Florian Enescu, Polynomials: Lecture notes Week 10.

1. Grace's Theorem; Resultants

In the proof of Grace's theorem we will use the following Lemma.

Lemma 1.1. *Let all roots $z_1, \ldots, z_n$ of $f(z)$ lie inside the circular domain K and let ξ lie outside K. Then all roots of $A_\xi f(z)$ lie inside K.*

Proof. Assume first that $\xi \neq \infty$. Then w is a root of $A_\xi f(z)$ then ξ is the center of mass of the roots of f with respect to w.

Theorem 1.6 from Lecture 6 says that since $z_1, \ldots, z_n$ are in K and the center is not in K, then w must be in K.

If $\xi = \infty$, then $A_\xi f(z) = f'(z)$. The condition that $\xi \notin K$ means that K is not the exterior of a circle so K is a convex set and hence it contains the convex hull of $z_1, \ldots, z_n$. Now, Gauss-Lucas Theorem shows that the critical points of $f(z)$ are in K as well.

□

We will prove now the last theorem stated in Lecture 9.

Theorem 1.2 (J. H. Grace, 1902). *Let f, g be two apolar polynomials. If all the roots of g belong to a circular domain K, then at least one of the roots of f also belongs to K.*

Proof. Suppose that all the roots $z_1, ..., z_n$ of f lie outside K.

By Lemma proven above, $A_{z_n} g(z)$ has all roots inside K. Repeated applications of the Lemma show that $A_{z_2} ... A_{z_n} f(z)$ has all roots in K. But this last expression is a polynomial of degree 1 hence of the form $c(z-a)$, $c \neq 0$. So, $a \in K$.

Let us compute remember that f, g are apolar so $0 = A_{z_1} ... A_{z_n} g(z) = A_{z_1}(c(z-a)) = c(z_1 - a)$.

Hence $z_1 = a$. But $z_1 \in K$, while $a \notin K$. Contradiction.

So, at least one z_i is in K.

□

The next topic we will examine is that of resultants. The resultant of two polynomials f, g whether they have a common root or not.

Let $f(x) = a_0 x^n + \cdots + a_n$, $g(x) = b_0 x^m + \cdots b_m$ with $a_i, b_j \in \mathbf{C}$, $a_0 \neq 0, b_0 \neq 0$.

Proposition 1.3. *The polynomials f, g have a common root if and only if there exist nonzero complex polynomials p, q of degrees at most $m-1$, respectively $n-1$ such that $pf = qg$.*

Proof. If f, g have a common root a, then $x - a$ divides both f and g, so $f = q(x-a)$ and $g = p(x-a)$ with q, p nonzero polynomials of degrees $n-1, m-1$ respectively.

But then $pf = qg$, as it can easily be checked.

For the converse, $pf = qg$ implies that if none of the linear factors of g appear in f then they all must appear in p but this shows that g divides p and hence the degree of p is at least m which is a contradiction.

So, at least one of the linear factors that appear in the factorization of g must appear in the factorization of f which is equivalent to saying that f, g have a common root.

□

Let us pursue what the previous result tells us.

Write $q(x) = u_0x^{m-1} + \cdots u_{m-1}, p(x) = v_0x^{n-1} + \cdots v_{n-1}$, where the coefficients are complex numbers not all zero.

Now, if we look at the equality

$$p(x)f(x) = q(x)g(x),$$

we note that if we identify the coefficients of the polynomial expression for the left and right hand terms we get that it is equivalent to finding $u_1, ..., u_{m-1}$ not all zero, and $v_0, ..., v_{n-1}$ not all zero such that they satisfy the following system of equations:

$$a_0u_0 - b_0v_0 = 0$$

$$a_1u_0 + a_0u_1 - b_1v_0 - b_0v_1 = 0$$

$$a_2u_0 + a_1u_1 + a_2u_2 - b_2v_0 - b_1v_1 - b_0v_2 = 0$$

$$\ldots$$

Note that the system has $m + n$ lines and $m + n$ unknown and it is homogeneous and linear.

Such a system admits a nonzero solution (our u's and v's) if and only if its determinant is zero.

After rearranging the rows and columns we get the following determinant denoted by $R(f, g)$. The matrix that gives this determinant is called the Sylvester matrix of f, g (it is of size $n + m$ by $n + m$).

$$R(f,g) = \begin{vmatrix} a_0 & a_1 & \cdots & a_n & 0 & \cdots \\ 0 & a_0 & \cdots & a_{n-1} & a_n & 0 \\ . & . & . & . & . & . \\ b_0 & b_1 & \cdots & b_m & 0 & \cdots \\ 0 & b_0 & \cdots & b_{m-1} & b_n & 0 \\ . & . & . & . & . & . \end{vmatrix}$$

Note that n, m are different that in fact the reader should not infer from the above presentation that a_n and b_m lie on same column. Also, the first m rows contain only a's and the remaining n rows contain only b's.

Hence we can conclude

Theorem 1.4. *The polynomials f, g share a root if and only if $R(f, g) = 0$.*

There is a more definite relation between the roots of f, g and the resultant.

Theorem 1.5. *Let $x_1, \cdots, x_n$, $y_1, ..., y_m$ be the roots of f respectively g.*
Then $R(f, g) = a_0^m b_0^n \Pi(x_i - y_j) = a_0^m \Pi_{i=1}^n g(x_i) = (-1)^{nm} b_0^n \Pi_{j=1}^m f(y_j)$.

Proof. First let us notice that $R(f, g)$ is a homogeneous polynomial of degree m in $a_0, ..., a_n$ and degree n in $b_0, ..., b_m$.

Since $a_i/a_0, b_j/b_0$ are symmetric expressions in $x_1, ..., x_n$ and $y_1, ..., y_m$ respectively by Viète, we get that $R(f, g) = a_0^m b_0^n P(x_1, ..., x_n, y_1, ..., y_m)$ with P a symmetric polynomial that vanishes whenever $x_i = y_j$ for some i and j.

But we can always use repeatedly the equality

$$x_i^k = (x_i - y_j)x_i^{k-1} + x_i^{k-1}y_j$$

to write
$P(x_1, ..., x_n, y_1, ..., y_m) = (x_i - y_j)Q(x_1, ..., x_n, y_1, ..., y_m) + U(x_1, ..., \hat{x_i}, ..., y_m)$.
where $\hat{x_i}$ symbolizes that x_i does not appear in U.

But letting $x_i = y_j$ makes $R(f, g) = 0$ so in fact $U \equiv 0$, and this shows that P must be divisible by all $x_i - y_j$ and so $R(f, g)$ is divisible by $S = a_0^m b_0^n \Pi(x_i - y_j)$.

Write $g(x) = b_0\Pi(x - y_j$. Hence $S = a_0^m \Pi g(x_i)$. Similarly, $f(y_j) = (y_j - x_1)\cdots(y_j - x_n) = (-1)^n(x_1 - y_j)\cdots(x_n - y_j)$.

So, $S = (-1)^{nm}\Pi_{j=1}^m f(y_j)$.

Now look at

$$S = a_0^m \Pi_{i=1}^n (b_0 x_i^m + \cdots b_m).$$

This is a polynomial expression that has degree exactly n in $b_0, ..., b_m$ (and homogenous). If we expand we see that S is a symmetric polynomial in $x_1, ..., x_n$ and in fact by using Viète's relations we see that S is homogenous of degree m is $a_0, ..., a_n$. But S divides $R(f, g)$ as polynomials so there must exist a constant λ such that

$$R(f, g) = \lambda S.$$

By looking at the coefficient of $x_1^m \cdots x_n^m$ in both sides we get that $\lambda = 1$. This finishes our proof.

□

Corollary 1.6. *$R(g, f) = (-1)^{deg(f)deg(g)}R(f, g)$*

Corollary 1.7. *If $f = qg + r$, then*

$$R(f, g) = b_0^{deg(f)-deg(r)}R(r, g).$$

Proof. Let y_j be the roots of g. Then $f(y_j) = r(y_j)$, since $g(y_j) = 0$.

So, $R(f, g) = b_0^{deg(f)}\pi f(y_j) = b_0^{deg(f)-deg(r)}\pi r(y_j) = b_0^{deg(f)-deg(r)}R(r, g)$.

□

Corollary 1.8.

$$R(f, g) = R(f, g)R(f, h).$$

Definition 1.9. Let $x_1, \dots, x_n$ be the roots of a degree n polynomial $f(x) = a_0x^n + \cdots a_n$.
The *discriminant* of f is

$$D(f) = a_0^{2n-2}\Pi_{i<j}(x_i - x_j)^2.$$

Theorem 1.10.

$$R(f, f') = (-1)^{n(n-1)/2}a_0 D(f).$$

Proof. First note that $R(f, f') = a_0^{n-1}\Pi f'(x_i)$.

However, $f'(x) = a_0 \sum_{i=1}^{n}(x - x_1)\cdots(x \hat{-} x_i)\cdots(x - x_n)$, so $f'(x_i) = a_0\Pi_{i\neq j}(x_i - x_j)$.

Therefore

$$R(f, f') = a_0^{2n-1}\Pi_{i\neq j}(x_i - x_j) = (-1)^{n(n-1)/2}a_0^{2n-1}\Pi_{i<j}(x_i - x_j)^2.$$

□

Corollary 1.11. *Let f, g, h be monic polynomials. Then*

$$D(fg) = D(f)D(g)R^2(f, g)$$

$$D(fgh) = D(f)D(g)D(h)R^2(f, g)R^2(g, h)R^2(h, f)$$

Theorem 1.12. *let f be a real polynomial of degree n without any real roots.*

Then the sign of $D(f)$ is the sign of $(-1)^{n/2}$.

Proof. Let $f(x) = a_0(x - x_1)\cdots(x - x_n)$.

We can verify that $D((x - a)f(x)) = D(f(x))[f(a)^2]$.

Now, let $a, \overline{a}$ be conjugate roots of f: $f(x) = (x - a)(x - \overline{a})g(x)$.

Then

$$D(f) = D(g)(a - \overline{a})(g(a)g(\overline{a})^2,$$

which shows that $D(f) = -D(g)$. Continuing like this until we exhaust all roots of f and we get our statement.

□