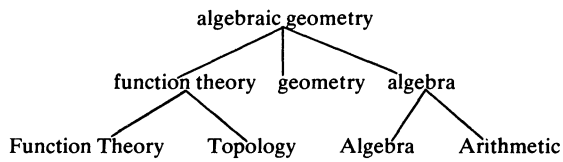
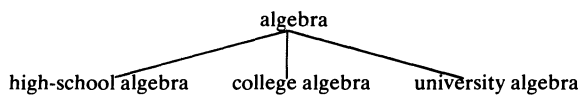
HISTORICAL RAMBLINGS IN ALGEBRAIC GEOMETRY AND RELATED ALGEBRA

SHREERAM S. ABHYANKAR

1. Various perspectives. Algebraic geometry has been approached from various perspectives:



The relevant algebra can be divided into 3 categories:



We associate different catchwords with these divisions:

function theory: function, integral;
geometry: curve, surface, variety;
high-school algebra: polynomial, power series;
college algebra: ring, field, ideal;
university algebra: functor.

We give a chronological tabulation of some of the distinguished proponents of the various divisions, together with one approximate date of work for each proponent; (for most of them, a publication — somewhat randomly chosen — is listed in the References; a large source list can be compiled by looking up the references in the References):

function theory: Euler (1748), Abel (1826), Jacobi (1832), RIEMANN (1857), Picard (1897), Poincaré (1910), Lefschetz (1921), Zariski (1929), Hodge (1941), Kodaira (1954), Hirzebruch (1956), Griffiths (1972).

geometry: Cremona (1850), M. NOETHER (1870), Bertini (1882), C. Segre (1894), Castelnuovo (1894), Enriques (1894), Zariski (1934).

high-school algebra: Bhaskara (1114), Cardano (1530), Ferrari (1540), NEWTON (1680), Tschirnhausen (1683), Euler (1748), Sylvester (1840), Cayley (1870), Kronecker (1882), Mertens (1886), König (1903), Perron (1905), Hurwitz (1913), Macaulay (1916), Zariski (1941), Hironaka (1964).

college algebra: DEDEKIND (1882), E. NOETHER (1925), Krull (1930), Zariski (1941), Chevalley (1943), Cohen (1946), Nagata (1960).

university algebra: Serre (1955), Cartan (1956), Eilenberg (1956), GROTHENDIECK (1960), Mumford (1965).

2. Fundamental thesis of this paper (obviously a partisan claim). The method of high-school algebra is powerful, beautiful, and accessible. So let us not be overwhelmed by the groups-rings-fields or the functorial arrows of the other two algebras and thereby lose sight of the power of the explicit algorithmic processes given to us by Newton, Tschirnhausen, Kronecker, and Sylvester.

Later on in the paper I shall relate four personal experiences to substantiate the above thesis.

Expanded version of an invited lecture at the November 30, 1974 Indianapolis meeting of the Indiana Section of the Mathematical Association of America. Extracts were also presented at the Workshop on the Evolution of Modern Mathematics held by the American Academy of Arts and Sciences at Boston on August 8–9, 1974. Excerpts of this paper appeared in the *Proceedings of the Workshop* in "Historia Mathematica," vol. 2, 1975, No. 4.

This work was supported by the National Science Foundation under Grant Number MPS–75–09090 at Purdue University.

Chapter I. Function Theory

3. Integral. After the invention of differential calculus by Newton and Leibnitz, at the end of the seventeenth century, attention was turned to the reverse process: integration. Rational functions of one variable (rational function = polynomial divided by polynomial) and then trigonometric functions were integrated. Then it was found that integrals of the form

$$\int [v(x)+w(x)\sqrt{(x-a)(x-b)(x-c)}]\,dx,$$

where $v(x)$ and $w(x)$ are rational functions and a, b, c are distinct constants, could not be integrated in terms of "known" functions. These integrals are called elliptic integrals since they first occurred in trying to find the arc length of an ellipse. Euler and others studied elliptic integrals. Then Abel initiated the study of general algebraic integrals of the form

$$\int r(x, y)\,dx,$$

where $r(x, y)$ is a rational function and where y is the n-valued function of x defined by

$$f(X, Y) = u_0(X)Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X) = 0$$

with polynomials $u_i(X)$ where $u_0(X) \neq 0$; it is assumed that f is irreducible, i.e., it cannot be factored into two polynomials. In honor of Abel, these integrals became known as abelian integrals.

The differential (= integrand = integral − integral sign) $r(x, y)\,dx$ is said to be of first kind if it has no pole. The integral $\int r(x, y)\,dx$ is said to be of first kind if the corresponding differential is so. It was shown that, for a fixed f, there are at most a finite number of linearly independent differentials of first kind; this number is called the genus of f and is denoted by g:

(g_1) g = number of linearly independent differentials of first kind.

It was also shown that (without restricting to first kind)

(g_2) $2g - 2$ = number of zeros of a differential − number of its poles.

That the right hand side depends only on f, and not on the particular differential, follows from the corresponding property which was proved about rational functions $s(x, y)$:

(1) number of zeros of a function = number of its poles.

Perhaps it may be preferable to regard (g_2) as the definition of the genus and (g_1) as a theorem. In fact, by and by, we shall write down several genus formulas and, depending on the mood, any one of them can be taken as the definition. At any rate, these formulas do provide a helpful thread running through various aspects of one-dimensional algebraic geometry.

Abel (1826, **[1]**) proved a far-reaching general theorem about sums of integrals of first kind. A little later, Jacobi (1832, **[39]**) inverted these integrals and thereby constructed periodic functions of g complex variables with $2g$ periods.

4. Riemann. The above matter needed clarification. Thus, already for elliptic integrals we have to know which of the two square-roots to start with; in the general case we have to know which "branch" of y to take. A picturesque clarification was provided by Riemann by constructing a surface S, called the Riemann surface of f, on which y "becomes" single-valued. The functions $r(x, y)$ and $s(x, y)$ and the differentials $r(x, y)\,dx$, together with the locations and orders of their zeros and poles, are now to be viewed as happenings on S, and so also the abelian integrals $\int r(x, y)\,dx$ are to be integrated along paths on S.

Using his surface, Riemann (1857, **[59]**) showed that

(g_3) $\left\{\begin{array}{l}\textit{number of linearly independent rational functions } s(x, y) \\ \qquad \textit{with prescribed } d \textit{ poles} = d + 1 - g + i,\end{array}\right.$

where i is a nonnegative integer which is zero if $d > 2g - 2$, and later Roch (1865, **[60]**) showed that

(*) $\begin{cases} i = \textit{number of linearly independent differentials of first kind} \\ \qquad\qquad \textit{having zeros at the prescribed poles.} \end{cases}$

Again, (g_3) may be regarded as giving a definition of the genus. In any case, the *Riemann-Roch theorem* $[(g_3)+(*)]$ shows that the resulting definition is equivalent to (g_1) as well as to (g_2).

Now any (closed) surface S^* is topologically equivalent to (i.e., by stretching and shrinking but without tearing, can be converted into) a sphere with a certain number of handles; say the number of handles is g^*. (See Fig. 1.)

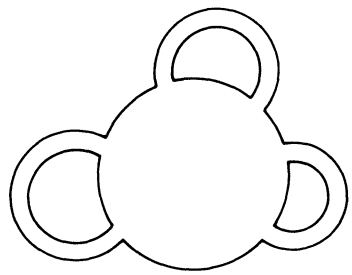
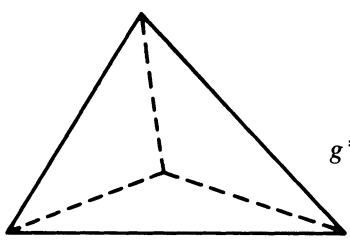


FIG. 1

Euler's theorem, which is one of the oldest theorems of topology, gives the following method of computing g^*: *On S^* draw any polyhedron Δ^* (all whose faces are curvilinear triangles); let Δ_0^*, Δ_1^*, and Δ_2^* be the number of vertices, edges, and faces of Δ^*; then*

$$\Delta_0^* - \Delta_1^* + \Delta_2^* = 2 - 2g^*.$$



$g^* = 0,\ \Delta_0^* = 4,\ \Delta_1^* = 6,\ \Delta_2^* = 4.$

FIG. 2

For verification for a tetrahedron see Fig. 2.

Using formula (g_2) and Euler's theorem, Riemann showed that for the Riemann surface S:

(g_4) $\qquad g$ = number of handles.

As a reference to the material described in these two sections, we may cite the excellent book of Stahl (1896, **[68]**).

Let us close our discussion with a philosophical remark:

Jacobi's approach seems to have been algorithmic: write down numerous explicit formulas! As a dividend he could deduce things such as: every integer is a sum of four squares and in exactly so many ways. Jacobi's approach received a big push in the hands of the arithmetician Kronecker.

On the other hand, Riemann's approach appears to have been existential: know that the formulas exist! Many of his existence arguments were based on "Dirichlet's principle," the genesis of which can be expressed by saying something like this: the real and imaginary parts of an analytic function of a

complex variable satisfy the same sort of partial differential equation as a gravitational potential or an electric potential (or a fluid flow or heat conduction...); so the existence of analytic functions or differentials with prescribed boundary behaviour or assigned singularities can be adduced from physical considerations. To carry the analogy further (or backwards), one could say that the mathematical formula "number of zeros of a function = number of its poles" corresponds to the physical fact that there is no gravitational pull inside a hollow shell, or that there is no electric intensity inside a hollow charged conductor. Riemann's approach was furthered by Hilbert (1904, **[34]**) who put the Dirichlet principle on firmer grounds.

5. Higher dimensions. A survey of the corresponding development in the hands of Picard, ..., Hirzebruch, can be found in the second edition of Zariski's book (1971, **[80]**). By extending and employing some of these developments, Clemens and Griffiths (1972, **[22]**) have proved the nonrationality of nonsingular hypersurfaces of degree 3 in 4-space, which was a problem of long standing.

Chapter II: Geometry

6. Spatial intuition. In the geometric viewpoint one studies the polynomial equation $f(X, Y) = 0$ by visualizing the (algebraic) plane curve consisting of all points (a, b) in the (X, Y)-plane satisfying $f(a, b) = 0$. To take care of "points at infinity" one introduces homogeneous coordinates and then the old point (a, b) is represented by all triples (ka, kb, k) with $k \neq 0$, and the plane curve is given by $F(X, Y, Z) = 0$ where F is the homogeneous polynomial (i.e., a polynomial all of whose terms have the same degree) obtained by "homogenizing" f; for example, homogenizing $Y^2 - X^3$ we get $Y^2Z - X^3$ and then dehomogenizing $Y^2Z - X^3$ (i.e., putting $Z = 1$) we get back $Y^2 - X^3$. The degree of f (or of F) is called the degree of the plane curve.

More generally, the totality of common solutions of a finite number of polynomial equations in r variables is visualized as a geometric configuration or "an (algebraic) variety" in r-dimensional space; again, to take care of "points at infinity" one may use homogeneous equations in $r + 1$ variables. A curve is a variety of "dimension" 1, a surface is a variety of "dimension" 2, ...

7. Max Noether. Noether deduced many properties of algebraic curves (including the geometric version of the Riemann-Roch theorem) as consequences of a theorem which acquired the designations "Noether's $AF + B\Phi$ Theorem" or "Noether's Fundamental Theorem." This is indeed one of the most proved theorems; already in Berzolari's 1906 encyclopedia article **[13]** one can find references to proofs by a dozen different authors; Noether's original proof appeared in (1870, **[54]**) and (1873, **[55]**).

Genesis. In elementary analytic geometry one may be given the following exercise: Let $F = 0$ and $\Phi = 0$ be two circles. Let it be required to find a third circle $H = 0$ which passes through the points of intersection of F and Φ, and which does something else. The pedantic way would be to first actually find the points of intersection — by solving the simultaneous equations $F = 0$ and $\Phi = 0$ — and then proceed to construct H. The clever way, given in good books of analytic geometry (e.g., Salmon (1852), **[62]**), would be thus. For any (constants) A and B, the conic

$$AF + B\Phi = 0$$

clearly passes through all the points of intersection of F and Φ. So it is reasonable to surmise that conversely any conic passing through the points of intersection of F and Φ can be so expressed.

Generalizing the above idea, let $F = 0$ and $\Phi = 0$ be plane curves of any degree. Then again, for any (homogeneous polynomials) A and B, the curve

$$AF + B\Phi = 0$$

clearly passes through all the points of intersection of F and Φ.

Noether's Fundamental Theorem says that the converse is true. That is, *if* $H = 0$ *is any plane curve such that "*H *passes through all the points of intersections of* F *and* Φ*" then*

$$H = AF + B\Phi \text{ for suitable } A \text{ and } B.$$

Of course, if, say, some point P is a double point of both F and Φ, then P must be a double (or higher) point of every plane curve of the form $AF + B\Phi = 0$.

In other words, Noether's theorem must be qualified by giving a "good" meaning to the phrase "H passes through ...." What meaning to give? ANSWER: That which will make Noether's theorem true.

Indeed the peculiar characteristic wisdom of geometric algebraic geometry is the

8. Dictum. Study of simple cases gives rise to a nice succinct statement. Take it as an axiom that the statement is true most generally. Make it true by the provision that we learn to "count properly" the intervening quantities. Or better still, have faith that god (or, if you prefer, nature) has a good meaning in mind, and march on!

9. Severi. After Max Noether, geometric algebraic geometry went to live in Italy which produced the great geometers: Bertini, C. Segre, Castelnuovo, Enriques, Severi. To quote from Zariski's address (1950, [**79**]): "The Italian geometers have erected ... a stupendous edifice: the theory of algebraic surfaces." This edifice was surveyed by Zariski (1934, [**80**]).

It is perhaps appropriate to elucidate the *Dictum* by quoting Severi: "Every good theorem must have a good counterexample." PARAPHRASE: don't be deterred if your formula is presently invalid in some cases; it only means that you have not yet completely deciphered god's mind.

At any rate, Severi's *Vorlesungen* (1921, [**66**]) is certainly one of the most readable books on algebraic geometry. Let me here venture to say that

10. Semple and Roth (1949, [**64**]) is perhaps the only book on algebraic geometry written after, say, 1940 which covers a lot of territory and is at the same time "readable" like an entertaining novel. However, it may not be the best introduction to the subject for the present-day student trained in contemporary rigor. May I, as a plea, suggest that someone please write a reasonably exhaustive book on algebraic geometry which is rigorous as well as "readable"!

11. Bryn Mawr. Before proceeding to give two more examples of the *Dictum*, let us note that a frequently cited proof of Noether's theorem was given by Charlotte Angas Scott of Bryn Mawr (1899, [**63**]). It is worth giving a long quote from her paper: "This (Noether's) theorem, discovered in the course of, and developed for the sake of, purely algebraic researches, is not however tabooed to the geometer ...; but it does not appear that any simple proof depending on geometrical conceptions has yet been given. Cayley (1887, [**18**]) regarded the theorem as intuitive for simple intersections. Zeuthen's proof (1888, [**81**]) depends on an elaborate determination of the number of conditions imposed by the intersections of two curves, when these are simple, the case of multiple intersections being then deduced by the somewhat dangerous process of proceeding to the limit."

12. Another example of the Dictum. Two lines meet in a point. A line and a conic meet in two points. Two conics meet in four points. So we are led to the very first theorem of algebraic geometry.

BEZOUT'S THEOREM. *Any two algebraic plane curves of degree* N *and* M*, devoid of common components, meet in exactly* MN *points (counted properly!).*

This was first asserted by Maclaurin (1720, [45]). It was discussed by Euler (1748, [26]) and Cramer (1750, [24]), and then more completely by Bezout (1770, [14]).

Algebraic Proof. See §21.

Geometric Proof. The curves can be degenerated into N lines and M lines. These clearly meet in MN points. Hence so do the original curves.

At any rate, we are led to the geometric characterization of the degree N of a plane curve C: $f(X, Y) = 0$ as the number of intersection of C with a line (either count properly or maximize over all lines). Furthermore: a point P of C is said to be a d-fold point (or a d-ple point, or a point of multiplicity d) if most lines through P meet C at P in d coincidental points, or alternatively, if they meet C, outside P, in $N - d$ points; the finite number of lines through P which meet C at P in more than d coincidental points are called the tangents to C at P; P is called a singular point (or a singularity) of C if $d > 1$; P is called a simple point of C if $d = 1$. C is said to be nonsingular if it has no singularities. Algebraically, the multiplicity d of a point $P = (a, b)$ of C at finite distance can be characterized as the smallest degree of a term occurring in the expansion of f around P; in other words

$$f(X + a, Y + b) = \zeta_d(X, Y) + \sum_{i+j>d} \lambda_{ij} X^i Y^j,$$

where $\zeta_d(X, Y)$ is a nonzero homogeneous polynomial of degree d; moreover, the d factors $\mu_i X + \nu_i Y$ of $\zeta_d(X, Y)$ give the tangents $\mu_i(X - a) + \nu_i(Y - b) = 0$ to C at P, which are not necessarily all distinct. So in particular: P is a singularity of C iff $f_Y(a, b) = 0 = f_X(a, b)$, where f_Y and f_X are the Y-derivative and X-derivative of f; moreover, if P is a simple point of C then: $f_Y(a, b) = 0$ iff the tangent to C at P is $X - a = 0$.

A similar situation prevails in higher dimensions.

The problem of resolution of singularities asks whether a given variety can be desingularized, i.e., can be transformed by an almost one-to-one (algebraic) transformation into a nonsingular variety.

For a plane curve $f(X, Y) = 0$, the Riemann surface of f is nothing but the desingularization of f. Geometrically (or algebraically) the same thing was achieved by Noether by means of "quadratic transformations." For example, the curve $Y^2 - X^3 = 0$ has a double point at the origin; the quadratic transformation (or substitution)

$$X = X' \text{ and } Y = X'Y'$$

yields

$$0 = Y^2 - X^3 = X'^2 Y'^2 - X'^3 = X'^2(Y'^2 - X')$$

and cancelling out the extraneous factor X'^2 we get the nonsingular curve $Y'^2 - X' = 0$. To desingularize the general plane curve $f = 0$ one has to make a finite succession of such transformations. Moreover, Noether also gave an "analysis" of the singularities of $f = 0$ in terms of these resolution steps. A brief exposition of related matter can be found in my Bloomington lecture [7].

13. Genus formula. As a final example of the *Dictum*, study of simple cases which we shall describe later, led the geometers from Maclaurin to Noether towards the following formula for the genus g of an irreducible plane curve C: $f(X, Y) = 0$ of degree N:

$$g = (1/2)(N - 1)(N - 2) - \text{number of double points counted properly}$$

i.e.,

$$(g_5) \qquad g = (1/2)(N - 1)(N - 2) - \Sigma\, \delta(P),$$

where the sum is over all the singular points P of C and where $\delta(P)$ is a certain positive integer which says that "P accounts for $\delta(P)$ ordinary double points."

CHAPTER III: HIGH-SCHOOL ALGEBRA

14. Brahmins. It might be said that (high-school) algebra started with the Brahmins (my ancestors) of India: Aryabhatta (476), Brahmagupta (598), Bhaskara (1114); quadratic equations were solved by completing the square; also rules were given for solving certain types of so-called diophantine equations. Bhaskaracharya (*acharya* = professor) was the director of an observatory at Ujjain; his treatise on astronomy (*Siddhantashiromani*) contains chapters on geometry (*Lilavati* — named after his daughter) and algebra (*Beejaganit*).

(See *Historical Notes* in Chrystal [**21**] and Krull [**44**].)

Now it is true that the history of mathematics should primarily consist of an account of the achievements of great men; but it may be of some interest to also see their impact on an average student like me. With this in mind I may be allowed some

15. Personal reminiscences. It is my fond memory that my father initiated me to mathematics — and at the same time to Sanskrit poetry — by teaching me portions of Bhaskara's treatise. After several years, during my last year in high-school, at my father's suggestion I studied Hobson's *Trigonometry* (1891, [**37**]), Hardy's *Pure Mathematics* (1908, [**31**]), and Chrystal's *Algebra* (1886, [**21**]). By the end of first-year college I had also studied Knopp's *Infinite Series* (1928, [**40**]), Burnside-Panton's *Theory of Equations* (1904, [**16**]), and Bôcher's *Higher Algebra* (1907, [**15**]). These books have served me well; they are all (except, perhaps, Hardy's) predominantly high-school-algebraic and algorithmic. After them my ways were set; after them I could not, (and didn't need to?), acquire much new technique. The goodly dose of analysis obtained during my last three years of college was all but forgotten.

As a graduate student at Harvard, I suppose I did acquire some mathematical sophistication. As a result, I always think high-school algebra but write college algebra. At any rate, Zariski's algebraic geometry gave ample scope to my early algorithmic training. However, what initially attracted me to Zariski was Math 103, Projective Geometry: von Staudt's algebra of throws and the resulting identity of geometry with algebra was indeed very exciting.

To continue with our story of high-school algebra:

16. Arabia and Italy. Then travelling through Arabia, algebra reached Europe around 1500. During 1500–1600, the Italians — Tartaglia, Cardano, Ferrari, et al — solved cubic and quartic equations.

Fermat's and Descartes' (1637) introduction of coordinates gave a big impetus to algebra.

(See Historical Notes in Chrystal [**21**] and Krull [**44**].)

Then there appeared the majestic genius of Newton (1680, [**53**]):

17. Newton. Truly the father of us all. Newton discovered the Binomial Theorem

$$(1+X)^n = \sum_i \frac{n(n-1)(n-2)\cdots(n-i+1)}{1\cdot 2\cdot 3\cdots i} X^i$$

first for positive integral n and then soon after for any fractional n. For fractional exponent this can be regarded as solving, for Y, the equation

$$Y^n - u(X) = 0, \tag{1}$$

where $u(X)$ is a power series (and where n is the denominator of the n of the Binomial Theorem).

What is a power series? Take the definition of a polynomial

$$a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$$

and from it delete the last term. This gives

$$a_0 + a_1X + a_2X^2 + \cdots$$

as the definition of a power series; likewise for several variables. So a power series is simpler than a polynomial. Like polynomials, power series can be added, subtracted, and multiplied.

Especially in case of one variable, we may even permit a finite number of negative exponents

$$u(X) = a_{-m}X^{-m} + a_{-m+1}X^{-m+1} + \cdots + a_0 + a_1X + a_2X^2 + \cdots$$

and call it a meromorphic series. Unlike a polynomial, a power series or a meromorphic series need not have a highest degree term; but it does have a lowest degree term; the degree of the lowest degree term in $u(X)$ is called the X-order of $u(X)$ and is denoted by $\mathrm{ord}_X u(X)$.

When we allow meromorphic series we can divide. Namely if $u(X) \neq 0$ and $q = \mathrm{ord}_X u(X)$, then

$$u(X) = a_q X^q[1 - v(X)] \text{ with } v(0) = 0;$$

hence, because of the geometric series

$$(1 - X)^{-1} = 1 + X + X^2 + \cdots$$

we have

$$u(X)^{-1} = a_q^{-1}X^{-q}[1 + v(X) + v(X)^2 + \cdots].$$

It is presumed that the "coefficients" a_i (or, in case of several variables, $a_{ij\cdots k}$) can be added, subtracted, multiplied, and divided. Thus they may be rational numbers, or real numbers, or complex numbers, or "integers modulo a prime number p," etc. In the case of rational numbers, or real numbers, or complex numbers, we are in zero (or infinite) characteristic, i.e., $1 + 1 + \cdots + 1$ is never zero. In the case of "integers modulo a prime number p" we are in the nonzero characteristic p, i.e.,

$$\underbrace{1 + 1 + \cdots + 1}_{p \text{ times}} \text{ equals } 0;$$

in this case the Binomial Theorem takes the amusing form: $(1 + X)^p = 1 + X^p$. Frequently, as in the case of complex numbers, every polynomial equation in one variable has a root; we shall usually (as in the theorem below) suppose this to be so.

Having solved the simpler equation (1), Newton proceeded to solve the most general equation of that type; to wit:

Newton's Theorem (on Puiseux expansion). Every polynomial

$$f(X, Y) = Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X)$$

with meromorphic series $u_i(X)$, where $n!$ is not divisible by the characteristic, can be completely factored if we allow series with fractional exponents. More precisely, we have

$$f(T^{n!}, Y) = \prod_{i=1}^{n} (Y - y_i(T))$$

with meromorphic series $y_i(T)$; the $y_i(T)$ are power series if all the $u_i(X)$ are. Moreover, if f is irreducible (when fractional exponents are not allowed), then $n!$ may be replaced by n, i.e., assuming that n is not divisible by the characteristic, we have

$$f(T^n, Y) = \prod_{i=1}^{n} (Y - y(\omega^i T)),$$

where $y(T)$ is a meromorphic series and ω is a primitive nth root of 1.

Not only did Newton prove the existence of the y_i, but he actually gave an algorithm for explicitly finding them by successive polynomial approximations. This algorithm is known as Newton's polygon (or, according to older literature, parallelogram). [The spirit of this algorithm is similar to Newton's iterative procedure of approximation of real roots of polynomials with constant (real) coefficients.]

To quote from the *Historical Note* on page 396 of Part II of Chrystal [**21**]: "... (Newton's) method

was well understood by Newton's followers, Stirling and Taylor; but seems to have been lost sight of ... after their time."

Newton's theorem was revived by Puiseux [**58**] in 1850. So it acquired the name "Puiseux expansion" which is a misnomer. What's more is that Puiseux's proof, being based upon Cauchy's integral theorems, applies only to convergent power series with complex coefficients. On the other hand, Newton's proof, being algorithmic, applies equally well to power series, whether they converge or not. Moreover, and that is the main point, Newton's algorithmic proof leads to numerous other existence theorems while Puiseux's existential proof does not do so.

Generally speaking, from Newton to Cauchy (1830), mathematicians used power series without regard to convergence. They were criticised for this and the matter was rectified by the analysts Cauchy and Abel who developed a rigorous theory of convergence. After another hundred years or so we were taught, say by Hensel [**32**], Krull [**43**] and Chevalley [**20**], that it really didn't matter, i.e., we may disregard convergence after all! So the algebraist was freed from the shackles of analysis, or rather (as in Vedanta philosophy) he was told that he always was free but had only forgotten it temporarily.

18. Tschirnhausen(1683, [**70**]). This contemporary of Newton tried to solve fifth degree equations. In that he failed. But his success was greater than his failure. He developed many transformations of equations.

If you pick up almost any book on algebra, written before (but none after) 1931, you will find numerous pages devoted to Tschirnhausen transformations. To cite some more meritorious books on algebra, in addition to those already cited in §15, we have: Weber (1894, [**72**]), König (1903, [**41**]), Macaulay (1916, [**46**]), and Perron (1927, [**57**]).

Then in 1931 came van der Waerden [**71**] drawing the line of demarcation between high-school algebra and college algebra; still, being on the boundary, van der Waerden does include some elimination theory.

As a simple sample of Tschirnhausen's transformations we have the trick of killing the coefficient of Y^{n-1} in

$$G(Y) = Y^n + b_1 Y^{n-1} + \cdots + b_n$$

by the substitution

$$Z = Y + \frac{b_1}{n}$$

thereby getting $G(Y) = Z^n + c_2 Z^{n-2} + \cdots + c_n$, because by the Binomial Theorem we have

$$(Y + X)^n = Y^n + nXY^{n-1} + \cdots.$$

For $n = 2$ this is simply the ancient method of solving a quadratic equation by completing the square.

Of course we can divide by n only if n is not divisible by the characteristic.

This killing of b_1 is indeed the basic fact underlying Zariski's [**76**], [**77**], [**78**] and Hironaka's [**35**] proofs of resolution of singularities of algebraic varieties of zero characteristic. For an analysis of this matter see my *Moscow lecture* [**6**], where I went on to say that "instead of killing b_1, Zariski used differentiation arguments; but then after all, the Binomial Theorem and differentiation are in essence one and the same thing."

In turn, the nonavailability of the above Tschirnhausen transformation is what makes the resolution problem quite different in nonzero characteristic. To put it differently, we have to grasp the divisibility properties of the binomial coefficients, and in other ways, to understand the Binomial Theorem better! In the *Moscow lecture* I also pointed out that as far as the resolution problem in nonzero characteristic is concerned, much critical information is lost by replacing a power series by the totality of all its multiples, i.e., by the "principal ideal generated by it"! This is an instance of the superiority of high-school algebra over college algebra and geometry, and so this brings us to:

19. Personal experience 1. In my Harvard dissertation (1956, [2]) I proved resolution of singularities of algebraic surfaces in nonzero characteristic. There I used a mixture of high-school and college algebra. After ten years, I understood the Binomial Theorem a little better and thereby learned how to replace some of the college algebra by high-school algebra; that enabled me to prove resolution for arithmetical surfaces (1965, [4]). Then replacing some more college algebra by high-school algebra enabled me to prove resolution for three-dimensional algebraic varieties in nonzero characteristic (1966, [5]). But still some college algebra has remained.

I am convinced that if one can decipher the mysteries of the Binomial Theorem and learn how to replace the remaining college algebra by high-school algebra, then one should be able to do the general resolution problem. Indeed, I could almost see a ray of light at the end of the tunnel. But this process of unlearning college algebra left me a bit exhausted; so I quit!

20. Euler (1748, [26]). Euler dealt with a mixture of algebra and analysis; the two fields started separating when Cauchy (1830) began to be concerned with questions of convergence. As an amusing evidence of Euler's algebraic technique let me quote his "proof" that

$$\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

Namely, if $p_1, p_2, \cdots$ are the roots of $a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m = 0$, then

$$\sum \frac{1}{p_i} = \frac{-a_1}{a_0}.$$

Now

$$\cos \sqrt{x} = 1 - \frac{x}{2} + \frac{x^2}{24} + \cdots$$

and the roots of $\cos \sqrt{x} = 0$ are

$$\frac{(2n+1)^2 \pi^2}{4} \text{ with } n = 0, 1, 2, \cdots$$

and hence

$$\sum \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

21. Elimination theory. This encompasses the explicit algorithmic procedures of solving several simultaneous polynomial equations in several variables. Here some of the prominent names are: Sylvester (1840, [69]), Kronecker (1882, [42]), Mertens (1886, [47]), König (1903, [41]), Hurwitz (1913, [38]), and Macaulay (1916, [46]). It is a vast theory. There used to be a belief, substantially justified, that elimination theory is capable of handling most problems of algebraic geometry in a rigorous and *constructive* manner. This is of course not surprising; after all, what is algebraic geometry but another name for systems of polynomial equations!

What is surprising is that under Bourbaki's influence it somehow became fashionable to bring elimination theory into disrepute. To quote from page 31 of Weil (1946, [74]): "The device that follows, which, it may be hoped, finally eliminates from algebraic geometry the last traces of elimination theory, is borrowed from C. Chevalley's *Princeton lectures*."

It seems to me, what Bourbaki achieved thereby was trading in constructive proofs for merely existence proofs.

I shall now describe only the first rudiment of elimination theory; namely the resultant of two polynomials in one variable T.

To solve two simultaneous equations

$$\xi = a_0T^n + a_1T^{n-1} + \cdots + a_n = 0, \tag{1}$$

$$\eta = b_0T^m + b_1T^{m-1} + \cdots + b_m = 0, \tag{2}$$

or rather to find a condition for them to have a common solution, we eliminate T between them. To wit, say $n \geqq m$; multiply (1) by b_0 and (2) by a_0T^{n-m} and then subtract; this cuts down $m+n$; now iterate; ultimately we get an expression dependent only on the a's and b's. Different (but substantially equivalent) ways of systematizing this procedure were given by Euler (1748, [**26**]), Bezout (1770, [**14**]), and Sylvester (1840, [**69**]); it seems (see Berzolari [**13**]) that a procedure was already known to Newton. Following Sylvester we define the T-resultant of ξ and η, denoted by $\mathrm{Res}_T(\xi, \eta)$, to be the $m+n$ by $m+n$ determinant

$$\begin{vmatrix}
a_0 & a_1 \cdots\cdots a_n & & & \\
 & a_0 & a_1 \cdots\cdots a_n & & \\
 & & \cdots\cdots\cdots\cdots & & \\
 & & & a_0 & a_1 \cdots\cdots a_n \\
b_0 & b_1 \cdots\cdots b_m & & & \\
 & b_0 & b_1 \cdots\cdots b_m & & \\
 & & \cdots\cdots\cdots\cdots & & \\
 & & & b_0 & b_1 \cdots\cdots b_m
\end{vmatrix}$$

with zeros in all blank spaces. We have the

RESULTANT THEOREM. $\mathrm{Res}_T(\xi, \eta) = 0$ *iff either* $a_0 = 0 = b_0$ *or* ξ *and* η *have a common solution.*

Now let

$$C_N\colon F(X, Y, Z) = 0 \text{ and } D_M\colon \Phi(X, Y, Z) = 0$$

be plane curves of degree N and M without common components, i.e., F and Φ are homogeneous polynomials of degree N and M respectively, and F and Φ do not have a common factor. Let

$$E(X, Z) = \mathrm{Res}_Y(F, \Phi).$$

By the determinantal expression we see that $E(X, Z)$ is a homogeneous polynomial of degree MN. By the above theorem we also see that the "roots" of $E(X, Z)$ are the homogeneous (X, Z)-coordinates of the points of intersection of C_N and D_M. This indicates that C_N and D_M meet in MN points; the indication is turned into

BEZOUT'S THEOREM. $\Sigma\, i(C, D; P) = MN$, where the sum is over all common points P of C and D and where the positive integer $i(C, D; P)$ is the *intersection multiplicity* of C and D at P, and it is also defined in terms of a resultant thus: Given P, we may choose our coordinates so that at P we have $Z = 1$ and $X = 0 = Y$. By the Preparation Theorem of Weierstrass (1860, [**73**]) (who, although sometimes called the prince of analysis, was indeed a most distinguished high-school algebraist) we can write

$$F(X, Y, 1) = f'(X, Y)f^*(X, Y) \text{ and } \Phi(X, Y, 1) = \phi'(X, Y)\phi^*(X, Y),$$

where $f'(X, Y)$ and $\phi'(X, Y)$ are polynomials in Y with coefficients power series in X such that $f'(0, 0) \neq 0 \neq \phi'(0, 0)$, and where

$$\left.\begin{aligned} f^*(X,Y) &= Y^e + v_1(X)Y^{e-1} + \cdots + v_e(X) \\ \phi^*(X,Y) &= Y^e + w_1(X)Y^{e-1} + \cdots + w_e(X) \end{aligned}\right\} \begin{array}{l}\text{where } v_i(X) \text{ and } w_j(X) \\ \text{are power series with} \\ v_i(0) = 0 = w_j(0).\end{array}$$

Now $\mathrm{Res}_Y(f^*, \phi^*)$ is a power series in X and we define

$$i(C, D; P) = X\text{-order of } \mathrm{Res}_Y(f^*, \phi^*).$$

One can see that then

$$i(C, D; P) = \text{number of power series in } X \text{ and } Y \text{ which are linearly independent over power series in } F(X, Y, 1) \text{ and } \Phi(X, Y, 1).$$

22. Personal experience 2. How many equations are needed for defining a curve C in affine 3-space? The history of this question goes back to Kronecker and is described in my *Montreal lectures* (1970, [**9**]); a brief exposition of related matter can also be found in my very elementary *Poona lectures* (1969, [**8**]). In the Montreal lectures I tried to revive elimination theory for studying this question. Among other things, for any nonsingular C I explicitly constructed three equations. Using these three explicit equations, Murthy and Towber (1974, [**49**]) have just now proved that (1) if C is rational or elliptic then two equations suffice, and that (2) every projective module over the polynomial ring in three variables is free (Serre's conjecture).* The funny thing is that Murthy (1972, [**48**]) had also independently proved the existence of three equations, but because he was using university-algebra (the functor Ext and such) his proof was only existential, and that was not enough for proving either (1) or (2).

So high-school algebra has come to the rescue of university algebra!

23. Discriminant. The two roots of the quadratic equation $T^2 + bT + c = 0$ coincide iff its discriminant $b^2 - 4c = 0$.

With this experience in mind, consider the plane curve

$$C\colon f(X, Y) = u_0(X)Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X) = 0,$$

where the $u_i(X)$ are polynomials with $u_0(X) \neq 0$. Let x and y be such that x is a variable and $f(x, y) = 0$. Let it be required to locate those values a of x for which at least one of the corresponding values b of y has multiplicity > 1, i.e., b is a multiple root of $f(a, Y) = 0$. To this end, we eliminate Y between $f(X, Y)$ and its Y-derivative $f_Y(X, Y)$, thereby obtaining a polynomial in X which, following Sylvester, we call the Y-discriminant of f and denote it by $\mathrm{Disc}_Y(f)$; i.e.,

$$\mathrm{Disc}_Y(f) = Y\text{-discriminant of } f = \mathrm{Res}_Y(f, f_Y).$$

Now the roots of $\mathrm{Disc}_Y(f)$ are precisely those finite values of x for which there are less than n finite values of y, either because one of the finite values of y has multiplicity > 1, or because one of the values of y has gone to infinity. All this follows from the Resultant Theorem and the following very simple but

FUNDAMENTAL PRINCIPLE. *The number of roots (or irreducible factors) of a polynomial $\zeta(T)$ in one variable, counted with their multiplicities (resp. degrees and multiplicities), equals the degree of $\zeta(T)$.*

SUPPLEMENT. *Moreover, if ξ is a root (or irreducible factor) of $\zeta(T)$ with multiplicity e then ξ is a root (resp. irreducible factor) of the T-derivative of $\zeta(T)$ with multiplicity $\geqq e - 1$, where equality holds iff e is not divisible by the characteristic.*

---

* This conjecture was proved for any number of variables by D. Quillen of MIT in February of 1976. (*Editor*)

Indeed, much algebraic geometry ultimately gets reduced to the Fundamental Principle, plain or supplemented. It is certainly the algebraical key to the various "counting properly."

Now a comment on "finite values." For a finite value of x, some value of y goes to infinity iff that value of x is a root of $u_0(X)$. So "no value of y goes to infinity whenever x is finite" iff $u_0(X)$ is a constant. Geometrically speaking, the roots of $u_0(X)$ give the X-coordinates of the asymptotes to C parallel to the Y-axis. For *Example*, see Fig. 3.

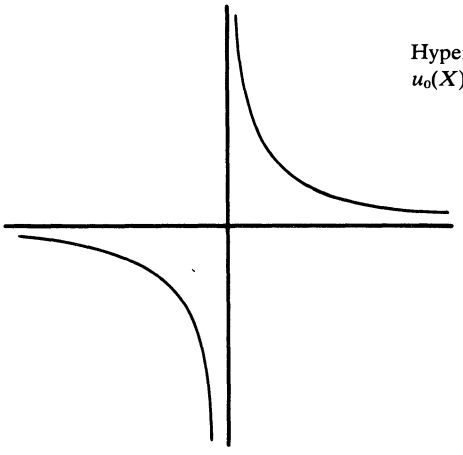


FIG. 3

24. Personal experience 3. In a recent paper (1973, [**10**]), Moh and I have made an analysis of: (1) Newton-Puiseux expansion, (2) Tschirnhausen transformations (3) resultants and discriminants, and (4) expansion of a polynomial in terms of another polynomial (analogous to decimal expansion). All these four items can be found in Chrystal [**21**] and Burnside-Panton [**16**], but not in any post-1931 algebra text. As an immediate corollary of this analysis we have proved [**11**] the following Epimorphism Theorem which can be stated in three versions thus:

HIGH-SCHOOL ALGEBRA VERSION. *Let $u(T)$ and $v(T)$ be two polynomials in one variable T such that T can be expressed as a polynomial in $u(T)$ and $v(T)$. Let degree of $u(T) = n > n' =$ degree of $v(T)$, and assume that either n or n' is not divisible by the characteristic. Then n' divides n.*

COLLEGE ALGEBRA VERSION. *Any two epimorphisms (whose degrees are nondivisible by the characteristic) of the polynomial ring in two variables onto the polynomial ring in one variable differ from each other by an automorphism of the polynomial ring in two variables. This continues to hold if the polynomial rings are replaced by the free associative (noncommutative) algebras.*

GEOMETRIC VERSION. *Any two embeddings of the affine line into the affine plane differ from each other by a biregular map of the affine plane, (in zero characteristic).*

Most of modern — post 1800 — mathematics seemed powerless to prove this kind of thing. But the 17th century stuff of Newton and Tschirnhausen enabled us to prove certain existence theorems, for instance the existence of the said automorphisms or of the said biregular map. Perhaps the matter is put in a better perspective by saying that what we obtain is a theorem about existence (or better, construction) of curves with prescribed singularities; to further illucidate this point consider:

JUNG-EVYATHAR-NAGATA PENCIL THEOREM (see Nagata [**52**]). *Let C be a nonsingular rational affine plane curve of degree n having only one point P at infinity, and suppose that C is analytically irreducible at P. Let m be the multiplicity of C at P. Let $P_1, P_2, \cdots, P_k$ be the points of C infinitely near P, and let m_i*

be the multiplicity of C at P_i. Assume that there exists another nonsingular rational plane curve D ($D \neq C$) of degree n such that: P is the only point of D at infinity, D is analytically irreducible at P, D has multiplicity m at P, and D has multiplicity m_i at P_i for all i. Then $n - m$ divides n.

In effect what we do is to show that, in case of zero characteristic, D automatically exists.

25. Hamburger. Hamburger (1871, [**30**]) and Noether (1890, [**56**]) gave an alternative expansion procedure for studying singularities of algebraic curves. In a forthcoming paper I shall present a comparative analysis of the Newton-Puiseux expansion and the Hamburger-Noether expansion. Roughly speaking, the HN expansion is cruder than the NP expansion, and so it works in any characteristic. For instance, upon replacing NP expansion by HN expansion in the last portion of our proof of the Epimorphism Theorem one gets a transparent proof of the above pencil theorem in nonzero characteristic (when the existence of D has to be assumed). But for more delicate questions, such as the automatic existence of D in zero characteristic, one has to use NP expansion.

26. Personal experience 4. I have found certain generalizations of Newton's polygon most helpful in obtaining partial results concerning the following problem which is still unproved even for two variables!

High-School Implicit Function Theorem. Let $Y_1, Y_2, \cdots, Y_n$ be polynomials in $X_1, X_2, \cdots, X_n$ (with coefficients in a field of zero characteristics) such that

$$\det\left(\frac{\partial Y_i}{\partial X_j}\right)_{\substack{i=1,\cdots,n \\ j=1,\cdots,n}} = \text{a nonzero constant.}$$

Show that then $X_1, X_2, \cdots, X_n$ are polynomials in $Y_1, Y_2, \cdots, Y_n$.

27. Hironaka. Newton's polygon has been generalized by Hironaka (1968, [**36**]) to "Newton's polyhedron" which is quite useful in understanding singularities of higher dimensional algebraic varieties.

Chapter IV: College Algebra

28. Dictionary. Telegraphically speaking:

college algebra	*high-school algebra or other suggestive idea*
set	collection of objects called its elements
group	set in which we can add and subtract
ring	also multiply
domain	also cancel nonzero common factors from an equation
field	also divide
algebraically closed field	also solve every one-variable polynomial equation
quotient field	from integers to rationals
R-module	group whose elements can be multiplied by elements of the ring R and still land in the group
ideals in R	R-module contained in R
ideal length	number of linearly independent elements
homomorphism	substitute
epimorphism	substitute to get everything
automorphsim	change of variables.

29. Conductor. Pythagoras (540 B.C.) realized that $\sqrt{2}$ was not a rational number. Gauss (1795) generalized this by showing that if a rational number satisfies an equation of the form

$$T^n + a_1 T^{n-1} + \cdots + a_n = 0$$

with integers a_i, then it must be an integer (see exercise 3 on page 7 of Hardy's *Pure Mathematics*

[31]). This gives rise to the following concepts: An element t is said to be integral over (or integrally dependent on) a domain R if t satisfies an equation of the above form with a_i in R. For any domain L containing R, by the integral closure of R in L we mean the totality of all elements of L which are integral over R; the said integral closure is seen to be a domain, i.e., sums and products of integrally dependent elements are themselves integrally dependent. By the normalization of R we mean the integral closure R^* of R in its quotient field. R is said to be normal if $R^* = R$; (Gauss' theorem says that the ring of ordinary integers is normal). An element r of R conducts R^* into R means rr^* is in R for every r^* in R^*, i.e., R serves as a common denominator for elements of R^*; the conductor of R, to be denoted by $\mathrm{Cond}(R)$ is the set of all such r. Alternatively, $\mathrm{Cond}(R)$ is the largest ideal in R which remains an ideal in R^*. Clearly: R is normal iff $\mathrm{Cond}(R) = R$, iff $\mathrm{Cond}(R) = R^*$. So, as two ways of estimating how far R is from being normal, we introduce the R-length of (the factor module) $R/\mathrm{Cond}(R)$ and the R-length of (the factor module) $R^*/\mathrm{Cond}(R)$; these lengths are nonnegative integers (or ∞) and the vanishing of either of them is a necessary and sufficient condition for R to be normal; the definition of these lengths is thus:

$$R\text{-length of } R/\mathrm{Cond}(R) = \text{the largest length } p \text{ of sequences of } R\text{-modules } \mathrm{Cond}(R) = I_0 \subset I_1 \subset \cdots \subset I_p = R$$

and

$$R\text{-length of } R^*/\mathrm{Cond}(R) = \text{the largest length } q \text{ of sequences of } R\text{-modules } \mathrm{Cond}(R) = I_0 \subset I_1 \subset \cdots \subset I_q = R^*$$

where $I_{j-1} \subset I_j$ means I_{j-1} is contained in but differs from I_j.

In addition to the above number-theoretic motivation for the concept of integral dependence, in §23 we have already hinted at the following function-theoretic and geometric motivation. Namely, consider an irreducible plane curve

$$C\colon f(X, Y) = u_0(X)Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X) = 0,$$

where the $u_i(X)$ are polynomials with $u_0(X) \neq 0$. Let x and y be elements such that x is a variable and $f(x, y) = 0$. Then y is integral over x, i.e., over the ring of polynomials in x, iff $u_0(X)$ is a constant. So, function-theoretically: y is integral over x iff "no value of y goes to infinity whenever x is finite." Geometrically: y is integral over x iff C has no asymptotes parallel to the Y-axis.

30. Dedekind. Co-inventor (with Kronecker) of ideal theory, Dedekind (1882, [25]) unified algebraic number theory and algebraic curve theory. To give a small sample of his work: Let $C\colon f(X, Y) = 0$ be an irreducible plane curve and let x and y be such that x is a variable and $f(x, y) = 0$. For a point P of C let $R(P)$ denote the local ring of P on C; in other words, $R(P)$ is the set of all rational functions on C (i.e., rational functions $r(x, y)$) which do not have a pole at P. [If $P = (a, b)$ is at finite distance, i.e., if a and b are both finite, then, more explicitly, $R(P)$ is the set of those rational functions which can be expressed in the form $v(x, y)/w(x, y)$ where $v(X, Y)$ and $w(X, Y)$ are polynomials with $w(a, b) \neq 0$.] Let $R^*(P)$ denote the normalization of $R(P)$. First we take note of

DEDEKIND'S LITTLE THEOREM: *P is a simple point of C iff $R(P)$ is normal.*

So now, as two ways of measuring how singular C is at P, in the sense of college algebra, we define

(') $$\delta(P)_{\mathrm{calg}} = R(P)\text{-length of } R(P)/\mathrm{cond}(R(P))$$

and

(*) $$\delta^*(P)_{\mathrm{calg}} = R(P)\text{-length of } R^*(P)/\mathrm{cond}(R(P)).$$

Then, when P is a singular point of C, with $\delta(P)$ as in §13, we can state

DEDEKIND'S CONDUCTOR THEOREM. $\delta^*(P)_{\text{calg}} = 2\delta(P)$.

In other words, upon letting N and g be the degree and genus of C, formula (g_5) of §13 acquires the more precise version

$$(g_6) \qquad g = (1/2)(N-1)(N-2) - (1/2)\Sigma\, \delta^*(P)_{\text{calg}}$$

where the sum is over all points P of C.

31. Macaulay. The pregnant work (1916, [**46**]) of this grand master of high-school algebra has inspired much research in ideal theory. For instance, with notation as in §30, Macaulay's profound analysis of Noether's Fundamental Theorem led another dozen authors, from Gröbner (1934, [**27**]) to Bass (1963, [**12**]), to give various proofs of

Noether's Fnd Th in new guise: $\delta^*(P)_{\text{calg}} = 2\delta(P)_{\text{calg}}$

which now converts (g_6) into

$$(g_7) \qquad g = (1/2)(N-1)(N-2) - \Sigma\, \delta(P)_{\text{calg}}.$$

Seeing how this makes two dozen proofs of Noether's Fundamental Theorem, I also got inspired to make the 25th variation which I outlined at the 1973 Calcutta Conference of the Indian Math. Soc. and which will appear in its Proceedings.

32. Krull. College algebraist par excellence. Each of Krull's papers, though ring-theoretic in nature, does valuable service to algebraic geometry. For example, he gave us (1938, [**43**]) local rings and their completions; briefly, this formalizes the process by which we "expand a given polynomial as a power series around given values of the variables."

33. Cohen and Chevalley. Cohen (1946, [**23**]) and Chevalley (1943, [**19**]) furthered the theory of local rings. In a profoundly beautiful paper (1945, [**20**]) Chevalley, taking the last formula of §21 as his starting point, developed an intersection theory for algebraic and algebroid varieties.

34. Zariski. Revolutionized algebraic geometry by harnessing algebra to its service. Since normal varieties have no singularities of codimension 1, it might be thought that Zariski (1939, [**75**]) introduced the normalization process to help out with resolution. The amusing thing is that while normalization has turned out to be quite useful in many aspects of algebraic geometry, for resolution perhaps it was actually a hindrance.

35. Nagata. (1962, [**51**]). A fitting successor to Krull.

CHAPTER V: UNIVERSITY ALGEBRA

36. Abstraction. The polynomials of high-school algebra are collected into a ring of college algebra. The rings of college algebra are collected into a category of university algebra, and then the categories are collected into a category of categories, and then ..., and then ...

37. Proportion and naturality. Our forefathers developed numerous techniques. Shall we pick up just a few of them, such as Hilbert's *syzygy*, blow them up out of all proportion, and forget all the remaining ones? When shall we attack those problems which are not amenable to the picked techniques?

What happened to Hilbert's man in the street? Is a theorem natural if it's *statement* requires over five hundred pages? Is a theorem natural if its statement is not readily comprehensible to Kummer?

38. Line of demarcation. That was the book of Cartan-Eilenberg (1956, [**17**]). Then came Grothendieck's *Elements* (1960..., [**28**]). Then Mumford's *Invariant Theorem* (1965, [**50**]). Then....

CHAPTER VI: MORE GEOMETRY

39. Infinitely near singularities. Noether's analysis of singularities of plane curves in terms of quadratic transformations, i.e., substitutions of the form

$$X = X' \text{ and } Y = X'Y' \tag{1}$$

alluded to in §13, goes something like this:

The reverse substitution is $X' = X$ and $Y' = Y/X$ and hence, as X and Y both approach zero, Y' takes all possible values. In other words, (1) explodes the point P: $X = Y = 0$ into the (projective) line E: $X' = 0$; moreover, the points of E corresponds to the tangent directions at P. We visualize this by thinking of points of E as the points in the first neighbourhood of P. In turn, any point Q of E can be exploded into a line $E(Q)$; we think of the points on the lines $E(Q)$, as Q varies on E, as the points in the second neighbourhood of P. And so on. Ultimately, we visualize the points in the successive neighbourhoods of P as the points infinitely near P. By analogy with optics, we may think of P as the source of a wave with wavefront E, every point Q of E as the source of a secondary wavefront $E(Q)$, and so on, thus getting a *Huygens diagram*: (see Fig. 4).

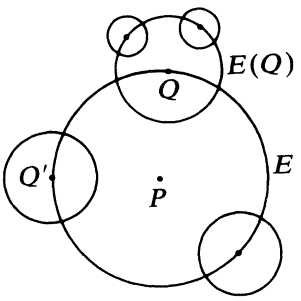


FIG. 4

Given a d-fold point P of a plane curve C we choose our coordinates to bring P to the origin and then apply (1). If now C: $f(X, Y) = 0$, then the substitution (1) transforms C into the curve C': $f'(X', Y') = 0$ given by

$$f(X', X'Y') = X'^d f'(X', Y').$$

C' will meet E in the points $P^1, \cdots, P^m$ which correspond to the tangents to C at P. If P^i is a d_i-fold point of C', then we shall have $d_1 + \cdots + d_m \leqq d$. We say that $P^1, \cdots, P^m$ are the points of C in the first neighbourhood of P, and the multiplicity of C at P_i is d_i. Now iterate this procedure. The points of C infinitely near P can be diagrammed by the *tree of C at P*: (see Fig. 5).

At every fork-point we hang as many fruits as the multiplicity of C at that point which will then be $\geqq$ the number of prongs at that point. It follows that every fork-point higher than a certain level will be unforked. So, although all the shoots of the tree will reach the sky, the number of shoots will eventually stabilize, i.e., the tree is truly infinite vertically but finite horizontally; the stabilized number of shoots may be called the geometric number of branches of C at P:

$$\beta(P)_{\text{geom}} = \text{stabilized number of shoots.} \tag{*}$$

The desingularization theorem says that at every fork-point higher than a certain level, there hangs exactly one fruit; in other words, C has only a finite number of singularities infinitely near P.

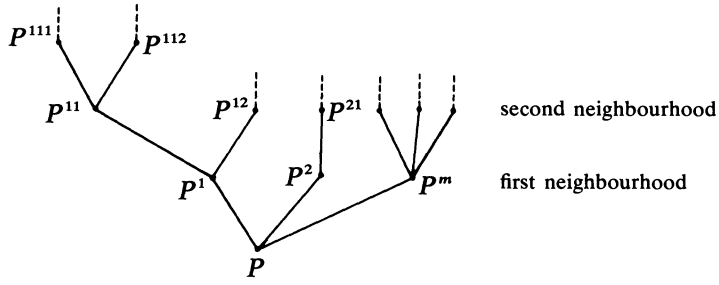


FIG. 5

Following the geometric literature, and only in this Chapter, to distinguish points in the usual sense from infinitely near points we may call the former "distinct" (another suggestive term would be "physically present"). Thus, since C has only finitely many distinct singularities, it follows that in toto C has only a finite number of singular points, distinct as well as infinitely near.

40. Adjoint. In most geometric writings on algebraic geometry the use of "adjoints" is pervasive. Here are a few random hints on this concept which might be helpful to a prospective reader of the geometric writings. Firstly, an adjoint to a plane curve C is a plane curve C^* which passes through all the singularities of C; secondly, we want this "counting properly." More precisely, if $(Q_j)_{j=1,2,\cdots}$ are all the singularities of C, distinct as well as infinitely near, then C^* is adjoint to C means

$$(\text{multiplicity of } C^* \text{ at } Q_j) \geqq (\text{multiplicity of } C \text{ at } Q_j) - 1 \text{ for all } j.$$

Alternatively, if the above condition is satisfied at a distinct singularity P of C and at every singularity Q of C infinitely near P, then we say that C^* is adjoint to C at P; and now C^* is adjoint to C means C^* is adjoint to C at every distinct singularity of C.

This brings us to

Aphorism 1. Adjoint is to geometry as derivative is to high-school algebra.

Commentary. By the Supplement of the Fundamental Principle, derivative drops the multiplicity of a root (at most) by one.

More commentary. But this refers to a one-variable polynomial, whereas C is given by a two-variable polynomial: $f(X, Y) = 0$. So which derivative to take? ANSWER: The derivative in a varying direction. But that is too hard. So we choose say f_Y. But now we get something extraneous because we are disregarding the additional condition $f_X = 0$ for singularities. So:

Aphorism 2. Adjoint is to geometry as derivative minus "something" is to college algebra.

41. Bezout and Noether. In terms of infinitely near points, Bezout's Theorem and Noether's Fundamental Theorem can be refined thus: Let C_N: $F = 0$ and D_M: $\Phi = 0$ be plane curves of degree N and M without common components. Let $(Q_j)_{j=1,2,\cdots}$ be all the common points, distinct as well as infinitely near, of C and D. Let s_j and r_j be the multiplicities of C and D at Q_j. Then:

Bezout's Theorem Refined. $\Sigma_j r_j s_j = MN$. Or, alternatively,

$$\Sigma\, i(C, D; P)_{\text{geom}} = MN,$$

where the sum is over all distinct common points P of C and D, and where the geometric intersection multiplicity of C and D at P is defined by the equation

$$(*) \qquad i(C, D; P)_{\text{geom}} = \Sigma^P r_j s_j$$

where Σ^P is the sum over those j for which either $Q_j = P$ or Q_j is infinitely near P.

NOETHER'S FUNDAMENTAL THEOREM REFINED. *A plane curve $H = 0$ can be expressed in the form*

$H = AF + B\Phi$, with plane curves $A = 0$ and $B = 0$ having multiplicities at least $r_j - 1$ and $s_j - 1$ at Q_j for all j, if and only if H has multiplicity at least $r_j + s_j - 1$ at Q_j for all j.

Warning. This does not make B adjoint to F. We can only say something like "B is adjoint to F on Φ, i.e., as far as common points of F with Φ are concerned." Similarly for A and Φ.

42. Genus formula refined. Let us now describe the study of special cases which we spoke of in §13 and which led the geometers to the formula

$$(g_5) \qquad g = (1/2)(N-1)(N-2) - \Sigma\,\delta(P)$$

for the genus g of an irreducible plane curve C_N: $f(X, Y) = 0$ of degree N, where the sum is over all distinct singularities P of C and where $\delta(P)$ is a certain positive integer which says that "P accounts for $\delta(P)$ ordinary double points."

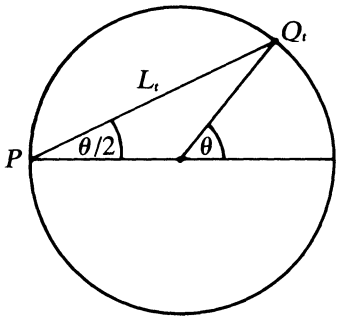


FIG. 6

(1) ($N = 2$). The circle C_2: $X^2 + Y^2 = 1$ has the usual parametrization: $X = \cos\theta$ and $Y = \sin\theta$. The familiar substitution $\tan\theta/2 = t$ converts it into the rational parametrization: $X = (1-t^2)/(1+t^2)$ and $Y = 2t/(1+t^2)$. See Fig. 6. Geometrically speaking, fix a point P on the circle, say $P = (-1, 0)$. Let L_t be the line through P of variable slope t. Then

$$Q_t = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$$

is the other point of intersection of L_t with the circle.

This explains why (rational functions of) the trigonometric functions can be integrated.

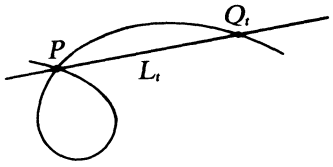


FIG. 7

FIG. 8

(2) ($N = 3$). Fix a point P on C_3. If P is a double point of C_3, then L_t meets C_3 in exactly one more point Q_t: (see Fig. 7), and hence the coordinates of Q_t are rational functions of t. If P was a simple point of C_3 then L_t would meet C_3 in two more points: (see Fig. 8), and this "indicates" why elliptic integrals cannot be integrated. If C_3 has two double points P_1 and P_2 then the line P_1P_2 would meet C_3 in at least $2+2>3$ points contradicting Bezout.

(3) ($N = 4$). Suppose C_4 has three double points P_1, P_2, P_3. Fix a simple point P_4 of C_4. Now a conic through P_1, P_2, P_3, P_4 meets C_4 in exactly one more point because $2 \cdot 4 - (2+2+2+1) = 1$.

Since 5 points determine a conic, we get a one-to-one correspondence between conics through P_1, P_2, P_3, P_4 and points of C_4; this gives a rational parametrization of C_4. If C_4 had another double point P_5 then the conic through $P_1, \cdots, P_5$ would meet C_4 in at least $2+2+2+1+2>2\cdot 4$ points contradicting Bezout.

(4) In this manner, by passing curves of degree $N-2$ through the double (and some simple) points of C_N, we see that

$$\text{number of double points of } C_N \leqq (1/2)(N-1)(N-2),$$

and that if equality holds then C_N is *rational*, i.e., C_N has a rational parametrization, i.e., the points of C_N can be put in an almost one-to-one (algebraic) correspondence with the points of a line.

(5) $(N>1)$. If C_N has a d-fold point P and an e-fold point Q, then the line PQ meets C_N in at least $d+e$ points and hence $d+e \leqq N$ by Bezout. It follows that the multiplicity of any point of C_N is $\leqq N-1$, and if C_N does have an $(N-1)$-fold point then it has no other singularity.

(6) In view of (4) and (5) we surmise that an $(N-1)$-fold point should be counted as $(1/2)(N-1)(N-2)$ double points. From this, and taking infinitely near points into account, we surmise that: If $P=P_1$ is a distinct singular point of C_N, if $(P_j)_{j=2,3,\cdots}$ are all the singularities of C_N infinitely near P, and if d_j is the multiplicity of C_N at P_j for $j=1,2,\cdots$, then upon defining

$$(*) \qquad \delta(P)_{\text{geom}} = (1/2)\Sigma\, d_j(d_j-1)$$

we have

$$\delta(P) = \delta(P)_{\text{geom}}.$$

(7) In view of (4) and (6) it appears (in effect, by considering adjoints to C of degree $N-2$) that summing over all the distinct singularities P of C_N we have

$$\Sigma\, \delta(P)_{\text{geom}} \leqq (1/2)(N-1)(N-2)$$

with equality if and only if C_N is rational; i.e., in words: C_N has at most $(1/2)(N-1)(N-2)$ double points counted properly, and the maximum is reached iff C_N is rational.

(8) Now the genus g of f is also a nonnegative integer whose vanishing is an iff condition for the Riemann surface of f to be without handles. So, in view of (7), we surmise that:

$$(g_8) \qquad g = (1/2)(N-1)(N-2) - \Sigma\, \delta(P)_{\text{geom}},$$

where the sum is over all the distinct singularities P of C. Or, equivalently, if $(Q_j)_{j=1,2,\cdots}$ are the singular points of C_N, distinct as well as infinitely near, and if s_j is the multiplicity of C_N at Q_j, then

$$(g_9) \qquad g = (1/2)(N-1)(N-2) - (1/2)\sum_j s_j(s_j-1).$$

Chapter VII: More High-School Algebra

43. Discriminant locus. Let

$$C\colon f(X,Y) = u_0(X)Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X) = 0$$

be a plane curve where the $u_i(X)$ are polynomials with $u_0(X) \neq 0$. Let x and y be such that x is a variable and $f(x,y)=0$.

By the *discriminant locus* (for the projection of C on the X-axis by lines parallel to the Y-axis) we mean the totality of the roots of $\text{Disc}_Y(f)$, each of which we call a *discriminant point.*

First let us remark that the equation $f(X,Y)=0$ is said to be "separable" if $\text{Disc}_Y(f)$ is not identically zero, i.e., if there are only a finite number of discriminant points; we shall usually suppose this to be the case; at any rate, by the Supplemented Fundamental Principle and the Resultant

Theorem one can see that this is automatically so when $f(X, Y)$ is free from multiple factors and the characteristic is zero.

Next let us recall that the discriminant points turn out to be precisely those (finite) values of x for which there are less than n finite values of y, either because one of the finite values of y has multiplicity >1, or because one of the values of y has gone to infinity; the latter happens exactly when the given value of x is a root of $u_0(X)$; in a sense, the roots of $u_0(X)$ may be regarded as accidental discriminant points.

Disregarding the roots of $u_0(X)$ and restricting to finite values of x and y, what causes the discriminant locus?

Aphorism 3. Discriminant locus equals branch locus plus projection of the singular locus.

Aphorism 4. Derivative locus equals "Different locus" plus singular locus.

NOTE. Count everything properly. Then these Aphorisms provide a guideline for estimating singularities.

Commentary. On the one hand, a point $P=(a, b)$ is a singularity of C iff $f(a, b)=f_Y(a, b)=f_X(a, b)=0$; on the other hand, since $\mathrm{Disc}_Y(f)=\mathrm{Res}_Y(f, f_Y)$, by the Resultant Theorem we see that a value a of x is a discriminant point iff for some b we have $f(a, b)=f_Y(a, b)=0$. So in particular, the projection of the singular locus is always part of the discriminant locus. Now when $P=(a, b)$ is a simple point of C then, $f_Y(a, b)=0$ iff the tangent to C at P is parallel to the Y-axis; a typical case of this is the parabola $(Y-b)^2-(X-a)=0$; generalizing this we call P a *Different point* if

$$f(X+a, Y+b)=f_0(X, Y)f_1(X, Y),$$

where $f_0(X, Y)$ is a polynomial in Y with coefficients power series in X, and where

$$f_1(X, Y)=Y^{e(1)}+v_{11}(X)Y^{e(1)-1}+\cdots+v_{1e(1)}(X)$$

is an irreducible polynomial in Y with $e(1)>1$ and with power series $v_{1j}(X)$ such that $v_{1j}(0)=0$ for all j; such an irreducible factor f_1, with $e(1)\geqq 1$, is called a *branch* of C centered at P. A value a of x is called a *branch point* if it is the projection of a Different point.

Clearly $f(a, b)=f_Y(a, b)=0$ for every Different point $P=(a, b)$. Therefore (1) for a point $P=(a, b)$ we have that: $f(a, b)=f_Y(a, b)=0$ iff P is either a Different point or a singular point. Hence, by the Resultant Theorem, (2) for a value a of x we have that: a is a discriminant point iff it is

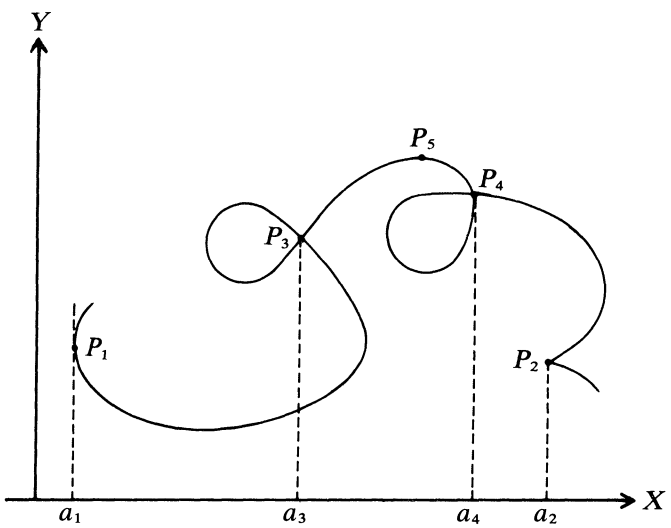


FIG. 9

either a branch point or a point in the projection of the singular locus. By the Fundamental Principle one can also see that: the number of branches of C lying above any given value a of x is always $\leqq n$; this number is $< n$ iff a is a branch point; (recall that we are disregarding the roots of $u_0(X)$).

EXAMPLE: Consider a point $P_i = (a_i, b_i)$. See Fig. 9.

(P_1) *Simple point with vertical tangent.*

$$f(X + a_1, Y + b_1) = X + \sum_{i+j>1} \lambda_{1ij} X^i Y^j.$$

P_1 is a simple point and a Different point. Number of branches centered at P_1 is 1; a_1 is a branch point but need not be in the projection of the singular locus. Near P_1, C could look like the origin of the parabola $Y^2 - X = 0$; (the two roots $y = \pm x^{1/2}$ have the same value at $x = 0$ but they are not power series in x; "in the complex case, analytic continuation around $x = 0$ permutes them").

(P_2) *Cusp.*

$$f(X + a_2, Y + b_2) = f'(X, Y)(Y^2 - X^3) \text{ with } f'(0,0) \neq 0.$$

P_2 is a singular point and a Different point. Number of branches centered at P_2 is 1; a_2 is a branch point and is in the projection of the singular locus. Near P_2, C looks like the origin of $Y^2 - X^3 = 0$; (the two roots $y = \pm x^{3/2}$ have the same value at $x = 0$ but they are not power series in x; "in the complex case, analytic continuation around $x = 0$ permutes them").

(P_3) *Node without vertical tangent.*

$$f(X + a_3, Y + b_3) = Y^2 - X^2 + \sum_{i+j>2} \lambda_{3ij} X^i Y^j.$$

P_3 is a singular point but not a Different point. Number of branches centered at P_3 is 2; a_3 need not be a branch point but is in the projection of the singular locus. Near P_3, C looks like the origin of $Y^2 - X^2(1 + X) = 0$; (the two roots $y = \pm x(1 + x)^{1/2}$ have the same value at $x = 0$ and they are power series in x (characteristic $\neq 2$); "in the complex case, analytic continuation around $x = 0$ does not permute them").

(P_4) *Node with a vertical tangent.*

$$f(X + a_4, Y + b_4) = XY + \sum_{i+j>2} \lambda_{4ij} X^i Y^j.$$

P_4 is a singular point and Different point; (composite of type (P_1) and (P_3)). Number of branches centered at P_4 is 2; a_4 is a branch point and is in the projection of the singular locus. Near P_4, C could look like the origin of $Y^3 + X^3 + XY = 0$.

(P_5) *Simple point with horizontal tangent.*

$$f(X + a_5, Y + b_5) = Y + \sum_{i+j>1} \lambda_{5ij} X^i Y^j.$$

P_5 is a simple point but not a Different point; a_5 need not be a discriminant point.

NOTE. Let us stress that the singular locus is always where it is, but the "Different locus" (i.e., the totality of Different points) and the "derivative locus $f = f_Y = 0$" vary according to the direction of projection.

For instance, if in the above example we project on the Y-axis by lines parallel to the X-axis then: P_2, P_3, P_4 are still singularities; again P_2, P_4 are Different points while P_3 is not; P_1 is no longer a Different point; P_5 is a new Different point;...

44. Characteristic pairs. To continue with the Newton-Puiseux expansion, suppose that we are in characteristic zero, and consider

$$f^*(X, Y) = Y^e + v_1(X)^{e-1} + \cdots + v_e(X), \tag{1}$$

where the $v_i(X)$ are power series with $v_i(0) = 0$ for all i.

Then, assuming f^* to be irreducible, by Newton's Theorem

$$f^*(T^e, Y) = \prod_{i=1}^{e} [Y - y(\rho^i T)], \tag{2}$$

where ρ is a primitive eth root of 1 and

$$y(T) = \Sigma\, \nu_k T^k \tag{3}$$

is a power series. Following Smith (1873, [**67**]) and Halphen (1884, [**29**]) we shall now pick out a finite number of significant exponents from the above expansion which give rise to a finite number of pairs of integers called the characteristic pairs of f^*. In terms of the characteristic pairs one can write down explicit formulas for the order of a resultant (and hence for intersection multiplicity) and for the order of a discriminant (and hence for the "measure of a singularity"). The exponents are picked thus:

$m_1 = \operatorname{ord}_T y(T) =$ smallest k such that $\nu_k \neq 0$.

$m_2 =$ smallest k such that $\nu_k \neq 0$ and $\operatorname{GCD}(e, m_1, m_2) < \operatorname{GCD}(e, m_1)$.

. . . . . .

$m_i =$ smallest k such that $\nu_k \neq 0$ and $\operatorname{GCD}(e, m_1, \cdots, m_i) < \operatorname{GCD}(e, m_1, \cdots, m_{i-1})$.

. . . . . .

$m_h =$ smallest k such that $\nu_k \neq 0$ and $1 = \operatorname{GCD}(e, m_1, \cdots, m_h) < \operatorname{GCD}(e, m_1, \cdots, m_{h-1})$.

Here GCD stands for the greatest common divisor. We also put:

$q_1 = m_1$	and	$d_1 = e$
$q_2 = m_2 - m_1$		$d_2 = \operatorname{GCD}(e, m_1)$
. . . . .		. . . . .
$q_i = m_i - m_{i-1}$		$d_i = \operatorname{GCD}(e, m_1, \cdots, m_{i-1})$
. . . . .		. . . . .
$q_h = m_h - m_{h-1}$		$d_h = \operatorname{GCD}(e, m_1, \cdots, m_{h-1})$
		$d_{h+1} = \operatorname{GCD}(e, m_1, \cdots, m_h) = 1.$

The pairs of integers $(q_i, d_i)_{i=1,\cdots,h}$ are called the *characteristic pairs* of f^*.

As an illustration of the significance of the characteristic pairs let us write down a formula for the order of a discriminant. So let $\operatorname{Disc}_Y(f^*)$ be the power series in X obtained by eliminating Y between $f^*(X, Y)$ and its Y-derivative $f^*_Y(X, Y)$, i.e.,

$$\operatorname{Disc}_Y(f^*) = Y\text{-discriminant of } f^* = \operatorname{Res}_Y(f^*, f^*_Y).$$

Then upon defining

$$\alpha(f^*) = (q_1 - 1)(d_1 - 1) + \sum_{i=2}^{h} q_i(d_i - 1)$$

one has the formula

$$[\operatorname{ord}_X \operatorname{Disc}_Y(f^*)] - e + 1 = \operatorname{ord}_T f^*_Y(T^e, y(T)) - \operatorname{ord}_T \frac{dT^e}{dT} = \alpha(f^*) \tag{4}$$

which one obtains by direct calculation; (d stands for derivative). [To obtain the second equation, substitute (3) in (2) and use the product formula for derivative; to prove the first equation, use the formula

$$\text{(5)}\qquad \mathrm{Res}_Y\Big(\prod_{i=1}^{e}(Y-y_i),\prod_{j=1}^{e}(Y-z_j)\Big)=\prod_{i,j}(y_i-z_j);$$

let us also record the following consequence of (5)

$$\text{(6)}\qquad \mathrm{Disc}_Y(f_1f_2\cdots f_\beta)=\prod_i \mathrm{Disc}_Y(f_i)\prod_{i\neq j}\mathrm{Res}_Y(f_i,f_j),$$

where the f_i are (monic) polynomials in Y.]

As another illustration of the significance of the characteristic pairs we shall write a formula for the intersection multiplicity, i.e., for the order of a resultant. So consider another polynomial

$$\phi^*(X,Y)=Y^\varepsilon+w_1(X)Y^{\varepsilon-1}+\cdots+w_\varepsilon(X),$$

where the $w_i(X)$ are power series with $w_i(0)=0$ for all i.

Assuming ϕ^* to be irreducible, let

$$\text{(2')}\qquad \phi^*(T^\varepsilon,Y)=\prod_{j=1}^{\varepsilon}[(Y-z(\sigma^jT)]$$

be the factorization of ϕ^* where σ is a primitive εth root of 1, and

$$\text{(3')}\qquad z(T)=\Sigma\,\nu'_kT^k$$

is a power series. Let $(m'_i,q'_i,d'_i)_{i=1,\cdots,h'}$ be the numbers which correspond to $(m_i,q_i,d_i)_{i=1,\cdots,h}$ when we replace f^* by ϕ^*. Let θ measure how close a root of f^* comes to being a root of ϕ^*, i.e., let

$$\theta=\max_{\substack{1\leqq i\leqq e\\ 1\leqq j\leqq\varepsilon}}\mathrm{ord}_X[y(\rho^iX^{1/e})-z(\sigma^jX^{1/e})]$$

and now let $p=$ largest integer such that $m_p/d_1=m'_p/d'_1\leqq\theta$. Then upon defining

$$\alpha(f^*,\phi^*)=\Big(\theta-\frac{m_p}{d_1}\Big)d'_1d_{p+1}+\sum_{i=1}^{p}q_id'_i$$

(and making a direct calculation using (2), (3), (5), (2'), (3')) one gets

$$\text{(7)}\qquad \mathrm{ord}_X\,\mathrm{Res}_Y(f^*,\phi^*)=\alpha(f^*,\phi^*).$$

Note that the above expression for $\alpha(f^*,\phi^*)$ is really symmetric because one can see that

$$\Big(\theta-\frac{m_p}{d_1}\Big)d'_1d_{p+1}=\Big(\theta-\frac{m'_p}{d'_1}\Big)d_1d'_{p+1}\ \text{ and }\ q_id'_i=q'_id_i\ \text{ for }\ 1\leqq i\leqq p.$$

Now drop the assumption of f^* being irreducible, and let

$$f^*(X,Y)=\prod_{i=1}^{\beta}f_i(X,Y)$$

be a factorization of f^* into irreducible polynomials

$$f_i(X,Y)=Y^{e(i)}+\sum_{k=1}^{e(i)}v_{ik}(X)Y^{e(i)-k}\ \text{ with power series }\ v_{ik}(X).$$

Let $y_i(X^{1/e(i)})$ be a root of f_i, i.e., let $y_i(T)$ be a power series with

$$f_i(T^{e(i)},y_i(T))=0.$$

Then upon defining

$$\text{(')}\qquad \alpha(f^*)=\sum_i\alpha(f_i)+\sum_{i\neq j}\alpha(f_i,f_j)$$

[by using (4), (5), (6), (7)] one can see that

$$(8)\qquad \left\{ \begin{aligned} [\mathrm{ord}_X \mathrm{Disc}_Y(f^*)] - e + \beta &= \sum_{i=1}^{\beta} \left[\mathrm{ord}_T f_Y^*(T^{e(i)}, y_i(T)) - \mathrm{ord}_T \frac{dT^{e(i)}}{dT} \right] \\ &= \alpha(f^*), \end{aligned} \right.$$

where d stands for derivative.

Finally, also drop the assumption of ϕ^* being irreducible, and let

$$\phi^*(X, Y) = \prod_{j=1}^{\gamma} \phi_j(X, Y)$$

be a factorization of ϕ^* into (monic) irreducible polynomials in Y with coefficients power series in X. Then upon defining

$$\alpha(f^*, \phi^*) = \sum_{i,j} \alpha(f_i, \phi_j)$$

[by (5) and (7)] one can see that

$$\mathrm{ord}_X \mathrm{Res}_Y(f^*, \phi^*) = \alpha(f^*, \phi^*).$$

Therefore, with notation as in the end of §21, upon defining the intersection multiplicity of the plane curves C and D at the point P, in the sense of high-school algebra, by the formula

$$(\prime\prime)\qquad i(C, D; P)_{\mathrm{halg}} = \alpha(f^*, \phi^*)$$

we get

$$(9)\qquad i(C, D; P)_{\mathrm{halg}} = i(C, D; P).$$

45. High-school genus formula. Let C_N: $F(X, Y, Z) = 0$ be an irreducible plane curve of degree N. Again assume that we are in characteristic zero. Given any point P of C_N we choose coordinates so that P is given by $X = Y = 0$ and $Z = 1$. By Weierstrass' Preparation Theorem (or by Newton's Theorem) we can write

$$F(X, Y, 1) = f'(X, Y) f^*(X, Y)$$

where $f'(X, Y)$ is a polynomial in Y with coefficients power series in X such that $f'(0, 0) \neq 0$, and where

$$f^*(X, Y) = Y^e + v_1(X) Y^{e-1} + \cdots + v_e(X)$$

with positive integer e and power series $v_i(X)$ such that $v_i(0) = 0$ for all i. As a measure of how singular C is at P in the sense of high-school algebra, (following the dictate of Aphorism 4 of §43 and the above formula (8)), we define

$$(*)\qquad \delta^*(P)_{\mathrm{halg}} = \alpha(f^*)$$

with $\alpha(f^*)$ as in the above formula ($'$); (we note that then $\delta^*(P)_{\mathrm{halg}}$ is a nonnegative integer which is zero iff P is in a simple point of C). By using the chain-rule for differentiation one shows that $\delta^*(P)_{\mathrm{halg}}$ is independent of the coordinate system. Taking formula (g_2) of §3 as the definition of the genus g of C, in §47 we shall outline the proof of the following formula:

$$(g_{10})\qquad g = (1/2)(N-1)(N-2) - (1/2) \Sigma \delta^*(P)_{\mathrm{halg}}$$

where the sum is over all points P of C.

46. Branches. Let C: $F(X, Y, Z) = 0$ be an irreducible plane curve of degree N; so F is an irreducible homogeneous polynomial of degree N. Let $f(X, Y) = F(X, Y, 1)$. In other words, "at finite

distance" C is given by

$$f(X, Y) = u_0(X)Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X) = 0,$$

where the $u_i(X)$ are polynomials with $u_0(X) \neq 0$; note that $n \leqq N$. Let x and y be such that x is a variable and $f(x, y) = 0$.

We shall now introduce the concept of branches of C centered at a point P of C; the totality of these as P varies over C are called the branches of C.

Let $P = (a, b)$ be a point of C at finite distance. By the Weierstrass Preparation Theorem we can write

$$f(X + a, Y + b) = f'(X, Y)f^*(X, Y), \tag{1}$$

where $f'(X, Y)$ is a polynomial in Y with coefficients power series in X such that $f'(0, 0) \neq 0$, and where

$$f^*(X, Y) = Y^e + v_1(X)Y^{e-1} + \cdots + v_e(X) \tag{2}$$

with positive integer e and power series $v_i(X)$ such that $v_i(0) = 0$ for all i. Let

$$f^*(X, Y) = \prod_{i=1}^{\beta} f_i(X, Y)$$

be a factorization of f^* into irreducible polynomials

$$f_i(X, Y) = Y^{e(i)} + \sum_{k=1}^{e(i)} v_{ik}(X)Y^{e(i)-k} \text{ with power series } v_{ik}(X).$$

Now we can take power series $x_i(T)$ and $y_i(T)$ such that

$$\operatorname{ord}_T x_i(T) = e(i), \quad \operatorname{ord}_T y_i(T) > 0, \quad \text{and } f_i(x_i(T), y_i(T)) = 0.$$

(In case of characteristic zero we can do all this by Newton's Theorem, and then we may choose $x_i(T)$ to be $T^{e(i)}$.) We define β to be the number of branches of C centered at P in the sense of high-school algebra:

$$\beta(P)_{\text{halg}} = \beta. \tag{3*}$$

With each of the "algebroid curves" $f_i(X, Y) = 0$ we associate a "branch" V_i of C centered at P. For every rational function $r(x, y)$ we define the *value* $r(V_i)$ of $r(x, y)$ at V_i, the *order* $V_i(r(x, y))$ of $r(x, y)$ at V_i, and the *degree* $\deg_P(r(x, y))$ of the divisor of $r(x, y)$ at P, thus:

$$V_i(r(x, y)) = \operatorname{ord}_T r(a + x_i(T), b + y_i(T))$$

$$\deg_P(r(x, y)) = \sum_{i=1}^{\beta} V_i(r(x, y))$$

and

$$r(V_i) = \begin{cases} \text{coefficient of } T^0 \text{ in } r(a + x_i(T), b + y_i(T)) & \text{if } V_i(r(x, y)) \geqq 0 \\ \infty & \text{if } V_i(r(x, y)) < 0. \end{cases}$$

For any differential $r(x, y)\,dx$ we define its *order* $V_i(r(x, y)\,dx)$ at V_i and the *degree* $\deg_P(r(x, y)\,dx)$ of its divisor at P, thus:

$$V_i(r(x, y)\,dx) = \operatorname{ord}_T\left(r(a + x_i(T), b + y_i(T))\frac{dx_i(T)}{dT}\right)$$

and

$$\deg_P(r(x, y)\,dx) = \sum_{i=1}^{\beta} V_i(r(x, y)\,dx).$$

We observe that by using formula (5) of §44 it can be shown that if $D\colon \phi(X, Y) = 0$ is any plane curve without common components with C, then

$$i(C, D; P) = \deg_P(\phi(x, y)). \tag{4}$$

As a measure of how singular C is at P, (following the dictate of Aphorism 4 of §43), we define

$$\delta^*(P) = \deg_P(f_Y(x, y)) - \deg_P(dx). \tag{5*}$$

(One can see that $\delta^*(P)$ is a nonnegative integer which is zero iff P is a simple point of C.) In view of formula (8) of §44 it can be shown that

$$\delta^*(P)_{\text{halg}} = \delta^*(P) \text{ in case of characteristic zero.} \tag{6}$$

It can be seen that the various concepts introduced above are independent of the particular parametrization $(x_i(T), y_i(T))$.

When P is a point of C at infinity, one introduces all the above concepts in a similar fashion after making a suitable change of coordinates. At any rate, let us note that if $P = (a, b)$ is any point of C at finite distance then the branches of C centered at P are exactly those branches V of C for which $x(V) = a$ and $y(V) = b$; moreover, for any branch V of C we have: $x(V) \neq \infty \neq y(V)$ if the center of V is a point of C at finite distance.

For any (nonzero) rational function $s(x, y)$ we define the *degree of its divisor* by the equation

$$\deg(s(x, y)) = \Sigma\, V(s(x, y)) \text{ with sum over all branches } V \text{ of } C$$

and we note that formula (1) of §3 is to be interpreted as the formula

$$\deg(s(x, y)) = 0 \tag{7}$$

a proof of which can be based on the above formula (4) and Bezout's Theorem as stated in §21.

For any (nonzero) differential $r(x, y)\,dx$ we define the *degree of its divisor* by the equation

$$\deg(r(x, y)\,dx) = \Sigma\, V(r(x, y)\,dx) \text{ with sum over all branches } V \text{ of } C$$

and note that the formula (g_2) of §3 for the genus g of C is to be interpreted as saying

$$2g - 2 = \deg(r(x, y)\,dx). \tag{g_2)*}$$

Let us regard the above equation as the definition of g.

For any branch V of C, the *branching degree* or the differential exponent $e(V)$ of V over x is defined by

$$e(V) = \begin{cases} V(x - x(V)) & \text{if } x(V) \neq \infty \\ V(x^{-1}) & \text{if } x(V) = \infty. \end{cases} \tag{8}$$

(In the above situation $e(i) = e(V_i)$.) By the Fundamental Principle it can be shown that

$$\begin{cases} \sum_{x(V)=a} e(V) = n & \text{for every value } a \text{ of } x, \text{ finite or infinite;} \\ & \text{the sum being over all branches } V \text{ of } C \text{ for which } x(V) = a. \end{cases} \tag{9}$$

As a measure of how different C is from the X-axis, at a branch V of C, we define the *order of the Different* of C over x at V by the equation

$$V(\text{Diff}(C, x)) = \begin{cases} V(dx) & \text{if } x(V) \neq \infty \\ V(dx^{-1}) & \text{if } x(V) = \infty. \end{cases} \tag{10}$$

Again with the same motivation, for any point P of C, we define the *degree of the Different* of C over x at P by the equation

$$\deg_P \mathrm{Diff}(C,x) = \Sigma\, V(\mathrm{Diff}(C,x)) \text{ with sum over all branches } V \text{ of } C \text{ centered at } P,$$

and we note then in view of (5*) we have

$$(11^*) \qquad \delta^*(P) = \deg_P(f_Y(x,y)) - \deg_P \mathrm{Diff}(C,x) \text{ if } P \text{ is at finite distance.}$$

Finally, to measure how different C is from the X-axis, globally, we define the *degree of the Different* of C over x by the equation

$$(12) \qquad \deg \mathrm{Diff}(C,x) = \Sigma\, V(\mathrm{Diff}(C,x)) \text{ with sum over all branches } V \text{ of } C.$$

Because of the equation $dx^{-1} = -x^{-2}dx$ by definitions (8), (10) and (12) we see that

$$\deg \mathrm{Diff}(C,x) = \deg(dx) + 2 \sum_{x(V)=\infty} e(V)$$

and hence by (g_2^*) and (9) we get

$$(g_{11}) \qquad 2-2g = 2n - \deg \mathrm{Diff}(C,x).$$

In view of definitions (8) and (10), for any branch V of C we have

$$(13) \qquad V(\mathrm{Diff}(C,x)) \geqq e(V)-1 \text{ with equality iff } e(V) \text{ is not divisible by the characteristic,}$$

and hence by (g_{11}) we get

$$(g_{12}) \qquad \begin{cases} 2-2g = 2n - \Sigma[e(V)-1] \text{ in case of characteristic zero,} \\ \qquad\qquad\qquad\qquad \text{with sum over all branches } V \text{ of } C. \end{cases}$$

Formulas (g_{11}) and (g_{12}) are variously ascribed to Riemann, Zeuthen, or Hurwitz.

A value a of x, finite or infinite, is called a *branch point* (for C over x, or for the projection of C on the X-axis by lines parallel to the Y-axis) if $e(V)>1$ for some branch V of C with $x(V)=a$. Upon letting

$$(14) \qquad \mu(a) = \text{number of branches } V \text{ of } C \text{ with } x(V) = a$$

by (9) we have

$$(15) \qquad \mu(a) = n - \sum_{x(V)=a} [e(V)-1]$$

and hence

$$(16) \qquad a \text{ is a branch point iff } \mu(a) < n.$$

To elaborate (9) in greater detail: Given any finite value a of x, let $b_1, \cdots, b_q$ be the finite roots of $f(a,Y)=0$ and let p_j be the multiplicity of the root b_j. Also let p_0 be the largest integers such that $u_i(a)=0$ whenever $0 \leqq i < p_0$, and let $b_0 = \infty$. Then

$$p_0 + p_1 + \cdots + p_q = n$$

and

$$\sum_{x(V)=a,\, y(V)=b_j} e(V) = p_j \text{ for } 0 \leqq j \leqq q.$$

In still greater detail, for $j = 1, \cdots, q$, let

$$f(X+a, Y+b_j) = f_j'(X,Y) f_j^*(X,Y) \text{ with } f_j^*(X,Y) = Y^{e_j} + \cdots$$

be as in (1) and (2) above. Then $e_j = p_j$ for $1 \leqq j \leqq q$ and

$$f(X + a, Y) = f''(X, Y) \prod_{j=1}^{q} f_j^*(X, Y - b_j),$$

where

$$f''(X, Y) = u_0(X + a)Y^{p_0} + \cdots$$

factors into the branches V of C for which $x(V) = a$ and $y(V) = \infty$.

Recalling that the roots of $\mathrm{Disc}_Y(f)$ constitute the discriminant locus, we conclude that the said locus consists of those finite values a of x for which, in the above notation, $q < n$.

47. Another genus formula. Again let C_N: $F(X, Y, Z) = 0$ be an irreducible plane curve of degree N. We can choose coordinates so that the point $(0, 1, 0)$ is not on C, i.e., upon letting $f(X, Y) = F(X, Y, 1)$ we may suppose

$$C: f(X, Y) = Y^N + v_1(X)Y^{N-1} + \cdots + v_N(X) = 0.$$

Let x and y be such that x is a variable and $f(x, y) = 0$. By definition (5*) of §46 we have

$$\sum_{x(V) \neq \infty} [V(f_Y(x, y)) - V(dx)] = \sum_{P \neq \infty} \delta^*(P), \tag{1}$$

where the right hand side indicates sum over all points P of C at finite distance. Now (x^{-1}, yx^{-1}) serve as coordinates for the points of C at infinity, and the relation between them is $\zeta(x^{-1}, yx^{-1}) = 0$ where

$$\zeta(X, Y) = X^N f(X^{-1}, YX^{-1}). \tag{2}$$

By definition (5*) of §46 we also get

$$\sum_{x(V) = \infty} [V(\zeta_Y(x^{-1}, yx^{-1})) - V(dx^{-1})] = \sum_{P = \infty} \delta^*(P), \tag{3}$$

where the right hand side indicates sum over all points P of C at infinity. Now

$$dx^{-1} = -x^{-2}dx \tag{4}$$

and by (2) we also have

$$\zeta_Y(x^{-1}, yx^{-1}) = x^{1-N} f_Y(x, y). \tag{5}$$

Substituting (4) and (5) in (3), and then adding (1) and (3) we get

$$\deg(f_Y(x, y)) - \deg(dx) + (N - 3) \sum_{x(V) = \infty} V(x^{-1}) = \sum \delta^*(P), \tag{6}$$

where the last sum is over all points P of C. Now

$$\deg(f_Y(x, y)) = 0 \qquad \text{by (7) of §46}$$

$$\sum_{x(V) = \infty} V(x^{-1}) = N \qquad \text{by (8) and (9) of §46,}$$

and hence by (6) we get

$$\deg(dx) = N(N - 3) - \Sigma \delta^*(P). \tag{7}$$

By the definition $(g_2)^*$ of §46, the genus g of C is given by

$$\deg(dx) = 2g - 2$$

and hence by (7) we have

(g_{13}) $$g = (1/2)(N-1)(N-2) - (1/2)\Sigma\,\delta^*(P),$$

where the sum is over all points P of C. In view of formula (6) of §46, this also yields formula (g_{10}) of §44.

Chapter VIII: More College Algebra

48. Lagrange Interpolation. From high-school algebra we have the following interpolation formula due to Lagrange (1760). Let

$$f(Y) = \prod_{i=1}^{n} (Y - y_i) \text{ with } y_1, \cdots, y_n \text{ all distinct.}$$

Then for any polynomial $\eta(Y)$ of degree $< n$ we have

$$\eta(Y) = \sum_{i=1}^{n} \frac{\eta(y_i) f(Y)}{f_Y(y_i)(Y - y_i)},$$

because both sides are polynomials of degree $< n$ and their values coincide for each of the n values $y_1, \cdots, y_n$ of Y.

49. Different and Conductor. Let C: $f(X, Y) = 0$ be an irreducible plane curve. Let x and y be elements such that x is a variable and $f(x, y) = 0$. Recall that for any point P of C, by $R(P)$ we denote the local ring of P on C and by $R^*(P)$ we denote the normalization of $R(P)$. So far we have introduced various integers $\delta^*(P), \delta^*(P)_{\text{calg}}, \cdots$ to measure how much C is singular at P; now we proceed to compare them. To begin with:

Aphorism 5. Derivative equals Different times conductor.

Precise Meaning. In view of the characterization of $\delta^*(P)$ given in formula (11*) of §46 and the definition of $\delta^*(P)_{\text{calg}}$ given in formula (*) of §30, we interpret this Aphorism to mean that at every point P of C we have

$$\delta^*(P) = \delta^*(P)_{\text{calg}}.$$

Commentary. Since both sides measure how much C is singular at P, one may guess them to be equal. Dedekind showed this to be the case. Namely, by Lagrange Interpolation and with the intervention of the concept of "complementary module" invented by him, Dedekind proved that for P at finite distance one has

$$\deg_P(f_Y(x, y)) = \deg_P \text{Diff}(C, x) + R(P)\text{-length of } R^*(P)/\text{Cond}(R(P)).$$

Aphorism 6. Adjoint is to geometry as conductor is to college algebra.

Precise Meaning. Let $\xi(X, Y)$ be any polynomial and let P be any point of C at finite distance. Then the plane curve $\xi(X, Y) = 0$ is adjoint to C at P iff $\xi(x, y)$ belongs to $\text{Cond}(R(P))$.

Commentary. This definitive form of Aphorism 2 of §40 is indicated by Aphorism 5. For another suggestive reasoning, let us consider, *at a point P of C* at finite distance, "Noether's Fnd Th Refined" (§41) and the "Warning" following it. We regard $F(X, Y, Z)$ as obtained by homogenizing $f(X, Y)$. The Refined Theorem roughly says that, when H satisfies certain conditions at P, we can write

(*) $$H(x, y, 1) = B(x, y, 1)\Phi(x, y, 1) \text{ in } R(P),$$

where B will then satisfy certain other conditions at P which, geometrically speaking, (and by taking the "warning" not too seriously) say that B is adjoint to F at P; and conversely. Symbolically: conditions on B = conditions on H/Φ = conditions on H − conditions on Φ. At any rate, we want to show that H/Φ is in $R(P)$; since $R^*(P)$ is a "nicer" ring, a standard algebraic procedure of doing this would be to first show that H/Φ is in $R^*(P)$ and then to hope that it is actually in $\text{Cond}(R(P))$ and

hence in $R(P)$; in other words, solve (*) first in $R^*(P)$ and then let the conductor bring it down to $R(P)$; now the conditions put on H seem to guarantee that H/Φ is in $R^*(P)$ and so we could carry out this procedure if we knew that: geometrically "$H/\Phi(=B)$ is adjoint to F at P" corresponds algebraically to "H/Φ is in Cond$(R(P))$." In any case, this reasoning at least indicates that Aphorism 6 and Noether's Fnd Th Refined are closely related, which bring us to:

MORE COMMENTARY. By induction on the number of quadratic transformations needed to resolve the singularity of C at a given point P, and by analyzing the effect of a single quadratic transformation, in one breath one can prove (for instance see my *Calcutta lecture* cited in §31) the following four things:

(1) Aphorism 6.
(2) Noether's Fnd Th Refined (§41).
(3) Noether's Fnd Th in new guise (§31): $2\delta(P)_{\text{calg}} = \delta^*(P)_{\text{calg}}$.
(4) With $\delta(P)_{\text{geom}}$ as defined in (*) of §42: $2\delta(P)_{\text{geom}} = \delta^*(P)_{\text{calg}}$.

CONCLUSION. Thus at every point P of C we have the *equation of concordance*

$$\delta^*(P)_{\text{halg}} = \delta^*(P) = \delta^*(P)_{\text{calg}} = 2\delta(P)_{\text{calg}} = 2\delta(P)_{\text{geom}}$$

where $\delta^*(P)_{\text{halg}}$ was defined in formula (*) of §45, and the first equality was noted in formula (6) of §46. Now let N be the degree of C. By regarding

$$(g_2)^* \qquad 2g-2 = \deg(r(x,y)\,dx)$$

as defining the genus g of C, in §46 we deduced certain formulas (g_{11}) and (g_{12}) for g and then in §47 we deduced the formula

$$(g_{13}) \qquad g = (1/2)(N-1)(N-2) - (1/2)\Sigma\,\delta^*(P)$$

where the sum is over all points P of C. Now the above equation of concordance converts (g_{13}) into the formulas

$$(g_{10}) \qquad g = (1/2)(N-1)(N-2) - (1/2)\Sigma\, g^*(P)_{\text{halg}}$$

$$(g_6) \qquad g = (1/2)(N-1)(N-2) - (1/2)\Sigma\,\delta^*(P)_{\text{calg}}$$

$$(g_7) \qquad g = (1/2)(N-1)(N-2) - \Sigma\,\delta(P)_{\text{calg}}$$

$$(g_8) \qquad g = (1/2)(N-1)(N-2) - \Sigma\,\delta(P)_{\text{geom}}$$

asserted in §45, §30, §31, §42, respectively; these four formulas may also be regarded as refinements of the heuristic formula (g_5) of §13. Finally, formula (g_9) of §42 was only an obvious variation of (g_8). Observe that we have dealt with the equality $\delta^*(P)_{\text{halg}} = \delta^*(P)$ and hence formula (g_{10}) only in zero characteristic; however, let it be remarked that, upon replacing Newton-Puiseux expansion by Hamburger-Noether expansion, the characteristic pairs can be suitably reinterpreted so that the equality $\delta^*(P)_{\text{halg}} = \delta^*(P)$ and formula (g_{10}) hold also in nonzero characteristic.

This completes an outline of the proof of most of the genus formulas. Needless to say that there are numerous other arrangements of proof, which is hardly surprising: Indeed, ours is an old and venerable subject which has been worked over by many masterminds!

50. Valuations. Let again C: $f(X, Y) = 0$ be an irreducible plane curve, and let x and y be elements such that x is a variable and $f(x, y) = 0$. By K we denote the "coefficient field" and we note that we are assuming K to be algebraically closed. We put

$$K(x) = \text{field of rational functions } r(x)$$

and

$$K(C) = K(x, y) = \text{function field of } C, \text{ i.e., the field of all rational functions } r(x, y).$$

For any branch V of C, the "mapping" which associates $V(r)$ to every element r in $K(C)$ is a *valuation* of $K(C)$, i.e.: as r ranges over $K(C)$, $V(r)$ ranges over all integers together with ∞; $V(r)=\infty$ iff $r=0$; and for all r and s in $K(C)$ we have $V(rs)=V(r)+V(s)$ and $V(r+s)\geqq \min(V(r), V(s))$.

It can be shown that in this manner we get all the valuations of $K(C)$. In other words, there is a one-to-one correspondence between all branches of C and all the valuations of $K(C)$. We shall continue to denote a branch of C and the corresponding valuation of $K(C)$ by the same letter.

For any point P of C we define the number of branches of C centered at P, in the sense of college algebra, by the equation

$$\beta(P)_{\text{calg}} = \text{number of valuations } V \text{ of } K(C) \text{ such that } V(r)\geqq 0 \text{ for all } r \text{ in } R(P);$$

one can show that these valuations precisely correspond to the branches of C centered at P, and hence $\beta(P)_{\text{halg}} = \beta(P)_{\text{calg}}$ where the former was defined in formula (3*) of §46; it can also be shown that

$$R^*(P) = \text{the set of all } s \text{ in } K(C) \text{ for which } V(s)\geqq 0$$
$$\text{for every valuation } V \text{ of } K(C) \text{ such that}$$
$$V(r)\geqq 0 \text{ for all } r \text{ in } R(P).$$

By induction on the number of quadratic transformations needed to resolve the singularity P one can also see that $\beta(P)_{\text{calg}} = \beta(P)_{\text{geom}}$ where the latter was defined in formula (*) of §39. Thus we have the following *equation of concordance*:

$$\beta(P)_{\text{halg}} = \beta(P)_{\text{calg}} = \beta(P)_{\text{geom}}.$$

If D: $\phi(X, Y)=0$ is any other plane curve then for any point P of C at finite distance we define the intersection multiplicity of C and D at P, in the sense of college algebra, by the equation

$$i(C, D; P)_{\text{calg}} = \Sigma(V(\phi(x, y))) \text{ with sum over all valuations } V \text{ of } K(C)$$
$$\text{such that } V(r)\geqq 0 \text{ for all } r \text{ in } R(P);$$

with suitable modification we define this also when P is a point of C at infinity. In view of formula (4) of §46 and what we have said above, it follows that $i(C, D; P) = i(C, D; P)_{\text{calg}}$. Again, by induction on the number of quadratic transformations needed to resolve the singularity P, one can also show that $i(C, D; P)_{\text{calg}} = i(C, D; P)_{\text{geom}}$ where the latter is as defined in formula (*) of §41. Thus we have the following *equation of concordance*:

$$i(C, D; P)_{\text{halg}} = i(C, D; P) = i(C, D; P)_{\text{calg}} = i(C, D; P)_{\text{geom}}$$

where the first equality holds in characteristic zero and was noted in formula (9) of §44.

Let J be a field and let L be a "finite algebraic field extension of J," i.e.: L is a field; L contains J; and there is a positive integer n such that if $z_1, \cdots, z_m$ are any elements of L which are linearly independent over J, then necessarily $m\leqq n$. For a valuation of W of L, by the *ramification index* of W over J is meant the smallest positive integer e such that $W(r)=e$ for some $r\in J$. For a valuation U of J by an *extension* of U to L is meant a valuation W of L such that: $W(r)\geqq 0$ for all r in J for which $U(r)\geqq 0$. A valuation U of J is said to be *ramified* in L if the ramification index, over J, of some extension of U to L is >1.

It can easily be seen that the valuations of $K(x)$ are in one-to-one correspondence with the values of x, finite (i.e., in K) or infinite. Namely, the valuation U of $K(x)$ and the value a of x correspond if: $U(x-a)=1$ in case $a\neq\infty$, and $U(x^{-1})=1$ if $a=\infty$; it is clear that a is a branch point iff U is ramified in $K(C)$. Also it is clear that for a valuation V of $K(C)$, the ramification index of V over $K(x)$ equals what, in §46, we called the branching degree of the corresponding branch of C.

51. Galois. The quadratic equation

$$T^2 + bT + c = 0$$

can be solved:

$$t = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

Similarly, equations of degree 3 and 4 were solved by radicals; i.e., by extractions of square roots, cube roots, etc.; (in this section we are supposing characteristic to be zero). Then attention was turned to the question whether general equations of degree 5 can be solved by radicals. Abel and Galois (1830) independently proved this to be impossible. Galois' proof went something like this: the general equation of degree 5 cannot be solved by radicals because the "galois group" of the "splitting field" of the said equation is "unsolvable." We shall now briefly explain the concepts of "galois group" and "splitting field" which were invented by Galois.

Let J be a field. By a galois extension of J is meant a finite algebraic field extension L of J such that: whenever a one-variable irreducible polynomial with coefficients in J has one root in L, then all its roots are in L. For a galois extension L of J, by the galois group of L over J is meant the group of all automorphism τ of L such that $\tau(z) = z$ for every z in J; Galois showed that this group is finite and its order (i.e., the numbers of elements in it) equals the (largest) number of elements in L which are linearly independent over J. Finally, for a one-variable polynomial f with coefficients in J, by the splitting field of f over J is meant the smallest field L containing J such that all the roots of f are in L; it is easily seen that L is then a galois extension of J.

Let J be a field, let L be a galois extension of J, and let W be a valuation of L. Following Hilbert (1897, [33]), we define the *inertia group* of W over J to be the subgroup of the galois group of L over J consisting of all those automorphisms τ such that: for every z in L with $W(z) \geqq 0$, we have $W(z - \tau(z)) > 0$. It turns out that the order of the inertia group equals the ramification index of W over J. It can also be shown that the inertia group is cyclic (i.e., is "generated" by a single element) if in addition to assuming J to be of characteristic zero we also assume $W(m) = 0$ for every positive integer m in J. In a sense, this theorem can be viewed as half of a college algebra proof of Newton's Theorem on Puiseux expansion.

Chapter IX: More Function Theory

52. Riemann surface. Let now the coefficient field K be the field of complex numbers, and consider the n-valued function y of x defined by

$$f(X, Y) = u_0(X)Y^n + u_1(X)Y^{n-1} + \cdots + u_n(X) = 0,$$

where the $u_i(X)$ are polynomials with $u_0(X) \not\equiv 0$, and f is irreducible. Also consider the plane curve C: $f(X, Y) = 0$. As said in §4, to clarify the meaning of abelian integrals, Riemann introduced a surface S, called the Riemann surface of f, on which y "becomes" single valued.

The field $K(C) = K(x, y)$ now becomes the field of all (single-valued meromorphic) functions on S. Moreover, for each point V of S, there is a function t on S, near V, such that t has a simple zero at V; i.e., near V, S looks like a disc around the origin in the complex t-plane.

It turns out that the points of S are in a natural one-to-one correspondence with the branches of C; so we shall identify them.

Let S_0 be the complex x-sphere, i.e., the complex x-plane together with ∞. Let us say that the point V of S lies above the point $x(V)$ of S_0. Then for any point a of S_0, upon letting

(*) $\mu(a)$ = the number of points V of S which lie above a (i.e., for which $x(V) = a$),

by (15) and (16) of §46 we have

$$(**) \qquad \mu(a) = n - \sum_{x(V)=a} [e(V)-1] \leqq n$$

and:

$$a \text{ is a branch point iff } \mu(a) < n.$$

Now if $a \neq \infty$ is a branch point then it is a discriminant point (but not conversely!); therefore, in all there are only a finite number of branch points. Finally, for any point V of S we have that: if $x(V) = a \neq \infty$ then $(x-a)^{1/e(V)}$ is a uniformizing parameter at V; and if $x(V) = \infty$ then $x^{-1/e(V)}$ is a uniformizing parameter at V. We summarize all this by saying that *S is an n-sheeted branched covering of S_0*.

The original construction of S went something like this: On S_0 draw an open connected polygon B_0 with the discriminant points, together with ∞, as vertices. (See Fig. 10.)

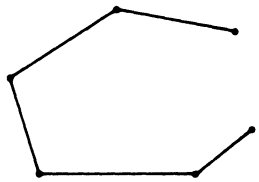


FIG. 10

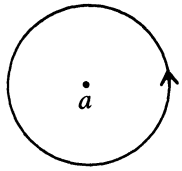


FIG. 11

Slit S_0 along B_0 to obtain S_0'; we suppose that S_0' is connected (i.e., does not fall into pieces) and simply connected (i.e., any closed path on S_0' can be shrunk to a point). Now as x varies in S_0', the n values of y can be sorted out into n analytic functions $y_1(x), \cdots, y_n(x)$; [by the "monodromy theorem" which says that one gets the same analytic continuation along two paths which can be "deformed" into each other]. Take n copies $S_1', \cdots, S_n'$ of S_0' and "put" $y_i(x)$ on S_i'. Now S is formed by suitably gluing together $S_1', \cdots, S_n'$ along the edges of $B_1, \cdots, B_n$ to reflect "the manner in which $y_1(x), \cdots, y_n(x)$ permute as we cross B_0."

Moreover: Given any point a of S_0, as we make analytic continuations along a small circle around a, (see Fig. 11), the functions $y_1(x), \cdots, y_n(x)$ will undergo a permutation. We write this permutation as a product of disjoint cycles. [For example

$$\begin{pmatrix} 123456 \\ 135624 \end{pmatrix} = (1)(235)(46);$$

the permutation group being "noncommutative" we write it "multiplicatively" instead of "additively."] It turns out that the number of these cycles is $\mu(a)$ and there is a natural one-to-one correspondence between them and the points $V_1, \cdots, V_{\mu(a)}$ of S lying above a, so that for $i = 1, \cdots, \mu(a)$ we have

$$e(V_i) = \text{the length of the corresponding cycle (or, the}$$
$$\text{number of sheets of } S \text{ which come together at } V_i).$$

[In the above example: $\mu(a) = 3$, $e(V_1) = 1$, $e(V_2) = 3$, $e(V_3) = 2$.]

53. Monodromy group. On S_0 mark a finite number of points $a_1, \cdots, a_m$ so as to include all the branch points. Also fix a point a different from the a_i. As we make analytic continuations along various closed paths beginning and ending at a, but without hitting any of the points a_i, we get various permutations of $(y_1(x), \cdots, y_n(x))$; the totality of these permutations form a group which is called the

monodromy group of f. The monodromy group is a certain subgroup of the group of all permutations on n symbols; we note that the latter is of order $n!$. The assumption of f being irreducible is reflected in the monodromy group being transitive; i.e., given any j with $1 \leqq j \leqq n$, some member of the monodromy group will carry $y_1(x)$ to $y_j(x)$.

To obtain generators for the monodromy group, we draw a suitable system of loops $A_1, \cdots, A_m$ beginning and ending at a, so that A_i makes a small circle around a_i. (See Fig. 12.)

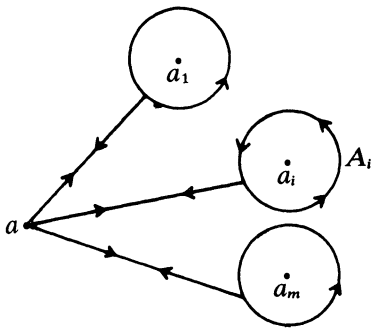


FIG. 12

Let τ_i be the permutation undergone by $(y_1(x), \cdots, y_n(x))$ as we make analytic continuation along A_i. By the monodromy theorem it follows that: (1) $\tau_1 \cdots \tau_m = 1$ (the identity permutation) and $\tau_1, \cdots, \tau_m$ generate the monodromy group.

One of Riemann's existence theorems says that conversely: (2) if any permutations $\tau_1, \cdots, \tau_m$ on n symbols are given such that $\tau_1 \cdots \tau_m = 1$ and the group generated by $\tau_1, \cdots, \tau_m$ is transitive, then there exists f, having no branch points outside the a_i, so that, for $i = 1, \cdots, m$, the roots of f permute according to τ_i when analytically continued along A_i.

It can be shown that the monodromy group is, in a natural manner, isomorphic to the galois group, over $K(x)$, of the splitting field of $f(x, Y)$ over $K(x)$. This in conjunction with (1) and (2) yields the following

Theorem about galois theory of algebraic curves. Let $U_1, \cdots, U_m$ be any valuations of $K(x)$. Then:

(1*) Let L be any galois extension of $K(x)$ such that no valuation of $K(x)$, different from the U_i, is ramified in L. Then for $i = 1, \cdots, m$, we can pick an extension W_i of U_i to L and also pick a generator τ_i of the splitting group of W_i over $K(x)$ such that $\tau_i \cdots \tau_m = 1$ and $\tau_1, \cdots, \tau_m$ generate the galois group of L over $K(x)$. Whence, in particular, the galois group of L over $K(x)$ is generated by $m - 1$ generators.

(2*) In this manner, by varying L, we can realize every finite group on $m - 1$ generators as the galois group of a galois extension of $K(x)$ in which no valuation other than $U_1, \cdots, U_m$ is ramified.

Now the *statement* of the above theorem is purely algebraic. However, the only available proof is the above sketched function-theoretic and topological argument of Riemann. To obtain a purely algebraic treatment of this theorem, is certainly one of the fundamental problems of algebraic geometry. (For some related matter see [**3**], and Chapter VIII of the second edition of [**80**] including the Appendices.)

In a sense, the following may be regarded as a number-theoretic analogue:

Problem about galois theory of algebraic numbers. Can any finite group be realized as the galois group of some galois extension of the field of rational numbers?

In a brilliant paper (1954, [**61**]), Šafarevič has answered this affirmatively for any solvable group. The general case still awaits solution.

While on the subject of analogy, on a simpler level, let us note that the topological simply-connectedness of the sphere and the plane have, respectively, the following algebraic consequences:

(3) $K(x)$ has no unramified extension, [i.e., in every finite algebraic extension of $K(x)$, other than $K(x)$, at least one valuation of $K(x)$ must be ramified].

(4) In every finite algebraic extension of $K(x)$, other than $K(x)$, at least two valuations of $K(x)$ must be ramified.

For this simpler situation a purely algebraic proof does already exist. Namely (4), and hence (3), immediately follow from formulas (g_{11}) and (13) of §46 by noting that the genus is always nonnegative. In fact the same formulas show that (3) remains valid also in nonzero characteristic.

This time the number-theoretic analogue is the following beautiful

THEOREM OF KRONECKER. *The field of rational numbers has no unramified extension.*

54. Sphere with handles. In §4 we said that S is topologically equivalent to a sphere with a certain number of handles. There we also said that, using formula (g_2) and Euler's Theorem, Riemann showed that for S:

(g_4) $$g = \text{number of handles.}$$

Now in §46, from formula (g_2) we have deduced formula

(g_{12}) $$2-2g = 2n - \Sigma[e(V)-1] \text{ with sum over all points } V \text{ of } S.$$

So our explanation of Riemann's deduction will be completed by deducing (g_4) from (g_{12}) and Euler's Theorem. To do this, draw a sufficiently fine polyhedron Δ on S_0 which includes all the branch points amongst its vertices; let Δ_0, Δ_1, and Δ_2 be the number of vertices, edges, and faces of Δ. Let Δ^* be the polyhedron on S obtained by "lifting" Δ; let Δ_0^*, Δ_1^*, and Δ_2^* be the number of vertices, edges, and faces of Δ^*. Let g^* be the number of handles of S. In view of (*) and (**) of §52 we see that

$$\Delta_0^* = n\Delta_0 - \Sigma[e(V)-1] \text{ with sum over all points } V \text{ of } S$$

and clearly $\Delta_1^* = n\Delta_1$ and $\Delta_2^* = n\Delta_2$. By Euler's Theorem we have

$$\Delta_0^* - \Delta_1^* + \Delta_2^* = 2-2g^* \text{ and } \Delta_0 - \Delta_1 + \Delta_2 = 2.$$

Combining the above three displayed lines we get

$$2-2g^* = 2n - \Sigma[e(V)-1]$$

and hence by (g_{12}) we get $g^* = g$ which proves (g_4).

55. Abel and Jacobi. Let $\Gamma_1, \cdots, \Gamma_{2g}$ be closed paths on S where Γ_i and Γ_{g+i} are a latitude and a longitude of the ith handle. (See Fig. 13.)

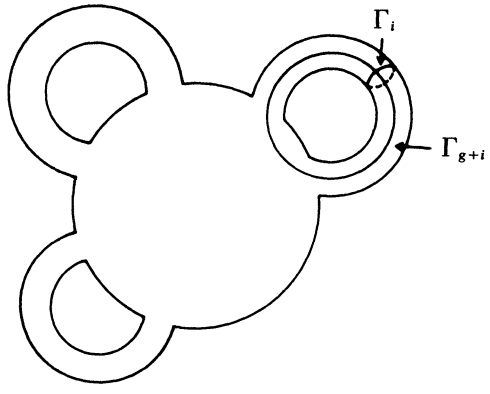


FIG. 13

Now any closed path Γ on S can be "deformed into $m_1\Gamma_1 + \cdots + m_{2g}\Gamma_{2g}$ with integers m_i" and then for any integral of first kind we have

$$\int_\Gamma r(x, y)\,dx = \sum m_i \int_{\Gamma_i} r(x, y)\,dx;$$

thus, for points Q and V of S, the value of the integral

$$\int_Q^V r(x, y)\,dx$$

is only determined up to a linear combination, with integer coefficients, of the $2g$ numbers

$$\int_{\Gamma_1} r(x, y)\,dx, \cdots, \int_{\Gamma_{2g}} r(x, y)\,dx.$$

To state the theorems of Abel and Jacobi, mentioned in §3, let

$$\Omega_1 = r_1(x, y)\,dx, \cdots, \Omega_g = r_g(x, y)\,dx$$

be linearly independent differentials of first kind, where the $r_i(x, y)$ are rational functions, and let

$$\gamma_{ij} = \int_{\Gamma_i} \Omega_j \text{ for } i = 1, \cdots, 2g \text{ and } j = 1, \cdots, g.$$

Also fix any points $Q_1, Q_2, \cdots$ on S.

ABEL'S THEOREM. *Let $V_1, \cdots, V_d, V_1', \cdots, V_d'$ be any points on S. Then $V_1, \cdots, V_d$ and $V_1', \cdots, V_d'$ are the zeros and poles of some rational function $s(x, y)$ iff for $j = 1, \cdots, g$ we have*

$$\sum_{k=1}^{d} \int_{Q_k}^{V_k} \Omega_j \equiv \sum_{k=1}^{d} \int_{Q_k}^{V_k'} \Omega_j \text{ modulo } ((\gamma_{ij})),$$

i.e., iff, upon taking any paths Λ_k and Λ_k' joining Q_k to V_k and V_k', for $j = 1, \cdots, g$ we have

$$\sum_{k=1}^{d} \int_{\Lambda_k} \Omega_j = \sum_{k=1}^{d} \int_{\Lambda_k'} \Omega_j + \sum_{k,i} n_{ki}\gamma_{ij} \text{ with integers } n_{ki}.$$

JACOBI'S INVERSION THEOREM. *Given any general g-tuple $(z_1, \cdots, z_g)$ of complex numbers, there is a unique g-tuple of points $(V_1, \cdots, V_g)$ of S such that for $j = 1, \cdots, g$ we have*

$$\sum_{k=1}^{g} \int_{Q_k}^{V_k} \Omega_j \equiv z_j \text{ modulo } ((\gamma_{ij})).$$

The resulting $2g$ functions

$$x(V_1(z_1, \cdots, z_g)), y(V_1(z_1, \cdots, z_g)), \cdots, x(V_g(z_1, \cdots, z_g)), y(V_g(z_1, \cdots, z_g))$$

of the g variables $z_1, \cdots, z_g$ are (single-valued meromorphic) periodic function with the $2g$ periods $(\gamma_{i1}, \cdots, \gamma_{ig})$, $i = 1, \cdots, 2g$.

These multiply periodic functions are called abelian functions. They can be viewed as functions on the g-dimensional (complex dimension g = real dimension $2g$) multi-torus obtained from the space of g complex variables by "dividing" it by the $2g$ periods. The said multi-torus is called the jacobian variety of f (or, of C; or, of $K(x, y)$):

(g_{14}) g = dimension of the jacobian variety.

In Abel's theorems, when we spoke of $V_1, \cdots, V_d$ and $V_1', \cdots, V_d'$ being the zeroes and poles of $s(x, y)$, what we precisely meant was that: given any point V of S, upon letting p to be the number of

times V occurs amongst $V_1, \cdots, V_d$ and upon letting p' to be the number of times V occurs amongst $V_1', \cdots, V_d'$ we have $V(s(x, y)) = p - p'$.

In analogy with number theory, the purport of the theorems of Abel and Jacobi is sometimes expressed by saying that the jacobian variety is the "divisor class group" of f. Roughly speaking: the various tuples of points $(V_1, \cdots, V_d; V_1', \cdots, V_d')$ of S constitute the "group of divisors of degree zero"; the subgroup of "principal divisors" or "divisors linearly equivalent to zero" consists of those tuples for which $V_1, \cdots, V_d$ and $V_1', \cdots, V_d'$ are the zeroes and poles of some rational function $s(x, y)$; finally, the divisor class group is the factor group of the group of divisors of degree zero modulo the subgroup of principal divisors.

To describe the analogous situation in number theory, let E be the integral closure of the ring of ordinary integers in a finite algebraic field extension M of the field of rational numbers. E is called the ring of (algebraic) integers in M. It was found that E need not be a UFD (= unique factorization domain), i.e., in E one may not have the unique factorization property of ordinary integers. To estimate how far the said property fails in E, one can consider the totality of all (nonzero) fractional E-ideals, i.e., finitely generated E-modules contained in M. The fractional E-ideals form a group under multiplication; the principal fractional E-ideals (i.e., those which consist of all the multiples of a single element) form a subgroup; the factor group of the former modulo the latter is called the divisor class group of M, and the order of the divisor class group is called the class number of M. Clearly the class number is 1 iff E is a UFD. The divisor class group of M is evidently the number-theoretic analogue of the jacobian variety of $K(x, y)$. Moreover, in a sense, the number-theoretic analogue of the finite dimensionality of the jacobian variety is the

THEOREM OF DEDEKIND AND KRONECKER which says that *the class number is always finite.*

Let us conclude this Chapter with the observation that the Riemann surface (sphere with g handles) and the jacobian variety (g-dimensional multi-torus) are two distinct generalizations of the surface of a torus (= doughnut or bicycle-tube).

CHAPTER X: MORE UNIVERSITY ALGEBRA

56. Serre. To again recall the theorem of Euler (and Riemann):

$$\textit{genus of a curve in the complex case} = \sum_{i=0}^{2} (-1)^i \Delta_i^*,$$

where Δ_0^*, Δ_1^*, and Δ_2^* are the number of vertices, edges and faces of a polyhedron drawn on the Riemann surface of the curve. Serre (1966, [65]) has made the following far-reaching generalization of this:

$$\textit{arithmetic genus of a variety of any dimension and in any characteristic} = \Sigma(-1)^i h^i,$$

where each of the individual h^i (unlike Euler's Δ_i^*) has an invariant meaning, namely, it is the number of linearly independent elements in the "ith cohomology group of the variety."

References

1. N. H. Abel, Mémoire sur une propriété générale d'une classe très-étendue de fonctions transcendantes, 1826, Oeuvres Complètes, New Edition (1881), vol. I, pp. 145–211; and Démonstration d'une propriété générale d'une certaine classe de fonctions transcendantes, Crelle Journal, 4 (1829) 200–201.

2. S. S. Abhyankar, Local uniformization on algebraic surfaces over ground fields of characteristic $p \neq 0$, Ann. of Math., 63 (1956) 491–526.

3. S. S. Abhyankar, Coverings of algebraic curves, Amer. J. Math., 79 (1957) 825–856.

4. ———, Resolution of singularities of arithmetical surfaces, Purdue conference on Arithmetical Algebraic Geometry, pp. 111–152, New York, 1965.

5. ———, Resolution of singularities of embedded algebraic surfaces, Academic Press, New York, 1966.

6. ———, On the problem of resolution of singularities of algebraic varieties, Moscow International Conference, Moscow, 1968, pp. 469–481.

7. ———, Singularities of algebraic curves, Analytic Methods in Mathematical Physics, New York, 1968, pp. 3–14.

8. ———, A glimpse of algebraic geometry, University of Poona, Poona, 1969.

9. ———, Algebraic space curves, Montreal, 1970.

10. S. S. Abhyankar and T. T. Moh, Newton-Puiseux expansion and generalized Tschirnhausen transformation, Crelle Journal, 260 (1973) 47–83 and 261 (1973) 29–54.

11. S. S. Abhyankar and T. T. Moh, Embeddings of the line in the plane, Crelle Journal, 276 (1975) 148–166.

12. H. Bass, On the ubiquity of Gorenstein rings, Math. Z., 82 (1963) 8–28.

13. L. Berzolari, Allgemeine Theorie der höheren ebenen algebraischen Kurven, Enzyklopädie der Math. Wiss., vol. III, Part 2^1, Leipzig, 1906.

14. É. Bézout, Théorie générale des équations algébriques, Paris, 1770.

15. M. Bocher, Introduction to higher algebra, Macmillan, New York, 1907.

16. W. S. Burnside and A. W. Panton, Theory of equations, vols. I and II, Dublin, 1904.

17. H. Cartan and S. Eilenberg, Homological algebra, Princeton Univ. Press, 1956.

18. A. Cayley, On the intersection of Curves, Math. Ann., 30 (1887) 85–90.

19. C. Chevalley, On the theory of local rings, Ann. of Math., 44 (1943) 690–708.

20. ———, Intersections of algebraic and algebroid varieties, Trans. A.M.S., 57 (1945) 1–85.

21. G. Chrystal, Algebra, Parts I and II, Edinburgh, 1886.

22. H. Clemens and P. Griffiths, The intermediate jacobian of the cubic threefold, Ann. of Math., 95 (1972) 281–356.

23. I. S. Cohen, On the structure and ideal theory of complete local rings, Trans. A.M.S., 59 (1946) 54–106.

24. G. Cramer, Introduction à l'analyse des lignes courbes, Geneva, 1750.

25. R. Dedekind and H. Weber, Theorie der algebraischen Funktionen einer Veränderlichen, Crelle Journal, 92 (1882) 181–290.

26. L. Euler, Introductio in analysin infinitorum, Berlin Academy, 1748.

27. W. Gröbner, Über irreduzible Ideale in kommutativen Ringen, Math. Ann., 110 (1934) 197–222.

28. A. Grothendieck, Eléments de géométrie algébrique, Chaps. I to IV, I.H.E.S., 1960–1967.

29. G. Halphen, Étude sur les points singuliers des courbes algébriques planes, Appendix (pp. 535–648) to French edition of Salmon [62], 1884.

30. M. Hamburger, Über die Entwicklungen algebraischen Funktionen in Reihen, Z. Math. Phys., 16 (1871) 461–491.

31. G. H. Hardy, A course of Pure Mathematics, Cambridge University Press, New York, 1908.

32. K. Hensel, Theorie der algebraischen Zahlen, Teubner, Leipzig, 1908.

33. D. Hilbert, Die Theorie der algebraischen Zahlkörper, Jahres. der Deutschen Math-Ver., 4 (1897) 175–546.

34. ———, Über das Dirichletsche Prinzip, Math. Ann., 59 (1904) 161–186.

35. H. Hironaka, Resolution of singularities of an algebraic variety over a field of characteristic 0, Ann. of Math., 79 (1964) 109–326.

36. ———, Characteristic polyhedra of singularities, J. Math. Kyoto Univ., 7 (1968) 251–293.

37. E. W. Hobson, Plane trigonometry, Cambridge Univ. Press, 1891. Repr. Chelsea, New York.

38. A. Hurwitz, Über die Trägheitsformen eines algebraischen Modulus, Ann. Mat. Pura Appl., 20 (1913) 113–151.

39. C. G. J. Jacobi, Considerationes generales de transcentibus Abelianis, Crelle Journal, 9 (1832) 394–403.

40. K. Knopp, Theory and application of infinite series, London, 1928. Hafner, New York.

41. J. König, Einleitung in die allgemeine Theorie der algebraischen Grössen, Leipzig, 1903.

42. L. Kronecker, Grundzüge einer arithmetischen Theorie der algebraischen Grössen, Crelle Journal, 92 (1882) 1–122.

43. W. Krull, Dimensionstheorie im Stellenringe, Crelle Journal, 179 (1938) 204–226.

44. ———, Elementare und klassische Algebra vom moderne Standpunkt, Parts I and II, De Gruyter, Berlin, 1952–1959.

45. C. Maclaurin, Geometria organica sive descripto linearum curverum universalis, London, 1720.

46. F. S. Macaulay, The algebraic theory of modular systems, Cambridge Univ. Press, London, 1916.

47. F. Mertens, Über die bestimmenden Eigenschaften der Resultante von n Formen mit n Veränderlichen, S.-B. Akad. Wissen. zu Wien, 93 (1886) 527–566.

48. M. P. Murthy, Generators for certain ideals in regular rings of dimension three, Comment. Math. Helv., 47 (1972) 179–184.

49. M. P. Murthy and J. Towber, Algebraic vector bundles over A^3 are trivial, Inventiones, 24 (1974) 173–189.

50. D. Mumford, Geometric invariant theory, New York, 1965.

51. M. Nagata, Local rings, Wiley, New York, 1962.

52. ———, A theorem of Gutwirth, J. Math. Kyoto Univ., 11 (1971) 149–154.

53. Isaac Newton, Geometria analytica, 1680.

54. M. Noether, Zur Theorie des eindeutigen Entsprechens algebraischer Gebilde von beliebig vielen Dimensionen, Math. Ann., 2 (1870) 293–316.

55. ———, Über einen Satz aus der Theorie der algebraischen Funktionen, Math. Ann., 6 (1873) 351–359.

56. ———, Les combinaisons caractéristiques dans la transformation d'un point singulier, Rend. Cir. Math. Palermo, 1 (1890) 89–108.

57. O. Perron, Algebra, 2 volumes, Springer, Berlin, 1927.

58. V. A. Puiseux, Recherches sur les fonctions algébriques, J. Math., 15 (1850) 365–480.

59. B. Riemann, Theorie der Abelschen Funktionen, Crelle Journal, 54 (1857) 115–155.

60. G. Roch, Über die Anzahl der willkürlichen Constanten in algebraischen Funktionen, Crelle Journal, 64 (1865) 372–376.

61. I. R. Šafarevič, Construction of fields of algebraic numbers with given solvable Galois group, Izvestia, 18 (1954) 525–578.

62. G. Salmon, A treatise on the higher plane curves, Dublin, 1852. Repr. Stechert, New York.

63. C. A. Scott, A proof of Noether's fundamental theorem, Math. Ann., 52 (1899) 593–597.

64. J. G. Semple and L. Roth, Introduction to algebraic geometry, Clarendon Press, Oxford, 1949.

65. J. P. Serre, Faisceaux algébriques cohérents, Ann. of Math., 61 (1955) 197–278.

66. F. Severi, Vorlesungen über algebraische Geometry, Teubner, Leipzig-Berlin, 1921.

67. H. J. S. Smith, On the higher singularities of plane curves, Proc. of London Math. Soc., 6 (1873) 153–182.

68. H. Stahl, Theorie der Abel'schen Funktionen, Leipzig, 1896.

69. J. J. Sylvester, On a general method of determining by mere inspection the derivations from two equations of any degree, Philosophical Magazine, 16 (1840) 132–135.

70. E. W. V. Tschirnhausen, Acta Eruditorum, Leipzig, 1683.

71. B. L. van der Waerden, Moderne Algebra, 2 volumes, Springer, Berlin, 1931.

72. H. Weber, Lehrbuch der Algebra, 3 volumes, Braunschweig, 1894–1908. Reprint: Stechert-Hafner, New York.

73. K. Weierstrass, Vorbereitungssatz, Berlin University Lecture of 1860, contained in: Einige auf die Theorie der analytischen Funktionen mehrerer Veränderlichen sich beziehende Sätze, Mathematische Werke, II (1895) 135–188.

74. A. Weil, Foundations of algebraic geometry, New York, 1946.

75. O. Zariski, Some results in the arithmetic theory of algebraic varieties, Amer. J. Math., 61 (1939) 249–294.

76. ———, The reduction of the singularities of an algebraic surface, Ann. of Math., 40 (1939) 639–689.

77. ———, Local uniformization on algebraic varieties, Ann. of Math., 41 (1940) 852–896.

78. ———, Reduction of the singularities of algebraic three dimensional varieties, Ann. of Math., 45 (1944) 472–542.

79. ———, The fundamental ideas of abstract algebraic geometry, Proc. Internat. Cong. Math., Cambridge, 1950, pp. 77–89.

80. ———, Algebraic surfaces, Springer, Berlin, 1935 (repr. Chelsea, New York); second supplemented edition, Springer, Berlin, 1971.

81. H. G. Zeuthen, Sur la détermination d'une courbe algébrique par des points donnés, Math. Ann., 31 (1889) 235–251.

Division of Mathematical Sciences, Purdue University, West Lafayette, IN 47907.