SCHUR'S THEOREM AND FERMAT'S LAST THEOREM MODULO A PRIME

SEBASTIEN VASEY

Recall the famous "last theorem" of Fermat: for any natural number $k \geq 3$, the equation $x^k + y^k = z^k$ does not have any solution in the natural numbers (by contrast, when $k = 2$ there are many solutions – for example $3^2 + 4^2 = 5^2$). Fermat wrote in the margin of his notebook that he knew a proof but that the margin was too small... It was only more than 350 years later, in 1994, that Andrew Wiles found a very long and complex proof that Fermat couldn't possibly have known.

The goal of these notes is to prove that Fermat's last theorem is not true if we do arithmetic modulo a prime. This is a simple application of Ramsey theory:

Theorem 1 (Schur)**.** For any natural number k, there exists a natural number n such that for any prime number p bigger than n, there exists natural numbers x, y, z so that $x^k + y^k = z^k$ modulo p.

For example, if $k = 4$, then modulo $p = 7$, we have $1^4 + 1^4 = 2^4$. In general, for any fixed exponent k, as long as the prime p is "high-enough" (bigger than a fixed threshold n, depending on k), then we can find natural numbers x, y, z so that $x^k + y^k = z^k$ modulo p.

In order to prove the theorem, we first recall Ramsey's theorem:

Theorem 2 (Ramsey's theorem, graph version)**.** For any natural number k, there exists a natural number n such that any graph with at least n vertices contains a k-clique or a k-independent set.

Instead of thinking in terms of k-cliques or k-independent sets, that is in terms of edges being or not being present, we can think instead of coloring the edges of complete graph on n vertices either blue or red. Ramsey's theorem then tells us we can find a monochromatic k-clique. This statement has the advantage of generalizing to a higher number of colors.

Theorem 3 (Ramsey's theorem, colorful version)**.** For any natural number k, and any number c of colors, there exists a natural number n such that for any coloring of the edges of a complete graph on at least n vertices with c colors, there is a monochromatic k-clique.

Proof. We proceed by induction on the number c of colors. When $c = 2$, this is just the usual Ramsey theorem. Assume now that $c \geq 2$ and the result is true for c. Let $r_c(k)$ denote the least natural number n satisfying the conclusion of the theorem. We claim that $r_{c+1}(k) \leq n := r_2(r_c(k))$. Indeed, let $f : E(K_n) \to \{1, \ldots, c+1\}$ be a coloring. Define an auxiliary coloring $g : E(K_n) \to \{1, 2\}$ as follows: $g(e) = 1$ if $f(e) = 1$, and $g(e) = 2$ otherwise. That is g only distinguishes the color 1 from the other c colors. By definition of n, there is a monochromatic clique $K \subseteq V(K_n)$ of size $r_c(k)$. If the color of K is 1, then we are done already: any k-element subset

of K gives the desired monochromatic clique for f. If the color of K is 2, then we know that f restricted to the edges of K takes value in $\{2, 3, \ldots, c+1\}$, which is only c colors. Since K has size $r_c(k)$, we can then find a monochromatic k-clique in K. □

The next application is interesting of its own, and will be the key result used in the proof of Theorem 1.

Theorem 4 (Schur's theorem)**.** For any number c of colors there exists a natural number n such that for any coloring of the numbers $1, 2, \ldots, n$ with c colors, we can find natural numbers $x, y, z \leq n$ of the same color so that $x + y = z$.

For example, let us say $c = 2$ and we color the numbers $1, 2, 3, 4$ as follows: 1 is blue, 2 is red (otherwise we can take $x = y = 1$, $z = 2$), 3 is blue, 4 is blue. Then we can take $x = 1$, $y = 3$, $z = 4$. This suggest (but is not a proof) that perhaps taking $n = 4$ would work in case $c = 2$. It may be fun for you to try to figure out whether $n = 4$ works in general or how big n has to be.

Proof of Theorem 4. Let n be as given by Ramsey's theorem for $k = 3$ and c colors. Let $f : \{1, \ldots, n\} \to \{1, \ldots, c\}$ be a coloring. Define a coloring $g : E(K_n) \to \{1, \ldots, c\}$ by $g(\{a, b\}) = f(|b - a|)$ (we think of the vertices of K_n as being $1, 2, \ldots, n$). By the assumption on n, we can find a g-monochromatic triangle with vertices $a < b < c$. Let $x := b - a$, $y := c - b$ and $z := c - a$. Then it is clear that $x + y = z$ and by definition of g, x, y, z all have the same f-color. □

Proof of Theorem 1. We fix k, and let n be as given by Schur's theorem for k colors. Let p be a prime bigger than n. We will use the following facts from algebra:

(1) The numbers $\{1, \ldots, p-1\}$ form a group under multiplication modulo p. This means in particular that for every natural number $a < p$, there is $b < p$ so that $ab = 1$ modulo p. We call b an *inverse of* a modulo p. You have already proven this fact in assignment 3.5c: you showed that p divides $a^p - a$, which is just $a(a^{p-1} - 1)$. Since by assumption p does not divide a, it must divide $a^{p-1} - 1$, and hence $a^{p-1} = 1$ modulo p. Thus one can take $b = a^{p-2}$.

(2) This group is cyclic: there is a natural number $g < p$ (called a generator) such that $\{1, 2, \ldots, p-1\} = \{g^1, g^2, \ldots, g^{p-1}\}$. In particular every $b < p$ is a power of g (of course modulo p). This is a little bit harder to prove, and require some algebra (at the level of Math-122). For each $a \in \{1, 2, \ldots, p-1\}$, the *order* of a is the least natural number d such that $a^d = 1$. It is a basic fact of group theory that the order of an element must divide the size of the group, here $p - 1$. Now for a fixed d dividing $p - 1$, how many elements of order exactly d are there? Well take any $a < p$ of order d. We have that any element of order d must be among $a^1, a^2, \ldots, a^{d-1}$, since these are distinct solutions to the polynomial $x^d - 1 = 0$, and this has at most d roots (here we are working in the field $\{0, 1, \ldots, p-1\}$ with addition and multiplication modulo p, but the fact that a polynomial of degree d has at most d roots is true for any field). Suppose now that a^i has order d. Then i and d must be coprime (if not, $i = \ell u$, $d = \ell u'$ for some common factor $\ell \geq 2$, and we then have that $(a^i)^{u'} = 1$, hence a^i has order at most u'). Thus there are at most $\phi(d)$-many elements of order d ($\phi(d)$ counts how many numbers below d are coprime to d).

Let us count the size of $\{1, \ldots, p-1\}$ in two ways. On the one hand, it is equal to $p-1$. On the other hand we can partition it according to the order of its elements: if O_d is the set of elements of order d, then $p-1 = \sum_{d|p-1} |O_d|$. We have just argued that $|O_d| \leq \phi(d)$ for all d, but you showed in assignment 3 that $\sum_{d|p-1} \phi(d) = p-1$. It follows that $|O_d| = \phi(d)$: for each possible order d, there must be *exactly* $\phi(d)$-many elements of order d. In particular, there is an element of order $p-1$, which will be the desired generator.

With these facts stated, we fix a generator g for $\{1, \ldots, p-1\}$. Dividing exponents by k, we get that any $a < p$ can be written as $a = g^{ki+j}$, where $i = i_a, j = j_a < k$ depend on a. Color each $a \in \{1, 2, \ldots, p-1\}$ according to the value of j_a. By Schur's theorem, we can find a, b, c distinct in $\{1, 2, \ldots, p-1\}$ such that $j := j_a = j_b = j_c$ and $a + b = c$. Expanding the definitions, $g^{ki_a+j} + g^{ki_b+j} = g^{ki_c+j}$. Multiplying by the inverse of g^j modulo p, we get that $g^{ki_a} + g^{ki_b} = g^{ki_c}$. Thus we can let $x := g^{i_a}$, $y := g^{i_b}$, and $z := g^{i_c}$ to get the desired solution to $x^k + y^k = z^k$ (modulo p). $\square$