Math 17
Winter 2015
Wednesday, January 14

Lemma:[1] If p is an odd prime, then there is a number b such that $0 < b < p$ and pb can be written as the sum of four squares.

Proof: You proved this for homework. (Actually, you proved that pb can be written as the sum of three squares, but by adding 0^2 you get the sum of four squares.)

Theorem (the Lagrange four square theorem)**:** Every natural number can be written as the sum of four squares.

Proof: You showed in homework that the theorem follows from the following proposition.

Proposition: Every odd prime can be written as the sum of four squares.

Proof (of the proposition)**:** Let p be an odd prime, and choose b as in the lemma, with $0 < b < p$ and

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = pb.$$

We can choose the x_i to all be nonnegative.

If $b = 1$ we are done. So suppose that $b > 1$. We will show that there is some c with $0 < c < b$ such that pc can be written as the sum of four squares.

Once we show this, we are done. (If $c = 1$, we have the result we want. If $c > 1$, we can apply the same argument to find d with $0 < d < c$ such that pd can be written as the sum of four squares. Continuing in this way, eventually we work our way down to getting $p \cdot 1$ (that is, p) written as the sum of four squares.)

We will make use of a few key facts, mostly about congruence relations, stated on the next page.

[1] A theorem or proposition is a significant result, a result we are interested in for its own sake. A lemma is a result that is proved as one step in proving a theorem or proposition. A corollary is a result that follows from a theorem or proposition; it may be called a corollary of that theorem. Lemmas generally don't have corollaries.

1. We can think of $x \equiv y \mod b$ in several ways:

   (a) x and y have the same remainder when divided by b;

   (b) $b|(x-y)$;

   (c) y is obtained from x by adding a (possibly negative) multiple of b.

2. If

   $$x_1 \equiv y_1 \mod b \quad \& \quad x_2 \equiv y_2 \mod b \quad \& \quad \cdots \quad \& \quad x_n \equiv y_n \mod b$$

   and $P(x_1, x_2 \ldots, x_n)$ is any polynomial with integer coefficients, then

   $$P(x_1, x_2 \ldots, x_n) \equiv P(y_1, y_2 \ldots, y_n) \mod b.$$

   (Since $x_i \equiv y_i \mod b$, we can write $y_i = x_i + k_i b$. Therefore, replacing x_i with y_i changes the value of the polynomial by adding some multiple of b; that is, it changes it to something congruent modulo b to the original value,)

3. If $b > 1$, every natural number x is congruent modulo b to some number in the interval $\left(-\frac{b}{2}, \frac{b}{2}\right]$. This is the number of smallest absolute value among all the numbers congruent to x mod b. (We discussed this in class.)

4. The Euler four square identity:

   $$(x_1^2 + x_2^2 + x_3^2 + x_4^2) \cdot (y_1^2 + y_2^2 + y_3^2 + y_4^2) =$$

   $$(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 +$$
   $$(x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 +$$
   $$(x_1y_3 - x_3y_1 - x_2y_4 + x_4y_2)^2 +$$
   $$(x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2.$$

5. If p is prime and $1 < b < p$, then pb is not a multiple of b^2. Therefore, if

   $$x_1^2 + x_2^2 + x_3^2 + x_4^2 = pb,$$

   then

   (a) It is not possible for every x_i to be a multiple of b.

   (b) It is not possible for every x_i to have the form $kb \pm \frac{b}{2}$.

   (You can check that in either case, $x_1^2 + x_2^2 + x_3^2 + x_4^2$ would be a multiple of b^2.)

Now that we've stated these useful facts, we continue with the proof. Remember, we have $1 < b < p$ and

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = pb,$$

and we must find c with $0 < c < b$ such that pc can be written as the sum of four squares.

Using fact (3) above, choose y_1, y_2, y_3, y_4 such that

$$-\frac{b}{2} < y_i \leq \frac{b}{2} \quad \& \quad y_i \equiv x_i \mod b.$$

Since

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 \equiv 0 \mod b,$$

by fact (2) above,

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 \equiv 0 \mod b.$$

That is, for some c, we have

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 = cb.$$

We will show this is the c we are looking for.

It is clear that $c \geq 0$. Because for every i we have $|y_i| \leq \frac{b}{2}$, we also have

$$cb = y_1^2 + y_2^2 + y_3^2 + y_4^2 \leq \frac{b^2}{4} + \frac{b^2}{4} + \frac{b^2}{4} + \frac{b^2}{4} = b^2,$$

so $c \leq b$.

If $c = 0$, then $y_1^2 + y_2^2 + y_3^2 + y_4^2 = 0$, so $y_i = 0$ for every i. Since $y_i \equiv x_i \mod b$, this means that every x_i is a multiple of b. But this contradicts fact (5a). Therefore, $0 < c$.

If $c = b$, then $y_1^2 + y_2^2 + y_3^2 + y_4^2$ has the maximum possible value. Therefore we must have $|y_i| = \frac{b}{2}$ for every i. Since $y_i \equiv x_i \mod b$, this means that every x_i is of the form $kb \pm \frac{b}{2}$. But this violates fact (5b). Therefore, $c < b$.

Now we have $0 < c < b$. It remains only to show that pc can be written as the sum of four squares.

To do this, we will use the Euler four square identity:

$$\begin{gathered} pc(b^2) = (pb)(cb) = \\ (x_1^2 + x_2^2 + x_3^2 + x_4^2) \cdot (y_1^2 + y_2^2 + y_3^2 + y_4^2) = \\ (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + \\ (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 + \\ (x_1y_3 - x_3y_1 - x_2y_4 + x_4y_2)^2 + \\ (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2 \end{gathered}$$

Consider the terms on the righthand side of this equation, and apply fact (3) above, and the fact that $x_i \equiv y_i \mod b$. For the first term, modulo b we have

$$(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) \equiv (x_1x_1 + x_2x_2 + x_3x_3 + x_4x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2 = bp \equiv 0,$$

so $(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)$ is a multiple of b, and we can write

$$(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) = z_1b.$$

For the second term, modulo b we have

$$(x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3) \equiv (x_1x_2 - x_2x_1 + x_3x_4 - x_4x_3) = 0,$$

so $(x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)$ is a multiple of b, and we can write

$$(x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3) = z_2b.$$

Exactly the same reasoning applies to the third and fourth terms:

$$(x_1y_3 - x_3y_1 - x_2y_4 + x_4y_2) = z_3b;$$

$$(x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2) = z_4b.$$

Plugging back in, we now have

$$pc(b^2) = (z_1b)^2 + (z_2b)^2 + (z_3b)^2 + (z_4b)^2.$$

Dividing by b^2 gives

$$pc = z_1^2 + z_2^2 + z_3^2 + z_4^2.$$

This is what we needed to show.

Definition: Let n and k be any natural numbers with $k \leq n$. We define the *binomial coefficient*

$$\binom{n}{k}$$

(read "n choose k") to be the number of k-element subsets of an n-element set. The terminology "n choose k" reflects that we are counting how many ways to choose k objects from a set of n objects (when order does not matter).

If you think about what happens when you expand out

$$(x+y)^n = (x+y)\,(x+y)\,\cdots\,(x+y),$$

you can convince yourself that the coefficient of the term $x^k y^{n-k}$ is $\binom{n}{k}$. The terminology "binomial coefficients" reflects that these are the coefficients we get when we raise a binomial (a sum of two monomials) to a power.

Lemma: If $0 < k < n+1$, then

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

(This is why the binomial coefficients are found in Pascal's triangle.)

Proof: The k-element subsets of an $(n+1)$-element set $\{a_1, a_2, \ldots, a_n, a_{n+1}\}$ can be classified into two classes.

1. The subsets that do not contain a_{n+1} are k-element subsets of $\{a_1, a_2, \ldots, a_n\}$. There are $\binom{n}{k}$ many of these.

2. The subsets that contain a_{n+1} are $(k-1)$-element subsets of $\{a_1, a_2, \ldots, a_n\}$ with a_{n+1} added in. There are $\binom{n}{k-1}$ many of these.

Therefore, the total number of k-element subsets is $\binom{n}{k} + \binom{n}{k-1}$.

Theorem: For all natural numbers n and $k \leq n$, we have

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Proof: One way to prove this is by combinatorial reasoning. We will prove it by induction.

Digresssion (proof by induction)**:** To prove that some property $\varphi(n)$ holds for every natural number n by induction, you need to do two things.

1. Prove $\varphi(0)$. This is called the base case.

2. Assume $\varphi(n)$ (this is called the inductive hypothesis) and prove $\varphi(n+1)$.

Why this works: Let

$$X = \{n \in \mathbb{N} \mid \varphi(n)\}.$$

We want to show that $X = \mathbb{N}$. We can do this if we show two things.

1. $0 \in X$.

2. X is closed under adding 1, that is, if we add 1 to some number in X we get another number in X.

To show $0 \in X$, we prove $\varphi(0)$. This is the base case of proof by induction.

To show X is closed under adding 1, we assume that n is some number in X, and show that $n+1$ is also a number in X. That is, we assume $\varphi(n)$ and show $\varphi(n+1)$. This is the inductive step.

An important note here: The inductive step can be confusing when you first start doing proofs by induction. You are trying to prove $\varphi(n)$, so isn't it circular to assume $\varphi(n)$? It is not, and we can see this if we pay attention to what we mean. You are trying to prove that $\varphi(n)$ is true *for all* n. You are assuming that n is *some particular number* for which φ is true. There is no problem here; there is at least one such number, because you just proved (in the base case) that φ is true of 0. Then you prove, from this assumption, that $n+1$ is another number for which φ is true.

If there is more than one variable floating around, we may say this is proof by induction "on n."

In our case, the property $\varphi(n)$ we want to prove, by induction on n, is

$$(\forall k \leq n) \left(\binom{n}{k} = \frac{n!}{k!(n-k)!} \right).$$

Proof (continued)**:** We prove

$$(\forall k \leq n)\left(\binom{n}{k} = \frac{n!}{k!(n-k)!}\right)$$

by induction on n.

Before starting, we note that an n-element set has only one 0-element subset (the empty set), and only one n-element subset (the entire set), so $\binom{n}{0} = \binom{n}{n} = 1$. We also note that, by convention, we define $0! = 1$.

Base Case: For $n = 0$, the only natural number $k \leq n$ is $k = 0$, so we must show that

$$\binom{0}{0} = \frac{0!}{0!\,0!}.$$

Since each side of this equation equals 1, this is true.

Inductive Step: Assume, as inductive hypothesis, that

$$(\forall k \leq n)\left(\binom{n}{k} = \frac{n!}{k!(n-k)!}\right).$$

We must show that

$$(\forall k \leq n+1)\left(\binom{n+1}{k} = \frac{(n+1)!}{k!((n+1)-k)!}\right).$$

For $k = 0$ we must show that

$$\left(\binom{n+1}{0} = \frac{(n+1)!}{0!((n+1)-0)!} = \frac{(n+1)!}{1 \cdot (n+1)!}.\right).$$

Since each side of this equation equals 1, this is true.

For $k = n+1$, we must show that

$$\left(\binom{n+1}{n+1} = \frac{(n+1)!}{(n+1)!((n+1)-(n+1)!} = \frac{(n+1)!}{(n+1)!0!} = \frac{(n+1)!}{(n+1)! \cdot 1}\right).$$

Since each side of this equation equals 1, this is true.

For $0 < k < n+1$, we must show that

$$\left(\binom{n+1}{k} = \frac{(n+1)!}{k!((n+1)-k)!}\right).$$

To do this we use the lemma and the inductive hypothesis. By the lemma, we have

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

By the inductive hypothesis, we can replace $\binom{n}{k-1}$ with $\dfrac{n!}{(k-1)! \cdot (n-(k-1))!}$ and $\binom{n}{k}$ with $\dfrac{n!}{k! \cdot (n-k)!}$, to get

$$\binom{n+1}{k} = \frac{n!}{(k-1)! \cdot (n-(k-1))!} + \frac{n!}{k! \cdot (n-k)!}.$$

Now we do some algebraic manipulation to get what we want.

$$\frac{n!}{(k-1)! \cdot (n-(k-1))!} + \frac{n!}{k! \cdot (n-k)!} = \frac{k \cdot n!}{(k)! \cdot (n-(k-1))!} + \frac{(n-(k-1))n!}{k! \cdot (n-(k-1))!} =$$

$$\frac{k \cdot n! + (n-(k-1))n!}{k! \cdot (n-(k-1))!} = \frac{n!(n+1)}{k!((n+1)-k)!} = \frac{(n+1)!}{k!((n+1)-k)!}.$$