Florian Enescu, Fall 2010 Polynomials: Lecture notes Week 12.

1. Descartes method; Approximation of roots of polynomials.

We will continue now by presenting *Lagrange's method* for solving degree four equations.
Let $x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0$.
Cosnider the following expression

$$A = X_1X_2 + X_3X_4.$$

If we permute the indices $\{1, 2, 3, 4\}$ in A we obtain two more expressions $B = X_1X_3 + X_12X_4, C = X_1X_4 + X_2X_3$.

Then A, B, C are the solutions of the degree three equation

$$z^3 - (A + B + C)z^2 + (AB + AC + BC)z - ABC = 0.$$

Since the expressions $A + B + C, AB + AC + BC, ABC$ are now symmetric in X_1, X_2, X_3, X_4, one can express them in terms of the symmetric polynomials in s_1, s_2, s_3, s_4.

But once we let $X_i = x_i, i = 1, 2, 3, 4$ where x_i represent the roots of our degree four equation, then we can see that $s_1 = -a_1, s_2 = a_2, s_3 = -a_3, s_4 = a_4$.

Hence the equation

$$z^3 - (A + B + C)z^2 + (AB + AC + BC)z - ABC = 0,$$

has its coefficents in terms of a_1, a_2, a_3, a_4, so we can apply what we know in the case of degree three equations to solve it and get the values for

$$x_1x_2 + x_3x_4, x_1x_3 + x_2x_4, x_1x_4 + x_2x_3.$$

Say that these values are a, b, c, respectively.

Using now that $x_1x_2x_3x_4 = a_4$ we can easily solve for $x_1x_2, x_3x_4, x_2x_4, x_1x_3, x_2x_3, x_1x_4$ and then get x_1, x_2, x_3, x_4.

Descartes Method

We would like to present now a method of solving degree four polynomial equations that is different than the previous two methods discussed in this course. However, this method still reduces the original equation to solving a degree three polynomial equation.

Let $x^4 + qx^2 + rx + s = 0$. We plan to factor the equations as follows

$$0 = x^4 + qx^2 + rx + s = (x^2 + jx + l)(x^2 - jx + m).$$

If we expand and identify the coefficients in both sides we get the following system

$$m + l - j^2 = q$$

$$j(m - l) = r$$

$$lm = s$$

Divide thru the second equation by j and then add and subtract the first two equations.
One gets:

$$2m = j^2 + q + r/j$$

$$2l = j^2 + q - r/j$$

$$lm = s$$

But then using $4s = 4lm = (j^2 + q - r/j)(j^2 + q - r/j)$ one gets after clearing denominators:

$$j^6 + 2qj^4 + (q^2 - 4s)j^2 - r^2 = 0,$$

a cubic equation in $u = j^2$.

In practice, once one gets a value for $u = j^2$ then one gets a value for j, and with that one gets l, m. After having j, l, m the original equation reduces to solving two degree two equations: $x^2 + jx + l = 0, x^2 - jx + m = 0$.

Example 1.1. Solve:

$$x^4 - 2x^2 + 8x - 3 = 0.$$

Note that $q = -2, r = 8, s = -3$.

Then the cubic equation in $u = j^2$ is

$$u^3 - 4u^2 + 16u - 64 = 0,$$

which has solution $u = 4$ as it can easily be seen. So one can take $j = 2$.

But then a easy computation leads to $m = 3, l = -1$ and so it remains to solve

$$x^2 + 2x - 1 = 0, x^2 - 2x + 3 = 0,$$

which the reader can easily solve.

Approximations of real roots

Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$.

We already know that if α is a root for f, then $\mid \alpha \mid < \max\{\mid a_0 \mid, \ldots, \mid a_{n-1} \mid\} + 1$.

If $f(\alpha) = 0$, then $a_n\alpha^n = -a_{n-1}\alpha^{n-1} - \cdots - a_1\alpha - a_0$.

Assume that k is such that a_k is negative and a_m is positive or zero for all $m > k$. If α is positive and f has real coefficients, then at least one coefficient must be negative.

Let M equal the maximum of the absolute value of the negative coefficients of f. Also assume that $a_n > 0$.

Hence

$$a_n\alpha^n = -a_{n-1}\alpha^{n-1} - \cdots - a_1\alpha - a_0 < M\alpha^k + \cdots + M\alpha + M = M\frac{\alpha^{k+1} - 1}{\alpha - 1}.$$

Let us assume that $\alpha > 1$.

Then

$$M\frac{\alpha^{k+1}-1}{\alpha-1} = M\frac{\alpha^{k+1}}{\alpha-1} - \frac{M}{\alpha-1} < M\frac{\alpha^{k+1}}{\alpha-1}.$$

So, $a_n < \frac{M}{a^{n-k-1}(\alpha-1)} < \frac{M}{(\alpha-1)^{n-k}}$.

Hence, $(\alpha-1)^{n-k} < \frac{M}{a_n}$ which gives

$$\alpha < \sqrt[n-k]{\frac{M}{a_n}} + 1.$$

In conclusion, we have proven the following

Theorem 1.2. *Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial with real coefficients such that $a_n > 0$. Let M be the maximum of the absolute value of the negative coefficients of f. Let k such that $a_k < 0$ but $a_m \geq 0$ for all $m > k$.*

Then

$$\alpha < \sqrt[n-k]{\frac{M}{a_n}} + 1,$$

where α is a root of f.

Proposition 1.3. *Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial with real coefficients such that $a_n > 0$.*

If a is such that $f^{(i)}(a) > 0$ for $i = 0, 1, 2, \ldots, n$, then for any root α of f, we have

$$\alpha < a.$$

Proof. Write Taylor's formula:

$$f(x) = \sum_{i=0}^{n} (x-a)^i \frac{f^{(i)}(a)}{i!}.$$

Then if $\alpha \geq a$, the hypothesis together with this formula shows that $f(\alpha) > 0$ which is a contradiction. □

Sturm sequences

Let f be a real polynomial such that f does not have multiple roots (that is, f and f' do no share a root).

A *Sturm sequence* on the interval $[a, b]$ is a finite sequence of polynomials

$$S = \{f_0, f_1, \ldots, f_t\},$$

such that $f_0 = f, f_1 = f'$ and the following properties are satisfied:

1) f_t has no real roots;

2) for all two polyonomials f_i, f_{i+1}, $0 \leq i \leq t-1$ do nt have common roots.

3) If $a \in \mathbf{R}$ with $f_i(a) = 0$ for some $i = 1, \ldots, i-1$, then $f_{i-1}(a)f_{i+1}(a) < 0$.

We will generally interested in Sturm sequences that satisfy the additional property that $f_i(a) \neq 0, f_i(b) \neq 0$ for all $i = 0, ..., t$.

Proposition 1.4. *For any polynomial f with real coefficients that does not have multiple roots there exists a Sturm sequence on an interval $[a, b]$.*

Proof. We take $f_0 = f, f_1 = f'$ and then for $i \geq 2$ we let f_{i+1} equal $-r_{i+1}$ where r_{i+1} is the remainder obtained by dividing f_{i-1} to f_i.

Hence

$$f_{i-1} = f_i q_i - f_{i+1}.$$

Obviously this procedure leads to $f_t = a$, where a is a constant.

First note that a cannot be zero since otherwise f_{t-1}, f_{t-2} will have a common root which gives that f_{t-2}, f_{t-3} have a common roots and so on until we get that $f_0 = f, f_1 = f'$ have a common root which we know that is impossible.

So, condition 1 is fulfilled.

Similarly, if f_i, f_{i+1} have common roots, then we recursively get that f_0, f_1 have common roots, which is impossible.

Now, assume that $f_i(a) = 0$, then $f_{i-1}(a) = -f_{i+1}(a)$ and so $f_{i\_1}(a)f_{i+1}(a) = -(f_{i+1}(a))^2 < 0$ (it cannot be equal to zero because then a would be a common root for f_i, f_{i+1}.

□

For a sequence of real numbers $a_1, ..., a_n$ we call its *sign variation* the integer denoted by N equal to the number of sign changes after deleting possible zeroes in the sequence.

For example, for $-1, 0, 1, 2, -4, 5$ we have $N = 3$.

If we have a Sturm sequence $\{f_0, f_1, \dots, f_t\}$ we denote by $w(x)$ the sign variation of the sequence $f_0(x), \dots, f_t(x)$.

Theorem 1.5. *Let f be a polynomial with real coefficients that does not have multiple roots.*

Then the number of real roots of f situated between $a < b$ equals $w(a) - w(b)$ for each Sturm sequence $S = \{f_0, f_1, \dots, f_t\}$ that satisfy the additional property that $f_i(a) \neq 0, f_i(b) \neq 0$ for all $i = 0, ..., t$.

Proof. Let $x_1, x_2, \dots, x_r$ be ALL the roots of the polynomials $f_0, f_1, ..., f_t$ ordered increasingly:

$$x_1 < x_2 < \cdots < x_r.$$

We will like to show that $w(x)$ will remain unchanged between x_k and x_{k+1}.

If there a variation in signs then there must exist $\alpha < \beta \in (x_k, x_{k+1})$ such that $f_i(\alpha)f_i(\beta) < 0$.

But then there exists $c \in (\alpha, \beta)$ such that $f_i(c) = 0$. But c is different from all x_i and this is a contradiction.

We will now show that if there exists a unique $y \in \mathbf{R}$ such that $a < y < b$ and $f_i(y) = 0$ for some $i = 0, 1, \dots, t$ then $w(a) - w(b) = 1$ if $f(y) = f_0(y) = 0$ and $w(a) - w(b) = 0$ otherwise.

Let us assume first $f_i(y) = 0$ for some $1 \leq i \leq t - 1$.

But then $f_{i-1}(y)$ and $f_{i+1}(y)$ has opposite signs. Moreover due to the uniqueness of y we know that f_{i-1}, f_{i+1} have constant sign between a and b, so the sign variation comes down to what happens with the two sequences:

$f_{i-1}(a), f_i(a), f_{i+1}(a)$ and $f_{i-1}(b), f_i(b), f_{i+1}(b)$.

But since the extremities have opposite signs, then no matter what it is in the middle, the sign variation is the same.

In the case that y is a root for f, then $f(a)$ must have different sign from $f(b)$ otherwise we contradict the uniqueness of y. Similarly, $f'(a), f'(b)$ have the same sign. Hence $w(a) - w(b) = 1$.

To conclude the proof of the Theorem we remark that as we move y from a to b we can apply the above paragraph for each (x_k, x_{k+1}). We get that the sign variation changes by 1 every time we "hit" a root of f.

□

Remark 1.6. We can apply the above for a polynomial with multiple roots as follows:

Divide f by the greatest common divisor between f, f' and then apply the above theorem. We obtain a count of the roots of f without taking in consideration their multiplicities.

Example 1.7. Let $f = X^3 + pX + q$.

Computing a Sturm sequence we get $f_0 = f, f_1 = 3X^2 + p, f_2 = -2pX - 3q, f_3 = -4p^3 - 27q^2$.

To find the number of real solutions we compute $w(-\infty) - w(\infty)$.

We need to discuss the sign of $f_3 = -4p^3 - 27q^2$.

If $f_3 \geq 0$, then $p \leq 0$. Let us first assume $p < 0$.

But $w(-\infty) = 3$ ($p < 0$) and $w(\infty) = 0$. So we have three real roots.

The case $p = 0, f_3 \geq 0$ gives $q = 0$ and hence we have a triple root equal to 0.

If $f_3 < 0$, then $w(-\infty) = 2$, $w(\infty) = 1$ regardless of the sign for p. So our polynomial has one real root (not counted with multiplicity).