LECTURE 10, MONDAY MARCH 15, 2004

FRANZ LEMMERMEYER

1. Minimal Polynomials

Let α and β be algebraic numbers, and let f and g denote their minimal polynomials. Consider the resultant $R(X)$ of the polynomials $f(X-Y)$ and $g(Y)$ with respect to Y; then we claim that $R(X)=0$ for $X=\alpha+\beta$. In fact, $Y=\beta$ is a common root of $f(X-Y)$ and $g(Y)$. Note that the minimal polynomial of $\alpha+\beta$ is in general not equal to $R(X)$ (try $\alpha=1+\sqrt{2}$ and $\beta=1-\sqrt{2}$) but to a factor of $R(X)$.

As an example, let us compute the minimal polynomial of $\sqrt{2}+\sqrt[3]{2}$. We know that $f(X)=x^2-2$ and $g(x)=x^3-2$, and typing

```
polresultant((x-y)^2-2,y^3-2,y)
```

into `pari` shows that $R(x)=x^6-6x^4-4x^3+12x^2-24x-4$.

As a more involved example, consider $\alpha=\sqrt{2}+\sqrt{3}$ and $\beta=\sqrt{5}+\sqrt{6}$; then we find that $\alpha+\beta$ is an element of $\mathbb{Q}(\sqrt{2},\sqrt{3},\sqrt{6}\,)$, hence should have a minimal polynomial of degree dividing 8. The resultant

```
r = polresultant((x-y)^4-10*(x-y)^2+1,y^4-22*y^2+1,y)
```

has degree 16, but factors into two polynomials of degree 8 each:

```
factor(r)
```

gives the two polynomials

$$a(x)=x^8-64x^6-96x^5+808x^4+1152x^3-2304x^2-1152x+144,$$
$$b(x)=x^8-64x^6+96x^5+808x^4-1152x^3-2304x^2+1152x+144.$$

Note that $b(x)=a(-x)$ here. In any case, $\alpha+\beta$ is the root of one of these polynomials. Now

$$\alpha+\beta=7.8318\ldots,$$

and evaluating a and b at this value shows

$$a(\alpha+\beta)\approx -6.5\cdot 10^{-22},\quad b(\alpha+\beta)\approx 4568619.30\ldots;$$

thus the minimal polynomial of $\alpha+\beta$ is $a(x)$.

Similarly we can compute the minimal polynomial of α/β: consider the polynomials $F(Y)=f(XY)$ and $G(Y)=g(Y)$: then for $X=\alpha/\beta$, F and G have a common root $Y=\beta$, hence $X=\alpha/\beta$ must be a root of $R_{F,G}=0$.

In our example above,

```
polresultant((xy)^2-2,y^3-2,y)
```

gives $R(X)=4X^6-8$, hence X^6-2 is a minimal polynomial (it is Eisenstein for the prime 2, hence irreducible) for $\sqrt{2}/\sqrt[3]{2}=\sqrt[6]{2}$.

Note that the theory of resultants gives a constructive proof of the fact that algebraic numbers form a field, and that algebraic integers form a ring.

2. Bezout

2.1. **Intersection in the Projective Plane.** Assume that we are given two curves $\mathcal{C}_F$ and $\mathcal{C}_G$, where F and G are homogeneous polynomials in $K[X,Y,Z]$ of degrees m and n, rspectively. Assume that the point $[0:0:1]$ is on neither of these curves. Then we can write

$$F(X,Y,Z) = Z^m + a_1 Z^{m-1} + \ldots + a_{m-1}Z + A_m,$$
$$G(X,Y,Z) = Z^n + b_1 Z^{n-1} + \ldots + b_{n-1}Z + B_n,$$

where the a_i and b_j are homogeneous polynomials of degree i and j in $K[X,Y]$. Assume that the point $[x:y:z]$ is a point of intersection. Then $F(x,y,Z)$ and $G(x,y,Z)$ have the common root $Z=z$, so the resultant $R_{F,G}$ of F and G with respect to Z (which is a polynomial in two variables X, Y) must vanish. In other words: the points of intersection of $\mathcal{C}_F$ and $\mathcal{C}_G$ come from the zeros of the resultant $R(X,Y) = R_{F,G}$.

In order to get an upper bound on the number of points of intersection we have to compute the degree of $R(X,Y)$. In fact we will show

Proposition 2.1. *The polynomial $R(X,Y)$ is homogeneous, and it is either 0 or has degree mn.*

Proof. We have to show that $R(tX,tY) = t^{mn}R(X,Y)$. We have

$$R(tX,tY) = \begin{vmatrix}
1 & ta_1 & t^2a_2 & \ldots & t^ma_m & & & & \\
 & 1 & ta_1 & t^2a_2 & \ldots & t^ma_m & & & \\
 & & 1 & ta_1 & t^2a_2 & \ldots & t^ma_m & & \\
 & & & \ldots & & & & & \\
 & & & & \ldots & & & & \\
 & & & & 1 & ta_1 & t^2a_2 & \ldots & t^ma_m \\
1 & tb_1 & t^2b_2 & \ldots & t^nb_n & & & & \\
 & 1 & tb_1 & t^2b_2 & \ldots & t^nb_n & & & \\
 & & 1 & tb_1 & t^2b_2 & \ldots & t^nb_n & & \\
 & & & \ldots & & & & & \\
 & & & & \ldots & & & & \\
 & & & & 1 & tb_1 & t^2b_2 & \ldots & t^nb_n
\end{vmatrix}$$

Now multiply the second row by t, the third by t^2, and the i-th row of a's by t^{i-1}; similarly, multiply the i-th row of b's by t^{i-1}. Then the k-th column has a common factor t^{k-1}, which we extract. The result is just $R(X,Y)$. Now we have multiplied the determinant by

$$t^{1+2+\ldots+n-1} \cdot t^{1+2+\ldots+m-1} = t^{\frac{m(m-1)+n(n-1)}{2}},$$

and we have extracted the factor

$$t^{1+2+\ldots+m+n-1} = t^{\frac{(m+n)(m+n-1)}{2}}.$$

But since

$$\frac{(m+n)(m+n-1)}{2} - \frac{m(m-1)+n(n-1)}{2} = mn,$$

this shows that $R(tX,tY) = t^{mn}R(X,Y)$ as claimed. □

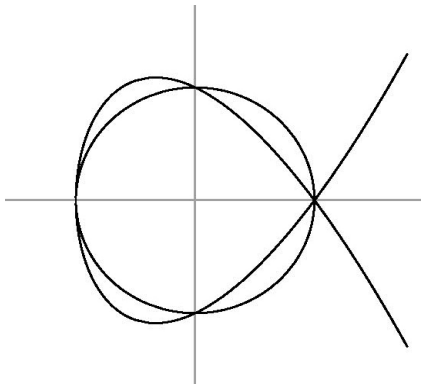

FIGURE 1. Intersection of $x^2 + y^2 = 1$ and $y^2 = x^3 - x^2 - x + 1$

Example. Consider the unit circle $X^2 + Y^2 - Z^2 = 0$ and the cubic $X^3 - X^2Z - XZ^2 + Z^3 - Y^2Z = 0$.

Since $[0 : 0 : 1]$ is not on these curves, we form the resultant with respect to Z and get

$$R(X,Y) = \begin{vmatrix} X^2+Y^2 & 0 & -1 & 0 & 0 \\ 0 & X^2+Y^2 & 0 & -1 & 0 \\ 0 & 0 & X^2+Y^2 & 0 & -1 \\ X^3 & -X^2-Y^2 & -X & 1 & 0 \\ 0 & X^3 & -X^2-Y^2 & -X & 1 \end{vmatrix} = -X^2Y^4.$$

Thus $X = 0$ or $Y = 0$. The first possibility leads to $0 = Y^2 - Z^2$ and $0 = Z(Y^2 - Z^2)$, so it gives the points $[0 : 1 : 1]$ and $[0 : -1 : 1]$. Similarly, $Y = 0$ leads to $0 = X^2 - Z^2$ and $0 = X^3 - X^2Z - XZ^2 + Z^3 = (X - Z)(X^2 - Z^2)$, which gives the points $[1 : 0 : 1]$ and $[-1 : 0 : 1]$. Thus we have four points of intersection in the affine plane, namely $(0, 1)$, $(0, -1)$, $(1, 0)$ and $(-1, 0)$.

Note that the factor $X^2 = 0$ corresponds to the two simple points of intersection $(0, 1)$ and $(0, -1)$, whereas the factor $Y^4 = 0$ corresponds to the two points $(-1, 0)$ and $(1, 0)$, both of which will turn out to have multiplicity 2. Using the above coordinate system, however, it is difficult to assign multiplicities, because one factor like $X = 0$ corresponds to two different points; this problem can be avoided by choosing a coordinate system in which $[0 : 0 : 1]$ is not on any line connecting two points of intersection.

Bezout's Theorem. It is clear that if two lines have two points in common, then they are equal. Similarly we have seen that if a line and a conic share three points, then they have a common component. These are very simple instances of the following Theorem of Bezout (actually it is only a weak version):

Theorem 2.2. *If two curves of order m and n have more than mn distinct points in common, then they have a common component.*

Proof. Suppose the curves have more than mn points in common. Pick $mn + 1$ of these points and choose coordinates in such a way that the point with coordinates $[0 : 0 : 1]$ is not collinear with any pair of them (we are working over an algebraically closed field K; so just pick a point not on the finitely many lines through pairs of

points in the finite set of $mn+1$ points) and is not on one of the curves. In these coordinates, the curves have equations

$$F(X,Y,Z) = Z^m + a_1 Z^{m-1} + \ldots + a_m = 0,$$
$$G(X,Y,Z) = Z^n + b_1 Z^{n-1} + \ldots + b_n = 0,$$

where the a_i and b_i are homogeneous polynomials of degree i in $K[X,Y]$.

Now let $[x:y:z]$ be the coordinate of one of the $mn+1$ points of intersection. We have seen that this implies that $R(x,y)=0$, where $R(X,Y)=R_{F,G}$ is the resultant of F and G with respect to Z. If $[x':y':z']$ is another point of intersection, and if we had $[x':y']=[x:y]$, then the three points $[x:y:z]$, $[x':y':z']$ and $[0:0:1]$ would be collinear (the line in question is $xX-yY=0$), contradicting our choice of coordinates.

Thus $R(X,Y)=0$ for $mn+1$ pairwise different ratios. Since R has degree $\le mn$, this implies that R must vanish completely, which implies that F and G have a common factor. □

Even the weak form of Bezout's Theorem has many important consequences:

Theorem 2.3. *If two curves of order n intersect in n^2 distinct points, and if exactly mn of these points lie on an irreducible curve of degree m, then the remaining n^2-mn points lie on a curve of degree $n-m$.*

Proof. Let the curves of order n be given by the equations $F=0$ and $G=0$, and let $H=0$ describe the irreducible curve of degree m. Pick a point P on $\mathcal{C}_H$ not on $\mathcal{C}_F$ or $\mathcal{C}_G$. Now choose nonzero elements $a,b\in K$ such that P lies on $aF+bG=0$, and consider the curve $\mathcal{C}: aF+bG=0$. Then $\mathcal{C}$ and $\mathcal{C}_H$ have at least $mn+1$ points of intersection, so by Bezout they have a common component. Since $\mathcal{C}_H$ is irreducible, we find that this common component is $\mathcal{C}_H$, hence H must divide $aF+bG$, and we have $aF+bG=HL$ for some polynomial L of degree $n-m$. Since the n^2 points of intersection $\mathcal{C}_F\cap\mathcal{C}_G$ all lie on $\mathcal{C}$, they must be on one of the components; we know that mn of them lie on H, so the rest must lie on L. □

Corollary 2.4 (Pascal's Theorem)**.** *Let $ABCA'B'C'$ be a hexagon on an irreducible conic. Then the points of intersection $AB'\cap A'B$, $AC'\cap A'C$ and $BC'\cap B'C$ are collinear.*

Proof. The triples of lines AC', BA', CB' and AB', BC', CA' define two cubics. They intersect in 9 points, six of which lie on the irreducible conic. Thus the remaining three lie on a curve of degree $3-2=1$. □

3. Strong Bezout

Let us now indicate briefly how to attach multiplicities to points of intersection in such a way that Bezout's theorem predicts exactly mn such points.

As in the proof of weak Bezout, we assume that $[0:0:1]$ is a point not on $\mathcal{C}_F$ or $\mathcal{C}_G$, or on any of the lines connecting two points of intersection. Then there is a bijection between points P of intersection and linear factors of $R(X,Y)$. The multiplicity $I_P(F,G)$ is defined to be the multiplicity of this factor. Our proof of the weak theorem then immediately implies

Theorem 3.1. *Let $F = 0$ and $G = 0$ be plane projective algebraic curves of degrees m and n without common component defined over an algebraically closed field K. Then they intersect in exactly mn points, counting multiplicities:*

$$\sum_P I_P(F, G) = mn.$$

The problem with this definition is that we have to show it does not depend on the choice of coordinates. This is a major pain in the neck. Instead of going through these technicalities, it is better to rethink our foundations and look for definitions of multiplicities that are independent of coordinate systems. We will do this eventually.

Another problem is: does this definition agree with our previous definition of multiplicity of intersections between curves and lines? In order to verify this, we will use

Lemma 3.2. *The resultant of $f(X)$ and $g(X) = X - a$ is $R_{f,g} = f(a)$.*

Proof. This is a simple calculation. □

This implies that the resultant of $F(X, Y, Z)$ and the line $Z - aX - bY = 0$ equals $G(X, Y) = F(X, Y, aX + bY)$; we have defined multiplicities of points of intersections of curves and lines using G, and the new definition agrees with the old.

The Example Revisited. We looked at the intersection of the unit circle $X^2 + Y^2 - Z^2 = 0$ and the elliptic curve $X^3 - X^2Z - XZ^2 + Z^3 - Y^2Z = 0$. The point $[0 : 0 : 1]$ is not on these curves, but lies on the line connecting the points $[0 : 1 : 1]$ and $[0 : -1 : 1]$. Thus in order to compute multiplicities, we have to choose a different coordinate system. Replacing X by $X - Z$ and Y by $Y - Z$ we get the equations

$$F(X, Y, Z) = X^2 + Y^2 - 2(X + Y)Z + Z^2,$$
$$G(X, Y, Z) = X^3 - (4X^2 + Y^2)Z + 2(2X + Y)Z^2 - Z^3,$$

as well as

$$R(X, Y) = X^2Y(Y - 2X)(X - 2Y)^2.$$

Now there are the following possibilities:

(1) $Y = 0$: then $X^2 - 2XZ + Z^2 = (X - Z)^2 = 0$ and $X^3 - 4X^2Z + 4XZ^2 - Z^3 = 0$, hence the corresponding point of intersection is $[1 : 0 : 1]$, and it has multiplicity 1.
(2) $Y - 2X = 0$: then $0 = 5X^2 - 6XZ + Z^2 = (5X - Z)(X - Z)$ and $X^3 - 8X^2Z + 8XZ^2 - Z^3 = 0$, hence $X = Z$, and the corresponding point of intersection is $[1 : 2 : 1]$.
(3) $X = 0$: this leads to $(Y - Z)^2 = 0$ and $-Y^2Z + 2YZ^2 - Z^3 = 0$, hence to the point $[0 : 1 : 1]$ with multiplicity 2.
(4) $X - 2Y = 0$: then $(5Y - Z)(Y - Z) = 0$ and $8Y^3 - 17Y^2Z + 10YZ^2 - Z^3 = 0$, hence to the point $[2 : 1 : 1]$ with multiplicity 2.

In the old (affine) coordinate system, we have two points of intersection with multiplicity 1 (namely $(0, 1)$ and $(0, -1)$), and two points with multiplicity 2 (namely $(-1, 0)$ and $(1, 0)$).