Florian Enescu, Polynomials Spring 2010: Lecture notes Week 4.

1. Estimates for roots of polynomials.

We will move now to general considerations on the roots of a polynomial and we will discuss an old result due to Cauchy.

Theorem 1.1. *Let $f(X) = X^n - b_1X^{n-1} - \cdots - b_n$, where $b_i \in \mathbf{R}_{\geq 0}$, for all $i = 1, \ldots, n$ and at least one of them is nonzero.*

Then $f(X)$ has exactly one positive root p, the root is simple, and all other roots have absolute value less or equal to p.

Proof. Let

$$F(x) = -\frac{f(x)}{x^n} = \frac{b_1}{x} + \cdots + \frac{b_n}{x^n} - 1.$$

Consider $F(x)$ as a function in $x > 0$. Due to our hypotheses $F(x)$ is decreasing from $= \infty$ to -1 and hence it crosses the $x - axis$ only once. So, $F(x) = 0$ has one positive root, say p. But $F(x) = 0$ implies $f(x) = 0$, so p is a unique positive root for f as well.

A root x is multiple for f is and only if $f(x) = f'(x) = 0$.

But

$$-\frac{f'(p)}{p^n} = F'(p) = -\frac{b_1}{p^2} - \cdots - \frac{nb_n}{p^{n+1}} < 0$$

which shows that p is a simple root for f.

It remains to prove that if x_o is a root for f then $\mid x_o \mid \leq p$.

We will prove the remaining statement by contradiction. Suppose $\mid x_o \mid > p$. Since $F(x)$ is decreasing we get that $F(\mid x_o \mid) < F(p) = 0$, so $f(\mid x_o \mid) > 0$ which gives

$$\mid x_o \mid^n > b_1 \mid x_o \mid^{n-1} + \cdots + b_n \mid x_o \mid, \quad (1)$$

However $f(x_0) = 0$ so

$$x_o^n = b_1 x_o^{n-1} + \cdots + b_n,$$

which leads to

$$\mid x_o \mid^n \leq b_1 \mid x_o \mid^{n-1} + \cdots + b_n \mid x_o \mid, \quad (2)$$

after applying the triangle inequality.

Obviously (1) contradicts (2) so $\mid x_o \mid \leq p$.

□

Does it have to be that all other roots different from p have absolute value strictly less than p? The answer is no, as illustrated by the polynomial $x^n - 1$. The roots of this polynomial are the nth roots of unity and they all lie on the unit circle, that is they all have absolute value equal to 1. Only one root is positive: $p = 1$.

However, something can still be said regarding this question in the presence of additional hypotheses.

Theorem 1.2 (Ostrovsky). *Let $f(X) = X^n - b_1X^{n-1} - \cdots - b_n$, where $b_i \in \mathbf{R}_{\geq 0}$, for all $i = 1, \ldots, n$ and at least one of them is nonzero.*

If the greatest common divisor of the indices of the negative coefficients equals 1, then f has a unique positive root p, simple, and the absolute values of the rest of the roots are strictly less than p.

Proof. Assume that the only positive coefficients are $b_{k_1}, \ldots, b_{k_m}$ with $k_1 < \cdots < k_m$.

Since $[k_1, \ldots, k_m] = 1$, then there exist integers $s_1, \ldots, s_m$ such that

$$s_1k_1 + \cdots + s_mk_m = 1.$$

We already know that

$$F(x) = -\frac{f(x)}{x^n} = \frac{b_{k_1}}{x^{k_1}} + \cdots + \frac{b_{k_m}}{x^{k_m}} - 1,$$

has a unique positive root p.

Let y be any other nonzero solution of $F(x)$ (hence of $f(x)$. Let $q =\mid y \mid$.

$$1 = \frac{b_{k_1}}{y^{k_1}} + \cdots + \frac{b_{k_m}}{y^{k_m}} \leq\mid \frac{b_{k_1}}{y^{k_1}} \mid + \cdots + \mid \frac{b_{k_m}}{y^{k_m}} \mid= \frac{b_{k_1}}{q^{k_1}} + \cdots + \frac{b_{k_m}}{q^{k_m}},$$

so $F(q) \geq 0$.

Clearly $F(q) = 0$ if and only if $\frac{b_{k_i}}{y^{k_i}} = \frac{b_{k_i}}{y^{k_i}} > 0$ for all i, or $x^{k_i} > 0$.

Note that in this case

$$(\frac{b_{k_1}}{y^{k_1}})^{s_1} \cdots (\frac{b_{k_m}}{y^{k_m}})^{s_m} > 0.$$

But

$$0 < (\frac{b_{k_1}}{y^{k_1}})^{s_1} \cdots (\frac{b_{k_m}}{y^{k_m}})^{s_m} = \frac{b_{k_1}^{s_1} \cdots b_{k_m}^{s_m}}{x},$$

and since all $b's$ are positive we get $y > 0$. But this contradicts the fact that $y \neq p$ (p is the unique positive solution for F).

So, $F(q) > 0 = F(p)$ and $q > 0$. F is strictly decreasing for $(0, \infty)$, hence $F(q) > F(p)$ gives $q < p$, which is what we wanted to show. □

We can use these two theorems to give estimates to the root of any polynomial with positive coefficients.

Theorem 1.3. *a)[Eneström-Kakeya]*

If all the coefficients of the polynomial $g(x) = a_0x^{n-1} + \cdots + a_{n-1}$ are positive then for any roots ξ of g we have

$$\min\{\frac{a_i}{a_{i-1}} : i = 1, ..., n-1\} \leq\mid \xi \mid\leq \gamma = \max\{\frac{a_i}{a_{i-1}} : i = 1, ..., n-1\}.$$

b) [Ostrovsky]

Let $\frac{a_k}{a_{k-1}} < \gamma$ for all $k_1, ..., k_m$ with notations exactly as in the preceding proof. If the greatest divisor of the numbers $n, k_1, \ldots, k_m$ is equal to 1, then $\mid \xi \mid< \gamma$.