PROOF OF THUE'S THEOREM – PART I

In this lecture we'll do the first half of the proof of Thue's theorem.

Suppose that P is a polynomial with integer coefficients. Let r be a rational point. We would like to understand the relationship between P vanishing to high order at r and the "size" of P in terms of its degree and size of its coefficients. Define $|P|$ to be the maximum absolute value of the coefficients of P.

1. Polynomials vanishing to high order at a rational point

Let us start with the simple case where P is a one-variable polynomial. Let $r = p/q$, written in lowest form. We suppose that P vanishes to ℓ orders at r, i.e.,

$$\partial^j P(r) = 0, \quad j = 0, 1, \ldots, \ell - 1.$$

Here is the first example of such a polynomial that comes to mind.

Example 1: $P(x) = (qx - p)^\ell$. Then $|P| \sim \|r\|^\ell$, where $\|r\| := \max(|p|, |q|)$.

Could we do better? In other words, can we find a polynomial P with smaller coefficients, i.e., $|P|$ smaller? As we shall see, the answer is no.

Proposition 1.1 (Gauss)**.** *If $P \in \mathbb{Z}[x]$ satisfies $\partial^j P(r) = 0$ for $j = 0, 1, \ldots, \ell - 1$, then $P(x) = (qx - p)^\ell P_1(x)$ for some $P_1 \in \mathbb{Z}[x]$.*

Proof. The vanishing condition tells us that $P(x) = (qx - p)^\ell P_2(x)$ for some polynomial $P_2 \in \mathbb{R}[x]$. It remains to show that the coefficients of P_2 are integers. By expanding and comparing coefficients, we see that we can solve for the coefficients of P_2 in terms of the coefficients of P, and we deduce that the coefficients of P_2 are at least rational.

Taking out the lowest common demoninator, we write $P(x) = \frac{1}{M}(qx - p)^\ell \tilde{P}_2(x)$, for some $\tilde{P}_2 \in \mathbb{Z}[x]$ so that there is no prime dividing all the coefficients of $\tilde{P}_2$ as well as M. So $MP(x) = (qx - p)^\ell \tilde{P}_2(x)$. If $M \neq \pm 1$, then let s be any prime divisor of M. Then we get a contradiction modulo s, since $qx - p$ is not 0 mod s as p/q was already given in lowest terms, and $\tilde{P}_2$ is also not 0 mod s. It follows that $M = \pm 1$ and hence $P_1 \in \mathbb{Z}[x]$. □

We're not quite done yet, as it's not always true that the norm of a polynomial is always at least as large as its factors.

Example 2: The polynomial $(x-1)^2 = x^2 - 2x + 1$ has norm 2, while $(x-1)^2(x+1) = x^3 - x^2 - x + 1$ has norm 1, which is smaller.

Fortunately, this example does not really pose an issue. The follow corollary answers our question for one variable polynomials.

Corollary 1.2. $|P| \geq \|r\|^\ell$.

Proof. In $P(x) = (qx - p)^\ell P_1(x)$, we see that q^ℓ divides the top coefficient and p^ℓ divides the bottom non-zero coefficient. □

Now, what about polynomials in two variables? Let $P \in \mathbb{Z}[x_1, x_2]$, and $r = (r_1, r_2) = (p_1/q_1, p_2/q_2) \in \mathbb{Q}^2$. We want to assume that P vanishes to high order at r. Let's say $\partial_j P(r) = 0$ for all $j \in J$, where J is some list of pairs, e.g., all $j = (j_1, j_2)$ with $|j| := j_1 + j_2 \leq \ell - 1$.

Define $\|r\| := \max(\|r_1\|, \|r_2\|)$. If $\|r\|$ is large, does $P(r) = 0$ imply something about the norm of P (as in the single variable case)? The following examples show that the answer is no.

Example 3: $P(x_1, x_2) = x_1 - x_2$ and $r = (r_1, r_1)$. Then $P(r) = 0$ but $|P| = 1$.

What if we assume P vanishes at r to high order? Say $\partial_j P(r) = 0$ for all j with $|j| \leq \ell - 1$? Still the answer is no.

Example 4: $P(x_1, x_2) = (x_1 - x_2)^\ell$ and $r = (r_1, r_1)$. Then $\partial_j P(r) = 0$ for all j with $|j| \leq \ell - 1$, but $|P| \leq 2^\ell$, independent of $\|r\|$.

These examples suggest that perhaps our notion of vanishing to high order at a point isn't very useful. It prompts us to modify the question. Let us consider polynomials of the form

$$P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1).$$

Suppose we have

$$\partial_1^j P(r) = 0, \quad j = 0, 1, \ldots, \ell - 1.$$

In this case, can we infer something about the size of P from the size of r?

Since we are only differentiating with respect to x_1, this condition is equivalent to

$$\partial^j [p_2 P_1 + q_2 P_0](r) = 0, \quad j = 0, 1, \ldots, \ell - 1.$$

It follows by Corollary 1.2 that $\|p_2 P_1 + q_2 P_0\| \geq \|r_1\|^\ell$. We see that $\|p_2 P_1 + q_2 P_0\| \leq \|p_2 P_1\| + \|q_2 P_0\| \leq 2\|r_2\||P|$. It follows that $|P| \geq \frac{1}{2} \frac{\|r_1\|^\ell}{\|r_2\|}$.

Let us look at some examples of polynomials that satisfy the above vanishing conditions.

Example 5: Let $P = q_2 x_2 - p_2$. Then $|P| = \|r_2\|$ and $\partial_1^j P(r) = 0$ for all j.

Example 6: Let $P = (q_2 x_2 - p_1)^\ell$. Then $|P| \geq \|r_1\|^\ell$.

2. Integer solutions to linear systems

So far we've been looking at explicit examples of polynomials that satisfy the vanishing to high order condition. This is somewhat reminiscent of the first lecture in the course where we wanted to know how big the degree of a polynomial P must be if $P(j, 2^j) = 0$ for $j = 1, 2, \ldots, 10^6$. There we also started by finding explicit examples, but at the end we arrived at our bound by counting dimensions. In a similar vein, we are going to find polynomials in $\mathbb{Z}[x_1, x_2]$ by parameter counting.

Proposition 2.1. *If $L\colon \mathbb{Z}^M \to \mathbb{Z}^N$ is a linear map, given by a matrix with integer coefficients, with $M > N$, then there exists a nonzero $x \in \mathbb{Z}^M$ such that $Lx = 0$.*

For real vector spaces, this result follows from elementary results from linear algebra. For integers, it's actually even more elementary — it's just pigeonhole principle.

Let's quickly sketch a proof first. Afterwards we'll be more careful quantitative bounds.

Proof. (Sketch) Let $Q_S^M := \{x \in \mathbb{Z}^M : |x_i| \leq S, i = 1, \ldots, M\}$. Since the map L restricted to $Q_S^M \to Q_{C\cdot S}^N$ where $C = C(L)$ is some sufficiently large constant. We have $|Q_S^M| \sim S^M$ and $|Q_{C\cdot S}| \sim C^N S^N$. So we can choose S so that $|Q_S^M| > |Q_{C\cdot S}^N|$. Then by pigeonhole, there are $x_1 \neq x_2 \in Q_S^M$ such that $L(x_1) = L(x_2)$, so that $L(x_1 - x_2) = 0$. □

How big is the x produced by the proof? Let us look for some quantitative bounds.

We can take $C = |L|_{op} := \max_{|x|_\infty = 1, x \in \mathbb{R}^M} |Lx|_\infty$ by the operator norm of L. In particular, $|L|_{op} \leq M \cdot \max|\text{coeff of } L|$. We need to take S so that $(2S+1)^M > (2|L|_{op}S+1)^N$. It suffices to have $(2S+1)^M > |L|_{op}(2S+1)^N$, or equivalently $2S+1 > |L|_{op}^{M/(M-N)}$. It follows that we can always find a nonzero $x \in \mathbb{Z}^M$ with $Lx = 0$ and $|x|_\infty \leq |L|_{op}^{N/(M-N)}$. So we can revise the proposition to a more quantitative version.

Proposition 2.2. *If $L\colon \mathbb{Z}^M \to \mathbb{Z}^N$ is a linear map, given by a matrix with integer coefficients, with $M > N$, then there exists a nonzero $x \in \mathbb{Z}^M$ with $|x|_\infty \leq |L|_{op}^{N/(M-N)}$ such that $Lx = 0$.*

Note that if $M = N + 1$, then our bound is $|L|_{op}^N$ which is not too great, where as if $M = 1.01N$ then our bound is $|L|_{op}^{100}$ which is pretty good.

Let's go back to the one-variable polynomial case for a moment. Recall that we already know that $(px - q)^\ell$ is the optimal polynomial vanishing to ℓ-th order at $r = p/q$. Nevertheless, let us try this counting machinery here and see how well it does in comparison.

Suppose we are looking for a polynomial P of degree D such that $\partial^j P(r) = 0$ for $j = 0, 1, \ldots, \ell - 1$. Let

$$P = \sum_{i=0}^{D} a_i x^i$$

We have

$$\partial^j P(x) = \sum_{i=0}^{D} a_i \frac{i!}{(i-j)!} x^{i-j} = j! \sum_{i=0}^{D} a_i \binom{i}{j} x^{i-j}.$$

(Extracting out the $j!$ factor in the last step is a useful trick of the trade that makes it easier to bound the coefficients.) Setting $\partial^j P(r) = 0$, we have

$$\sum_{i=0}^{D} a_i \binom{i}{j} q^{D-(i-j)} p^{i-j} = 0.$$

The coefficients of the a_i's are all bounded in absolute value by $2^D \|r\|^D$. Viewing $(a_0, \ldots, a_D) \in \mathbb{Z}^{D+1}$ as our unknowns, it follows from Proposition 2.2 that we can find a polynomial P of degree D with $\partial^j P(r) = 0$ for $j = 0, 1, \ldots, \ell - 1$ such that

$$|P| \leq (2^D \|r\|^D)^{\ell/(D-\ell)} \sim \|r\|^{\ell D/(D-\ell)}.$$

So we could take, for example, $D = 100\ell$ to get $|P| \sim \|r\|^{1.01\ell}$. For comparison, the optimal example $(qx - p)^\ell$ has $D = \ell$ and $|P| \sim \|r\|^\ell$.