p_r) = (q_1)(q_2)(....)(q_s)$, be any two prime factorizations of n. We will show that these factorizations must be equivalent in the sense of the theorem, so that the theorem is in fact not false for n. This contradiction will prove there is no integer for which the theorem is false.

Since $n = (p_1)(p_2)(....)(p_r) = (q_1)(q_2)(....)(q_s)$, then p_1 divides both sides, so by the prime divisibility property p_1 divides some q_j, which by renumbering we may call q_1. Then since q_1 is prime we have $q_1 = \pm p_1$. Then by canceling p_1 from both sides (we may cancel any non zero integer since Z is a domain), we get $(p_2)(....)(p_r) = (\pm q_2)(....)(q_s) = m$, where $m = n/p_1$ has smaller absolute value than n, since $|m| = |n|/|p_1|$ and p_1 prime implies $|p_1| > 1$.

If $m = \pm 1$ we are done, since then there are no more primes p or q on either side, hence r = s = 1. If $|m| \geq 2$, the uniqueness statement in the theorem is true for m, so there are the same number of primes on both sides, i.e. r-1 = s-1, (hence r = s), and after renumbering, we have $p_2 = \pm q_2$, ... ,$p_r = \pm q_r$. Combining that with the fact that $q_1 = \pm p_1$, we are done. **QED.**

As usual now, we get a similar theorem for Gaussian integers.
Theorem: Unique factorization of (non zero, non unit) Gaussian integers. If z is a non zero, non unit, Gaussian integer, and if $z = (w_1)(....)(w_r) = (u_1)(....)(u_s)$, where all w's and u's are

Gaussian primes, then r = s, and after possibly renumbering the u's we have $w_1 = (\text{unit})u_1$, $w_2 = (\text{unit})u_2$, ..., $w_r = (\text{unit})u_r$.

Proof: If the statement is false for some Gaussian integers, let z be a Gaussian integer of smallest norm for which the statement is false. Then z must have two factorizations which are not equivalent in the sense of the theorem. Let $z = (w_1)(w_2)(....)(w_r) = (u_1)(u_2)(....)(u_s)$, be any two prime factorizations of z. We will show that these factorizations are equivalent in the sense of the theorem, so that the theorem is in fact not false for z. This contradiction will prove there is no z for which the theorem is false.

Since $z = (w_1)(w_2)(....)(w_r) = (u_1)(u_2)(....)(u_s)$, then w_1 divides both sides, so by the prime divisibility property w_1 divides some u_j, which by renumbering we may call u_1. Then since u_1 is prime we have $u_1 = (\text{unit})w_1$. Then by cancelling w_1 from both sides, we get $(w_2)(....)(w_r) = (\text{unit})(u_2)(....)(u_s) = y$ = a Gaussian integer with smaller norm than z, since $Nm(y) = Nm(z)/Nm(w_1)$, and w_1 prime implies $Nm(w_1) > 1$.

If y is a unit we are done, since then there are no more prime w's or u's on either side, so r = s = 1. If y is not a unit, then since y has smaller norm than z, the uniqueness statement in the theorem is true for y, so the number of factors on both sides is the same, i.e. r-1 = s-1, (hence r = s), and after renumbering we have $w_2 = (\text{unit})u_2$,....,$w_r = (\text{unit})\ u_r$. Combining that with the fact $w_1 = (\text{unit})u_1$, we are done. **QED.**