IRREDUCIBILITY OF POLYNOMIALS
(MA2316, EIGHTH WEEK)

VLADIMIR DOTSENKO

This week we shall discuss two different stories which have some relevance for number theory. First, we shall talk about irreducibility of polynomials in $\mathbb{Z}[x]$ and various methods to prove it, second, we discuss Diophantine equations for polynomials, and draw some parallels between $\mathbb{Z}$ and $\mathbb{C}[x]$, two different Euclidean rings which share some things in common.

In this lecture, we shall work with polynomials from $\mathbb{Z}[x]$ only, and moreover shall assume everywhere that we are dealing with *primitive* polynomials, that is polynomials whose coefficients have no simultaneous common divisors. Such a polynomial is irreducible if and only if it cannot be factorised as a product of two factors of smaller degrees.

There are some very well known methods to prove irreducibility of polynomials. One method is very naïve: if a polynomial $f(x)$ is irreducible when considered as a polynomial with coefficients modulo p, then of course it is irreducible over integers. This is quite alright, but useless for many cases. For example, as you will see in the next tutorial, the polynomial x^4+1 is irreducible over integers, but becomes reducible modulo p for every p.

Another famous method involves the Eisenstein's criterion. It states that if a polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ satisfies, for some prime p the conditions $\gcd(a_n, p) = 1$, $p \mid a_i$ for $i < n$, and $p^2 \nmid a_0$, then $f(x)$ is irreducible over integers. Sometimes, this method is not directly applicable, but becomes applicable after some transformation. For example, $x^{p-1}+x^{p-2}+\cdots+x+1$ satisfies these conditions after the change of variables $x = y + 1$ as a polynomial in y.

Let us explain a criterion that generalises that of Eisenstein proved by Gustav Dumas. For that, we shall assign to a polynomial $f(x)$ some combinatorial data. Let us write that polynomial as

$$f(x) = a'_n p^{\alpha_n} x^n + a'_{n-1} p^{\alpha_{n-1}} x^{n-1} + \cdots + a'_1 p^{\alpha_1} x + a'_0 p^{\alpha_0},$$

where $\gcd(a'_k, p) = 1$. Furthermore, let us mark in the 2D plane all points (k, α_k). These points give rise to the *Newton diagram* of $f(x)$ modulo p, which is defined as follows. Let $P_0 = (0, \alpha_0)$, and let $P_1 = (i_1, \alpha_{i_1})$, where i_1 is the maximal integer i for which there are no marked points below the line connecting $(0, \alpha_0)$ and (i, α_i). Further, let $P_2 = (i_2, \alpha_{i_2})$, where i_2 is the maximal integer i for which there are no marked points below the line connecting (i_1, α_{i_1}) and (i, α_i), etc., the last point P_r being (n, α_n). If a side P_iP_{i+1} of the Newton diagram contains points with integer coordinates, let us also mark all these points, thus obtaining points $Q_0 = P_0$, Q_1, …, $Q_{r+s} = P_r$. Each segment Q_iQ_{i+1} is a *primitive* one, that is there are no integer points on it. We call these segments edges of the Newton diagram. To each polynomial, one can associate its *edge diagram* obtained by translating all the edges to the origin, and keeping its edge with its multiplicity. For example, if the diagram has $P_0 = (0, 0)$, and $P_1 = (2, 2)$, that gives rise to an edge of the same direction but half-length, taken with multiplicity two.

Theorem 1 (Dumas). (1) *Suppose that $f(x), g(x), h(x) \in \mathbb{Z}[x]$, and $f(x) = g(x)h(x)$. Then the edge diagram of $f(x)$ is the union of the edge diagram of $g(x)$ and the edge diagram of $h(x)$ (with multiplicities).*

(2) *Suppose that the edge diagram of $f(x)$ consists of one single edge. Then $f(x)$ is irreducible over integers.*

Proof. The second part follows trivially from the first one, so let us prove the first part.

Let

$$f(x) = a'_n p^{\alpha_n} x^n + a'_{n-1} p^{\alpha_{n-1}} x^{n-1} + \cdots + a'_1 p^{\alpha_1} x + a'_0 p^{\alpha_0},$$
$$g(x) = b'_m p^{\beta_m} x^n + b'_{n-1} p^{\beta_{m-1}} x^{m-1} + \cdots + b'_1 p^{\beta_1} x + b'_0 p^{\beta_0},$$
$$h(x) = c'_{n-m} p^{\gamma_{n-m}} x^n + c'_{n-m-1} p^{\gamma_{n-m-1}} x^{n-m-1} + \cdots + c'_1 p^{\gamma_1} x + c'_0 p^{\gamma_0},$$

where a'_i, b'_j, c'_k are not divisible by p.

Let us take one of the sides $P_l P_{l+1}$ of the Newton diagram of $f(x)$ (possibly consisting of several edges). Let the coordinates of P_l and P_{l+1} be (i_-, α_{i_-}) and (i_+, α_{i_+}) respectively. The slope of the line $P_l P_{l+1}$ is

$$M = \frac{\alpha_{i_+} - \alpha_{i_-}}{i_+ - i_-}.$$

We shall write M as a fraction in lower terms, $M = \frac{A}{I}$, where $I > 0$, $\gcd(A, I) = 1$. In the (i, α) coordinates the equation of the line $P_l P_{l+1}$ is

$$I\alpha - Ai = F, \text{ where } F = I\alpha_+ - Ai_+ = I\alpha_- - Ai_-.$$

By our construction, all the points (i, α_i) lie on or above that line, that is $I\alpha_i - Ai \geq F$, and the inequality is strict for $i < i_-$ and for $i > i_+$.

Let us call the quantity $I\alpha - Ai$ the weight of a monomial $a' p^\alpha x^i$, where $\gcd(a', p) = 1$. The numbers i_- and i_+ are the smallest and the largest exponents of monomials of the minimal weight appearing in $f(x)$.

Let us, using the same definition of weight, find "candidates" among sides of Newton diagrams $g(x)$ and $h(x)$. Namely, let us put G to be the minimal weight of monomials appearing in $g(x)$, and H the minimal weight of monomials appearing in $h(x)$. Also, we define j_- and j_+ as the smallest and the largest exponents of monomials of the minimal weight appearing in $g(x)$, and define k_- and k_+ as the smallest and the largest exponents of monomials of the minimal weight appearing in $h(x)$.

Let us examine the coefficient of $x^{j_- + k_-}$ in $f(x)$. On the one hand, it is equal to $a'_{j_- + k_-} p^{\alpha_{j_- + k_-}}$. On the other hand, it is given by the formula

$$\sum_{j+k=j_-+k_-} (b'_j p^{\beta_j})(c'_k p^{\gamma_k}).$$

Note that the weight of the product of two monomials is equal to the sum of their weights. This implies that for the term with $j = j_-$ and $k = k_-$, the weight is equal to $G + H$. The weights of all other monomials that are used to create $x^{j_- + k_-}$ is strictly greater than $G + H$, since for them either $j < j_-$ or $k < k_-$.

If $j + k$ is constant, the weight of $(b'_j p^{\beta_j} x^j)(c'_k p^{\gamma_k} x^k)$ increases as $\beta_j + \gamma_k$ increases, since $I > 0$. Since in our case $j + k = j_- + k_-$, this implies that the sum $\beta_j + \gamma_k$ is minimal for $j = j_-$ and $k = k_-$. Therefore, the maximal power of p by which

$$\sum_{j+k=j_-+k_-} (b'_j p^{\beta_j})(c'_k p^{\gamma_k})$$

is divisible is equal to $p^{j_- + k_-}$, and the weight of the monomial $a' p^\alpha x^i$ is equal to $G + H$. It is also clear that for $i < j_+ k_-$ the weight of $a'_i p^{\alpha_i} x^i$ is strictly greater than $G + H$, and for $i \geq j_+ k_-$ the weight of $a'_i p^{\alpha_i} x^i$ is at least $G + H$. Therefore, $F = G + H$, and $i_- = j_- + k_-$. Similarly, one can prove $i_+ = j_+ + k_+$. Therefore,

$$i_+ - i_- = (j_+ - j_-) + (k_+ - k_-).$$

In particular, at least one of the numbers $(j_+ - j_-)$ and $(k_+ - k_-)$ is positive. Note that the slopes of respective sides of Newton diagrams are all $M = \frac{A}{I}$ since

$$\frac{\beta_{j_+} - \beta_{j_-}}{j_+ - j_-} = \frac{A}{I} = \frac{\gamma_{k_+} - \gamma_{k_-}}{k_+ - k_-}$$

because of the way we define the weight. This shows that the sum of the lengths of sides of slope M for Newton diagrams of $g(x)$ and $h(x)$ is equal to the length of the side of that slope for $f(x)$. □

Remark. *Applying this result to a polynomial satisfying conditions of the Eisenstein's criterion, we see that the edge diagram manifestly consists of one edge, and therefore such a polynomial must be irreducible.*