Math Club Notes: 2007 August 22

1 Cauchy's theorem and the cubes mod p

Previously, in math club:

We have considered the cubes mod p, observing that, for example, with $p = 13$ every cube has three cube roots:

k	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
k^3 mod 13	5	5	1	-1	5	-1	0	1	-5	1	-1	-5	-5

On the other hand, with $p = 17$ every cube has exactly one cube root:

k	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
k^3 mod 17	-2	-3	5	-6	4	7	-8	-1	0	1	8	-7	-4	6	-5	3	2

We have shown (see our notes for April 4) that these are the only two possibilities, that is, in any field either every cube has one cube root or every cube has three cube roots. We also noted by a simple counting argument that if cubes in $\mathbb{Z}_p$ have three cube roots then p is of the form $3k + 1$. We conjectured the converse as well, that is, that if p is of the form $3k + 1$ then cubes in $\mathbb{Z}_p$ have three cube roots. Today we proved that conjecture using Cauchy's theorem.[1]

Cauchy's theorem says this:

> If the order of a group G is a multiple of a prime q, then the number of solutions to $x^q = 1$ in G (where 1 is the identity of G) is also a multiple of q.

Note that the equation $x^q = 1$ has at least one solution, namely $x = 1$, so an immediate corollary is that the number of solutions is at least q. Our conjecture is the special case $q = 3$ and $G = U(p)$ (that is, $\{1, \dots, p-1\}$ under multiplication mod p).

Now to prove Cauchy's theorem.[2]

Consider sequences $(x_1, x_2, \dots, x_q)$ of q elements from G, having the property that $x_1 x_2 \cdots x_q = 1$. To count such sequences, note that we can choose the first $q-1$ elements of the sequence arbitrarily, then set $x_q = (x_1 x_2 \cdots x_{q-1})^{-1}$. Thus there are $|G|^{q-1}$ such sequences. This is a multiple of q, since $|G|$ is.

[1] This theorem, and its relevance to this problem, was pointed out to me by Dr. Weiss. Note, incidentally, that Cauchy's theorem can be restated thus: if the prime q divides $|G|$, then G has a subgroup of order q. There are several theorems about the existence of subgroups of certain orders; the big ones were, I think, proven by Sylow.

[2] The following proof is from [1]. My presentation is a lot more verbose than McKay's.

Now, note that if we move the last element of such a sequence to the beginning, we obtain a new sequence with the same property, since

$$\begin{aligned} x_q x_1 x_2 \cdots x_{q-1} &= x_q x_1 x_2 \cdots x_{q-1} (x_q x_q^{-1}) \\ &= x_q (x_1 x_2 \cdots x_{q-1} x_q) x_q^{-1} \\ &= x_q x_q^{-1} \\ &= 1 \, . \end{aligned}$$

(Note that we're not assuming G to be Abelian, so we can't just rearrange the elements in the product.) Thus any cyclic permutation of such a sequence is another such sequence.

Now, fix some such sequence $\bar{x} = (x_1, \dots, x_q)$. Let σ denote the permutation that moves the last element to the beginning. Applying σ repeatedly to our sequence yields a sequence of cyclic permutations of our sequence,

$$\sigma\bar{x}, \sigma^2\bar{x}, \sigma^3\bar{x}, \dots \, .$$

Now, suppose that $\sigma^n\bar{x} = \bar{x}$. (This is certainly true for $n = q$, and might be true for other n.) From number theory we know that, for suitable integers s and t,

$$\gcd(n, q) = ns + qt \, .$$

Thus

$$\sigma^{\gcd(n,q)}\bar{x} = \sigma^{ns+qt}\bar{x} = \sigma^{ns}\sigma^{qt}\bar{x} = (\sigma^n)^s(\sigma^q)^t\bar{x} = \bar{x} \, .$$

So $\sigma^{\gcd(n,q)}$ also fixes $\bar{x}$. Thus if n is the least n such that $\sigma^n\bar{x} = \bar{x}$, then n is a divisor of q, that is, either $n = 1$ or $n = q$. In the case $n = 1$, we have $\sigma\bar{x} = \bar{x}$, and so all the elements of $\bar{x}$ are equal. In the case $n = q$, we have that the cyclic permutations $\bar{x}, \sigma\bar{x}, \sigma^2\bar{x}, \dots, \sigma^{q-1}\bar{x}$ are all distinct.

So, if we consider sequences $(x_1, \dots, x_q)$ to be equivalent if one can be obtained from the other by a cyclic permutation, then the set of the $|G|^{q-1}$ sequences under discussion is partitioned into equivalence classes of two types: some classes have just one element, a sequence with all its elements equal — say there are a classes of this type; the other classes have q elements, being distinct cyclic permutations of some sequence — say there are b classes of this type. Then we have

$$|G|^{q-1} = a + bq \, ,$$

whence a is a multiple of q. And that's what we wanted to show.

(I have seen one or two other proofs like this — that is, combinatorial proofs of group-theoretic results. I quite like them. Maybe I'll bring more to future meetings.)

2 Infinitely many congruence classes

Previously, in math club:

Definition 1 Let A be a subring of $\mathbb{R}$, and let p be a polynomial with coefficients in $\mathbb{R}$. We say that p *fixes* A if $p(t) \in A$ for all $t \in A$.

Definition 2 Let A be a subring of $\mathbb{R}$. We say that A *pins coefficients* if every polynomial which has real coefficients and fixes A must have coefficients which are all in A.

On January 24 we observed that $\mathbb{Z}$ doesn't pin coefficients; for example, $\frac{1}{2}t(t+1)$ fixes $\mathbb{Z}$ but has coefficients not in $\mathbb{Z}$.

On February 28 we observed that any ring is fixed by the identity polynomial $p(t) = t$, and so any ring that pins coefficients must contain 1; by closure under addition, any ring that pins coefficients must contain all of $\mathbb{Z}$.

On March 7 we proved that every subfield of $\mathbb{R}$ pins coefficients. (So the remaining question is whether there exist any rings which are not fields but do pin coefficients.)

On May 30 we generalized the previous result on $\mathbb{Z}$ to show that if a ring pins coefficients, then for every uninvertible element m in that ring, there are infinitely many congruence classes modulo m in that ring.

This last result seemed at the time like an extremely strong constraint on a ring, so strong that I doubted there were any such rings (other than fields). Following up a suggestion by Dr. Weiss, however, I quickly found one: $\mathbb{Q}[e]$. This ring consists of numbers that can be written in the form

$$a_0 + a_1 e + a_2 e^2 + \cdots + a_n e^n \tag{1}$$

for some rational numbers a_i and some nonnegative integer n. That is, this ring consists of rational linear combinations of powers of e.

The important thing about e for our purposes is that it is transcendental, that is, it is not a zero of any polynomial with rational coefficients (except the zero polynomial). One consequence is that every number in $\mathbb{Q}[e]$ has exactly one representation in the form (1). Indeed, suppose that

$$a_0 + a_1 e + a_2 e^2 + \cdots + a_n e^n = b_0 + b_1 e + b_2 e^2 + \cdots + b_n e^n \,.$$

(We can assume these two representations have the same length, since if one is shorter we can just add some zeroes at the end.) Then

$$(a_0 - b_0) + (a_1 - b_1)e + (a_2 - b_2)e^2 + \cdots + (a_n - b_n)e^n = 0 \,,$$

showing that e is a zero of the polynomial on the left. Therefore that polynomial is the zero polynomial, whence $a_i = b_i$ for all i.

Since such representations are unique, we can define the *degree* $\deg x$ of a number $x \in \mathbb{Q}[e]$ to be the highest power of e that occurs in its representation with a nonzero coefficient. (We define $\deg 0 = -\infty$.) This notion of degree is totally analogous to the notion of the degree of a polynomial (which is not much of a surprise). In particular, we have the log-like rule[3]

$$\deg(xy) = \deg x + \deg y \,.$$

(The definition of $\deg 0$ was chosen to make this rule hold even when $x = 0$ or $y = 0$.) In particular,

$$\deg x \geq 1 \text{ and } \deg y \geq 1 \implies \deg(xy) \geq 2 \,.$$

Thus, by contraposition,

$$\deg(xy) \leq 1 \implies \deg x \leq 0 \text{ or } \deg y \leq 0 \,,$$

and in particular, since it is rational numbers that have degree ≤ 0,

$$xy \in \mathbb{Q} \implies x \in \mathbb{Q} \text{ or } y \in \mathbb{Q} \,.$$

Moreover, if $xy \neq 0$, then $x \neq 0$, and so $x \in \mathbb{Q}$ and $xy \in \mathbb{Q}$ together imply $y = xy/x \in \mathbb{Q}$; and likewise for y. Thus

$$xy \in \mathbb{Q} \text{ and } xy \neq 0 \implies x \in \mathbb{Q} \text{ and } y \in \mathbb{Q} \,.$$

This semi-obvious fact has several useful consequences. For one, it lets us characterize the invertible elements of this ring. Indeed, suppose $xy = 1$. Since $1 \in \mathbb{Q}$ and $1 \neq 0$, both $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$. Thus the only invertible elements in $\mathbb{Q}[e]$ are the rationals.

For another, suppose that $x, y \in \mathbb{Q}$ and $x \neq y$. If $x \equiv y \pmod{m}$, then for some s, $ms = x - y$, which is rational and nonzero, whence $m \in \mathbb{Q}$. By contraposition, if m is uninvertible (hence not rational), then distinct rational x and y are incongruent modulo m; thus there are at least as many congruence classes modulo m as there are rational numbers.

So this ring $\mathbb{Q}[e]$ has the desired property: every uninvertible element gives rise to infinitely many congruence classes.

Consequently, our previous construction fails in this ring; we cannot construct by the methods we already know a polynomial which fixes this ring but has coefficients not in it. In other words, we don't know whether this ring pins coefficients or not. Determining that will require a new technique.

[3] It is somewhat instructive to consider why this rule can't be made to work in rings $\mathbb{Q}[a]$ where a is algebraic; consider $\mathbb{Q}[\sqrt{2}]$, for example.

3 The determinant of a Vandermonde matrix

One of our outstanding problems (following up on an argument in our notes of March 7) is to show that

$$\begin{vmatrix} 1 & r_0 & r_0^2 & \dots & r_0^n \\ 1 & r_1 & r_1^2 & \dots & r_1^n \\ 1 & r_2 & r_2^2 & \dots & r_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & r_n & r_n^2 & \dots & r_n^n \end{vmatrix} = \prod_{0\le i<j\le n} (r_j - r_i)\,. \tag{2}$$

During the meeting we came up with the following proof, the main idea of which is to think of the two sides of this equality as polynomials in r_0.

Case 1: For some i, $r_i = 0$.

Without loss of generality (why?), $r_0 = 0$. Thus the determinant in question is

$$\begin{aligned}
&\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & r_1 & r_1^2 & \dots & r_1^n \\ 1 & r_2 & r_2^2 & \dots & r_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & r_n & r_n^2 & \dots & r_n^n \end{vmatrix} \\
&= \begin{vmatrix} r_1 & r_1^2 & \dots & r_1^n \\ r_2 & r_2^2 & \dots & r_2^n \\ \vdots & \vdots & \ddots & \vdots \\ r_n & r_n^2 & \dots & r_n^n \end{vmatrix} && \text{(expansion along first row)} \\
&= r_1 r_2 \cdots r_n \begin{vmatrix} 1 & r_1 & \dots & r_1^{n-1} \\ 1 & r_2 & \dots & r_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & r_n & \dots & r_n^{n-1} \end{vmatrix} && \text{(multilinearity of determinant)} \\
&= r_1 r_2 \cdots r_n \prod_{1\le i<j\le n} (r_j - r_i) && \text{(by induction)} \\
&= \prod_{0\le i<j\le n} (r_j - r_i) && (r_0 = 0)
\end{aligned}$$

Case 2: Some two of the r_i are equal.

Suppose $r_i = r_j$ and $i < j$. Then the ith and jth rows of the matrix are equal, so its determinant is zero. On the other hand, the product on the right-hand side of (2) contains a factor $(r_i - r_j)$, so it too is zero.

Case 3: The general case.

By the previous cases we may assume that none of the r_i is zero and that they are all distinct.

Define the functions

$$f(x) = \begin{vmatrix} 1 & x & x^2 & \dots & x^n \\ 1 & r_1 & r_1^2 & \dots & r_1^n \\ 1 & r_2 & r_2^2 & \dots & r_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & r_n & r_n^2 & \dots & r_n^n \end{vmatrix}$$

and

$$g(x) = \prod_{1 \le i \le n} (x - r_i) \prod_{1 \le i < j \le n} (r_j - r_i) \, .$$

We wish to show that $f = g$. Since f and g are polynomials of degree at most n (why?), it suffices to show that they agree at $n + 1$ points.

The first n points are $(r_i)_1^n$. (As assumed above, these are n distinct points.) Indeed, $f(r_i) = 0$ for $i \in [1..n]$ since for such an argument, the first row of the matrix equals some later row; and $g(r_i) = 0$ for $i \in [1..n]$ since for such an argument, the product contains a factor $(r_i - r_i)$.

The last point is 0. (As assumed above, this is a distinct point from all the r_i.) Indeed, that $f(0) = g(0)$ is exactly case 1.

And that completes the proof.

(I think I read somewhere that this result can also be proved by manipulating the determinant with row operations. I'll look that up and report back.)

References

[1] James H. McKay. Another proof of Cauchy's group theorem. *Am. Math. Mon.*, 66:119, 1959. (Cited on page 1.)