These are old notes. There are some typos and proof is incomplete in 16. These are not corrected since I don't have the original TeX file and need to retype. The last section needs to expand further too.

Topics in Algebra

Notes **Part 1** **Avinash Sathaye**

1. **Curves** In general, we will assume that our ground field k is algebraically closed and characteristic zero. This assumption is not always needed and we may usually remark when it is not needed. A curve for us is usually given as a plane curve $f(x,y)=0$. The curve is irreducible if the polynomial $f(x,y)$ is irreducible and usually we want the polynomial not to have multiple factors, the curve is then said to be reduced.

   Another view of a curve is something parametrized by a single variable, so a curve might be given by a parametrization $x=p(t), y=q(t)$ where p,q are rational functions. It is a well known theorem (Lüroth's) which says that we can always find a rational function $h(t)$ such that $k(p,q)=k(h)$. Being parametrized by rational functions is, however, a very special property. Most curves are not parametrizable! Simplest example is $y^2=x+x^3$ which is doable from scratch!

2. **Parametrizations, places and valuations** Even though curves are not parametrizable, they have power series solutions centered at various points. A power series solution may be thought of as $x=u(t), y=v(t)$ with $u,v\in k((t))$. As before, we can take an optimal t by assuming that $k((u,v))=k((t))$. Moreover, we can take advantage of our assumptions on the field to normalize the parametrization as follows.

   First of all, we either have the $\operatorname{ord}_t(u), \operatorname{ord}_t(v)$ as both nonnegative or at least one of them might be negative. In the first case, we can find constants a,b so that $\operatorname{ord}_t(x-a)>0, \operatorname{ord}_t(y-b)>0$. We can then further normalize the parametrization by arranging $u=a+t^d, y=b+\eta(t)$ with positive integer d, using the special assumption on the field k. The series $\eta(t)\in k[[t]]$ is still not unique, but is well defined only up to a change $t\to \omega t$ where ω is a d-th root of unity. This parametrization $(t^d,\eta(t))$ leads to a "valuation" defined by $v(h(x,y))=\operatorname{ord}_t(h(t^d,\eta(t))$ defined for all rational functions $h(x,y)$ which are well defined under the substitution, this means that the numerator and the denominator of h don't become zero after the substitution. Note that the alternate forms $\eta(\omega t)$ do not cause a change in the valuation. The valuation has the usual properties $v(gh)=v(g)+v(h)$, if $v(g)\neq v(h)$, then $v(g+h)=\min(v(g),v(h))$ and if $v(g)=v(h)=m$, then $v(g+h)\geq m$. Moreover the nonzero elements $a\in k$ have $v(a)=0$. Since k is algebraically closed, it is easy to deduce that $v(g)=v(h)=m$ implies that there is a unique $c\in k$ such that $v(g+ch)>m$.

   We describe the above situation by saying that v is a valuation (or a branch) of the curve centered at the point (a,b). The parametrization (or its conjugacy class) $t^d,\eta(t)$ is said to define a place of the curve at the point (a,b). Note that the description of the place also depends on the choice of coordinates and we really think of the valuation as the main object, the place being a convenient way of evaluating the function v. the resulting valuations or places are said to be at finite distance.

   In case, the parametrization gives negative orders for either x or y, then we have to arrange things a bit differently. If x has negative order, then we can set $x=t^d$ where $d<0$. Then $y=\eta(t)$ as before, except $\eta(t)\in k((t))$ might be a meromorphic series. We again get the equivalence class of places and a valuation defined just as before, except now, even polynomials may have negative v-values. Such valuations are said to be valuations or branches at infinity.

   The power series $\eta(t)$ is also known as a Puiseux series in either case.

3. **Finding the places.** One way of finding the branches at a point concretely is thru Weierstrass Preparation Theorem and Newton's Theorem on places (Puiseux series). This goes as follows.

First assume that we wish to find the branches at finite distance and arrange the center to be the origin $(0,0)$, by a suitable translation. Let d be the order of the polynomial f, so we can write $f = f_d + f_{d+1} + \cdots$ where f_i denote homogeneous pieces of degree i and we assume that $f_d \neq 0$. Then by a further linear change of coordinates we can assume f_d to be monic in y and can write $f = a_0 y^d + a_1 y^{d-1} + \cdots a_d$ where $a_0 \in k[[x,y]]$ is a unit and $a_i \in k[[x]]$ with $\operatorname{ord}_x(a_i) \geq i$. The Weierstrass Preparation Theorem says that we can factor out some unit a so that $f = a(y^d + b_1 y^{d-1} + \cdots + b_d)$ where $b_i \in k[[x]]$ have order $\geq i$. The idea is to first factor out a_0 and rearrange the resulting expression in a similar form, then factor out the new coefficient of y^d and iterate! The process can be shown to converge; both formally and in the usual sense of convergence if you start with convergent coefficients.

Newton's Theorem then guarantees that the polynomial $F = y^d + b_1 y^{d-1} + \cdots + b_d$ can be factored as $\prod(y - \alpha_i(x)$ where α_i are fractional power series in x. The resulting branches are then obvious! Explicitly, write $F = \prod(F_i)$ where F_i is a monic polynomial of degree d_i in y. Then F_i gives a parametrization $x = t^{d_i}, y = \eta_i(t)$ such that

$$F_i = \prod_{\omega^{d_i}=1} (y - \eta(\omega t))$$

The resulting valuation is as described above and does not depend on the choice of the conjugate power series $\eta(\omega t)$.

For branches at infinity, the process is similar. We arrange f to be monic in y (for convenience) and factor in the algebraic closure of the formal power series ring in x^{-1}. So, the parametrization is now of the form $x = t^{-d_i}, y = \eta(\omega t) \in k((t))$.

4. **Intersection multiplicities.** Given any point $P(a,b)$ in the plane, we need to consider the so-called intersection multiplicity of f with another curve g. In case f has no multiple factors and no common factors with g, this is defined as

$$\langle f, g \rangle_P = \sum_v v(g) \text{ where the sum extends over all valuations of } f \text{ centered at } P$$

In case f has multiple factors, say $f = \prod f_i^{r_i}$ we define

$$\langle f, g \rangle_P = \sum r_i \langle f_i, g \rangle_P$$

It can be shown that the intersection multiplicity is equal to the length of the ideal generated by f, g in the local ring of the point P. and hence is symmetric in f, g.

We will extend this definition to various other natural situations thus:

$$\langle f, g \rangle_\infty = \sum v(g)$$

where the sum is extended over all the valuations of f at infinity. Suitable adjustment is made for the case of multiple factors and again we omit the case of common factors. This does not have a direct interpretation as a length, but it is still symmetric in f, g. (This becomes clear after the Bezout Theorem below.)

- a set of points S

$$\langle f, g \rangle_S = \sum_{P \in S} \langle f, g \rangle_P$$

In case the set S consists of all points in the plane, we use the word fin. In case the set S is the set of points on some curve h, we simply write h, instead of making a notation for the set of points of h.

5. **Bezout Theorem** A useful formula for calculating the intersection multiplicities is

$$\langle f, g\rangle_{fin} + \langle f, g\rangle_{\infty} = 0$$

Of course, we assume that f, g don't have common factors. In the trivial case when g is a nonzero constant modulo f, the formula has no nonzero terms and is obviously true.

This is equivalent to what is usually known as the Bezout Theorem for plane curves. We now outline a proof. First note that it is enough to prove it for the case when f is irreducible, since all terms are additive with respect to factors of f. Also, we can assume that g is not a constant modulo f.

Next, we note that we can assume without loss of generality that at each valuation of f at infinity, we have $w(g) < 0$ or in other words the coordinate ring $A = k[x, y]/(f)$ is integral over $k[\bar{g}]$ where $\bar{g}$ denotes the image of g modulo f. First note that by taking a general linear change of coordinates in x, y we can arrange f to be monic in y and hence x satisfies $w(x) < 0$ for all w. Now, for a large enough d, gx^d and x^d will both satisfy the desired condition and the desired formula for g can be proved by subtraction.

Also note that **under our assumption, we can write**

$$\langle f, g\rangle_{fin} = \sum v(g)$$

where the sum extends over all valuations v of f for which $v(g) \geq 0$.

Next we consider the simple case when $g = x$ and f is monic in y of degree n.

In this case, it is easy to verify from the definitions above that

$$\langle f, x + c\rangle_{x+c} = -\langle f, x\rangle_{\infty} = n$$

for every constant c. In fact, we use the symmetry for $\langle\ ,\ \rangle$ at finite points to calculate the term $\langle f, x + c\rangle_{x+c}$ by looking modulo the curve $x + c = 0$ and note that we simply get the sum of the multiplicities of the factors of $f(-c, y)$ which adds up to the degree n. It is easy to check that $-n = \langle f, x + c\rangle_{\infty}$ from the definition. Thus the case when $g = x + c$ is completely proved.

Now, for the general case, let L denote the function field of f (The quotient field of the coordinate ring A.) If $\bar{g}$ denotes the image of g modulo f, we can find some polynomial h such that $\bar{h} \in L$ satisfies $k(\bar{g}, \bar{h}) = L$ and the minimum polynomial of $\bar{h}$ over $k[\bar{g}]$ is a monic polynomial Φ of some degree m which naturally coincides with the field degree $[L : k(\bar{g})]$. Thus $\Phi(\bar{g}, \bar{h}) = 0$ is monic of degree m in $\bar{h}$. Note that the set of all valuations v of f for which $v(g) \geq 0$ coincides with the set of all valuations for Φ at finite distance and from what we have proved already, we get that

$$\langle f, g\rangle_{fin} = m = -\sum w(g)$$

where the sum extends over all valuations w of f for which $w(g) < 0$.

Thus we have proved the desired Theorem.

6. **Comments on the formula.** Assume that g is not a constant modulo any factors of f.

   Let us list the valuations of f at infinity as $v_1, \cdots, v_l, v_{l+1}, \cdots v_s$ where $v_i(g) \geq 0$ for $i \leq l$ and $v_i(g) < 0$ for $i > l$. It is clear that for most c, we get that $v_i(g+c) = 0$ for $i \leq l$. Indeed there are at most l values of c for which this fails (one for each v_i). Note that for $i > l$ we have that $v_i(g+c)$ is independent of c. For a general value of c, we have

$$\langle f, g+c \rangle_{fin} = -\sum_{i>l} v_i(g+c) = -\sum_{i>l} v_i(g) = m$$

   where this common number is the field degree $[L : k(\bar{g})]$ in case f is irreducible. In the general case it can be interpreted as the sum of the field degrees modulo various factors of f with proper adjustment for multiple factors.

   The formula can be interpreted as saying that the sum of zeros of a function $g + c$ modulo the curve f coincides with the sum of its poles, either one being the covering degree of the function field over $k(\bar{g})$.

   Since, we have avoided the projective space altogether, the usual interpretation of Bezout Theorem concerning plane projective curves is not given and is left as an exercise! One interpretation is given below, however!

7. **A Bertini Theorem** Now assume that f has no multiple factors. Let us define $N(f, g)$ to be the number of common points of the curves f, g at finite distance. We also assume that g is not a constant modulo any factors of f. We show that for all but finitely many c, the number $N(f, g+c)$, coincides with the quantity $m = [L : k(\bar{g})]$ discussed above.

   As before, we may assume that f is irreducible, since the theorem can be verified separately for each component.

   Further, as before, we first prove the case when $g = x$. In this case, our claim reduces to proving that the polynomial $f(-c, y)$ has distinct roots for all but finitely many c. This is evidently true since a necessary condition for $f(-c, y)$ to have a multiple root is that the discriminant of f with respect to y is divisible by $(x + c)$. Since the discriminant cannot be identically zero, there are only finitely many such factors. For a general g, we can either go to the curve Φ as shown above, or we can show that a necessary condition for $\langle f, g+c \rangle_P$ to be greater than 1 is that the Jacobian $J(f, g)$ must have a common point with f at P. Now, the $J(f, g)$ can have only finitely many points in common with f, since otherwise it will be identically zero modulo f. This will prove that $\langle f, g+c \rangle_P = 1$ for all but finitely many c.

   To see that $J(f, g)$ cannot be identically zero modulo f, we could proceed thus. Consider an irreducible relation between x, g modulo f given by $\psi(x, g) = 0$.. By the usual differential calculus, it is easy to see that the image of $J(f, g)$ modulo f is equal to the image of $-f_y \frac{\psi_x}{\psi_g}$ and this cannot be identically zero modulo f. The details are left to the reader.

8. **Differentials** The above also leads to a quick definition to what is termed as a canonical divisor, or the divisor of a differential. Differentials are formal objects which behave like derivatives with respect to an unspecified variable. For an irreducible f and polynomials g_1, g_2 which are non constant modulo f, we get a relation Ψ satisfied by the images $\bar{g_1}, \bar{g_2}$ modulo f. Then the differentials are connected by the relation

$$\Psi_{g_1} d\bar{g_1} + \Psi_{g_2} d\bar{g_2} = 0$$

where all expressions are considered modulo f. If g is a constant modulo f, then we may declare $d\bar{g} = 0$. **We may often write dg for $d\bar{g}$, if we are clearly working on some curve f.**

Given any valuation v of f and a non constant function g, we can define $v(dg) = \text{ord}_t(\frac{dG}{dt})$ where $G = g(t^d, \eta(t))$ is the result of substitution of the parametrization defining the valuation. If u is some function with $v(u) = 1$, then it is easy to show that $v(dg) = v(\frac{dg}{du})$ where $\frac{dg}{du}$ is computed as explained above. We simply have to use the fact that the parametrization defining the valuation makes our original function f vanish.

Let us now observe a simple but useful fact: If $v(g) \neq 0$ then $v(g) = v(dg) + 1$. If $v(g) = 0$, then there a unique c depending on v and g such that $v(g + c) > 0$ and for such a c, we get that $v(g + c) = v(dg) + 1$. Many mistakes are caused by ignoring this situation of $v(g) = 0$.

Let us also derive a useful formula for $\langle f, g \rangle_P$ at a finite common point of f, g.

$$\langle f, g \rangle_P = \sum_{v \to P} v(g) = \sum_{v \to P} (v(dg) + 1) = \langle f, dg \rangle_P + \nu(f, P)$$

where the term $\langle f, dg \rangle_P$ is an obvious generalization of our notation and $\nu(f, P)$ denotes the number of branches of f centered at P. We will also extend the notation to $\nu(f, \infty)$ to mean the number of branches of f at infinity.

All of this runs into technical problems when the characteristic is positive or the field is not algebraically closed! The reader should look up the details elsewhere!

9. **Divisors** We can now define formal divisors on an irreducible curve f. A divisor D is a formal sum of valuations $D = \sum_v v(D)v$ where the coefficient $v(D)$ is an integer which is zero except for finitely many valuations v. We take all valuations of the function field of f for this consideration. If f is reducible, but without multiple components, then the divisors can be considered separately on each irreducible component.

   The degree of a divisor D is defined to be $\deg(D) = \sum_v v(D)$.

   For convenience assume that f is irreducible and g is non constant modulo f. As above, assume that the valuations of f at infinity are listed as $v_1, \cdots, v_l, \cdots, v_s$ where $v_i(g) \geq 0$ for $i \leq l$ and $v_i(g) < 0$ for $i > l$.

   We define the divisor of zeros as $< g >_+ = \sum_{v(g)>0} v(g)v$. Note that then the degree of the divisor $Z(g)$ coincides with

$$\sum_{v(g) \geq 0} v(g) = \langle f, g \rangle_{fin} + \sum_{i \leq l} v_i(g)$$

   The polar divisor of g is defined as $< g >_- = -\sum_{v(g)<0} v(g)v$.

   The Bezout Theorem now can be interpreted that the degree $\deg(< (g + c) >_+)$ is constant for all c and coincides with the degree of the polar divisor $\deg(< g >_-)$, which in turn gives the covering degree of the function field of f over $k(\bar{g})$.

   We define the divisor $< g >$ to be simply the full sum $< g >= \sum_v v(g)v$, for a non constant g. Then the Bezout Theorem also gives that the degree $\deg(< g >)$ is zero. The definition of $< g >$ can be easily extended to rational functions g with the same conclusions. We also extend it to nonzero constants by declaring $< c >$ to be simply $0 = \sum_v 0v$ or the divisor with all components zero. Naturally, its degree is also zero.

This is the promised interpretation of the number of zeros and poles. The calculations need careful modifications for reducible f, which are left to the reader.

A canonical divisor is defined to be any divisor of the form $< dg >= \sum_v v(dg)v$ for a non constant function g. Note that given two non constant functions g_1, g_2 the divisors, then we clearly get,

$$< dg_1 >=< dg_2 >< \frac{dg_1}{dg_2} >$$

In particular the degree of a canonical divisor does not depend on the choice of g and is one of the important natural constants associated with the curve.

Suppose f is as general as possible so that it may be reducible and possibly with multiple factors. We propose to define a canonical divisor for it by taking a g which is not a constant modulo any of its factors and by considering the divisor $\sum_v v(dg)$ which extends to all the valuations of the functions fields of various irreducible components of f and we think of it as a divisor on the reduced curve f^* associated with f.

The geometric genus $P_g(f)$ of f is then defined to be the same as that for its reduced curve f^* and defined by stipulating that the degree of a canonical divisor is equal to $2P_g(f) - 2$.

Unlike the case of irreducible curves, $P_g(f)$ may come out negative and is introduced mostly as a notational convenience.

10. **A calculation.** Let us continue with the above notation and further assume that f has no multiple factors. Set $N(f, g)$ to be the number of common points of f, g at finite distance. Note that we know:

$$m = \sum_{f \bigcap g+c} v(g+c) + \sum_{1 \le i \le l} v_i(g+c)$$

for any c. Here the first sum is over the common points of f and $g + c$.

Thus we can write

$$m - N(f, g+c) = \sum_{P \text{ on } f \bigcap g+c} (\langle f, g+c \rangle_P - 1) + \sum_{v_i(g+c)>0} v_i(g+c)$$

Also note that by the Bertini Theorem above, for a general value of c, $m = N(f, g + c)$.

Using the calculations with differentials above, we get

$$m - N(f, g+c) = \sum_{P \text{ on } f \bigcap g+c} (\langle f, dg \rangle_P + \nu(f, P) - 1) + \sum_{v_i(g+c)>0} v_i(g+c)$$

Note that for any fixed i appearing in the last summation, there is a unique constant c_i for which $v_i(g + c_i) = v_i(dg) + 1$ and for all $c \neq c_i$ the term $v_i(g + c) = 0$ and hence does not appear in the formula.

Adding up all the expressions for various c, we get:

$$\sum_c (m - N(f, g+c)) = \sum_{P \text{ on } f} (\langle f, dg \rangle_P + \nu(f, P) - 1) + \sum_{1 \le i \le l} (v_i(dg) + 1)$$

Latest revision on January 5, 1998

Or, simplifying:

$$\sum_c (m - N(f, g+c)) = \sum_{P \text{ on } f} (\langle f, dg\rangle_P + \nu(f,P) - 1) + l + \sum_{1 \le i \le l} v_i(dg)$$

Noting that $m = -\sum_{l<i\le s} v_i(g) = -\sum_{l<i\le s}(v_i(dg)+1)$, we get that $\sum_{l<i\le s} v_i(dg) = -(m+s-l)$. Combining this with the above expression we get

$$\sum_c (m - N(f, g+c)) = \sum_v v(dg) + \sum_{P \text{ on } f} (\nu(f,P) - 1) + l + m + s - l$$

where the first sum extends over **all the valuations** of f and hence gives the degree of a canonical divisor, as explained above. Note that s is the number of branches of f at infinity and hence we can conveniently define $\nu(f, \infty) = s$.

Thus we get the formula:

$$\sum_c (m - N(f, g+c)) = 2P_g(f) - 2 + m + s + \sum_{P \text{ on } f} (\nu(f,P) - 1)$$

or in a better arrangement:

$$\sum_c (m - N(f, g+c)) = 2P_g(f) + (m-1) + (\nu(f,\infty) - 1) + \sum_{P \text{ on } f} (\nu(f,P) - 1)$$

The quantity $1 - m + \sum_c (m - N(f, g+c))$ is now easily seen to be independent of our choice of g (subject to the condition about being non constant modulo all factors of f) and we define the invariant

$$r(f) = 2P_g(f) + (\nu(f,\infty) - 1) + \sum_{P \text{ on } f} (\nu(f,P) - 1)$$

It can be shown that this coincides with the Euler Characteristic of the plane minus the curve, but the point of the above discussion is to make a clear computable definition of this invariant.

If f is a curve with possible multiple components, then we define $r(f) = r(f^*)$, where f^* is obtained by keeping only one copy of each multiple factors. We will develop more explicit formulas for $r(f)$ next.

11. **A simple formula for $r(f)$ in the reduced case.** First consider the case when $f = f^*$, i.e. when f is reduced. Without loss of generality, we can arrange f to be monic in y of degree n. Taking $x = g$ and working as above, we note that $m = n$ and $l = 0$.

    Working as above, we see that

$$n - N(f, x+c) = \sum_{P \text{ on } f \bigcap x+c} (\langle f, x+c\rangle_P - 1)$$

    It is clear that at a point P given by $x = -c, y = -c'$ we have $\langle y + c', x + c\rangle_P = 1$ and thus we can write the term in the above summation as

$$\langle x + c, f\rangle_P - \langle x + c, y + c'\rangle_P = \langle x + c, f_y\rangle_P$$

    The last change is obtained by arguing thus: We know that modulo $x + c$, we have a unique valuation, say w at P. Also, modulo $x + c$, we have $f_y = \frac{df}{dy} = \frac{df}{d(y+c')}$. Moreover, since $w(f), w(y + c') > 0$, we get that

$$w(f_y) = w(\frac{df}{d(y+c')}) = w(df) - w(d(y+c')) = w(f) - 1 - (w(y+c') - 1) = w(f) - 1 = \langle x+c, f_y\rangle_P$$

Now we look modulo f_y and note that $f_x = \frac{df}{dx}$ modulo f_y. By an argument similar to the above, we get that:

$$\langle x + c, f_y \rangle_P = \langle f, f_y \rangle_P - \langle f_x, f_y \rangle_P$$

We remark that this calculation is more involved since f_y may have many valuations at P and may have multiple factors, so some contributions may have to be counted multiply. But the argument for each valuation is the same! This calculation will fail if f has multiple factors passing thru P, since then for some valuations w of f_y, the value $w(f), w(f_x)$ might be both undefined! This is where we use that f is reduced.

Now we get a new formula for $r(f)$ as follows:

$$r(f) = 1 - n + \sum_{P \text{ on } f} (\langle f, f_y \rangle_P - \langle f_x, f_y \rangle_P)$$

and this simplifies to

$$r(f) = 1 - n + \langle f, f_y \rangle_{fin} - \langle f_x, f_y \rangle_f$$

12. **A simple formula for $r(f)$ in the general case.** Now let f be general. Write $f = \prod f_i^{p_i}$ and let $u = \prod_i f_i^{p_i - 1}$. Set $g = \sum_i (p_i f_{iy} \prod_{j \neq i} f_j) = f_y/u$.

We wish to note two facts about g. One obvious thing is that modulo any factor of g, we have $f_y = 0$, so $f_x = \frac{df}{dx}$. Thus, at any common point P of $g, f, x + c$, we have

$$\langle x + c, g \rangle_P = \langle f, g \rangle_P - \langle f_x, g \rangle_P$$

The other fact is that at the point P as above,

$$\langle x + c, f^*{}_y \rangle_P = \langle x + c, g \rangle_P$$

For this, let the point be $(x + c, y + c')$ as above and look modulo $x + c$. Assume that modulo $(x+c)$, the polynomials f_i have the leading terms $c_i(y+c')^{\mu_i}$, when expanded in powers of $(y+c')$. Note that under our assumption, $\sum_i \mu_i > 0$.

Then the leading term of f^* modulo $(x + c)$ is $(\prod_i c_i)(y + c')^{\sum_i \mu_i}$ and the left hand side our our equation is then $(\sum_i \mu_i) - 1$. By a simple calculation, we see that the leading term of g modulo $(x + c)$ is

$$(\prod_i c_i)(\sum_i p_i \mu_i)(y + c')^{(\sum_i \mu_i) - 1}$$

This clearly leads to the same value for the right hand side and our formula is proved.

Thus we can use g in place f_y in the simple formula for $r(f)$ to give

$$r(f) = 1 - n + \deg_y(u) + \langle f, f_y/u \rangle_{fin} - \langle f_x, f_y/u \rangle_f$$

where as before, we assume that f is monic in y of some degree n.

13. **Variation of $r(f)$.**

We now come to the main part of the theory of $r(f)$ which describes the variation of $r(f + c)$ as c varies. We prove the main Zeuthen-Segre formula in our setup which states that for all but finitely many values of λ, we have:

$$r(f + \lambda) = \sum_c (r(f + \lambda) - r(f + c))$$

Moreover, the terms on the right hand side are always non negative as long as the general curve $f+\lambda$ is irreducible.

We first prove the equation and then the non negativity of the terms.

14. **The proof of the formula** First we set u to be the GCD of the partial derivatives of f_x, f_y where we assume f to be monic of degree n in y as usual. Set $g = f_y/u$. For every constant c, set u_c to be the products of the components of u which divide $f+c$ and note that $u = \prod_c u_c$ where the product is has non-unit terms for only finitely many c. From what we have seen above, we can deduce that

$$r(f+c) = 1 - n + \deg_y(u_c) + \langle f+c, g\rangle_{fin} - \langle f_x, g\rangle_{f+c}$$

We only need to note that every p-fold factor of $f+c$ divides u_c exactly $p-1$ times, when $p>0$ and that factors of u which are not factors of u_c do not meet the curve $f+c$, so the last two terms of the above formula satisfy

$$\langle f+c, g\rangle_{fin} - \langle f_x, g\rangle_{f+c} = \langle f+c, f_y/u_c\rangle_{fin} - \langle f_x, f_y/u_c\rangle_{f+c}$$

Now we wish to study the variation

$$\sum_c (r(f+\lambda) - r(f+c))$$

as c varies and λ stands for a general value.

Let us consider the valuations of g at infinity arranged so that the first $v_1, \cdots, v_l$ are such that $v_i(f) \geq 0$ and the last $v_{l+1} \cdots, v_s$ are such that $v_i(f) < 0$. Here we have to take the valuations for all components of g and in case g has multiple factors, we need to repeat them with appropriate multiplicity. Under our assumption that the general $f+\lambda$ is irreducible, we may be able to prearrange that g is irreducible by using another Bertini theorem. We are, however, avoiding a detailed discussion of this, since it is not essential for the proof.

Note that for $1 \leq i \leq l$ there is a unique c_i so that $0 < v_i(f+c_i) = v_i(df) + 1$ and $v_i(f+c) = 0$ if $c \neq c_i$. For $l < i \leq s$, we have $v_i(f+c) < 0$ does not depend on the choice of c.

It is clear that the variation of the first part $1 - n + \deg_y(u_c)$ of our formula for $r(f+c)$ will give $\sum_c -\deg_y(u_c) = -\deg_y(u)$.

For a general value λ, we clearly get

$$r(f+\lambda) = 1 - n + \langle f+\lambda, g\rangle_{fin} = 1 - n - \sum_{l<i\leq s} v_i(f)$$

Note that we are claiming that the last term $\langle f_x, g\rangle_{f+\lambda}$ is zero for a general λ. To see this, note that f_x, g don't have any common factors by assumption and hence meet in finitely many points in the plane. Clearly, these points do not lie on $f+\lambda = 0$ for a general λ. In fact, we get that $f+\lambda$ is nonsingular for a general λ. This is yet another Bertini Theorem!

Thus, we easily see that the variation of the second part

$$\langle f+c, g\rangle_{fin} - \langle f_x, g\rangle_{f+c}$$

is equal to

$$\sum_{1\le i\le l} v_i(f+c_i) - \sum_i v_i(f_x)$$

As before, for for $1 \le i \le l$, we get $v_i(f+c_i) - v_i(f_x) = v_i(x)$. So the total variation of the second part comes out to be

$$\sum_i v_i(x) - \sum_{l<i\le s} v_i(f) = -\deg_y(g) + n - 1 + r(f+\lambda)$$

Since $\deg_y(g) = \deg_y(f_y) - \deg_y(u) = n - 1 - \deg_y(u)$, this reduces to: $r(f+\lambda) - \deg_y(u)$.

Combining this with the variation of the first part of the formula, we get the desired formula that the total variation coincides with the general value $r(f+\lambda)$.

15. **Positivity of individual terms** Now we come to the proof that each term of the variation is nonnegative in case $f+\lambda$ is irreducible for a general λ is irreducible for a general λ. In fact, we have a stronger result, which says that the only situation when the general value $r(f+lambda)$ is smaller than a special value $r(f+c)$ occurs when the original polynomial f can be expressed as a polynomial $\phi(f_1)$ where ϕ has degree bigger than 1 and f_1 is a "line". This last condition is well know to mean that f_1 is a nonsingular polynomial curve or equivalently can be transformed to just y after an automorphism of the ring $k[x,y]$.

First note that in case $u = 1$, we have nothing to prove. Clearly,

$$r(f+\lambda) - r(f+c) = \sum_{1\le i\le l} (v_i(f+c)) + \langle f_x, f_y\rangle_{f+c}$$

and each term is clearly non negative.

In fact, the same argument can be made when $f+c$ has no multiple factors since then $(f+c)^* = f+c$. Thus, we are only concerned with the situation when $f+c$ has multiple factors. In this case, the term of interest is:

$$r(f+\lambda) - r(f+c) = -\deg_y(u_c) + \sum_{v_i(f+c)>0} (v_i(f+c)) + \langle f_x, g\rangle_{f+c}$$

As observed earlier, we can ignore the contributions from the factors $u_{c'}$ with $c' \neq c$ and hence **without loss of generality, we may change $g = f_y/u$ to f_y/u_c. We will assume that we have done this and adjust all our notations accordingly.**

Note that u_c is a factor of f_x as well as $f+c$ and hence the last term is at least as big as $\langle u_c, g\rangle_{f+c}$.

By recycling our old notation, we write $f+c = \prod_i f_i^{p_i}$. The expression which needs to be non negative is bigger than or equal to the sum of contributions from each f_i given as:

$$\sum_{v_j(f_i)>0} (v_j(f_i)) + \langle f_i, g\rangle_{fin} - \deg_y(f_i)$$

and clearly it is enough to show that each of these is non negative.

Now let d_i be the degree in y of f_i and let us note that

$$\langle f_i, g\rangle_{fin} = \langle f_i, f_{iy}\rangle_{f_i} + \sum_{w\neq i} \langle f_i, f_w\rangle_{f_i}$$

which is evident from the formula

$$g = \sum_i p_i f_{iy} \prod_{w \neq i} f_w$$

Also note that each f_i is an irreducible monic polynomial in y of degree d_i and so we have

$$r(f_i) \geq 0 \text{ and } r(f_i) = 1 - d_i + \langle f_i, f_{iy} \rangle_{f_i} - \langle f_{ix}, f_{iy} \rangle_{f_i}$$

For the first part, note that we can use the formula

$$r(f_i) = 2P_g(f_i) + \nu(f_i, \infty) - 1 + \sum_{P \text{ on } f_i} (\nu(f_i, P) - 1)$$

which is clearly a sum of non negative terms.

Thus we see that $\langle f_i, f_{iy} \rangle_{f_i} \geq d_i - 1$ and we get equality in the case when f_i s nonsingular and has $r(f_i) = 0$. This, in turn, means that f_i is a nonsingular curve with one place at infinity and genus zero. This is also seen to be a polynomial nonsingular curve. Such curves are well known to be isomorphic to "lines" and the famous Abhyankar-Moh-Suzuki theorem says that by an automorphism, the f_i can be made to be just y.

In our case, we fail to get the desired result only when we have the following:

- some f_i is a polynomial nonsingular curve.
- f_i does not meet f_w for $w \neq i$ and so $f_w = c_w + f_i h_w$ for some polynomial h_w and constant $c_w \neq 0$.

  Note that at least one of the h_w **is not** a polynomial in f_i, for otherwise, the whole $f + c$ and hence any general $f + \lambda$ will become a polynomial in f_i of degree greater than one. This contradicts the assumption on f not being a polynomial expression of degree at least 2 in a line f_1.
- $v_j(f_i) = 0$ for every valuation v_j with $1 \leq j \leq l$. This in turn means that $\langle f_i + c, g \rangle_{fin}$ is independent of c.

We will show that the last condition above cannot hold, if the earlier ones are valid. Thus proving the result!

16. **A calculation with lines.** Our notation has gotten very complicated by now. So, let us isolate the situation and set up new notation to state and prove our result.

Let $F = \prod_i f_i^{p_i}$ be a polynomial and let f_1 be a line (or nonsingular polynomial curve). We are considering a finite product but have left out the number of factors to conserve symbols. Let $d_1 = \deg_y(f_1)$.

By using the AMS theorem, we will assume that there is some polynomial G such that $k[f_1, G] = k[x, y]$. Let us further expand the polynomials f_i for $i > 2$ as polynomials in f_1, G and assume that

$$f_i = G^{r_i} \phi_i(f_1) + \cdots$$

where f_1 degree of ϕ_i is $\delta_i \geq 1$ and the remaining terms have smaller G-degree. To simplify calculations, we will arrange $\phi_i(f_1)$ to be monic in f_1. Note that $r_1 = 0, \delta_1 = 1$. We further assume that $\sum_i r_i > 0$, so that F does not reduce to a polynomial in f_1 alone. Also, we note that each non constant term in the expression of each of the f_i is divisible by f_1.

Our polynomial g above can now be recast as:

$$\Psi = (\prod_i f_i)(\sum_i \frac{p_i f_{iy}}{f_i})$$

This translates the setup in the previous section except for the last condition on the constancy of $\langle f_i + c, g \rangle_{fin}$. It translates thus:

Let us define the degree function θ_c by setting $\theta_c(H(f_1, G)) = \deg_G(H(c, G))$. Now the last condition above means that:

$$\theta_c(\Psi) \text{ is a constant equal to } \theta_c(f_{1y}) \text{ or } \deg_y(f_1) - 1 = d_1 - 1$$

We will deduce a contradiction!

We need to determine $\theta_c(\Psi)$ and in turn need to estimate $\theta_c((f_1^a G^b)_y$.

Note that $(f_1^a G^b)_y = a f_1^{a-1} G^b f_{1y} + b f_1^a G^{b-1} G_y$. We will show that $\theta_c(G_y) < \theta_c(f_y)$ for all c and hence we get that the highest G-degree terms of $(f_1^a G^b)_y$ come from $a f_1^{a-1} G^b f_{1y}$ whenever $a > 0$.

Write the derivative $f_{1y} = \alpha G^{d_1 - 1} +$ smaller terms where we are assuming the known fact that the $\theta_c(f_y) = d_1 - 1$ is independent of c.

Calculation of Ψ leads to the term with the highest possible degree:

$$\sum f_1^{(\sum \delta_i) - 1)} G^{\sum r_i} (p_i \delta_i)(\alpha G^{d_1 - 1})$$

and since this is clearly nonzero, we get that

$$\theta_c(\Psi) = d_1 - 1 + \sum r_i$$

This is a contradiction, since $\sum_i r_i > 0$.

Thus, it remains to prove the calculation of θ_c. Indeed, the heart of Abhyankar-Moh theory can be described as a calculation of such functions in a more general setting. We describe this next.

17. **Calculation with general curves with one place at infinity.** Now let f be an irreducible curve with one place at infinity. It is easy to see that highest degree coefficients of f in any coordinate system are nonzero constants (in other words f is essentially monic in all variables).

Thus we can assume that $f = y^n + a_1 y^{n-1} + \cdots a_n$ where $a_i \in k[x]$.

The one place at infinity gives rise to a valuation v and this is the only valuation of the function field of f which has negative values for at least some polynomials. In fact, every polynomial which is non constant modulo f has to have a negative value at v, since that is the only possible pole for it!

We sometimes prefer to consider the negative of the v-function and define it as $\delta(h) = -v(h)$ which is defined for all rational functions h for which f does not divide the numerator or the denominator (in the reduced form of h). Let D denote the derivative $\frac{\partial}{\partial y}$.

The Abhyankar-Moh theory of curves with one place at infinity produces a very special basis for the polynomial ring $k[x, y]$. Briefly, there is a set S of $n + 1$ polynomials $1 = f_0, f_1, \cdots, f_{n-1}$ such that f_i are monic of degree i in y and all the polynomials $\{x^j f_i\}$ have distinct v-values.

17.··· The set of all polynomials can now be expanded uniquely as polynomials in f with coefficients as combinations of $\{x^j f_i\}$.

Thus for a polynomial $h(x,y)$, the value $v(h)$ can be determined by simply expanding it in this way and picking up the least value element among the terms not divisible by f.

Moreover, the translates of f have all one place at infinity with the same set S giving the same values in the corresponding valuations.

Finally, for each of the terms f_i, we have explicit formulas for $v(f_i) - v(D(f_i))$ and we can deduce the value of $v(D(h))$ using this information.

In case of a line, the function θ_c mentioned above is the same as the negative of the valuation at infinity for a translate by c and we get the desired result.

So, now we proceed to the:

18. **The detailed formulas of the Abhyankar-Moh Theory.** We are interested in a plane curve $f \in k[X,Y]$ where k is as usual an algebraically closed field of characteristic zero. The theory works in more general cases when one of the degrees is not divisible by the characteristic, but we will not worry about that.

We need to make a distinction between polynomials and their images modulo f, so we have switched to capital letters to denote our polynomial variables. Let $\phi_f : k[X,Y] \longrightarrow k[x,y] = A$ be the canonical map where we go modulo f and we will denote the image $\phi_f(h(X,Y))$ simply by $h(x,y)$ whenever convenient. The ring A is, of course, the coordinate ring of the plane curve f.

The main assumption is that the curve f is assumed **to have one place at infinity**. This implies that the curve is already irreducible and so A is an integral domain. We let K denote the quotient field of A and let V denote the unique valuation ring of K/k which does not contain A. Let v denote the corresponding valuation and let δ denote the "degree" function defined by $\delta(h) = -v(h)$. Let $\Gamma(f) = \Gamma(A)$ denote the degree-semigroup $\{\delta(h(x,y)) | 0 \neq h(x,y) \in A\}$.

Also, it is easy to see that the polynomial $f(X,Y)$ is essentially monic in X, Y, meaning that the coefficient is a nonzero constant. We may use Abhyankar's "nonzero" θ to denote any unspecified nonzero constant, which may be used several times in an expression with the understanding that it may represent different values in different locations.

Clearly, essentially monic polynomials can be actually made monic either by a convenient change of variables or simply by changing f itself by a constant multiple.

Thus let

$$f(X,Y) = Y^n + f_1(X)Y^{n-1} + \cdots + f_n(X) \in k[X][Y].$$

We assume that $n > 0$, since otherwise, f is just a linear expression in X and needs no explanation!

By Newton's theorem, there is a Puiseux series expansion $x = t^{-n}, y = \eta(t) \in k((t))$ such that

$$f(t^{-n}, Y) = \prod_{\omega^n = 1} (Y - \eta(\omega t)).$$

As already noted, the power series $\eta(t)$ is unique up to the change $t \to \omega t$. Let $S = \text{Supp}\,(\eta(t))$ denote the support of the power series, meaning the set of exponents with nonzero coefficients. Clearly S does not depend on the choice of $\eta(t)$.

Inductively, define a sequence of h characteristic pairs as follows.

- If $n = 1$, then we set $d_1 = 1, h = 0$ and stop.
- Let $q_1 = \min(S)$ and $d_1 = n$. Let $d_2 = \text{GCD}\,(q_1, d_1)$. Note that q_1 is also equal to $-m = -\deg_X(f(X,Y))$. For convenience, as well as historical reasons, let $m_1 = -m = q_1$.
  If $d_2 = 1$, then set $h = 1$ and stop!
- If $d_2 > 1$, then let

$$m_2 = \min\{s \in S | s \text{ is not divisible by } d_2\} \text{ and let } q_2 = m_2 - m_1.$$

  Set $d_3 = \text{GCD}\,(q_2, d_2)$. If $d_3 = 1$, then set $h = 2$ and stop.
- Having defined sequences q_i, d_i, m_i from $i = 1 \cdots, l-1$, as well as d_l define:

$$m_l = \min\{s \in S | s \text{ is not divisible by } d_l\} \text{ and let } q_l = m_l - m_{l-1}.$$

  Set $d_{l+1} = \text{GCD}\,(q_l, d_l)$. If $d_{l+1} = 1$, then set $h = l$ and stop.
- Finally, for convenience we display the expression $\eta(t)$ to highlight the characteristic terms as:

$$\eta(t) = \alpha_1 t^{m_1} + \cdots + \alpha_2 t^{m_2} + \cdots + alpha_h t^{m_h} + \cdots.$$

By Newton's Theorem, the GCD of n and members of S must be 1, for otherwise $f(X,Y)$ acquires multiple factors! Thus the process terminates in a finite number h of steps and we say that the curve f has h characteristic pairs at infinity.

Now we define several auxiliary expressions based on the above sequences which will be used in calculations.

For $i = 1 \cdots h$ define:

$$s_i = \sum_{j=1}^{i} q_j d_j \ , \ r_i = s_i/d_i \ , \ \delta_i = -r_i \ , \ n_i = \frac{d_i}{d_{i+1}}.$$

We also define $r_0 = -n, \delta_0 = n$.

19. Polynomials associated with a one place curve.

Note that the subring $B = k[x]$ of A is isomorphic to a polynomial ring and that A is a free module of rank n over B with a basis $1, y, \cdots, y^{n-1}$. We wish to change this to another free basis with the property that its values (and hence the degrees) give distinct residues modulo n. Such a basis has the advantage that the value of any polynomial $u(x,y)$ can be simply read off from the lowest value term in its expansion and leads to explicit formulas for the degree-semigroup $\Gamma(f)$.

What we will construct is a sequence of polynomials $G_0, G_1, \cdots, G_h$ in $k[X,Y]$ and let $g_0, g_1, \cdots, g_h$ be their images modulo f.

Let

$$\Omega = \{g^u = g_0^{u_0} g_1^{u_1} \cdots g_h^{u_h} | 0 \leq u_0 \in \mathbb{Z} \text{ and } u_i \in \mathbb{Z} \text{ with } 0 \leq u_i < n_i \text{ for } 1 \leq i \leq h\}$$

It is an easy exercise to check that the monomials g^u in Ω with $u_0 = 0$ are n in number and will give the desired basis.

The polynomials G_i are constructed as follows.

Let $G_0 = X$, $G_1 = Y$. For each i from 2 to h define G_i to be the approximate d_i-th root of f. By this we mean that G_i is the unique polynomial of Y-degree n/d_i satisfying the condition that $\deg_Y(f - G_i^{n/d_i}) < n - n/d_i$.

The existence of such an approximate root is proved by iterating the so-called Tschirnhausen operations and is explained thus:

Let d be any factor of n and start by setting $G = Y^{n/d}$. Write the G-adic expansion of f as $f = G^d + a_1 G^{d-1} + \cdots a_d$ where a_i are polynomials with Y-degrees less than n/d. If $a_1 = 0$ then we have the desired approximate d-th root. Otherwise replace G by $G + a_1/d$. It is easy to show that the Y-degree of a_1 steadily decreases until a_1 actually becomes 0 (and then the polynomial G stabilizes!).

The major point of the Abhyankar-Moh theory is the following set of results:

(a) The polynomials $G_1, \cdots, G_h$ are themselves polynomials with one place at infinity having respectively $0, \cdots, h-1$ places at infinity.
Moreover, if we let $\eta_i(t)$ be the corresponding Puiseux series for G_i with $i \geq 1$, then we can compare them to $\eta(t)$ as:

$$\eta(t) = \eta_i(t^{d_i}) + \theta t^{m_i} + \text{ higher terms.}$$

(b) $v(g_i) = r_i$ for $0 \leq i \leq h$. This is easily checked using the above comparison of branches.

(c) If we take distinct monomials g^u and $g^{u'}$ in Ω such that $u_i \neq u_i'$ for some $i > 0$, then $v(g^u)$ and $v(g^{u'})$ distinct modulo n. To see this, write:

$$v(g^u) - v(g^{u'}) = r_0(u_0 - u_0') + cdots + r_j(u_j - u_j')$$

where j is the last integer with $u_j \neq u_j'$. It is easy to check that all terms on the right hand side of the equation except the last are divisible by d_j while the absolute value of the last term satisfies:

$$0 < |r_j(u_j - u_j')| < |r_j n_j| = |r_j d_j / d_{j+1}|$$

since the GCD of r_j, d_j is easily seen to be the same as that of q_j, d_j or d_{j+1}, it follows that the last term is not divisible by d_j. It follows that the left hand side cannot be divisible by d_j and hence by $d_1 = n$.
This establishes the necessary basis.

(d) It can be shown that the r_i satisfy an additional property that $r_i n_i$ is in the semigroup generated by $r_0, \cdots, r_{i-1}$ and indeed can be written as

$$n_i r_i = r_0 u_0 + \cdots r_{i-1} u_{i-1}$$

where $0 \leq u_0$ and $0 \leq u_j < n_j$ for $1 \leq j \leq i-1$.

(e) Moreover the r_i satisfy inequalities $0 > r_{i+1} > n_i r_i$ for $i \geq 1$. The first part comes from the fact that r_i being a value of some non constant element(namely g_i in A), must be negative, for otherwise, it has no poles! The second part comes from an iterative construction of the g_i.

(f) In turn, I have shown elsewhere that a sequence of r_i having the above properties comes from a curve with one place at infinity. Explicitly, the conditions can be described thus:
Any sequence of negative integers $r_0, \cdots, r_h$ leads to a d-sequence defined by $d_1 = |r_0|$ and inductively, $d_{i+1} = \text{GCD}\,(r_i, d_i)$. Naturally, this gives $n_i = d_i/d_{i+1}$ for $i = 1, \cdot, h$.
Assume that:

- $d_{h+1} = 1$.
- $0 > r_{i+1} > n_i r_i$ for $1 \leq i \leq h-1$
- For $1 \leq i \leq h$, we have an expression of the form:

$$n_i r_i = r_0 u_0 + \cdots r_{i-1} u_{i-1}$$

where $0 \leq u_0$ and $0 \leq u_j < n_j$ for $1 \leq j \leq i-1$.

- **The main irreducibility lemma.** The most important lemma is a test for a curve to have one place at infinity or for a polynomial $f = Y^n + \cdots \in k[X][Y]$ to be irreducible as an element of $k((X^{-1}))[Y]$. The test is simply this.
Take some test function $\sigma(t)$ having h characteristic terms and find the order $\text{ord}_t(f(t^{-n}, \sigma(t))$ If this order is bigger than s_h as calculated from the characteristic terms, then f has one place at infinity.
The idea of the proof is this. Let $\sigma_1, \cdots, \sigma_n$ be the n-distinct conjugate expressions of σ obtained by replacing t by ωt, where $\omega^n = 1$. Build the polynomial

$$f^*(Y) = \prod_1^n (Y - \sigma_i) \in k((t^n))[Y] = k((X^{-1}))[Y]$$

It is not hard to see that every root of f^* substituted in f gives the same order (where we are substituting t^{-n} for X as well - or simply thinking of t as a suitable $X^{-1/n}$)
By multiplying all these $f(\sigma_i)$ we get an expression of the form $\prod(\sigma_i - p_j)$ whose order is bigger than ns_h. By collecting terms for a fixed root p_j, we find that one of them, say $p = p_j$ gives that $\prod(\sigma_i - p)$ has order bigger than s_h.
The order of the product can be explicitly evaluated in terms of the characteristic terms and it is easy to deduce that p must coincide thru m_h with at least one of the conjugates σ_i.
Nut now the root p of f clearly has n distinct conjugates (by looking at the part thru m_h) and so the polynomial $f(Y)$ of degree n must be irreducible!
- **Idea of the proof of the rest of the details.** Now the theory is built up by starting with $G_0 = X, G_1 = Y$ and building them one at a time while proving all the properties as we go along.
We illustrate how G_2 is built.
Start with a brand new indeterminate Z and a test function $X = t^{-n}, Y = \sigma = Zt^{m_1}$. From the known factorization $f(t^{-n}, Y) = \prod(Y - \eta(\omega t))$, we can determine the leading term of the substitution:

$$f(t^{-n}, \sigma) = \theta(Z^{n/d_2} - (\theta)^{n/d_2})^{d_2} t^{q_1 d_1} + \cdots \text{ higher terms}$$

If we collect the terms giving the leading form, they come out as P^{n/d_2} where P is some polynomial of degree n/d_2. It follows that if we replace Z by a root of the above leading form, then the resulting "test function" gives irreducibility for P. (It may be necessary to divide the test function by d_2 to match the above criterion, but the test comes out all right!)
Now we a potential G_2 and by using test functions matching with η between m_1 and m_2 we can improve it so that it matches η thru m_2. On the other hand, it is easy to check that all such test functions give leading forms of the type $\theta(Z - c)^{d_2}$ and by comparing the expansion with an approximate d_2-th root, it is possible to prove that the approximate d_2-th root has one place at infinity and expansion coinciding up to m_2. The above process is then repeated to construct G_3 etc.