Chapter 4
The Genus of an Algebraic Curve

Essay 4.1 Abel's Memoir

> *It appears to me that if one wants to make progress in the study of mathematics one should study the masters and not the pupils.*—Niels Henrik Abel, quoted from an unpublished source in [72, p. 138]

Niels Henrik Abel's submission of his *Mémoire sur une propriété générale d'une class très-étendue de functions transcendantes* to the Paris Academy in October 1826 should have been a high point in the history of mathematics. Instead, it was a low point in the history of the Paris Academy.

Abel, lonely and unknown, was temporarily in Paris thanks to a travel grant from the government of Norway, and he hoped to win recognition in the city that was then the mathematical capital of Europe. Unfortunately, he naively believed that recognition could be won by submitting a work of undeniable genius to Europe's leading mathematical institution. He did not understand that works of undeniable genius are inherently difficult to read, even for the most learned readers, and he did not understand that the members of Europe's leading mathematical institution would not devote the needed time and thought to the work of a 24-year-old mathematician who was unknown to them and who came from a country they had scarcely heard of.

Of course, one of the famous men of the Academy might by some lucky accident have taken notice of the memoir long enough to realize that it was worth pursuing, but none did. In 1837, eight years after Abel's untimely death, the Norwegian scholars charged with publishing Abel's collected works applied to the Academy *via the Norwegian government and its diplomatic representatives in Paris* for a copy of the memoir—Abel had apparently not kept a copy for himself—but the effort did not succeed, and the memoir is absent from the first publication of Abel's works in 1839. Finally, the Academy did publish the memoir in 1841 as [2], making it available to eager readers like C. G. J. Jacobi for the first time.

In the two and a half years Abel (shown in Fig. 4.1) lived after submitting the memoir, he enjoyed a growing reputation based on his publications in Crelle's

H. M. Edwards, *Essays in Constructive Mathematics*,
https://doi.org/10.1007/978-3-030-98558-5_4

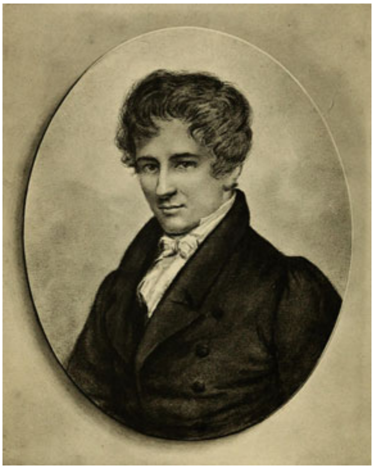

Fig. 4.1 Abel

Journal, but he patiently awaited publication of the Paris memoir, believing it would ensure his fame. He even alluded to the memoir in one of his published works, piquing the curiosity and indignation of Jacobi, who read the allusion too late to write to Abel about it. That Abel's memoir remained unpublished in his lifetime[1] deprived him of the challenge and encouragement of readers' responses and therefore probably deprived mathematics of important further work.

(Incredibly, the tragedy was repeated only three years after Abel died when Galois went to an early grave ignored by the same Paris Academy.)

Abel's memoir deals with *integrals of algebraic differentials*, a topic that is not at all easy to understand from the point of view of naive geometry and integration along a curve. Because an algebraic differential like $dx/\sqrt{1-x^4}$ is "many-valued" and because, moreover, an integral of such a differential depends on choosing both a path and a constant of integration, modern readers may well despair of understanding even what Abel *means* by the sum of a finite number of integrals of a given algebraic differential, much less why questions about such sums might be interesting or significant.

But there is another way to describe the main idea that makes better sense to modern readers and explains the main theorem of the memoir more clearly. Abel's "algebraic differentials" are differentials of the form $f(x, y)\,dx$, where f is a rational function of two variables and where y is an "algebraic function" of x. The notion of an "algebraic function" has become a source of unease for modern readers because an algebraic function is normally "many-valued" and the property of being single-valued is the essence of the set-theoretic notion of a "function." But of course there

[1] The last work Abel published was a brief note that contained a theorem from the memoir. Abel's biographer Oystein Ore says that the theorem of that last brief note is the theorem of the memoir [72, p. 219], but it is far short of the theorem in the introduction of the memoir that I am discussing in this essay and that I take to be, in Ore's phrase, "the main theorem from the Paris memoir."

are modern ways to deal with algebraic functions. One is to give the functions their own special domain; this is the source of the theory of Riemann surfaces. The other is to regard an "algebraic function" not as a function at all, but simply as an element of an algebraic function field, which is to say an algebraic field whose transcendence degree is positive (see Essay 2.2). The subject of Abel's memoir is algebraic functions of *one variable*, which is to say, in the terminology of Essay 2.2, elements of an algebraic field of transcendence degree 1. In other words, Abel is dealing with *the field of rational functions on an algebraic curve defined over the rationals*.

The concept that I propose as an aid to understanding Abel's memoir is that of an **algebraic variation** of a set of points on an algebraic curve. Abel describes such a variation as the solutions of a pair of equations

$$\begin{aligned} \chi(x, y) &= 0, \\ \theta(x, y, a, a', a'', \ldots) &= 0, \end{aligned}$$

where $\chi(x, y)$ is the irreducible polynomial with integer coefficients, monic in y, that defines the algebraic curve under discussion, and $\theta(x, y, a)$ is an auxiliary polynomial in x and y whose coefficients $a, a', a'', \ldots$ are indeterminates. For each fixed value of the coefficients $a, a', a'', \ldots$ the pair of equations determines a set of points $\{(x_k, y_k)\}$ on the curve $\chi = 0$, and as the coefficients vary, these points vary along the curve. A variation of points on the curve that can be generated in this way is an *algebraic variation*.

Somewhat more precisely, let C^N denote the set of all N-tuples of points on the curve C defined by $\chi(x, y) = 0$. An algebraic variation of a point of C^N is determined by choosing a $\theta(x, y, a)$ of the form $\theta(x, y, a) = \sum a_{ij} x^i y^j$, where the exponent pairs (i, j) are in some specified finite set. To say that $\theta(x, y, a) = 0$ at a particular point of the curve $\chi(x, y) = 0$ means that the parameters a_{ij} in θ satisfy a certain (linear) condition. Choose values for the a_{ij} that make $\theta = 0$ at all N of the given points. There will be *other* points of $\chi(x, y) = 0$ where $\theta = 0$ for these values of a_{ij}, say there are M of them. An algebraic variation of the N given points is one that results when the a_{ij} are allowed to vary from their fixed values in such a way that the M additional zeros all remain at zero while the N original ones are allowed to move. For each point of C^N, the points of C^N that can be reached from it by a sequence of algebraic variations lie on an algebraic subvariety of C^N.

Abel probably had some geometric conception of such variations of sets of points on $\chi(x, y) = 0$, but exactly what it might have been can only be guessed. Today one would never discuss intersection points without first specifying an algebraically closed ground field, but Abel would probably not have thought of curves as ordered pairs of complex numbers in anything like the modern way. More likely, he would have just imagined sets of points of intersection of an ordinary plane curve with an auxiliary curve and considered constraints on variations of the intersection points produced by varying the auxiliary curve. In modern terms, the number of *constraints* on the variation of N points of a curve is the *codimension* of the subvariety of algebraic variations within the N-dimensional variety C^N of all variations. This codimension is very nearly the same as the **genus** of the curve, and whatever his geometric conception of the problem setting may have been, it is this number that Abel successfully investigated.

In terms somewhat closer to Abel's, if $f(x, y)\,dx$ is an algebraic differential (which is to say, a rational function f on the curve C times the symbol dx), and if an infinitesimal algebraic variation of the points $\{(x_k, y_k)\}$ is performed, Abel asserts that the resulting variation of $\sum f(x_k, y_k)\,dx_k$ is a differential that can be expressed *rationally* i*n* terms of the parameters a_{ij} and their differentials.[2] Thus, if the point $(P_1, P_2, \ldots, P_N)$ can be moved to the point $(Q_1, Q_2, \ldots, Q_N)$ of C^N by an algebraic variation, then

$$\int_{P_1}^{Q_1} f(x, y)\,dx + \int_{P_2}^{Q_2} f(x, y)\,dx + \cdots + \int_{P_N}^{Q_N} f(x, y)\,dx$$

is equal to the integral of a *rational* differential in the a_{ij} and can therefore be expressed in terms of elementary functions—logarithms and trigonometric functions, as well as rational functions—of the a_{ij}.

Now let g be the codimension of the subvarieties of algebraic variations. Then $N - g$ of the points $(P_1, P_2, \ldots, P_N)$ can be moved in arbitrary ways by an algebraic variation, provided the remaining g points move in such a way as to keep the new $(P_1', P_2', \ldots, P_N')$ on the same subvariety. Thus, if O is a chosen base point on the curve C, there is an algebraic variation—or at least a succession of algebraic variations—of a point $(P_1, P_2, \ldots, P_N)$ of C^N that connects it to a point of C^N of the form $(O, O, \ldots, O, Q_1, Q_2, \ldots Q_g)$. Then

$$\int_O^{P_1} f(x, y)\,dx + \int_O^{P_2} f(x, y)\,dx + \cdots + \int_O^{P_{N-g}} f(x, y)\,dx$$
$$+ \int_{Q_1}^{P_{N-g+1}} f(x, y)\,dx + \cdots + \int_{Q_g}^{P_N} f(x, y)\,dx$$

can be expressed in terms of elementary functions of the parameters used in the variation, so that when the g integrals from O to Q_i are added, one obtains

$$\int_O^{P_1} f(x, y)\,dx + \int_O^{P_2} f(x, y)\,dx + \cdots + \int_O^{P_N} f(x, y)\,dx$$
$$= \int_O^{Q_1} f(x, y)\,dx + \cdots + \int_O^{Q_g} f(x, y)\,dx + E,$$

where E can be expressed in terms of elementary functions of the parameters of the variation. (The paths of integration are, of course, the ones determined by the algebraic variation from $(O, O, \ldots, O, Q_1, Q_2, \ldots Q_g)$ to $(P_1, P_2, \ldots, P_N)$ that is assumed.) Thus, disregarding elementary functions, *a sum of any number N of integrals of $f(x, y)\,dx$ can be expressed as a sum of just g integrals*, where g depends

[2] This, in essence, is the theorem of Abel's last published note that Ore mistook for the main theorem of the memoir. See the note above.

only on the differential $f(x, y)\,dx$ being integrated–and in fact depends only on the algebraic curve $\chi(x, y) = 0$ on which the differential has its existence—not on N.

This is the main theorem of Abel's Paris memoir. In Abel's own words, "The number of these conditions [the number g above] does not depend at all on the number of summands, but only on the nature of the particular integrands that one considers. Thus, for example, for an elliptic integrand this number is 1; for an integrand that contains no irrationalities but a radical of the second degree, under which the variable has degree at most six, the number of necessary conditions is 2, and so forth."[3]

I have said above that the crucial number of conditions g is "roughly" the genus of the curve C. Abel's statement that g is 1 in the elliptic case, 2 in case y^2 is a polynomial of degree 5 or 6 in x, and so forth, of course suggests that g is connected to the genus and *is* the genus in many cases. It fails to be the genus only because Abel bases his variation of the points on the variation of parameters a_{ij} in functions of the form $\theta(x, y, a) = \sum a_{ij}x^i y^j$, which is not quite general enough and in some cases gives too large a value for g because it omits certain variations that deserve to be called algebraic variations. When θ is instead taken to have the form $\theta(x, y, a) = \sum a_i\theta_i(x, y)$ where the "functions" $\theta_i(x, y)$ are integral over x—which may reduce g because it may include more variations—g becomes the actual genus, as will be shown in Essay 4.6.

Essay 4.2 Euler's Addition Formula

> *Man sollte weniger danach streben, die Grenzen der mathematischen Wissenschafflen zu erweitern, als vielmehr danach, den bereits vorhandenen Stoff aus umfassenderen Gesichtspunkten zu betrachten.* (One should strive less to extend the boundaries of the mathematical sciences and much more to treat the already available material from more comprehensive viewpoints.)—Eduard Study [38, p. 140]

Euler (shown in Fig. 4.2) stated his addition formula for elliptic integrals in a variety of ways, none of which shed enough light on the formula to suggest a generalization to other kinds of integrands. The great achievement of Abel's Paris memoir was to describe Euler's formula as the case $g = 1$ of a more general phenomenon.

It is customary today to describe an elliptic curve by a formula of the form $y^2 = x^3 + g_2x + g_3$, in which g_2 and g_3 are rational numbers, called its "Weierstrass normal form." When the curve is written in this form, the "addition" or "group law"

[3] In an effort to clarify Abel's statement, I have taken some liberties with the translation. His actual words were, "Le nombre de ces relations ne dépend nullement du nombre des fonctions, mais seulement de la nature des fonctions particulière qu'on considère. Ainsi, par exemple, pour une fonction elliptique ce nombre est 1; pour une fonction dont la dérivée ne contient d'autres irrationalités qu'un radical du second degré, sous lequel la variable ne passe pas le cinquième ou sixième degré, le nombre des relations nécessaires est 2, et ainsi de suite." His "fonctions" are the integrals above, and his "dérivées" are the integrands. What he is calling "une fonction elliptique" is what is today called an elliptic integral.

Fig. 4.2 Euler

on the curve is described as follows: Let P and Q be given points on the curve, and let S be the third point in which the line through P and Q intersects the curve. (The curve, being a cubic, intersects a line in the xy-plane in three points when they are counted in the right way.) The sum $R = P + Q$ of P and Q is defined to be the third point in which the line through S and the point at infinity intersects the curve, as shown in Fig. 4.3. (The lines through the point at infinity are the lines $x =$ constant—these are the lines that intersect the curve in only two finite points—so R is the point whose x-coordinate is the same as that of S and whose y-coordinate is the y-coordinate of S with the sign reversed.)

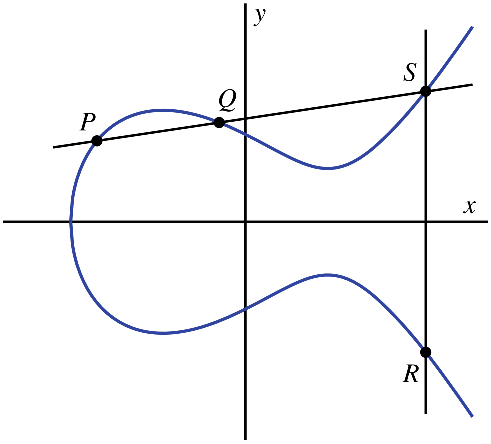

Fig. 4.3 The operation of addition on an elliptic curve

This construction is connected to the theorem of Abel's memoir in the following way: Let $\theta(x, y, a, b, c) = ax + by + c$. Algebraic variations of the given points P and Q are obtained by choosing initial values for a, b, and c that make θ zero at P and Q and allowing a, b, and c to vary in such a way that the third point of intersection of the line with the curve, call it S, remains fixed. In other words, the algebraic variations of the pair (P, Q) are the pairs of points (P', Q') on the curve for which P', Q', and S are colinear. In particular, if a point O of the curve is chosen as the origin—or the identity of the group law—then the algebraic variation of (P, Q) that carries P to O carries Q to the third point R in which the line through O and S intersects the curve. When O is chosen to be the point at infinity, R is the point $P + Q$ described above.

Abel's point of view explains why this "addition" is useful and shows that it is intrinsic to the curve. According to Abel, for any rational function $f(x, y)$ of x and $y = \sqrt{x^3 + g_2x + g_3}$, the sum $\int_O^P f(x, y)\,dx + \int_R^Q f(x, y)\,dx$ can be expressed in terms of integrals of rational functions, or, what is the same, $\int_O^P f(x, y)\,dx + \int_O^Q f(x, y)\,dx = \int_O^R f(x, y)\,dx$ plus an integral of rational functions. In particular, in the special case $f(x, y) = \frac{1}{y}$, in which the integrand is *holomorphic* in the sense explained in Essay 4.6, the formula is

$$\int_O^P \frac{dx}{y} + \int_O^Q \frac{dx}{y} = \int_O^R \frac{dx}{y},$$

which is one form of Euler's addition theorem. More precisely, these integrals depend on the paths of integration, and for the formula to hold, these paths must be chosen correctly. Thus, the sum of two integrals of dx/y can be expressed as just one integral of the same integrand, provided the limits of integration satisfy a certain algebraic relation and the paths of integration are chosen correctly.

Once it is known that two such integrals can be reduced to one, it follows that any number of such integrals can similarly be reduced to one. Abel's construction describes this all at once, rather than as a step-by-step reduction. An algebraic variation of a set of points $(P_1, P_2, \ldots, P_N)$ on the curve is described by a function $\theta(x, y, a)$ of the form $\sum a_{ij}x^iy^j$ for some selection of exponent pairs (i, j). Since y^2 is a polynomial in x, it is natural to assume that all of the chosen values of j are less than 2, so that θ takes the form $\phi_1(x) + \phi_2(x)y$, where ϕ_1 and ϕ_2 are polynomials in x containing terms of certain specified degrees whose coefficients are indeterminates a_{i1} and a_{i2}. The procedure is to give θ enough terms that values can be chosen for the parameters a_{ij} that make θ zero at the given points $P_1, P_2, \ldots, P_N$, and then to allow the parameters to vary from their chosen values in such a way that the value of θ remains at zero for the zeros of θ *other than* $P_1, P_2, \ldots, P_N$, while the N given zeros of θ are allowed to vary. The main question is, How many conditions are imposed on the variation of the N points along the curve by the requirement that the variation be describable in this way? That the answer is 1—that the genus of this curve is 1—can be seen in the following way.

The crucial step is to determine the number of zeros of $\theta(x, y) = \phi_1(x) + \phi_2(x)y$ on the curve. A simple way to do this is to make use of the idea that a rational function on an algebraic curve assumes each value the same number of times, when they are counted properly, and in particular that *the number of zeros is equal to the number of poles*. The function x assumes every value twice, and in particular, it has a double pole at the one point where $x = \infty$. The function y, on the other hand, assumes every value three times and has a triple pole at the one point where $x = \infty$. (These statements can be justified in various ways, but since they are used here only as heuristic devices, no formal justification will be given.) It follows that a polynomial $\phi(x)$ of degree ν has 2ν poles, all of them at the point where $x = \infty$, and that $\phi(x)y$ has $2\nu + 3$ poles, all at that same point. Consequently, if $\phi_1(x)$ has degree ν and $\phi_2(x)$ has degree $\nu - 2$, then $\theta(x, y) = \phi_1(x) + \phi_2(x)y$ has 2ν poles, so it also has that number of zeros. Since ϕ_1 has $\nu + 1$ variable coefficients and $\phi_2(x)$ has $\nu - 1$, θ has 2ν coefficients. If $2\nu > N$, the N conditions on the 2ν coefficients of θ imposed by the requirement that θ be zero at N given points can be satisfied by some choice of θ. Since θ has 2ν zeros, it has $2\nu - N$ zeros other than the N required ones, and the algebraic variations of the N given points are found by varying the 2ν coefficients of θ in such a way that these $2\nu - N$ extra zeros remain as zeros. The $2\nu - N$ conditions stating that θ must have these zeros are *independent*, so the coefficients of θ then vary with $2\nu - (2\nu - N) = N$ degrees of freedom. However, multiplication of θ by a constant does not change its zeros, so varying the coefficients of θ with N degrees of freedom varies its zeros with only $N - 1$ degrees of freedom. In short, an algebraic variation of N given points moves them in only $N - 1$ different directions, which is to say that algebraic variations satisfy *one* constraint in this case. Otherwise stated, algebraic variations describe subvarieties of codimension 1 in C^N.

This description of the phenomenon is in no way tied to the Weierstrass normal form. Gauss alludes indirectly to the elliptic curve $y^2 = 1 - x^4$ in the introduction to Section 7 of the *Disquisitiones Arithmeticae* [43] when he mentions the transcendental functions related to integrals of $dx/\sqrt{1 - x^4}$. Euler too dealt with the curve $y^2 = 1 - x^4$ [40], for which explicit and beautiful formulas can be developed for the addition law, and it is clear from Abel's published papers that this particular curve is one that he studied intensely. To require that it be put in Weierstrass normal form before the group law is described loses certain symmetries that deserve to be kept. But the above heuristic derivation of the fact that a curve in Weierstrass normal form has genus 1 also proves that $y^2 = 1 - x^4$ has genus 1, because i*n* this case x is ∞ at two points, both of them simple poles, whereas y has double poles at these points ($(\frac{y}{x^2})^2 = (\frac{1}{x})^4 - 1$ is finite when $x = \infty$), so $\theta(x, y) = \phi_1(x) + \phi_2(x)y$ has a ν-fold pole at each—and therefore 2ν zeros—when $\deg \phi_1 = \nu$ and $\deg \phi_2 = \nu - 2$. Again the number of parameters in such a function $\phi_1(x) + \phi_2(x)y$ is 2ν, and the same arguments then show that the algebraic variation of N points on the curve moves them with only $N - 1$ degrees of freedom and therefore determines subvarieties of C^N of codimension 1.

In the same way, Abel's construction generalizes the Euler addition formula to any curve C for which the algebraic variations describe subvarieties of C^N of codimension 1. If $(P_1, P_2, \ldots, P_N)$ is moved to $(O, O, \ldots, O, R)$ by means of an algebraic variation, then, as before,

$$(1)\quad \int_O^{P_1} f(x,y)\,dx + \int_O^{P_2} f(x,y)\,dx + \cdots + \int_O^{P_N} f(x,y)\,dx = \int_O^{R} f(x,y)\,dx + E,$$

where E is a quantity that can be expressed in terms of integrals of rational functions. (Moreover, E is zero when the integrand is *holomorphic* in the sense defined in Essay 4.6. This is true, as will be shown, of the integrand dx/y for curves in Weierstrass normal form or for the curve $y^2 = 1 - x^4$.)

Essay 4.3 An Algebraic Definition of the Genus

Modern treatments of the genus of a curve normally describe it in terms of the topology of the associated Riemann surface. Therefore, modern mathematicians are usually amazed to learn that the idea stems from Abel, who lived and worked at a time when even the notion of a complex function of a complex variable was in its early infancy and the notion of a Riemann surface was still in the future. (Riemann surfaces first appeared in Riemann's dissertation [77] of 1851.) But as the discussion in the preceding essay shows, Abel's point of view does not depend on complex numbers.

The geometric picture of N points on the curve varying with $N - g$ degrees of freedom that was presented in the preceding essays does depend on complex numbers, because the coordinates of the intersection points defined by $\chi = 0$, $\theta = 0$ exist only in some algebraically closed field, and the notion of continuous variation requires something like real numbers. But the actual determination of the genus depends on purely algebraic considerations, at least in the examples of the preceding essay. All that is needed is to construct, for a large number ν, a formula $\theta(x, y, a, a', a'', \ldots)$ for the most general "function" in the field that has poles only at points where $x = \infty$ and no longer has poles at those points when it is divided by x^ν (although this division will probably cause it to have poles at $x = 0$). The **genus** g is determined by the condition that the number of zeros of θ is $g - 1$ greater than the number of arbitrary constants in the formula for θ.

Of course elements of a function field are not really functions in the usual sense, so they do not really have zeros and poles, and the condition that an element have poles only where $x = \infty$ is far from rigorous. Therefore, this description of the genus needs more explanation. Starting with the field of rational functions on an algebraic curve $\chi(x, y) = 0$—which is simply the root field of a monic, irreducible polynomial $\chi(x, y)$ in y with coefficients in $\mathbf{Z}[x]$—one needs to define what it means to say that an element θ of the field has no poles where x is finite and that θ/x^ν has no poles where $1/x$ is finite, and then one needs to determine how many zeros such a θ has and how many arbitrary constants there are in the formula for the most general such θ.

The idea of an element θ having no poles where x is finite has a standard algebraic formulation: An element θ of the field of rational functions on a curve $\chi(x, y) = 0$ is **integral over** x if some power of θ is equal to a sum of multiples of lower powers in which the multipliers are elements of the ring $\mathbf{Q}[x]$ of polynomials in x with rational coefficients.[4] The justification of this definition is, in the last analysis, pragmatic—it *works* in the sense that it suggests correct theorems and is useful in proofs.

(The analogous definition of an **algebraic integer** in an algebraic number field—see Essay 2.5—emerged in the work of Kronecker and Dedekind in the 1860s and 1870s. Bourbaki [11] claims it is in the work of Eisenstein as early as 1852, but I do not find it there. Kronecker [55, §1] used the above definition of integrality over x of an "algebraic function" in his study of function fields, but as far as I have found he does not explain or motivate it.)

It is easy to see that the elements of the field of rational functions on $\chi(x, y)$ that are integral over x form a ring in the field and that this ring contains $\mathbf{Q}[x]$.[5]

If θ is integral over x, then dividing an equation that demonstrates its integrality, say $\theta^n = a_1(x)\theta^{n-1} + \cdots + a_n(x)$, by $x^{n\mu}$ for μ larger than the maximum degree of the $a_i(x)$ shows that θ/x^μ is integral over $1/x$ for all such values of μ.[6] The **order** of θ at $x = \infty$ is by definition the smallest ν for which θ/x^ν is integral over $1/x$.

Let $\Theta(x^\nu)$ denote the elements θ of the field of rational functions on $\chi(x, y) = 0$ that are integral over x and have order at most ν where $x = \infty$. The goal is to find a formula for the most general element θ of $\Theta(x^\nu)$, and to compare the "number of zeros" of θ to the "number of arbitrary constants it contains."

The "number of zeros" of such a θ has a very plausible meaning. By assumption, $\chi(x, y)$ is monic in y, say of degree n in y. Then the values of y for a given x are the roots of a monic polynomial of degree n, so there are n of them, counted with multiplicities. For this reason, there are n points on the curve for each x, so x assumes each value exactly n times on the curve, counted with multiplicities. For this reason, it is reasonable to take the view that x also assumes the value ∞ exactly n times—that x has n poles on the curve, counted with multiplicities. Therefore, x^ν has n poles of order ν for a total of $n\nu$ poles. Since θ has the same poles as x^ν (except when θ is in the subset of $\Theta(x^\nu)$ containing "functions" that have fewer than

[4] One could also use the more restrictive, but perhaps more natural, definition in which the multipliers are required to be in $\mathbf{Z}[x]$. Then an element would be integral over x in the sense defined above if and only if some integer multiple of it was integral in the more restrictive sense. Since $\frac{1}{2}$ is certainly a "function" without poles, the definition given above is the one that describes "rational functions without poles" on an algebraic curve.

[5] If $z_1^{n_1}$ can be expressed as a sum of multiples of lower powers of z_1 in which the multipliers are in $\mathbf{Q}[x]$, and $z_2^{n_2}$ can be expressed as a sum of multiples of lower powers of z_2 in which the multipliers are in $\mathbf{Q}[x]$, then every polynomial in z_1 and z_2 with coefficients in $\mathbf{Q}[x]$ can be expressed as a sum of multiples of $z_1^i z_2^j$ with coefficients in $\mathbf{Q}[x]$, where $i < n_1$ and $j < n_2$. Therefore, multiplication by any such polynomial in z_1 and z_2—in particular multiplication by $z_1 + z_2$ and $z_1 z_2$—can be represented by the $n_1 n_2 \times n_1 n_2$ matrix of elements of $\mathbf{Q}[x]$ that gives its effect on these $n_1 n_2$ monomials $z_1^i z_2^j$. Therefore, since the polynomial in z_1 and z_2 is a root of the (monic) characteristic polynomial of this $n_1 n_2 \times n_1 n_2$ matrix by the Cayley–Hamilton theorem, $z_1 + z_2$ and $z_1 z_2$, and, in the same way, all polynomials in z_1 and z_2 with coefficients in $\mathbf{Q}[x]$, are integral over x.

[6] Note that $\chi(x, y)/x^{n\lambda}$ has the form $\chi_1(1/x, y/x^\lambda)$, where χ_1 is irreducible with integer coefficients and monic in its second variable, when λ is large enough, and that x and y can be expressed rationally in terms of $u = 1/x$ and $v = y/x^\lambda$, so the field of rational functions on $\chi(x, y) = 0$ can also be regarded as the field of rational functions on the curve $\chi_1(u, v) = 0$. To say that an element of this field is integral over $1/x$ means, of course, that it is integral over u.

the maximum number of poles allowed, which is sparse in $\Theta(x^\nu)$), it follows that θ should be regarded as having $n\nu$ poles; reversing the above reasoning then leads to the conclusion that θ assumes each value $n\nu$ times, including the value zero. In short, it is plausible to take $n\nu$ to be the number of zeros of a typical element of $\Theta(x^\nu)$.

This analysis of the number of zeros of a typical element of $\Theta(x^\nu)$ for large ν overlooks, however, a phenomenon that is exhibited by the example $\chi(x, y) = (x^2 + y^2)^2 - 2(x + y)^2$. The field of rational functions on this "curve"—the root field of this $\chi(x, y)$—contains the element $\frac{x^2+y^2}{x+y}$, which is a root of $X^2 - 2$ (by definition, the square of $x^2 + y^2$ is $2(x + y)^2$). Therefore, it is reasonable to let $\sqrt{2}$ denote this element of the field. Then $\chi(x, y) = (x^2 + y^2 - \sqrt{2}(x + y))(x^2 + y^2 + \sqrt{2}(x + y))$, which shows that the field is an extension of degree *two*, not four, of the field of rational functions in x with coefficients in the number field $\mathbf{Q}(\sqrt{2})$. Geometrically, the curve $(x^2 + y^2)^2 - 2(x + y)^2 = 0$ is quite simple, because the reduction $(x^2 + y^2 - \sqrt{2}(x + y))(x^2 + y^2 + \sqrt{2}(x + y)) = 0$ shows that it is a *union of two circles*, namely, the circle whose diameter is the line from the origin to $(\sqrt{2}, \sqrt{2})$ and the one whose diameter is the line from the origin to $(-\sqrt{2}, -\sqrt{2})$. Geometers traditionally use algebraically closed ground fields in part to avoid situations like this in which a curve described by a simple irreducible polynomial becomes a union of two curves when the field of constants is extended.

The simple constructive solution to this difficulty is not to make the giant leap to an algebraically closed ground field—the usual choice being the field of complex numbers, which is not an algebraic but a transcendental extension—but to *adjoin new constants as needed*. In the example, the constant $\sqrt{2}$ is not just needed, it is already present as $\frac{x^2+y^2}{x+y}$, and when it is used the curve is reducible, and the geometric picture of the ("curve" whose field of rational functions is the root field of $(x^2 + y^2)^2 - 2(x + y)^2$ is a single circle $x^2 + y^2 = \sqrt{2}(x + y)$. This revision of the picture makes it clear that the number of zeros of x on the curve is two, not four. Consequently, the number of poles of x^ν, which is the same as the number of zeros, is 2ν, not 4ν, and a typical element of $\Theta(x^\nu)$ has 2ν zeros, not 4ν.

More generally, one needs to take into consideration the possibility that the root field of $\chi(x, y)$ may contain constants other than the obvious constants in $\mathbf{Q}$. Here a "constant" is an element of the root field that is a root of a polynomial with integer coefficients, or, what is the same, an element of $\Theta(x^0)$. (A polynomial in x with rational coefficients is equal to a polynomial in $1/x$ with rational coefficients if and only if it is a rational number, and a root of a monic polynomial with rational coefficients is a root of a polynomial with integer coefficients.) For this reason, $\Theta(x^0)$ will be called the **field of constants** of the root field of $\chi(x, y)$. (Note that $\Theta(x^\nu)$ is a vector space over $\mathbf{Q}$ for all ν, and that $\Theta(x^0)$ is in fact a *field*.)

The example then suggests the following definition: The **number of zeros** of a typical element of $\Theta(x^\nu)$ for large values of ν is $n_0\nu$, where n_0 is the degree of the root field of χ as an extension, not of the field $\mathbf{Q}(x)$ of rational functions in x with integer coefficients, but of the field of rational functions of x with coefficients in the field of constants $\Theta(x^0)$. When $\Theta(x^0) = \mathbf{Q}$, n_0 is simply the degree of χ in y, but in the general case it is this degree divided by $[\Theta(x^0) : \mathbf{Q}]$.

Similarly, when one counts the "number of constants" in a formula for a typical element of $\Theta(x^\nu)$, one thinks of the constants as being in $\Theta(x^0)$, not **Q**. Because the product of an element of $\Theta(x^\nu)$ and an element of $\Theta(x^\mu)$ is an element of $\Theta(x^{\nu+\mu})$, $\Theta(x^\nu)$ is a vector space over $\Theta(x^0)$. The number of constants in a formula for a typical element of $\Theta(x^\nu)$ is quite simply the *dimension* of $\Theta(x^\nu)$ as a vector space over the field $\Theta(x^0)$.

In this way, Abel's conception of the number of integrals to which a sum of integrals of an algebraic integrand can be reduced leads to the definition of the **genus** of the root field of $\chi(x, y)$ as the number g for which $\dim \Theta(x^\nu) = n_0\nu - g + 1$, where ν is a large integer, where n_0 is the degree n of χ in y divided by the degree c of the field of constants $\Theta(x^0)$ as an extension of **Q**, where $\Theta(x^\nu)$ denotes the subset of the root field containing elements θ that are integral over x and have order at most ν where $x = \infty$, and where the dimension is the dimension of $\Theta(x^\nu)$ as a vector space over the field of constants $\Theta(x^0)$. The underlying idea is that an element of $\Theta(x^\nu)$ has $n_0\nu$ zeros and contains $\dim \Theta(x^\nu)$ parameters; variation of all $\dim \Theta(x^\nu)$ parameters in such a θ varies its zeros with only $\dim \Theta(x^\nu) - 1$ degrees of freedom—one degree of freedom is lost because multiplication of a function by a constant does not change its zeros—so the number of constraints g on the motion of the $n_0\nu$ zeros under an algebraic variation is determined by the equation $\dim \Theta(x^\nu) - 1 = n_0\nu - g$.

The main theorem will be to show that this genus is **intrinsic** to the curve $\chi(x, y) = 0$ in the sense that if the root fields of two polynomials $\chi(x, y)$ are isomorphic—if the two corresponding curves are **birationally equivalent**—then the fields have the same genus. Although the proof will be somewhat long, the underlying reason that the genus is intrinsic stems from the above discussions: It is the codimension of the subvarieties of C^N determined by the algebraic variations of N points on the curve.

Essay 4.4 Newton's Polygon

> ... *ses* [Newton's] *principaux Guides dans ces Recherches* [on cubic curves] *ont été la Doctrine des* Séries infinies, *qui lui doit presque tout, & l'usage du* Parallélogramme analytique, *dont il est l'Inventeur. ... Il est facheux que Mr. Newton se soit contenté d'étaler ses découvertes sans y joindre les Démonstrations, et qu'il ait préféré le plaisir de se faire admirer à celui d'instruire.* (Newton's main guides in his researches on cubic curves were the doctrine of *infinite series*, which owes him practically everything, and the use of the *analytic parallelogram*, of which he is the inventor. ... It is annoying that Mr. Newton contented himself with laying out his discoveries without accompanying them with proofs, and that he preferred the pleasure of making himself admired to that of instructing.)—Gabriel Cramer [19, Preface]

The program outlined at the end of the last essay for constructing the genus of an algebraic curve—or, more precisely, the genus of the root field of a given $\chi(x, y)$—will be carried out in the following essays using an algorithm of Isaac Newton (shown in Fig. 4.4) for expanding an algebraic function of x as a power series in *fractional*

Fig. 4.4 Newton

powers of x.[7] Known as Newton's *polygon*, or sometimes Newton's *parallelogram*, it constructs, for a given polynomial equation $\chi(x, y) = 0$, infinite series expansions of y in fractional powers of x. It involves *choices* and results in n *different* expansions, where $n = \deg_y \chi$.

It will be useful to expand y not only in powers of x but also in powers of $x - \alpha$ for various algebraic numbers α, something that can be accomplished by the same method, since setting $x_1 = x - \alpha$ and $\chi_1(x_1, y) = \chi(x - \alpha, y)$ gives an algebraic relation between x_1 and y that can be used to expand y in (fractional) powers of x_1 using Newton's polygon.

Let $\chi(x, y)$ be an irreducible polynomial with integer coefficients that has positive degree in both x and y and is monic in y, and let α be a given value for x. The objective is to find infinite series "solutions" $y = \theta_0(x-\alpha)^{\varepsilon_0} + \theta_1(x-\alpha)^{\varepsilon_1} + \theta_2(x-\alpha)^{\varepsilon_2} + \cdots$ of $\chi(x, y) = 0$ in which the coefficients θ_i are algebraic numbers and the exponents $\varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_k < \cdots$ are an increasing sequence of rational numbers. It will also be assumed that the exponents increase without bound in the sense that for any given N one can find a value of k for which $\varepsilon_k > N$. The meaning of the statement that such a series "solves" $\chi(x, y) = 0$ is clear, if somewhat nonconstructive: Such

[7] Newton's presentation is quite sketchy. My main source was Walker [82]. See also Newton [71, vol. 3, p. 50 and p. 360, vol. 4, p. 629], Hensel–Landsberg [46], and Chebotarev [16]. Chebotarev cites (end of §2) the Hensel–Landsberg book as his basic source, but he examines the Newton polygon much more fully than that book does, dealing thoroughly, for instance, with the history of the method. Unfortunately, his article is available only in Russian. Chebotarev advocates calling it Newton's "diagram" as Hensel and Landsberg do, saying that the "polygon" was not present in Newton's formulation, but the name "Newton's polygon" now seems firmly established.

series can be added and multiplied term by term, and $x = \alpha + (x - \alpha)$ is such a series (a terminating one), so $\chi(x, y)$ represents such a series, the coefficients of which can be computed in an open-ended way by finding, for any given upper bound, all terms of the series $\chi(x, y)$ in which the exponent is less than that bound. To say that $\chi(x, y) = 0$ means simply that the result is always zero.

Since y is integral over x, the exponents ε_i are to be expected to be nonnegative. Therefore, a knowledge of the terms of the series for y through the term $\theta_k(x - \alpha)^{\varepsilon_k}$ is all that is needed to compute all terms of $\chi(x, y)$ in which the exponents are less than or equal to ε_k, because all omitted terms contain $(x - \alpha)^{\varepsilon_{k+i}}$ for some $i > 0$ and $\varepsilon_{k+i} > \varepsilon_k$. What is sought, then, are infinite sequences $\theta_0, \theta_1, \theta_2, \dots$ and $0 \le \varepsilon_0 < \varepsilon_1 < \cdots$ for which all terms of the *terminating* sequence $\chi(x, \theta_0(x-\alpha)^{\varepsilon_0} + \theta_1(x-\alpha)^{\varepsilon_1} + \cdots + \theta_k(x-\alpha)^{\varepsilon_k})$ have exponents greater than ε_k. A constructive solution of this problem must of course be an *algorithm* for *generating* such sequences. "Newton's polygon" is such an algorithm.

More specifically, given the initial terms $\theta_0(x-\alpha)^{\varepsilon_0} + \theta_1(x-\alpha)^{\varepsilon_1} + \cdots + \theta_k(x-\alpha)^{\varepsilon_k}$ of an infinite series solution y of $\chi(x, y) = 0$ in the sense just described, the algorithm should give *all possible* values $\theta_{k+1}(x-\alpha)^{\varepsilon_{k+1}}$ for the next term of the sequence. They can be completely described in the following way: To avoid fractional exponents, let m be the least common denominator of $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_k$ and let $s = (x - \alpha)^{1/m}$, so that the initial terms that are assumed to be known take the form $\beta_0 + \beta_1 s + \cdots + \beta_h s^h$, where h is the integer $m\varepsilon_k$ and where β_i is zero unless i is of the form $m\varepsilon_j$ for some j, in which case $\beta_i = \theta_j$. Let the term following $\beta_h s^h$ be $\gamma s^{\rho+h}$, so that the required equation is $\chi(\alpha + s^m, \beta_0 + \beta_1 h + \cdots + \beta_h s^h + \gamma s^{h+\rho} + \cdots) = 0$, where ρ is a positive rational number. To determine the possible values of γ and ρ expand $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + t s^h)$, a polynomial in s and t whose coefficients are algebraic numbers (because they are polynomials in $\beta_0, \beta_1, \dots, \beta_h, \alpha$, and the coefficients of χ), as a polynomial in t, $\Phi_0(s) + \Phi_1(s)t + \Phi_2(s)t^2 + \cdots + \Phi_n(s)t^n$, whose coefficients $\Phi_i(s)$ are polynomials in s. Again, to avoid fractional exponents, let ρ be written $\rho = \frac{\sigma}{\tau}$, where σ and τ are positive integers, and let $s_1 = s^{1/\tau}$, so that the required identity becomes $\chi(\alpha + s_1^{m\tau}, \beta_0 + \beta_1 s_1^{\tau} + \cdots + \beta_h s_1^{h\tau} + \gamma s_1^{h\tau+\sigma} + \cdots) = 0$, which is to say

$$\Phi_0(s_1^{\tau}) + \Phi_1(s_1^{\tau})(\gamma s_1^{\sigma} + \cdots) + \Phi_2(s_1^{\tau})(\gamma s_1^{\sigma} + \cdots)^2 + \cdots + \Phi_n(s_1^{\tau})(\gamma s_1^{\sigma} + \cdots)^n = 0.$$

The simple idea that underlies Newton's polygon is the observation that this infinite series in s_1 with algebraic number coefficients, which is a sum of $n + 1$ such series, can be identically zero only if all terms in the sum cancel, and, in particular, only if the *lowest-order terms* of these series cancel. If the polynomial $\Phi_i(s)$ is nonzero, it has the form $\zeta_i s^{j_i} + \cdots$, where $\zeta_i \neq 0$ and the omitted terms all have degree greater than j_i. With this notation, the term of $\Phi_i(s_1^{\tau})(\gamma s_1^{\sigma} + \cdots)^i$ of lowest degree, when $\Phi_i(s) \neq 0$, is $\zeta_i \gamma^i s_1^{\sigma i + \tau j_i}$. The required cancellation dictates that the positive integers σ and τ must have the property that $\sigma i + \tau j_i$ assumes its minimum value for at least two different pairs (i, j_i) (note that these pairs are determined by χ, m, and $\beta_0 + \beta_1 s + \cdots + \beta_h s^h$). These conditions limit the pairs (σ, τ) to a finite number of possibilities—the geometrical picture is the one described below—and even gives strong information about the coefficient γ of the next term, namely, that it is a nonzero root of the polynomial $\sum \zeta_i \gamma^i$, where the sum is extended over just those values of i for which $\Phi_i(s) \neq 0$ and $\sigma i + \tau j_i$ assumes its minimum value.

Some term of some series $\Phi_i(s_1^\tau)(\gamma s_1^\tau + \cdots)^i$ for $i > 0$ must cancel the first term $\zeta_0 s_1^{\tau j_0}$ of $\Phi_0(s_1^\tau)$, so $\tau j_0 \geq \sigma i + \tau j_i$ for some $i > 0$. Since σ and τ are both positive, j_0 must be greater than j_i for at least one $i > 0$. Therefore, the above discussion shows that the series $\beta_0 + \beta_1 s + \cdots + \beta_h s^h$ can be extended to be an infinite series solution y of $\chi(x, y) = 0$ when $x = \alpha + s^m$ only if the polynomial in two variables $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + ts^h) = \Phi_0(s) + \Phi_1(s)t + \Phi_2(s)t^2 + \cdots + \Phi_n(s)t^n$ has the property that s divides $\Phi_0(s)$ more times than it divides $\Phi_i(s)$ for at least one $i > 0$. Otherwise stated, the term or terms of this polynomial of lowest degree in s must all contain t.

As will be shown, these *necessary* conditions on the constants that describe the next term when a certain number of terms are known permit one to construct *all possible* solutions y of $\chi(x, y) = 0$ as infinite series of fractional powers of $x - \alpha$.

A **truncated solution** y of $\chi(x, y) = 0$ at $x = \alpha$ will by definition consist of (1) an algebraic number field $\mathbf{A}$ containing α, (2) a positive integer m, and (3) a finite sequence $\beta_0, \beta_1, \ldots, \beta_h$ in $\mathbf{A}$ with the property that the term or terms of $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + ts^h)$ of lowest degree in s all contain t. In addition, it will be assumed that the result $\Phi_0(s)$ of setting $t = 0$ in this polynomial is not zero; otherwise, $y = \beta_0 + \beta_1(x - \alpha)^{1/m} + \cdots + \beta_h(x - \alpha)^{h/m}$ is an **actual solution** of $\chi(x, y) = 0$ and there is no need to use higher powers of $x - \alpha$.

Newton's Polygon

Input: A truncated solution y of $\chi(x, y) = 0$ at $x = \alpha$, as that term was just defined.

Algorithm: *As above, let* $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + ts^h)$ *be written in the form* $\Phi_0(s) + \Phi_1(s)t + \Phi_2(s)t^2 + \cdots + \Phi_n(s)t^n$ *of a polynomial in t whose coefficients* $\Phi_i(s)$ *are polynomials in s with coefficients in the field* $\mathbf{A}$ *specified by the input. Consider the set of pairs* (i, j_i) *of integers, where i is in the range* $0 \leq i \leq n$, *where* $\Phi_i(s) \neq 0$, *and where* j_i *is the number of times that s divides* $\Phi_i(s)$. *By assumption,* j_0 *is defined and greater than at least one other* j_i. *The* **segments of the Newton polygon** *corresponding to this input are the line segments that join two points* (i, j_i), *say those corresponding to the indices* i_1 *and* $i_2 > i_1$, *in such a way that* (1) *the segment has negative slope, so it is described by the equation* $\sigma i + \tau j = k$ *where* $\sigma = j_{i_1} - j_{i_2}$ *and* $\tau = i_2 - i_1$ *are both positive and where k is the common value of* $\sigma i + \tau j$ *for these two indices,* (2) $\sigma i + \tau j_i \geq k$ *for all indices i for which* j_i *is defined, and* (3) $\sigma i + \tau j_i > k$ *whenever* j_i *is defined and* $i < i_1$ *or* $i > i_2$. *With each such segment, associate the polynomial*

$$\eta(c) = \sum_{\sigma i + \tau j_i = k} \zeta_i c^i$$

with coefficients in $\mathbf{A}$, *where the sum is over just those values of i for which* (i, j_i) *lies on the segment, of which there are at least two, and where* ζ_i *is the coefficient of*

s^{j_i} in $\Phi_i(s)$. Extend the input field **A**, *if necessary, to split all polynomials $\eta(c)$ that result in this way from segments of the polygon.*

(Geometrically, the segments join to form a polygonal path that joins the point $(0, j_0)$ to the first point, call it (I, J), of the form (i, j_i) at which j_i assumes its minimum value. This path is determined by the fact that it joins points of the form (i, j_i) in such a way that none of these points are in the interior of the closed polygon formed by it and the segments from $(0, j_0)$ to $(0, J)$ and from $(0, J)$ to (I, J). See Fig. 4.5.)

Output: A truncated solution of $\chi(x, y) = 0$ for $x = \alpha$ for each *nonzero* root γ of each polynomial $\eta(c)$ in the extended field constructed by the algorithm, namely, the truncated solution

$$\begin{aligned} y = \beta_0 + \beta_1(x-\alpha)^{1/m} + \beta_2(x-\alpha)^{2/m} \\ + \cdots + \beta_h(x-\alpha)^{h/m} + \gamma(x-\alpha)^{(h\tau+\sigma)/\tau m} \end{aligned} \tag{1}$$

in which one term with coefficient γ and exponent $\frac{h\tau+\sigma}{\tau m} = \frac{h+\rho}{m}$ is added to the input truncated solution where $\rho = \frac{\sigma}{\tau}$. In other words, the output truncated solution corresponding to γ consists of the (possibly) extended field **A** constructed by the algorithm, the positive integer τm, and the sequence $\beta'_0, \beta'_1, \beta'_2, \ldots, \beta'_{h\tau+\sigma}$ in which $\beta'_{i\tau} = \beta_i$ for $i = 0, 1, \ldots, h$ and $\beta'_{h\tau+\sigma} = \gamma$ but all other coefficients β'_i are zero.

That each output (1) is a truncated solution—unless, of course, it is an *actual* solution—can be proved as follows: Set $s = s_1^\tau$ and $t = s_1^\sigma(\gamma + t_1)$ in the definition of $\Phi_0(s), \Phi_1(s), \ldots, \Phi_n(s)$ to put the new equation in the form

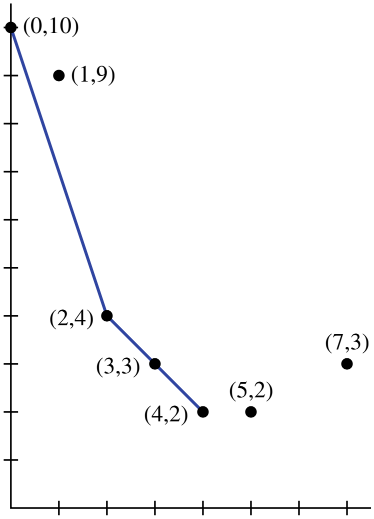

Fig. 4.5 When there are seven points $(i, j_i) = (0, 10), (1, 9), (2, 4), (3, 3), (4, 2), (5, 2), (7, 3)$, Newton's polygon has two segments. They join (0, 10) and (4, 2) via (2, 4).

$$
\begin{aligned}
&\chi(\alpha + s_1^{\tau m}, \beta_0 + \beta_1 s_{\mathrm{J}}^{\tau} + \beta_2 s_1^{2\tau} + \cdots + \beta_h s_1^{h\tau} + \gamma s_1^{h\tau+\sigma} + t_1 s_1^{h\tau+\sigma}) \\
&\quad = \sum_{i=0}^{n} \Phi_i(s_1^{\tau})(s_1^{\sigma})^i(\gamma + t_1)^i
\end{aligned} \tag{2}
$$

By the choice of σ and τ, no term on the right contains s_1 to a power less than the minimum value of $\sigma i + \tau j_i$, call it k, and the terms that contain s_1 to the power k exactly are $s_1^k \eta(\gamma + t_1)$ by the definition of η. Since $\eta(\gamma + t_1)$ is a nonzero polynomial (its degree in t_1 is the largest value of i for which the point (i, j_i) lies on the corresponding segment of the Newton polygon) with constant term zero (by the choice of γ), (2) is a polynomial in which the terms of lowest degree k in s_1 all contain t_1, as was to be shown.

The **ambiguity** of a truncated solution is, in the notation used above, the least index i for which j_i attains its minimum. Otherwise stated, it is the i-coordinate I of the endpoint (I, J) of the Newton polygon other than $(0, j_0)$. A truncated solution will be called **unambiguous** if its ambiguity is 1. In this case, the polygon consists of a single line segment, and $\eta(c)$ is a polynomial of degree 1 whose single root is nonzero, so the algorithm produces a single output; moreover, the algorithm does not increase m, and it results in no extension of **A** because the root of $\eta(c)$ is already in **A**.

The ambiguity of an output solution is the multiplicity of its γ as a root of its $\eta(c)$, as follows from the above observation that the terms of lowest degree in s_1 are $s_1^k \eta(\gamma + t_1)$, because the multiplicity of γ as a root of $\eta(c)$ is the number of times t_1 divides $\eta(\gamma + t_1)$. Thus, among the nonzero terms $\epsilon s_1^p t_1^q$ of (2) in which p assumes its minimum value k, the one in which q has its least value is the one in which q is the multiplicity of γ as a root of its $\eta(c)$.

In particular, *if the input truncated solution is unambiguous, so is the output truncated solution.* Thus, if it begins with an unambiguous truncated solution, the algorithm constructs an infinite series solution (which may in rare cases be an actual terminating solution) with coefficients in the same **A**. In short, the construction of infinite series solutions is reduced by the Newton's polygon algorithm to the construction of unambiguous truncated solutions.

Theorem 1 *Construct n distinct infinite series solutions y of $\chi(x, y) = 0$ at $x = \alpha$.*

As above, α is a given algebraic number and $\chi(x, y)$ is a given polynomial with integer coefficients that is irreducible, contains both x and y, and is monic of degree n in y. For the reason just stated, an infinite series solution can be regarded as having been constructed when an unambiguous truncated solution has been constructed. The proof of the theorem will follow an example:

Example Let $\chi(x, y) = y^3 - xy + x^3$ (the curve $\chi = 0$ is the folium of Descartes—see Fig 4.6 in Essay 4.5) and let $\alpha = 0$. If one begins with the truncated solution $m = 1$, $y = 0$, one begins with $\chi(0 + s, 0 + t) = s^3 - st + t^3$, and Newton's polygon joins the points (0, 3) and (3, 0) via the point (1, 1). The two segments of the polygon are described by the equations $2i + j = 3$ and $i + 2j = 3$.

The first segment gives just one output truncated solution $y = x^2$, because $m = 1$, $h = 0$, $\sigma = 2$, $\tau = 1$ and because the polynomial $\eta(c) = 1 - c$ in this case has just one nonzero root 1. It is a simple root, so this output solution is unambiguous.

The second segment gives two output truncated solutions $y = \gamma\sqrt{x}$ (because $m = 1$, $h = 0$, $\sigma = 1$, $\tau = 2$), where γ is a nonzero root of $\eta(c) = -c + c^3$. Thus $\gamma = \pm 1$, and the output consists of two truncated solutions $y = \pm\sqrt{x}$. Both are unambiguous.

Thus, Newton's polygon constructs three unambiguous truncated solutions and therefore constructs the three required infinite series solutions of $y^3 - xy + x^3 = 0$ for $x = 0$. These infinite series solutions can be found by repeated application of the Newton polygon algorithm, but the first few terms can be found more easily by the following method.

The truncated solution $y = \pm\sqrt{x}$. calls for the computation of $\chi(s^2, \pm s + st) = (s^2)^3 - s^2(\pm s + st) + (\pm s + st)^3 = s^3(s^3 - (\pm 1 + t) + (\pm 1 + t)^3) = s^3(s^3 + 2t \pm 3t^2 + t^3)$. The term $2t$ shows that this truncated solution is unambiguous. The continuation of the truncated solution $y = \pm\sqrt{x} + \cdots$ can be found using the equation $s^3 + 2t \pm 3t^2 + t^3 = 0$ to express t as a power series in $s = \sqrt{x}$ and substituting the result in $y = \pm s + st$. Consider first the case in which the sign is plus. The relation $s^3 + 2t + 3t^2 + t^3 = 0$ can be written $t = -\frac{1}{2}s^3 + t^2\big(-\frac{3}{2} - \frac{t}{2}\big)$ to find that $t = -\frac{1}{2}s^3 + \big(-\frac{1}{2}s^3 + t^2\big(-\frac{3}{2} - \frac{t}{2}\big)\big)^2\big(-\frac{3}{2} - \frac{t}{2}\big) = -\frac{1}{2}s^3 + \big(\frac{1}{4}s^6 + \frac{3}{2}s^3t^2 + \cdots\big)\big(-\frac{3}{2} - \frac{t}{2}\big) = -\frac{1}{2}s^3 - \frac{3}{8}s^6 - \frac{9}{4}s^3t^2 - \frac{1}{8}s^6t + \cdots = -\frac{1}{2}s^3 - \frac{3}{8}s^6 - \frac{9}{4}\cdot\frac{1}{4}\cdot s^9 + \frac{1}{8}\cdot\frac{1}{2}s^9 + \cdots = -\frac{1}{2}s^3 - \frac{3}{8}s^6 - \frac{1}{2}s^9 + \cdots$, where the omitted terms all contain s^{12}, from which $y = s + st = s - \frac{1}{2}s^4 - \frac{3}{8}s^7 - \frac{1}{2}s^{10} + \cdots$. When $s^3 + 2t + 3t^2 + t^3 = 0$ is changed to $s^3 + 2t - 3t^2 + t^3 = 0$, the corresponding solution is found by changing s to $-s$ and t to $-t$. In summary, the second segment $i + 2j = 3$ corresponds to two infinite series solutions of $y^3 - xy + x^3 = 0$; they begin

$$y = \pm\sqrt{x} - \frac{1}{2}x^2 \mp \frac{3}{8}x^3\sqrt{x} - \frac{1}{2}x^5 + \cdots$$

The infinite series solution $y = x^2 + \cdots$ corresponding to the first segment $2i + j = 3$ calls for computing the polynomial $\chi(s, s^2 + s^2t) = s^3 - s^3(1 + t) + s^6(1 + t)^3 = s^3(-t + s^3(1 + t)^3)$. The term $-t$ shows that the truncated solution $y = x^2$ is unambiguous. The expansion of y in powers of x can be found by using the relation $-t + s^3(1 + t)^3 = 0$ to expand t in powers of $s = x$ and substituting the result in $y = s^2 + s^2t$. Now, $t = s^3(1 + t)^3$ implies

$$\begin{aligned} t &= s^3(1 + s^3(1 + t)^3)^3 = s^3(1 + 3s^3(1 + t)^3 + 3s^6(1 + t)^6 + s^9(1 + t)^9) \\ &= s^3 + 3s^6(1 + t)^3 + 3s^9(1 + t)^6 + s^{12}(1 + t)^9 \\ &= s^3 + 3s^6 + 9s^6t + 9s^6t^2 + \cdots + 3s^9 + 18s^9t + \cdots + s^{12} + \cdots \\ &= s^3 + 3s^6 + 12s^9 + 28s^{12} + \cdots, \end{aligned}$$

so

$$y = x^2 + x^5 + 3x^8 + 12x^{11} + 28x^{14} + \cdots$$

is the beginning of this infinite series solution of $y^3 - xy + x^3 = 0$.

(Note that the sum of the three series is zero, at least up to the terms in x^5, in accord with the fact that the coefficient of y^2 in $y^3 - xy + x^3$ is zero.)

***Proof of Theorem* 1** A truncated solution of $\chi(x, y)$ at $x = \alpha$ in which $m = 1$ and $h = 0$ is an algebraic number β_0 for which the terms of $\chi(\alpha + s, \beta_0 + t)$ of lowest degree in s all contain t; since $\chi(\alpha + s, \beta_0 + t)$ contains the term t^n with no s at all, $y = \beta_0$ is a truncated solution if and only if $\chi(\alpha + s, \beta_0)$ does not contain a term without s or, to put it more simply, if and only if $\chi(\alpha, \beta_0) = 0$. In short, these truncated solutions $y = \beta_0$ are the roots of $\chi(\alpha, y)$.

The *ambiguity* of such a truncated solution $y = \beta_0$ of $\chi(x, y)$ at $x = \alpha$ is equal to the *multiplicity* of β_0 as a root of $\chi(\alpha, y)$, because the ambiguity of the truncated solution is by definition the least index i for which $\Phi_i(0) \neq 0$, where $\chi(\alpha + s, \beta_0 + t) = \Phi_0(s) + \Phi_1(s)t + \Phi_2(s)t^2 + \cdots + t^n$, which is the multiplicity of β_0 as a root of $\chi(\alpha, y)$. In particular, if all roots of $\chi(\alpha, y)$ are simple, the Newton polygon algorithm applied to any one of the n unambiguous truncated solutions $y = \beta_0$ generates an infinite series solution y of $\chi(x, y) = 0$ at $x = \alpha$, which proves the theorem in this case. In the general case, one can apply the following algorithm:

Input: A set of truncated solutions of $\chi(x, y) = 0$ at $x = \alpha$.

Algorithm: *While the set contains a truncated solution whose ambiguity is greater than 1, let the Newton polygon algorithm be used to replace one such truncated solution with one or more longer truncated solutions.*

The theorem will be proved by proving that this algorithm *terminates*—that is, it reaches a stage at which all truncated solutions in the set that has been found are unambiguous—and by proving that each step leaves the sum of the ambiguities unchanged, so that if the algorithm starts with the truncated solutions $y = \beta_0$, the sum of whose ambiguities is $\deg_y \chi(\alpha, y) = n$ (because this sum is the sum of the multiplicities of the roots β_0 of $\chi(\alpha, y)$), it terminates with a set of n unambiguous truncated solutions, which then imply n infinite series solutions.

That the sum of the ambiguities does not change can be seen as follows: Let the notation be as in the description of Newton's polygon. The ambiguity of the input truncated solution is the least index I for which j_i attains its minimum value J. Since the segments of the Newton polygon join $(0, j_0)$ to (I, J) and since the number of nonzero roots—counted with multiplicities—of any $\eta(c)$ is its degree minus the number of times c divides it, which is the difference $i_2 - i_1$ of the i-coordinates of the endpoints of the corresponding segment, the ambiguity of the input truncated solution is the total number of nonzero roots, counted with multiplicities, of the polynomials $\eta(c)$ corresponding to segments of the polygon. Since, as was noted above, the multiplicity of γ as a root of $\eta(c)$ is the ambiguity of the output truncated solution corresponding to γ, the sum of the ambiguities of the output solutions is the ambiguity of the input solution, as was to be shown.

Each step of the above algorithm increases the number of truncated solutions in the list unless the input truncated solution, which has ambiguity greater than 1 by assumption, yields a single output truncated solution, which means that $\eta(c)$ is a constant times $(c - \gamma)^\mu$ for some nonzero algebraic number γ, where μ is the ambiguity of the input solution. It will be shown that the number of steps of this type is bounded above, so that repeated application of the Newton polygon algorithm eventually must increase the number of truncated solutions in the set. Since the total of their ambiguities is n at each step, it will follow that the process must terminate with n unambiguous truncated solutions, and the theorem will be proved.

Suppose, therefore, that the (ambiguous) input truncated solution is one that produces a single output truncated solution. It is to be shown that iteration of the algorithm eventually produces more than one output truncated solution. Let μ be the ambiguity of the input solution, which is therefore the ambiguity of each subsequent output solution as long as there is only one of them. As was just noted, when there is only one output truncated solution, $\eta(c)$ is a constant times $(c-\gamma)^\mu$ for some algebraic number γ, which implies that Newton's polygon consists of a single segment that passes through pairs (i, j_i) in which i has all values from 0 to μ because $\eta(c)$ contains terms in which c has all of these exponents. The single segment of the polygon is $(j_0 - j_\mu)i + \mu j = k$, where $k = \mu j_0$. Then $(j_0 - j_\mu)1 + \mu j_1 = \mu j_0$, which shows that $j_0 - j_1$ is divisible by μ. Therefore, the segment can also be written $\sigma i + j = j_0$, where $\sigma = \frac{j_0 - j_\mu}{\mu}$; that is, τ can be taken to be 1, so that $s_1 = s$. Then (2) is divisible at least j_0 times by s (because $k = j_0$), whereas $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \gamma s^h + ts^h)$ is, by the definition of j_μ, divisible exactly j_μ times by s. In other words, adding the next term $\gamma s^{h+\sigma}$ to the truncated solution increases the number of times s divides $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + ts^h)$ from j_μ to at least $j_0 = j_\mu + \sigma\mu$.

Thus, if ν successive steps repeat the phenomenon of producing a single output truncated solution, it produces a truncated solution $y = \beta_0 + \cdots + \beta_h s^h + \gamma_1 s^{h+\sigma_1} + \gamma_2 s^{h+\sigma_1+\sigma_2} + \cdots + \gamma_\nu s^{h+\Sigma}$, where $\Sigma = \sigma_1 + \sigma_2 + \cdots + \sigma_\nu$, for which $\chi(\alpha + s^m, y + ts^{h+\Sigma})$ is divisible $j_\mu + \mu\Sigma$ times by s, say

$$\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + \cdots + \gamma_\nu s^{h+\Sigma} + ts^{h+\Sigma}) = s^{j_\mu + \mu\Sigma} q(s, t).$$

Differentiation with respect to t gives

$$\begin{aligned} &s^{h+\Sigma} \frac{\partial \chi}{\partial y}(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h + \cdots + \gamma_\nu s^{h+\Sigma} + ts^{h+\Sigma}) \\ &\quad = s^{j_\mu + \mu\Sigma} \frac{\partial q}{\partial t}(s, t). \end{aligned}$$

On the other hand, elimination of y between $\chi(x, y)$ and $\frac{\partial \chi}{\partial y}(x, y)$ (see Essay 1.3) gives, because the irreducibility of χ implies that these polynomials are relatively prime, an equation of the form

$$A(x, y)\chi(x, y) + B(x, y)\frac{\partial \chi}{\partial y}(x, y) = D(x),$$

in which $A(x, y)$, $B(x, y)$, and $D(x)$ are polynomials with integer coefficients. Substitution of $x = \alpha + s^m$ and $y = \beta_0 + \beta_1 s + \cdots + \beta_h s^h + \cdots + \gamma_\nu s^{h+\Sigma} + ts^{h+\Sigma}$ in $A(x,\ y)\chi(x, y) + B(x, y)\frac{\partial \chi}{\partial y}(x, y) = D(x)$ gives $D(\alpha + s^m)$ on the right and on the left gives a polynomial in s and t that is divisible at least $(j_\mu + \mu\Sigma) - (h + \Sigma)$ times by s. Thus $j_\mu + (\mu - 1)\Sigma - h$ is bounded above by the number of times s divides $D(\alpha + s^m)$. Since $\mu > 1$, this implies an upper bound on Σ; but $\Sigma \geq \nu$, because Σ is a sum of ν terms, each of which is at least 1, so ν is bounded above, and the proof of Theorem 1 is complete.[8] □

Theorem 2 *Every truncated solution of $\chi(x, y) = 0$ for $x = \alpha$ is a truncation of one of the infinite series solutions constructed by Theorem* 1.

Proof As was shown prior to the statement of the Newton polygon algorithm, if an infinite series solution is truncated, and the algorithm is applied to the result, one of the outputs is the truncated series with the next nonzero term after the truncation added. Therefore, any truncated solution is among the outputs if one starts with the truncated solutions $y = \beta_0$ in which $\chi(\alpha, \beta_0) = 0$ and repeatedly applies the algorithm. Since these are the truncated solutions constructed by Theorem 1, Theorem 2 follows. □

Essay 4.5 Determination of the Genus

> *On doit donner au problème une forme telle qu'il soit toujours possible de le résoudre, ce qu'on peut toujours faire d'un problème quelquonque.* (One should give the problem a form in which it will always be possible to solve it, which can always be done for any problem whatever.)—Niels Henrik Abel [6, p. 217]

I confess that the meaning of this dictum of Abel's is not altogether clear to me. Certainly it sounds like good advice, if one can understand what it means. My best guess is that he means something like what Kronecker meant when he said that one should require of one's definitions that one be able to determine by a finite calculation whether the definition is fulfilled in any given case. In the case of the determination of the genus of an algebraic field of transcendence degree one—the genus of a given algebraic curve—I believe both men would focus on constructive techniques like the ones given in this essay.

The genus was described in Essay 4.3 in terms of the dimensions of the spaces $\Theta(x^\nu)$ of elements of the root field of $\chi(x, y)$ (as always, an irreducible polynomial with integer coefficients that contains both x and y and is monic in y) that are integral over x and become integral over $\frac{1}{x}$ when they are divided by x^ν These dimensions can be determined easily once one constructs what Dedekind and Weber [22] called a **normal basis** of the root field.

Theorem *For a given $\chi(x, y)$, construct a subset $y_1, y_2, \ldots, y_n$ of its root field with the property that $y_1, y_2, \ldots, y_n$ is a* **basis** *of the field over the field of rational functions*

[8] Walker's proof of this point [82, p. 102] is not constructive, because he jumps from the observation that the ambiguity can never increase and can never go below 1 to the conclusion that he can find a step beyond which the ambiguity never decreases.

in x in the sense that each element w of the root field has a unique representation in the form

$$(1) \qquad w = \phi_1(x)y_1 + \phi_2(x)y_2 + \cdots + \phi_n(x)y_n,$$

where the coefficients $\phi_i(x)$ are rational functions of x, and is an **integral basis** *in the sense that w is integral over x if and only if each coefficient $\phi_i(x)$ in its representation (1) is a polynomial with rational coefficients, and further is a* **normal basis** *in the sense that w is in $\Theta(x^\nu)$ if and only if each coefficient $\phi_i(x)$ in its representation (1) is not only a polynomial but also satisfies $\deg \phi_i + \lambda_i \leq \nu$, where λ_i is the order of y_i at $x = \infty$ for each i, that is, the least integer for which y_i is in $\Theta(x^{\lambda_i})$.*

Proof Dedekind and Weber gave what appears to be an algorithm for constructing an integral basis (their §3), but their construction relies on the assumption that for a given constant α one can either find an element y that is integral over x and remains integral over x when it is divided by $x - \alpha$ or prove that there is no such y. The proof below uses, in essence, the method of Newton's polygon to justify this assumption and then constructs an integral basis using a method similar to theirs. However, they also assume that a polynomial with rational coefficients can be written as a product of linear factors—they assume complex number coefficients—and the proof below is a modified version of theirs that adjoins only the constants that are needed. The first step will be to find a *common denominator* of the elements integral over x.

The operation of multiplication by an element of the field can be described by the $n \times n$ matrix of rational functions of x that describes it with respect to the basis $1, y, \ldots, y^{n-1}$ of the root field as a vector space over the field of rational functions in x. In other words, an element z of the root field of $\chi(x, y)$ can be described by the matrix whose entry m_{ij} in the ith row of the jth column is the rational function of x that is the coefficient of y^{j-1} in the representation of zy^{i-1} with respect to the basis $1, y, \ldots, y^{n-1}$. The trace $\sum_{i=1}^{n} m_{ii}$ of the matrix obtained in this way is the **trace** of z with respect to x.

Lemma *If an element of the root field of $\chi(x, y)$ is integral over x, its trace is a polynomial in x with rational coefficients.*

Proof Let $\psi = p(x, y)/q(x)$ be integral over x. Then, by the definition of integrality, there is a relation of the form $F(\psi) = 0$, in which F is a monic polynomial with coefficients in $\mathbf{Q}[x]$. Since F can be written as a product of irreducible, monic polynomials with coefficients in $\mathbf{Q}[x]$, ψ must be a root of an *irreducible* monic polynomial with coefficients in $\mathbf{Q}[x]$; call it F_1.

By the proposition of Essay 2.3, the root field of $\chi(x, y)$, because it contains ψ and is generated over $\mathbf{Q}(x)$ by y, can be described by two adjunction relations $f_1(\psi) = 0$ and $f_2(y, \psi) = 0$, where f_1 and f_2 have coefficients that are rational functions of x, f_1 is monic of degree ν_1, say, and is irreducible, while f_2 is monic of degree ν_2, say, in y and is irreducible as a polynomial in y with coefficients in the field of rational functions in x with ψ adjoined. Because f_1 and F_1 both have ψ as a root, because both are monic with coefficients that are rational functions of x, and because both are irreducible over the field of rational functions in x (F_1 is irreducible in this sense by virtue of Gauss's lemma), $f_1 = F_1$. In particular, the coefficients of f_1 are not just rational functions of x, they are polynomials in x with rational coefficients.

The trace of ψ is by definition the trace of the matrix that represents multiplication by ψ relative to the basis $1, y, \ldots, y^{n-1}$ of the root field over the field of rational functions in x. Therefore (because $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, so $\mathrm{tr}(M^{-1}AM) = \mathrm{tr}(AMM^{-1}) = \mathrm{tr}(A)$), it is the trace of the matrix that represents multiplication by ψ relative to *any* basis. In particular, it is the trace of the matrix that represents multiplication by ψ relative to the basis $\psi^i y^j$, $0 \le i < \nu_1$, $0 \le j < \nu_2$. When the elements of this basis are suitably ordered, the matrix that represents multiplication by ψ becomes a $\nu_2 \times \nu_2$ matrix of $\nu_1 \times \nu_1$ blocks (note that $\nu_1\nu_2 = n$) in which the blocks off the diagonal are all 0 and the blocks on the diagonal are all the same matrix: Its first $\nu_1 - 1$ rows are the last $\nu_1 - 1$ rows of I_{ν_1}, and its last row contains the negatives of the coefficients (after the first) of the polynomial $f_1 = F_1$, listed in reverse order. In particular, its entries are all in $\mathbf{Q}[x]$, so its trace is in $\mathbf{Q}[x]$. (In fact, its trace—and therefore the trace of ψ—is simply $-\nu_2$ times the second coefficient of F_1.) □

The matrix, call it S, whose entry in the ith row of the jth column is the trace of y^{i+j-2}, is a matrix of polynomials in x with integer coefficients. Therefore, its determinant, call it $D(x)$, is a polynomial in x with integer coefficients. The lemma implies that $D(x)$ *is a common denominator of the elements of the root field integral over* x. In fact, if $p(x, y)/q(x)$ is integral over x, where $p(x, y)$ and $q(x)$ are polynomials with rational coefficients and $q(x) \neq 0$, and if it is in lowest terms, then not only does $q(x)$ divide $D(x)$, but so does $q(x)^2$. This can be proved as follows:

The matrix S of which $D(x)$ is the determinant represents the bilinear form "the trace of the product" on the root field of $\chi(x, y)$ relative to the basis $1, y, \ldots, y^{n-1}$. This observation implies that $D(x) \neq 0$, because if $D(x)$ were zero, there would be a solution $v(x)$ of $S \cdot v(x) = 0$ that was a nonzero column matrix whose entries $v_i(x)$ were rational functions of x, and this would imply $\mathrm{tr}_x(\hat{w}\hat{v}) = 0$ for all elements $\hat{w}$ of the root field, where $\hat{v} = \sum_{i=1}^{n} v_i(x)y^{i-1}$, contrary to the fact that $\mathrm{tr}_x(\hat{w}\hat{v}) = n$ when $\hat{w}$ is the reciprocal of $\hat{v}$.

If $p(x, y)/q(x)$ is integral over x and in lowest terms in the sense that $q(x)$ and the coefficients $p_i(x)$ of $p(x, y) = p_0(x) + p_1(x)y + \cdots + p_{n-1}(x)y^{n-1}$ have no common divisor of positive degree, and if one of the coefficients $p_i(x)$ is nonzero, then a new basis of integral elements is obtained by replacing y^i with $p(x, y)/q(x)$ in the basis $1, y, \ldots, y^{n-1}$ The entries of the matrix that represents the bilinear form "the trace of the product" relative to this new basis are polynomials in x with rational coefficients, because they are traces of elements integral over x. Therefore, its determinant is a polynomial in x. On the other hand, its determinant is $\left(\frac{p_i(x)}{q(x)}\right)^2 D(x)$, because the matrix that makes the transition from one basis to the other is the identity matrix with the $(i + 1)$st row replaced by a new row consisting of the coefficients of $p(x, y)/q(x)$ and which therefore has $\frac{p_i(x)}{q(x)}$ in its $(i + 1)$st column, so both the transition matrix and its transpose have determinant $\frac{p_i(x)}{q(x)}$. Therefore, $q(x)^2$ divides $p_i(x)^2 D(x)$ for each i (trivially so when $p_i(x) = 0$). Thus, $q(x)^2$ divides the greatest common divisor of these polynomials $p_i(x)^2 D(x)$, which is the greatest common divisor of the $p_i(x)^2$ times $D(x)$. Since $p(x, y)/q(x)$ is in lowest terms by assumption, $q(x)^2$ is relatively prime to the greatest common divisor of the $p_i(x)^2$, so $q(x)^2$ divides $D(x)$, as was to be shown.

Since every element $p(x, y)/q(x)$ of the root field can be written in the form $P(x, y) + \frac{r(x,y)}{q(x)}$, where $P(x, y)$ is a polynomial in x and y and where $\frac{r(x,y)}{q(x)}$ is a proper fraction in the sense that $\deg_x r < \deg q$ (and, as it is natural to assume, $\deg_y r < n = \deg_y \chi$), in order to determine which elements $p(x, y)/q(x)$ are integral over x it will suffice to determine which elements $r(x, y)/q(x)$ are integral over x, because polynomials $P(x, y)$ are always integral over x. But since $r(x, y)/q(x)$ for a given $q(x)$ contains just $n \cdot \deg q$ unknown rational coefficients, the following proposition reduces this determination to the solution of a system of homogeneous linear equations.

Proposition *A rational function $p(x, y)/q(x)$ with rational coefficients is integral over x if and only if for each algebraic number α that is a root of $q(x)$ and for each infinite series solution y of $\chi(x, y) = 0$ in fractional powers of $x - \alpha$ given by Newton's polygon, all terms of the power series in s that results from substituting the series for y in $p(x, y)$ and then substituting $\alpha + s^m$ for x, where m clears the denominators in the functional exponents, are divisible by s at least as many times as the polynomial $q(\alpha + s^m)$ is.*

Loosely speaking, the condition is that each expression of $p(x, y)/q(x)$ obtained by using an expansion of y as a power series in fractional powers of $x - \alpha$, where α is a root of $q(x)$, and writing the reciprocal of $q(x)$ as a negative power of $x - \alpha$ times a power series in $x - \alpha$ with nonzero constant term, contains nonnegative exponents exclusively; in short, $p(x, y)/q(x)$ *has no poles where x is finite*.

Proof Let $\Psi(x, p) = p^\nu + c_1(x)p^{\nu-1} + \cdots + c_\nu(x)$ be the irreducible, monic polynomial[9] whose coefficients $c_i(x)$ are rational functions of x of which $p(x, y)$—regarded as an element of the root field—is a root. Because $p(x, y)$ is integral over x (it is a polynomial in y with coefficients in $\mathbf{Q}[x]$), the coefficients $c_i(x)$ are polynomials in x with rational coefficients. To say that $p(x, y)/q(x)$ is integral over x means that $c_i(x)$ is divisible by $q(x)^i$ for each i. It is to be shown that this is true if and only if $p(\alpha + s^m, \beta_0 + \beta_1 s + \beta_2 s^2 + \cdots) \equiv 0 \bmod s^e$ for all infinite series solutions $y = \beta_0 + \beta_1 s + \beta_2 s^2 + \cdots$ of $\chi(x, y) = 0$, where α is a root of $q(x)$, where $s = (x - \alpha)^{1/m}$, and where e is the number of times that s divides $q(\alpha + s^m)$.

By definition, to say that $\Psi(x, p) = 0$, where x and p are regarded as elements of the root field, means that $\Psi(x, p(x, y)) \equiv 0 \bmod \chi(x, y)$. In other words, it means that $\Psi(x, p(x, y)) = q(x, y)\chi(x, y)$ for some polynomial $q(x, y)$ with rational coefficients. Since $\chi(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h) \equiv 0 \bmod s^{h+1}$ for each h, it follows that $\Psi(\alpha + s^m, p(\alpha + s^m, \beta_0 + \beta_1 s + \cdots + \beta_h s^h)) \equiv 0 \bmod s^{h+1}$ for each h. Therefore, $p(\alpha + (x - \alpha), \beta_0 + \beta_1(x - \alpha)^{1/m} + \cdots + \beta_h(x - \alpha)^{h/m})$, when it is regarded as a polynomial in $(x - \alpha)^{1/m}$ and truncated by omitting all terms in which the exponent is larger than h/m, is a truncated solution of $\Psi(x, p) = 0$ for $x = \alpha$. By Theorem 2

[9] This polynomial can be found because $\Psi(x, p)$ is a factor of the characteristic polynomial of the matrix that represents multiplication by $p(x, y)$ relative to the basis $1, y, \ldots, y^{n-1}$.

of the last essay, it is therefore a truncation of one of the infinite series solutions p of $\Psi(x, p) = 0$ for $x = \alpha$ found by the construction[10] of Theorem 1 of the last essay. It is to be shown, therefore, that $q(x)^i$ divides $c_i(x)$ for each i if and only if for each root α of $q(x)$ in an algebraic number field, every infinite series solution p of $\Psi(x, p) = 0$ in fractional powers of $x - \alpha$ is divisible by the highest power of $x - \alpha$ that divides $q(x) = q(\alpha + (x - \alpha))$.

Suppose first that $q(x)^i$ divides $c_i(x)$ for each i. For a given root α of $q(x)$ whose multiplicity is e, $(x - \alpha)^{ei}$ then divides $c_i(x)$. The initial term of any infinite series solution p of $\Psi(x, p) = 0$ in fractional powers of $x - \alpha$ can be found using the method by which the Newton's polygon algorithm finds the next term of a truncated series solution. Specifically, the equation $\Psi(\alpha + s, p) = c_\nu(\alpha + s) + c_{\nu-1}(\alpha + s)p + \cdots + c_1(\alpha + s)p^{\nu-1} + p^\nu = 0$ shows, because the terms of lowest degree cancel, that the lowest order term of a series expansion $p = \gamma s^{\sigma/\tau} + \cdots$ corresponds to a segment of the "Newton polygon" dictated by the points (i, j_i), where j_i for $i = 0, 1, \ldots, \nu$ is the number of times $s = x - \alpha$ divides $c_{\nu-i}(\alpha + s)$, except that j_i is undefined when $c_{\nu-i}(x) = 0$. Since $(x - \alpha)^{e(\nu-i)}$ divides $c_{\nu-i}(x)$, j_i is at least $e(\nu - i)$ whenever it is defined. In particular, the minimum value 0 of j_i occurs only for $i = \nu$. The rightmost segment of the polygon, call it $\sigma i + \tau j = k$, therefore has $(\nu, 0)$ as its right end; its other end is at a point (i, j_i) for which $\sigma i + \tau j_i = k = \sigma\nu + \tau \cdot 0$. For this index i, both $j_i = \frac{\sigma}{\tau}(\nu - i)$ and $j_i \geq e(\nu - i)$ hold. Therefore, for this segment of the polygon, $\frac{\sigma}{\tau} \geq e$. All infinite series solutions $p = \gamma(x - \alpha)^{\sigma/\tau} + \cdots$ that correspond to this segment of the polygon are therefore divisible by $(x - \alpha)^e$. As is easily shown, the ratio $\frac{\sigma}{\tau}$ is *smallest* for this rightmost segment,[11] so all solutions $p = \gamma(x - \alpha)^{\sigma/\tau} + \cdots$ are divisible by $(x - \alpha)^e$, as was to be shown.

[10] Strictly speaking, this construction does not apply to $\Psi(x, p)$, because its coefficients are rational and the description of the Newton polygon algorithm in Essay 4.4 assumes that the coefficients of the given equation $\chi(x, y) = 0$ are integers, but the algorithm applies without modification to the case of rational coefficients. Moreover, $\chi(x, y)$ was assumed in Essay 4.4 to be irreducible. The series expansions of a reducible polynomial can be found by finding the expansions of its irreducible factors.

[11] What is to be shown is that the ratio σ/τ for any segment of the polygon is *larger* than the ratio σ/τ for the segment to its right. Since σ/τ is minus the slope of the segment, this is the statement that the slopes of the segments *increase* as one moves from left to right, which is evident. In actual inequalities, the three endpoints of two successive segments of Newton's polygon, call them (r, j_r), (s, j_s), (t, j_t), satisfy

$$\sigma r + \tau j_r = \sigma s + \tau j_s < \sigma t + \tau j_t,$$
$$\sigma' r + \tau' j_r > \sigma' s + \tau' j_s = \sigma' t + \tau' j_t,$$

where σ' and τ' pertain to the segment from (s, j_s) to (t, j_t), from which

$$\tau(j_r - j_s) = \sigma(s - r) \qquad \text{and} \qquad \tau'(j_r - j_s) > \sigma'(s - r)$$

follow. Therefore

$$\frac{\sigma}{\tau} = \frac{j_r - j_s}{s - r} > \frac{\sigma'}{\tau'},$$

as was to be shown.

Conversely, if $q(x)^i$ fails to divide $c_i(x)$ for some i, then $(x-\alpha)^{ei}$ fails to divide $c_i(x)$ for some root α of multiplicity e of $q(x)$ and some index i. For such an α the points (i, j_i) of the polygon arising from $\Psi(\alpha+s, p) = c_\nu(\alpha+s)+c_{\nu-1}(\alpha+s)p+\ldots+p^\nu$ include at least one for which $e(\nu-i) > j_i$. If $j_i = 0$ for some $i < \nu$, then $\Psi(\alpha, p)$ contains a term of degree less than ν in p, so this polynomial in p has a nonzero root, call it β_0, and there is a solution $p = \beta_0 + \cdots$ of $\Psi(\alpha, p) = 0$ that is not divisible by $s = x - \alpha$, and therefore not divisible by $(x-\alpha)^e$. Otherwise, as before, the rightmost segment of the polygon, call it $\sigma i + \tau j = k$, passes through $(\nu, 0)$ and at least one other point of the form (i, j_i). At least one point (i, j_i) lies below the line $j = e(\nu - i)$ of slope $-e$ passing through $(\nu, 0)$; since all points (i, j_i) lie on or above any segment of the polygon, the rightmost segment $j = \frac{\sigma}{\tau}(\nu - i)$ must lie under the line $j = e(\nu - i)$ for $i < \nu$. Thus, $\frac{\sigma}{\tau} < e$, so no solution $p = \gamma(x-\alpha)^{\sigma/\tau} + \cdots$ arising from this segment of the polygon is divisible by $(x-\alpha)^e$, and the proof of the proposition is complete. □

Thus, in a proper fraction $r(x, y)/q(x)$ that is integral over x, the coefficients of $r(x, y)$ satisfy a homogeneous system of linear equations, so the most general such fraction can be written as a linear combination of a finite number of them, say of $\xi_1, \xi_2, \ldots, \xi_k$, with rational coefficients. When these elements $\xi_1, \xi_2, \ldots, \xi_k$ together with $1, y, y^2, \ldots, y^{n-1}$ are taken as input to the following algorithm of Kronecker [56, §7], the algorithm produces an *integral basis* of the root field of $\chi(x, y)$ as described in the statement of the theorem.

Construction of an Integral Basis

Input: Elements $y_1, y_2, \ldots, y_l$ of the root field of $\chi(x, y)$ integral over x that span the elements integral over x in the sense that each element integral over x can be expressed in the form $\sum_{i=1}^{l} \phi_i(x) y_i$ where the coefficients $\phi_i(x)$ are polynomials in x with rational coefficients. (At the outset, $l = n + k$, and the coefficients of the ξ_i can be taken to be rational numbers.)

Algorithm: *As long as the number l of elements in the spanning set is greater than n, carry out the following operations. Consider the $l \times l$ symmetric matrix* $[\mathrm{tr}_x(y_i y_j)]$ *and consider its symmetric $n \times n$ minor determinants—those $n \times n$ minor determinants in which the indices of the n columns selected coincide with those of n the rows selected. Each such minor determinant is a polynomial in x with rational coefficients because all of its entries are. Rearrange* y_1, $y_2, \ldots,$ y_l, *if necessary, to make the first such minor—the one formed by selecting the first n rows and columns—nonzero and of degree no greater than that of any other nonzero symmetric $n \times n$ minor. Then the first n entries of* y_1, $y_2, \ldots,$ y_l *are linearly independent over* $\mathbf{Q}(x)$, *which means that each remaining entry* y_{n+1}, $y_{n+2}, \ldots,$ y_l *can be expressed as a sum of multiples of the first n in which the multipliers are rational functions of x. Each multiplier in each of these expressions can be written as a polynomial in x plus a proper rational function of x, one in which the degree of the numerator is less than the degree of the denominator. Let polynomial multiples of the first n of the y's be subtracted from the later y's in order to make the multipliers in the representations of the later y's* *in terms of the first n all proper rational functions. Delete any y's that have become zero as a result of these subtractions, rearrange the list again, and repeat.*

Output: A list $y_1, y_2, \dots, y_n$ of just n elements integral over x that span, over $\mathbf{Q}[x]$, the set of all elements integral over x.

The operations of the algorithm—rearrange the y's, delete zeros, and subtract one y times a polynomial in x with rational coefficients from another y—do not change the conditions satisfied by the original set of y's that they span the elements integral over x when coefficients that are polynomials in x with rational coefficients are used.

An argument like the one above that proves that $D(x)$ is a common denominator of the elements integral over x proves that *each iteration of the algorithm reduces the degree of the determinant of the first $n \times n$ symmetric minor.* Specifically, if, after the multipliers in the representations of $y_{n+1}, y_{n+2}, \dots, y_l$ as sums of multiples of $y_1, y_2, \dots, y_n$ have been reduced so that they are proper rational functions, and after zeros have been deleted, there are more than n items in the list, then one of the coefficients—say the coefficient of y_1—in the representation of y_{n+1} is a nonzero proper fraction, call it $\frac{p(x)}{q(x)}$, where $\deg p < \deg q$. The symmetric $n \times n$ minor for any selection of n indices is a polynomial. As before, $M_1 = \left(\frac{p(x)}{q(x)}\right)^2 M_0$ when M_1 is the minor in which the selected indices are $2, 3, \dots, n+1$ and M_0 is the one in which they are $1, 2, \dots, n$. Thus, $q(x)^2 M_1 = p(x)^2 M_0$, which shows that $\deg M_1 < \deg M_0$. Thus, the minor of *least* degree has degree less than $\deg M_0$, and $\deg M_0$ decreases with each step, as was to be shown.

In this way, the algorithm continues to reduce the degree of the first $n \times n$ minor. By the principle of infinite descent, the algorithm must terminate. In other words, a stage must be reached at which the list contains only n elements. Clearly, they are an integral basis of the root field.

The proof of the theorem will be completed by a second algorithm, which starts with an integral basis and produces a normal basis. It requires that one also construct an integral basis relative to the parameter $u = \frac{1}{x}$; in other words, it uses a set $z_1, z_2, \dots, z_n$ of elements of the root field of $\chi(x, y)$ with the property that every element of the root field has a unique representation in the form $\sum \psi_i(x) z_i$, where the coefficients $\psi_i(x)$ are rational functions of x, and that the element is integral over $u = \frac{1}{x}$ if and only if each $\psi_i(x)$ is a polynomial in $\frac{1}{x}$. The algorithm just given can be used to construct such a set $z_1, z_2, \dots, z_n$; simply describe the root field as the root field of $\chi_1(u, v) = \chi(x, y)/x^{n\lambda}$, where $u = \frac{1}{x}$, $v = \frac{y}{x^\lambda}$, and λ is large enough to make χ_1 a polynomial in u and v.

Such an integral basis $z_1, z_2, \dots, z_n$ relative to $\frac{1}{x}$ will be used to determine, given an integral basis $y_1, y_2, \dots, y_n$, whether the basis

$$\frac{y_1}{x^{\lambda_1}}, \quad \frac{y_2}{x^{\lambda_2}}, \quad \dots, \quad \frac{y_n}{x^{\lambda_n}},$$

is an integral basis relative to $\frac{1}{x}$, where λ_i, for each i, is the order of y_i at $x = \infty$; that is, λ_i is the least integer for which y_i/x^{λ_i} is integral over $\frac{1}{x}$.

Construction of a Normal Basis

Input: An integral basis $y_1, y_2, \ldots, y_n$ of the root field of $\chi(x, y)$ relative to x.

Algorithm: *Find the orders $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $y_1, y_2, \ldots, y_n$ at $x = \infty$. As long as $\frac{y_1}{x^{\lambda_1}}, \frac{y_2}{x^{\lambda_2}}, \ldots, \frac{y_n}{x^{\lambda_n}}$ (which is a basis consisting of elements integral over $\frac{1}{x}$) is not an integral basis relative to $\frac{1}{x}$, construct a new integral basis in which one y_k is replaced by a new y'_k whose order λ'_k at $x = \infty$ is less than λ_k in the following way. Write each z_i of an integral basis relative to $\frac{1}{x}$ in the form $\sum_j \psi_{ij}(x)\frac{y_j}{x^{\lambda_j}}$, where the $\psi_{ij}(x)$ are rational functions of x. By assumption, at least one $\psi_{ij}(x)$ is not a polynomial in $\frac{1}{x}$. (If all were polynomials in $\frac{1}{x}$, then each z_i and therefore each element integral over $\frac{1}{x}$ would be a sum of multiples of the $\frac{y_i}{x^{\lambda_i}}$ with coefficients that were polynomials in $\frac{1}{x}$.) Choose a value of i for which at least one $\psi_{ij}(x)$ is not a polynomial in $\frac{1}{x}$. Since $x^\nu z_i = \sum \psi_{ij}(x) x^{\nu-\lambda_j} y_j$ is integral over x for sufficiently large ν, and $y_1, y_2, \ldots, y_n$ is an integral basis, the denominator of $\psi_{ij}(x)$ is a power of x for each $j = 1, 2, \ldots, n$, say $\psi_{ij}(x) = x\xi_j(x) + \theta_j\left(\frac{1}{x}\right)$, where $\xi_j(x)$ is a polynomial in x, and $\theta_j\left(\frac{1}{x}\right)$ is a polynomial in $\frac{1}{x}$. By the choice of i, $\xi_j(x) \neq 0$ for at least one j. Let $\sigma > 0$ be the maximum of the degrees of $\xi_1(x), \xi_2(x), \ldots, \xi_n(x)$. Among those indices j for which* $\deg \xi_j = \sigma$, *let k be one for which λ_k is as large as possible and set $y'_k = \sum c_j x^{\lambda_k - \lambda_j} y_j$, where c_j is the coefficient of x^σ in $\xi_j(x)$ (which is zero if* $\deg \xi_j \neq \sigma$).

Output: An integral basis $y_1, y_2, \ldots, y_n$ with the property that

$$\frac{y_1}{x^{\lambda_1}}, \frac{y_2}{x^{\lambda_2}}, \ldots, \frac{y_n}{x^{\lambda_n}},$$

is an integral basis relative to $\frac{1}{x}$.

Justification Replacement of y_k with y'_k gives an integral basis, as is shown by the two formulas $y'_k = \sum_j c_j x^{\lambda_k - \lambda_j} y_j$ (note that $\lambda_k \geq \lambda_j$ for all j by the choice of k) and $y_k = \frac{1}{c_k} y'_k - \sum_{j \neq k} \frac{c_j}{c_k} y_j$ (note that $c_k \neq 0$ by the choice of k). All that is to be shown, then, is that $\lambda'_k < \lambda_k$. To this end, note that $\frac{z_i}{x^{\sigma+1}} = \sum (c_j + \cdots) \cdot \frac{y_j}{x^{\lambda_j}}$, where the omitted terms contain $\frac{1}{x}, \frac{1}{x^2}, \frac{1}{x^3}, \ldots$. Multiply by x and use the definition of y'_k to obtain $\frac{z_i}{x^\sigma} = x \cdot \frac{y'_k}{x^{\lambda_k}} + \sum \eta_j\left(\frac{1}{x}\right) \cdot \frac{y_j}{x^{\lambda_j}}$, where $\eta_j\left(\frac{1}{x}\right)$ for each j is $x \cdot \frac{\psi_{ij}(x) - c_j x^{\sigma+1}}{x^{\sigma+1}}$, which is a polynomial in $\frac{1}{x}$. Thus, $x \cdot \frac{y'_k}{x^{\lambda_k}}$ is a difference of elements integral over $\frac{1}{x}$, which implies that the order of y'_k at $x = \infty$ is at most $\lambda_k - 1$, as was to be shown.

Since the algorithm reduces the sum of the λ_i at each step, it must terminate by the principle of infinite descent. When it terminates, the integral basis $y_1, y_2, \ldots, y_n$ is a normal basis, because $w = \sum \phi_i(x) y_i$ has order at most ν if and only if all coefficients of $\frac{w}{x^\nu} = \sum \frac{\phi_i(x)}{x^{\nu - \lambda_i}} \cdot \frac{y_i}{x^{\lambda_i}}$ are polynomials in $\frac{1}{x}$, which is true if and only if $\deg \phi_i \leq \nu - \lambda_i$, and the proof of the theorem is complete. □

If $y_1, y_2, \ldots, y_n$ is a normal basis of the root field of $\chi(x, y)$, the elements of $\Theta(x^\nu)$ are those whose representations in the form $\sum_i \phi_i(x) y_i$ have coefficients $\phi_i(x)$

that are polynomials in x, with rational coefficients, of degree at most $\nu - \lambda_i$ for each i. When $\nu < \lambda_i$ this condition of course means that $\phi_i(x) = 0$. Therefore, the dimension of $\Theta(x^\nu)$ as a vector space over $\mathbf{Q}$ is the sum of the numbers $\nu - \lambda_i + 1$ over all indices i for which $\lambda_i \leq \nu$. For large ν, then, the dimension of $\Theta(x^\nu)$ as a vector space over $\mathbf{Q}$ is exactly $(\nu + 1)n - \sum \lambda_i$. At the other extreme, when $\nu = 0$ this dimension—which is the degree of the field of constants $\Theta(x^0)$ as an extension of $\mathbf{Q}$, denoted by c in Essay 4.3—is simply the number of indices i for which $\lambda_i = 0$.

In the notation of Essay 4.3, the genus of the root field of $\chi(x, y)$ is $g = n_0\nu - \dim \Theta(x^\nu) + 1$ for all sufficiently large ν, where $n_0 = n/c$ and the dimension is the dimension as a vector space over the field of constants, which is the dimension as a vector space over $\mathbf{Q}$ divided by c; thus,

$$g = n_0\nu - \frac{1}{c}\Big((\nu + 1)n - \sum \lambda_i\Big) + 1 = \frac{\sum \lambda_i}{c} - (n_0 - 1).$$

In particular, when $\mathbf{Q}$ is the field of constants of the root field of $\chi(x, y)$, the genus of the root field is simply

$$\Big(\sum \lambda_i\Big) - (n - 1),$$

where $n = \deg_y \chi$ and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the orders of the elements $y_1, y_2, \dots, y_n$ of a normal basis of the field.

As the discussion of Essay 4.3 already shows, the natural description of the genus uses the field of constants of the root field under consideration instead of the field of rational numbers:

Determination of the Genus *As was just explained, the construction of the theorem gives a basis over* $\mathbf{Q}$ *of the field of constants of the root field of* $\chi(x, y)$, *namely, the elements* y_i *of order zero in a normal basis. When the field* $\mathbf{Q}$ *is replaced by the (possibly) larger field of constants in the theorem, the construction gives a subset* $y_1, y_2, \dots, y_{n_0}$ *of the root field of* $\chi(x, y)$ *and nonnegative integers* $\mu_1, \mu_2, \dots, \mu_{n_0}$ *with the property that the elements of* $\Theta(x^\nu)$ *for any given* ν *are precisely those of the form*

$$\phi_1(x)y_1 + \phi_2(x)y_2 + \cdots + \phi_{n_0}(x)y_{n_0}$$

where $\phi_i(x)$ *is a polynomial of degree at most* $\nu - \mu_i$ *in* x *whose coefficients are in the field of constants of the root field of* $\chi(x, y)$. *Thus, for large* ν *the dimension of* $\Theta(x^\nu)$ *as a vector space over the field of constants is* $\sum_{i=1}^{n_0}(\nu - \mu_i + 1) = n_0\nu - \sum \mu_i + n_0$. *By the definition of the genus, this dimension is* $n_0\nu - g + 1$, *from which it follows that*

$$g = \Big(\sum_{i=1}^{n_0} \mu_i\Big) - (n_0 - 1).$$

In particular, $\sum \mu_i \geq n_0 - 1$.

Example 1 $\chi(x, y) = y^3 - xy + x^3$ (the folium of Descartes, shown in Fig. 4.6).

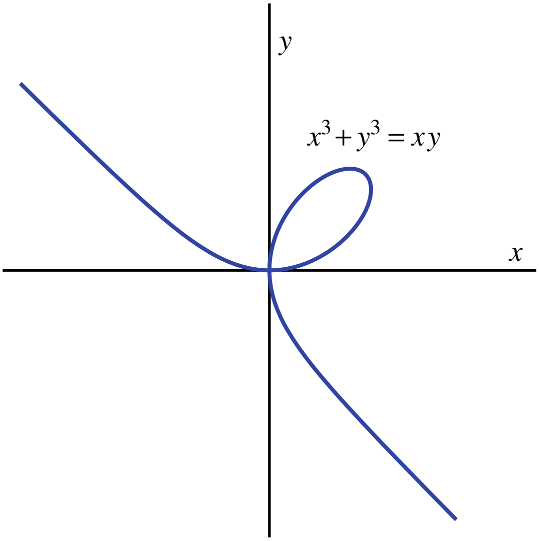

Fig. 4.6 The folium of Descartes

Multiplication by y is represented by

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -x^3 & x & 0 \end{bmatrix}$$

relative to the basis $1, y, y^2$ of the root field over $\mathbf{Q}(x)$. Therefore, the trace of y is 0. The trace of y^2 is the trace of

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -x^3 & x & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 1 \\ -x^3 & x & 0 \\ 0 & -x^3 & x \end{bmatrix},$$

which is $2x$. The trace of $y^3 = xy - x^3$ is x times the trace of y plus $-x^3$ times the trace of 1, which is $x \cdot 0 - x^3 \cdot 3 = -3x^3$. Similarly, the trace of $y^4 = xy^2 - x^3y$ is $2x^2$, from which it follows that

$$S = \begin{bmatrix} 3 & 0 & 2x \\ 0 & 2x & -3x^3 \\ 2x & -3x^3 & 2x^2 \end{bmatrix} \quad \text{and} \quad D(x) = 12x^3 - 8x^3 - 27x^6 = x^3(4 - 27x^3).$$

The square of the denominator $q(x)$ of an element of the root field integral over x must divide $x^3(4 - 27x^3)$, so x is a common denominator of these integral elements.

A proper fraction integral over x must therefore be of the form $\frac{a+by+cy^2}{x}$, where a, b, and c are rational numbers.

By the proposition, and by the fact that $y = \pm\sqrt{x} - \cdots$ and $y = x^2 + \cdots$ are the series expansions of y in fractional powers of x, such an expression is integral over x if and only if $a + b(\pm s) + c(\pm s)^2 \equiv 0 \bmod s^2$ and $a + bs^2 + cs^4 \equiv 0 \bmod s$. These conditions hold if and only if $a = b = 0$, so the proper fractions integral over x are the rational multiples of $\frac{y^2}{x}$. Thus, $1, y, \frac{y^2}{x}$ are an integral basis. For this basis, $\lambda_1 = 0$ and $\lambda_2 = 1$. To find the order λ_3 of $y_3 = \frac{y^2}{x}$ at $x = \infty$, one needs to find the equation of which it is a root, which is the characteristic polynomial of $\frac{1}{x}\begin{bmatrix} 0 & 0 & 1 \\ -x^3 & x & 0 \\ 0 & -x^3 & x \end{bmatrix}$. This characteristic polynomial is $X^3 - 2X^2 + X - x^3$, so $y_3^3 - 2y_3^2 + y_3 - x^3 = 0$, and $\left(\frac{y_3}{x}\right)^3 - 2 \cdot \frac{1}{x} \cdot \left(\frac{y_3}{x}\right)^2 + \frac{1}{x^2} \cdot \left(\frac{y_3}{x}\right) - 1 = 0$, which makes it clear that $\lambda_3 = 1$.

With $u = \frac{1}{x}$ and $v = \frac{y}{x}$ the equation $v^3 - uv + 1 = 0$ holds. That $1, v, v^2$ is an integral basis of the root field of $v^3 - uv + 1$ follows from the fact that in this case

$$S = \begin{bmatrix} 3 & 0 & 2u \\ 0 & 2u & -3 \\ 2u & -3 & 2u^2 \end{bmatrix}, \quad \text{from which} \quad D(u) = 4u^3 - 27.$$

Since $D(u)$ is square-free, $1, v, v^2$ is an integral basis over u. Thus, $1, y, y^2/x$ is a normal basis, because $1, \frac{y}{x}, \frac{y^2/x}{x}$ is the integral basis $1, v, v^2$ over u.

In this case, then, $\mathbf{Q}$ is the field of constants, and the genus is $(0+1+1)-(3-1) = 0$.

Example 2 $\chi(x, y) = y^3 + x^3y + x$ (the Klein curve).

In this case, $D(x) = -4x^9 - 27x^2$, whose only square factor is x^2, so again the proper fractions integral over x have the form $\frac{a+by+cy^2}{x}$, where a, b, and c are rational numbers. Application of Newton's polygon in the case $\alpha = 0$ leads easily to three unambiguous truncated solutions of $y^3 + x^3y + x = 0$, namely, $y = \gamma\sqrt[3]{x}$, where γ is a cube root of -1. Substitution of $y = -s + \cdots$ for y and of s^3 for x in $a + by + cy^2$ gives a series divisible by $x = s^3$ only if $a = b = c = 0$, so $1, y, y^2$ is an integral basis over x. The orders of the first two are 0 and 2, respectively. The third, call it $w = y^2$, is a root of the characteristic polynomial of

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -x & -x^3 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 1 \\ -x & -x^3 & 0 \\ 0 & -x & -x^3 \end{bmatrix};$$

therefore, $w^3 + 2x^3w^2 + x^6w - x^2 = 0$, from which it is clear that the order of w at $x = \infty$ is 3. (Division by x^9 gives an equation showing that w/x^3 is integral over $1/x$, but division by x^6 gives one that shows that w/x^2 is not integral over $1/x$.) That $1, y, y^2$ is a normal basis follows from the observation that $1, \frac{y}{x^2}, \frac{y^2}{x^3}$ is an integral basis over $u = \frac{1}{x}$, because division of $y^3 + x^3y + x = 0$ by x^6 gives $v^3 + uv + u^5 = 0$, where $v = \frac{y}{x^2}$, and because, as is easily shown, $1, v = \frac{y}{x^2}, \frac{v^2}{u} = \frac{y^2}{x^3}$ is an integral basis over u. Since $\lambda_1 = 0$, $\lambda_2 = 2$, and $\lambda_3 = 3$, it follows that $\mathbf{Q}$ is the field of constants, and the genus is $(0 + 2 + 3) - (3 - 1) = 3$.

Example 3 $\chi(x, y) = (x^2 + y^2)^2 - 2(x + y)^2$ (see Essay 4.3).

As was noted in Essay 4.3, the algebraic analysis of this example should begin with the observation that the root field of $\chi(x, y)$ contains a square root of 2 in the form of the element $\frac{x^2+y^2}{x+y}$, which enables one to treat the root field as the root field of the polynomial $x^2 + y^2 - \sqrt{2}(x + y)$, whose degree in y is 2 instead of 4.

(The irrational constants in the root field, if there are any, can be found by constructing one solution y of $\chi(x, y) = 0$ for one rational value α of x; the field of constants $\mathbf{A}$ needed to express such a solution must contain all constants in the root field, because the solution makes it possible to express any element of the root field as a power series—possibly with some negative exponents—with coefficients in $\mathbf{A}$, and in particular to express any constant in the root field as an element of $\mathbf{A}$. For example, when $\alpha = 0$ the roots of $\chi(0, y) = y^4 - 2y^2$ yield two unambiguous truncated solutions $y = \pm\sqrt{2} \cdot x^0$ and the truncated solution $y = 0 \cdot x^0$, whose ambiguity is 2. If the ambiguous solution $y = 0 \cdot x^0$ is used as an input to Newton's polygon, the output is the truncated solution $y = -1 \cdot x$, with ambiguity 2. If this truncated solution is the input, there are two unambiguous outputs $y = -x \pm \sqrt{2} \cdot x^2$, for each of which $\sqrt{2}$ must be adjoined. Thus, $\mathbf{A} = \mathbf{Q}(\sqrt{2})$ for any one of the four infinite series solutions for $\alpha = 0$, and no cleverness is needed to discover the irrational constant $\sqrt{2}$ in the root field. For any $\chi(x, y)$, the construction of a single unambiguous solution $x = \alpha + s^m$, $y = \beta_0 + \beta_1 s + \cdots$ of $\chi(x, y) = 0$ gives a number field $\mathbf{A}$ that contains, for the same reason, all constants in the root field of $\chi(x, y)$; factorization of $\chi(x, y)$ over such an $\mathbf{A}$ will then show the extent to which the adjunction of constants can reduce the degree in y of $\chi(x, y)$, or, more precisely, will determine the degree of the root field as an extension of $\mathbf{A}(x)$.)

The elements $1, \sqrt{2}, y, \sqrt{2}y$ are easily shown to be a normal basis in which the λ's are $0, 0, 1, 1$, respectively, so the genus is $\frac{1}{c}\sum \lambda_i - (n_0 - 1) = \frac{1}{2}(0+0+1+1) - (2-1) = 0$. When $\mathbf{Q}$ is replaced by $\mathbf{Q}(\sqrt{2})$, 1 and y are a normal basis in which the λ's are 0 and 1 respectively, and the genus is $(0 + 1) - (2 - 1) = 0$.

Of course, the genus is 0 geometrically, because the curve is a circle, which is birationally equivalent to a line.

Example 4 $\chi(x, y) = y^2 + x^4 - 1$ (the elliptic curve mentioned in Essay 4.2).

Here the trace of 1 is 2, and the trace of y is 0, so the trace of $y^2 = 1 - x^4$ is $2(1 - x^4)$ and $D(x) = 4(1 - x^4)$. Since this polynomial has distinct roots, $1, y$ is an integral basis. The order of 1 at $x = \infty$ is of course 0, and the order of y is 2 (because division of $y^2 + x^4 - 1$ by x^4 gives a polynomial in $\frac{y}{x^2}$ and $\frac{1}{x}$). Since 1 and $\frac{y}{x^2}$ are an integral basis relative to $\frac{1}{x}$, as is easily shown, 1 and y are a *normal* basis and the field of constants is $\mathbf{Q}$, which implies that the genus is $(0 + 2) - (2 - 1) = 1$.

Example 5 $\chi(x, y) = y^2 + x^6 - 1$ (a frequently cited hyperelliptic curve).

By considerations similar to those in the last example, $D(x) = 4(1-x^6)$ has distinct roots, so 1 and y form an integral basis. The orders at $x = \infty$ are 0 and 3 respectively, and this basis is a normal basis. Therefore the genus is $(0 + 3) - (2 - 1) = 2$.

Example 6 $\chi(x, y) = y^2 - f(x)$, where $f(x)$ is a polynomial of degree $2n - 1$ or $2n$ with distinct roots (a general hyperelliptic curve).

As in the previous examples, 1 and y are a normal basis for which the orders at $x = \infty$ are 0 and n, so the genus is $(0 + n) - (2 - 1) = n - 1$, as is implied by the passage from Abel's memoir quoted in Essay 4.1.

Essay 4.6 Holomorphic Differentials

Given an algebraic curve C, the method of the preceding essay determines its g regarded as in Essay 4.3 as the codimension of the subvarieties of C^N swept out by algebraic variations of N points on the curve. The objective of the present essay is to express this idea in terms of differential equations

$$(1) \qquad \sum_{i=1}^{N} h_j(x_i, y_i)\,dx_i = 0 \qquad (j = 1, 2, \ldots, g)$$

describing these subvarieties of C^N. Here the differentials $h_j(x, y)\,dx$ for $j = 1, 2, \ldots, g$ are to be a basis, over the field of constants, of the space of **holomorphic differentials** on the curve, a concept that is to be defined. The equations (1) state that algebraic variations satisfy g infinitesimal conditions, where g is the dimension of the space of holomorphic differentials; therefore, not only do the algebraic variations partition C^N into subvarieties of codimension g, but this partition is expressed by g explicit differential equations.

In these equations, (x_i, y_i) for $i = 1, 2, \ldots, N$ are given solutions of $\chi(x_i, y_i) = 0$, where $\chi(x, y) = 0$ is the equation of the curve C. The heuristic idea of (1) is the following: If (1) correctly describes the possible algebraic variations of N points, it certainly describes the possible algebraic variations of *fewer* than N points: Just add conditions $dx_i = 0$ for a certain number of the points. Therefore, there is no loss of generality in assuming that N is the number $n_0\nu$ of zeros of an element of $\Theta(x^\nu)$ for some large ν. (Here n_0 again denotes the degree of the root field as an extension of the field obtained by adjoining all its constants to $\mathbf{Q}(x)$, or, in the notation used before, $n_0 = n/c$.) As has been shown, the most general element of $\Theta(x^\nu)$ is given by an explicit formula θ that contains $N - g + 1$ unknown constants, call them $a_1, a_2, \ldots, a_{N-g+1}$ (and in fact contains them linearly); the conditions $\chi(x, y) = 0$ and $\theta(x, y, a_1, a_2, \ldots, a_{N-g+1}) = 0$ define, implicitly, N solutions (x_i, y_i) of $\chi(x_i, y_i) = 0$ as functions of $a_1, a_2, \ldots, a_{N-g+1}$, where $N = n_0\nu$. Since multiplication of θ by a constant does not change its common zeros with χ, one of the parameters in θ, say a_{N-g+1}, can be set equal to 1. Then the N moving points depend on $N-g$ parameters, and they sweep out a subvariety of codimension g. In principle, the equations (1) result from *implicit differentiation* of the defining equations $\chi = 0$, $\theta = 0$ of the N moving points (x_i, y_i) in the following way.

For fixed values of the x's, y's and a's, the $2N$ relations $d\chi(x_i, y_i) = 0$, $d\theta(x_i, y_i, a) = 0$ give $2N$ homogeneous, linear equations in the $3N - g$ differentials

dx_i, dy_i, da_j, whose coefficients are rational functions of the $3N - g$ variables. The relation $d\chi(x_i, y_i) = 0$ involves just one pair of values (x_i, y_i) and $a_1, a_2, \ldots, a_{N-g}$, so it can be used (provided x_i is a local parameter at (x_i, y_i)) to express each dy_i in terms of the corresponding dx_i and $da_1, da_2, \ldots, da_{N-g}$ and in this way to reduce the differential equations to just N equations in $2N - g$ differentials dx_i and da_i. These equations can be solved, in the generic case, to express $da_1, da_2, \ldots, da_{N-g}$ in terms of the dx_i and to eliminate them, leaving g relations among the dx_i. *These g relations are the required differential equations* (1) because they describe the relations satisfied by the dx_i when the parameters a_i are allowed to vary. In other words, these are the infinitesimal relations satisfied by algebraic variations of the N points (x_i, y_i).

In practice, the actual elimination of the da_i to find the relations among the dx_i seems impractical, even in the simplest examples. Instead, the derivation of the equations (1) will depend on observing that the holomorphic differentials, the ones that express the crucial relations (1), are the differentials that *have no poles*. Heuristically, such differentials lead to relations (1) in the following way.

If $\theta(x, y)$ has $n_0\nu$ zeros on the curve $\chi(x, y) = 0$ and no zeros where $x = \infty$, and if $h(x, y)\,dx$ has no poles—even when $x = \infty$—then the differential $\frac{h(x,y)\,dx}{\theta(x,y)}$ has poles only at the $n_0\nu$ zeros of $\theta(x, y)$. Thus, one can make use of the fact that *the sum of the residues of a differential is zero* to find that

$$\sum_{\text{zeros of }\theta} \left(\text{residue of } \frac{h(x, y)\,dx}{\theta(x, y)} \text{ at that zero of } \theta\right) = 0.$$

As a function on the curve $\chi(x, y) = 0$, $\theta(x, y)$ can be regarded, locally, as a function of x near each of its zeros (provided these zeros avoid places on the curve where x is not a local parameter), so $\frac{d\theta}{dx}$ is meaningful at each zero of $\theta(x, y)$ on the curve. When this derivative is not zero, its reciprocal is the residue of $\frac{dx}{\theta(x,y)}$ at the zero of θ because $\theta(x, y) = a_1 x + a_2 x^2 + \cdots$ implies that this residue is, by definition, $\frac{1}{a_1}$ when $a_1 \neq 0$. Thus

$$0 = \sum_{\text{zeros of }\theta} \left(\text{residue of } \frac{h(x, y)\,dx}{\theta(x, y)} \text{ at that zero of } \theta\right) = \sum_{\text{zeros of }\theta} h(x_i, y_i)\frac{dx_i}{d\theta},$$

where dx_i for each i is the infinitesimal change in x_i that results from an infinitesimal change $d\theta$ in θ. In other words, if the $n_0\nu$ points where θ is zero are moved to the nearby points where θ is $d\theta$, then the $n_0\nu$ changes dx_i in the x-coordinates of the intersection points satisfy $\sum h(x_i, y_i)\,dx_i = 0$, as was to be shown, provided the zeros of $\theta(x, y)$ are at points where both x and θ are local parameters on the curve. That this necessary condition for the dx_i to result from an algebraic variation of the intersection points is also a sufficient condition follows from—or at any rate is made plausible by—the fact that *the number of linearly independent holomorphic differentials is g*, so that the system of differential equations (1) describes a subvariety containing the algebraic variations that has the same dimension $N - g$ (at generic points) as the subvariety of algebraic variations and that therefore must coincide with it.

With this geometric motivation, the remainder of this essay will (a) define the notion of "holomorphic differential" in a precise algebraic way that accords with the notion of "no poles," (b) prove that the dimension of the holomorphic differentials as a vector space over the field of constants is g, (c) prove that the sum of the residues of a differential is zero, and (d) flesh out the implicit differentiation sketched above to reach the conclusion $\sum h_j(x_i, y_i)\,dx_i = 0$.

As before, let $\chi(x, y)$ be an irreducible polynomial in two indeterminates with integer coefficients that contains both indeterminates and is monic in the indeterminate y. Let K denote the root field of $\chi(x, y)$. A **differential** in K is an expression of the form $f(x, y)\,dx$, where $f(x, y)$ is an element of K and dx is merely a symbol. More precisely, $f(x, y)\,dx$ is a differential *expressed with respect to the parameter* x; it is easy to guess how a differential expressed with respect to the parameter x might be expressed with respect to another parameter of K, but in this essay all differentials will be expressed with respect to the preferred parameter x.

As before, the root field K of $\chi(x, y)$ will be regarded as an extension not of $\mathbf{Q}(x)$, the field of rational functions in x, but of $K_0(x)$, the field of rational functions in x with coefficients in the field of constants K_0 of K. (The symbol K_0 thus replaces the symbol $\Theta(x^0)$ as the notation for the field of constants of the root field.) As before, let n_0 be the degree of K as an extension of $K_0(x)$. By the definition of the genus, the dimension of $\Theta(x^\nu)$ as a vector space over K_0 is $n_0\nu - g + 1$ for all sufficiently large ν.

In this essay, instead of differentials $f(x, y)\,dx$ themselves, their *trace* will be considered; the trace of $f(x, y)\,dx$ is by definition the differential $\mathrm{tr}_x(f(x, y))\,dx$, where dx is a symbol and $\mathrm{tr}_x(f(x, y))$ is the element of $K_0(x)$ that is the trace of $f(x, y)$ with respect to the field extension $K \supset K_0(x)$; in other words, $\mathrm{tr}_x(f(x, y))$ is the trace of the $n_0 \times n_0$ matrix that represents multiplication by $f(x, y)$ with respect to the basis $1, y, y^2, \ldots, y^{n_0-1}$ of K over $K_0(x)$ (or, for that matter, with respect to any basis of K over $K_0(x)$). The heuristic idea behind this definition is that $\mathrm{tr}_x(f(x, y)\,dx)$ is the sum over all n_0 values at x of the differential $f(x, y)\,dx$, which, being symmetric in the n_0 values of y for any given x, is a rational function of x alone.

Holomorphic differentials were described above as differentials without poles. Certainly, the trace of a holomorphic differential must therefore be dx times an element of $K_0(x)$ without poles; in other words, if $h(x, y)\,dx$ is holomorphic, then $\mathrm{tr}_x(h(x, y))$ must be a polynomial in x. However, this necessary condition should not be expected to be sufficient, because $f(x, y)\,dx$ might have two poles at the same value of x that cancel when the sum is taken over all y. In the case of two canceling poles, however, one would expect to be able to choose an element $\theta(x, y)$ of K that was zero at just one of the poles and that had no poles for finite values of x, so that $\theta(x, y)f(x, y)\,dx$ would be a differential that had no poles where $f(x, y)\,dx$ did not and for which the poles that canceled in $\mathrm{tr}_x(f(x, y)\,dx)$ no longer canceled in $\mathrm{tr}_x(\theta(x, y)f(x, y)\,dx)$. Therefore, such a differential would not satisfy the *stronger* necessary condition for a differential $f(x, y)\,dx$ to be holomorphic: For every $\theta(x, y)$ that is integral over x, $\mathrm{tr}_x(\theta(x, y)f(x, y))$ is a polynomial in x. But this necessary condition, too, should not be expected to be sufficient, because it would not detect poles of $f(x, y)\,dx$ at places where $x = \infty$.

For these reasons, a differential $f(x, y)\,dx$ will be said to be **holomorphic for finite** x if $\mathrm{tr}_x(\theta(x, y)f(x, y))$ is a polynomial in x whenever $\theta(x, y)$ is integral over x, and will be said to be **holomorphic** if it is holomorphic for finite x and if $f\,d\left(\frac{1}{u}\right) = -\frac{f}{u^2}\,du$ is holomorphic for finite u. (Here $d\left(\frac{1}{u}\right) = -\frac{du}{u^2}$ is a definition. It will be justified by Corollary 1 below. Since tr_u is the same as tr_x when $u = \frac{1}{x}$—both are found using a basis of the field over $\mathbf{Q}(x) = \mathbf{Q}(u)$—to say that $-\frac{f}{u^2}\,du$ is holomorphic for finite u means that $\mathrm{tr}_u(\theta \cdot \frac{f}{u^2}) = \mathrm{tr}_x(x^2\theta f)$ is a polynomial in $u = \frac{1}{x}$ whenever θ is integral over u.)

Theorem *Construct the holomorphic differentials for a given $\chi(x, y)$ and prove that their dimension as a vector space over the field of constants K_0 is the genus of the root field of $\chi(x, y)$.*

Proof Let the construction that was used to determine the genus in Essay 4.5 be used to construct a normal basis $y_1, y_2, \ldots, y_{n_0}$ of the root field as an extension of the field $K_0(x)$ of rational functions of x with coefficients in the field of constants K_0 and to construct nonnegative integers $\mu_1, \mu_2, \ldots, \mu_{n_0}$ for which an element of the root field is in $\Theta(x^\nu)$ if and only if its unique expression in the form $\phi_1(x)y_1 + \phi_2(x)y_2 + \cdots + \phi_{n_0}(x)y_{n_0}$, where the coefficients are in $K_0(x)$, has coefficients that are *polynomials* in x with coefficients in K_0 and the degrees of these polynomials satisfy $\deg \phi_i + \mu_i \leq \nu$.

Let S_1 denote the symmetric $n_0 \times n_0$ matrix of polynomials in x with coefficients in K_0 whose entry in the ith row of the jth column is $\mathrm{tr}_x(y_i y_j)$, where the trace is taken relative to the extension $K \supset K_0(x)$. (In other words, this entry is the trace of the matrix that represents multiplication by $y_i y_j$ relative to the basis y_1, $y_2, \ldots, y_{n_0}$ of K over $K_0(x)$.) The symmetric bilinear form "the trace of the product" is represented by S_1 in the sense that if $h = h_1y_1 + h_2y_2 + \cdots + h_{n_0}y_{n_0}$ and $\theta = \theta_1y_1 + \theta_2y_2 + \cdots + \theta_{n_0}y_{n_0}$ are the representations of two elements h and θ of K relative to this basis, then $\mathrm{tr}_x(h\theta) = [h]S_1[\theta]$ where $[h]$ represents the row matrix whose entries are $h_1, h_2, \ldots, h_{n_0}$, and $[\theta]$ represents the column matrix whose entries are $\theta_1, \theta_2, \ldots, \theta_{n_0}$.

With this notation, to say that $h\,dx$ is holomorphic is to say that $[h]S_1[\theta]$ is a polynomial of degree at most $\nu - 2$ whenever the ith entry θ_i of the column matrix $[\theta]$ is a polynomial whose degree is at most $\nu - \mu_i$, because $\mathrm{tr}_x(h\theta)$ must be a polynomial in x, while $\mathrm{tr}_x(-x^2h \cdot \frac{\theta}{x^\nu}) = -\frac{\mathrm{tr}_x(h\theta)}{x^{\nu-2}}$ must be a polynomial in $\frac{1}{x}$. (Note that the h_i need not be polynomials.) In other words, the row matrix $[h]S_1$ has the property that its product with a column matrix $[\theta]$ is a polynomial of degree at most $\nu - 2$ when the ith entry of θ is a polynomial of degree at most $\nu - \mu_i$. If one takes all entries but one of $[\theta]$ to be zero and that one to be a polynomial of degree $\nu - \mu_i$ for some large ν, one sees that the ith entry of $[h]S_1$ must be a rational function whose product with any polynomial of degree $\nu - \mu_i$ is a polynomial of degree at most $\nu - 2$. Thus, the ith entry of $[h]S_1$ must be a polynomial of degree at most $\mu_i - 2$ when $\mu_i \geq 2$ and must be zero if μ_i is 0 or 1. In other words, $[h]$ must have the form $[c]S_1^{-1}$, where c is a row matrix whose ith entry is a polynomial in x of degree at most $\mu_i - 2$ with coefficients in K_0. (In particular, the ith entry is zero when μ_i is 0 or 1.) This formula $[h] = [c]S_1^{-1}$ completely describes the holomorphic differentials

$h\,dx$. The number of constants in the coefficients of the entries of $[c]$ is the sum of the numbers $\mu_i - 1$ over all values of i for which $\mu_i > 0$. Since exactly one μ_i is zero (because $K_0 = \Theta(x^0)$ consists of all elements $\phi_1 y_1 + \phi_2 y_2 + \cdots + \phi_{n_0} y_{n_0}$ in which $\phi_i = 0$ when $\mu_i > 0$ and ϕ_i is constant when $\mu_i = 0$), it follows that the number of arbitrary constants in this formula for h is $(\sum \mu_i) - (n_0 - 1)$, which is the genus, as was to be shown. □

The proof that the sum of the residues of any differential $f(x, y)\,dx$ is zero reduces, by virtue of the *definition* of the sum of the residues as the sum of the residues of the rational differential $\mathrm{tr}_x(f(x, y))\,dx$, to the same statement for *rational* differentials $\frac{p(x)}{q(x)}\,dx$, where $p(x)$ and $q(x)$ are polynomials with coefficients in some algebraic number field K_0 and $q(x) \neq 0$. To define the sum of the residues of such a differential, it will be convenient to assume that the denominator $q(x)$ splits into linear factors over K_0, although, as will be seen, the *sum* of the residues can be expressed rationally in terms of the coefficients of $p(x)$ and $q(x)$ even when this condition is not fulfilled. By the method of partial fractions, one can see that if $q(x) = \prod (x - a_i)^{e_i}$, where the a_i are distinct constants, then

$$\frac{p(x)}{q(x)} = P(x) + \sum_i \sum_{\sigma=1}^{e_i} \frac{\rho_{i\sigma}}{(x - a_i)^{\sigma}}$$

for suitable constants $\rho_{i\sigma}$. (One can assume without loss of generality that $\deg p < \deg q$, so that $P(x) = 0$. Multiplication of both sides of the required equation

$$\frac{p(x)}{q(x)} = \sum_i \sum_{\sigma=1}^{e_i} \frac{\rho_{i\sigma}}{(x - a_i)^{\sigma}}$$

by $q(x) = \prod (x - a_i)^{e_i}$ gives an equation of the form $p(x) = \sum \rho_{i\sigma} A_{i\sigma}(x)$ in which the polynomials $A_{i\sigma}(x)$ have degree less than $k = \deg q$ and depend only on $q(x)$. This gives an inhomogeneous $k \times k$ system of linear equations satisfied by the k required coefficients $\rho_{i\sigma}$. When $p(x) = 0$, these equations have only the trivial solution,[12] so for any $p(x)$ of degree less than k they have a unique solution.) The **residue at** $x = a$ of $\frac{p(x)}{q(x)}\,dx$ is defined to be ρ_{i1}, the coefficient of $\frac{1}{x - a_i}$ in the partial fractions expansion of $\frac{p(x)}{q(x)}$, when a is one of the roots a_i of $q(x)$; otherwise, the residue at $x = a$ of $\frac{p(x)}{q(x)}\,dx$ is zero.

Note that the residue at $x = a$ of $\left(\frac{p(x)}{q(x)} + \frac{p_1(x)}{q_1(x)}\right)dx$ is the sum of the residues at $x = a$ of $\frac{p(x)}{q(x)}\,dx$ and $\frac{p_1(x)}{q_1(x)}\,dx$. (The partial fractions decomposition of a sum is the sum of the partial fractions decompositions when terms with the same denominators are combined.)

[12] Multiplication of $\sum_{i=1}^{\mu} \frac{\phi_i(x)}{(x - a_i)^{\nu_i}} = 0$, where $a_1, a_2, \ldots, a_\mu$ are distinct algebraic numbers and $\deg \phi_i < \nu_i$ for each i, by $\prod_{i=1}^{\mu} (x - a_i)^{\nu_i}$ gives an equation $\sum_{i=1}^{\mu} \psi_i(x) = 0$ in which the $\psi_i(x)$ are polynomials. All but one of these polynomials is divisible by $(x - a_1)^{\nu_1}$, so the remaining one must also be divisible by $(x - a_1)^{\nu_1}$, from which it follows that $\phi_1(x)$ must be zero. In the same way, $\phi_i(x) = 0$ for each i.

The conventional statement that *the sum of the residues of a rational differential is zero* assumes that the "residue at $x = \infty$" is included in the sum. In this way, the conventional statement can be seen, as the corollary below shows, as a method of *evaluating* the sum of the residues of $\frac{p(x)}{q(x)}dx$ over all *finite* values of a. This evaluation is in fact quite easy:

Proposition *The sum of the residues at a of a rational differential* $\frac{p(x)}{q(x)}dx$ *over all finite values of a is*

$$\lim_{x\to\infty}\frac{x\cdot r(x)}{q(x)} = \left.\frac{u^{e-1}r\left(\frac{1}{u}\right)}{u^e q\left(\frac{1}{u}\right)}\right|_{u=0} \qquad (e = \deg q),$$

where $r(x)$ is the remainder when $p(x)$ is divided by $q(x)$, where the limit on the left is merely a mnemonic standing for the expression on the right, and the expression on the right denotes the quotient of constants in which the denominator is the leading coefficient of q and the numerator is the coefficient of x^{e-1} in $r(x)$.

Proof Since the residues of $\left(P(x) + \frac{r(x)}{q(x)}\right)dx$ are the same as those of $\frac{r(x)}{q(x)}dx$, one can assume without loss of generality that the quotient $\frac{p(x)}{q(x)}$ in the given differential is a proper fraction; i.e., one can assume $r(x) = p(x)$. The residue of $\frac{\rho}{(x-a)^e}$ is ρ if $e = 1$ and 0 if $e > 1$, so for fractions of the particular form $\frac{r(x)}{q(x)} = \frac{\rho}{(x-a)^e}$ the residue is given by the formula $\lim_{x\to\infty}\frac{x\cdot r(x)}{q(x)}$. The theorem therefore follows from the observation that if $\frac{r(x)}{q(x)}$ and $\frac{r_1(x)}{q_1(x)}$ are proper fractions, then

$$\frac{x\cdot r(x)}{q(x)} + \frac{x\cdot r_1(x)}{q_1(x)} = \frac{x\cdot r(x)q_1(x) + x\cdot r_1(x)q(x)}{q(x)q_1(x)},$$

so the same is true of their limits as $x \to \infty$, interpreted as in the statement of the theorem. (Note also that $\lim_{x\to\infty}\frac{x\cdot r(x)}{q(x)}$ is unchanged if a common factor is canceled from numerator and denominator.) □

Corollary 1 *The sum of the residues of* $\frac{p(x)}{q(x)}dx$ *over all finite values of x is minus the residue at $x = \infty$, which residue is by definition the residue at $u = 0$ of*

$$\frac{p\left(\frac{1}{u}\right)}{q\left(\frac{1}{u}\right)}d\left(\frac{1}{u}\right) = -\frac{p\left(\frac{1}{u}\right)}{u^2 q\left(\frac{1}{u}\right)}du.$$

(The expression on the left is a mere mnemonic that takes advantage of the formula $d\left(\frac{1}{u}\right) = -\frac{du}{u^2}$ of elementary calculus.)

Deduction What is to be shown is that the value of

$$\frac{u^{e-1}r\left(\frac{1}{u}\right)}{u^e q\left(\frac{1}{u}\right)}$$

at $u = 0$ is the residue at $u = 0$ of

$$\frac{p\left(\frac{1}{u}\right)}{u^2 q\left(\frac{1}{u}\right)} du = \frac{u^{e-1} p\left(\frac{1}{u}\right)}{u \cdot u^e q\left(\frac{1}{u}\right)} du.$$

Since this differential has the form

$$\frac{P(u)}{uQ(u)} du, \quad \text{where} \quad \frac{P(u)}{Q(u)}$$

is a proper fraction in which $Q(0) \neq 0$, this conclusion follows immediately from the definition. □

Corollary 2 *When the residue of $\frac{p(x)}{q(x)} dx$ at $x = \infty$ is defined as in Corollary 1, the sum of the residues of a rational differential is zero.*

These algebraic facts make possible a plausible implicit differentiation of $\chi(x, y) = 0$ and $\theta(x, y, a_1, a_2, \ldots, a_{N-g}) = 0$ that leads to

$$\sum_{i=1}^{N} h_j(x_i, y_i)\, dx_i = 0 \qquad (j = 1, 2, \ldots, g) \tag{2}$$

when the dy's and da's are eliminated.

As before, there is no loss of generality in assuming that $N = n_0\nu$ for some large ν and that the (x_i, y_i) are the intersection points $\chi(x_i, y_i) = 0$, $\theta(x_i, y_i) = 0$ for some fixed $\theta = a_1\theta_1 + a_2\theta_2 + \cdots + a_{N-g+1}\theta_{N-g+1}$, where the θ_i are a basis of $\Theta(x^\nu)$ over K_0 and $a_1, a_2, \ldots, a_{N-g+1}$ are fixed constants. In addition, it will be assumed that the chosen θ is in "general position" in the sense that x is a local parameter at each of the N intersection points (x_i, y_i) and θ has poles of order ν at each of the n points where $x = \infty$.

Each of the N intersection points (x_i, y_i) implies a pair of differential equations

$$\begin{aligned} \chi_x\, dx_i + \chi_y\, dy_i &= 0, \\ \theta_x\, dx_i + \theta_y\, dy_i + \theta_1\, da_1 + \theta_2\, da_2 + \cdots + \theta_{N-g+1}\, da_{N-g+1} &= 0, \end{aligned} \tag{3}$$

where subscripts x and y denote partial derivatives, and these partial derivatives are to be evaluated at the point $(x_i, y_i, a_1, a_2, \ldots, a_{N-g+1})$ at which $a_1, a_2, \ldots, a_{N-g+1}$ have the given values that determine the N points (x_i, y_i), and x_i and y_i are the coordinates of one of these points.

Elimination of dy_i from the pair of equations (3) gives the single equation $dx_i + Q(\theta_1 da_1 + \theta_2 da_2 + \cdots + \theta_{N-g+1} da_{N-g+1}) = 0$ in which Q denotes the quotient $\frac{\chi_y}{\chi_y\theta_x - \theta_y\chi_x}$. This quotient is in fact the reciprocal of the derivative of θ with respect to x (eliminate dy from the equations $\chi_x\, dx + \chi_y\, dy = 0$ and $\theta_x\, dx + \theta_y\, dy = d\theta$, a computation that assumes x is a parameter on the curve at the point in question). Otherwise stated, it is the residue of the differential $\frac{dx}{\theta}$ at this zero of the denominator θ, because it is the value of the quotient $\frac{x - x_i}{\theta(x,y)}$ at the point (x_i, y_i) where numerator and denominator, taken separately, are both zero. (It is natural to think of this number as a limit, but of course it can be described algebraically as the value of the rational function of x and y when it is put in canonical form—a numerator in which y has degree less than $n = \deg_y \chi$ and a denominator that is a polynomial in x alone that is relatively prime to the numerator.)

Therefore, if each equation $dx_i + Q(\theta_1 da_1 + \theta_2 da_2 + \cdots + \theta_{N-g+1} da_{N-g+1}) = 0$ is multiplied by the value $h(x_i, y_i)$ of h at the corresponding point (x_i, y_i) and all N of these equations are added, the result is $\sum h(x_i, y_i)\,dx_i + C_1\,da_1 + C_2\,da_2 + \cdots + C_{N-g+1}\,da_{N-g+1} = 0$, where the coefficient C_j of da_j is the sum over all N zeros of θ on the curve $\chi = 0$ of $h\theta_j$ times the residue of $\frac{dx}{\theta}$ at that point. It is to be shown that each such coefficient C_j is zero.

Since neither θ_j nor $h\,dx$ has poles for finite x, the differential $\theta_j h\,dx/\theta$ has residues for finite x *only* at the zeros (x_i, y_i) of θ, and these residues are the values at (x_i, y_i) of $\theta_j h$ times the residue of dx/θ at (x_i, y_i). In short, C_j is the sum of all residues of the differential $\theta_j h\,dx/\theta$ at points where x is finite. Therefore, it is minus the sum of the residues at $x = \infty$ of the differential $\theta_j h\,dx/\theta$. Since θ_j has order at most ν at $x = \infty$ (it is in $\Theta(x^\nu)$) and θ has order ν at $x = \infty$ (by assumption), θ_j/θ is finite at $x = \infty$, so $\theta_j h\,dx/\theta$ has no pole at $x = \infty$, which implies $C_j = 0$ and $\sum h(x_i, y_i)\,dx_i = 0$, as was to be shown.

Example 7 $\chi(x, y) = y^3 + x^3 y + x$ (the Klein curve).

As was seen in Example 2 of Essay 4.5, $1, y, y^2$ are a normal basis over $\mathbf{Q}(x)$ for which the λ's are 0, 2, 3. Therefore $(h_1 + h_2 y + h_3 y^2)\,dx$ is a holomorphic differential if and only if $[h_1 \;\; h_2 \;\; h_3]S = [0 \;\; a \;\; bx + c]$, where a, b, and c are rational numbers and the matrix S, which has $\mathrm{tr}(y^{i+j-2})$ in the ith row of the jth column, is easily found to be

$$\begin{bmatrix} 3 & 0 & -2x^3 \\ 0 & -2x^3 & -3x \\ -2x^3 & -3x & 2x^6 \end{bmatrix}.$$

When $c = 1$ and $a = b = 0$, this gives a 3×3 homogeneous linear system whose solution is $[h_1 \;\; h_2 \;\; h_3] = \frac{1}{4x^9+27x^2}[4x^6 \;\; -9x \;\; 6x^3]$. Thus, $h = \frac{4x^6-9xy+6x^3y^2}{4x^9+27x^2}$, which can be written more simply as $h = \frac{1}{3y^2+x^3}$. It is easy to see that the solution in which $b = 1$ and $c = a = 0$ is x times this one, and the solution in which $a = 1$ and $b = c = 0$ is y times this one, which leads to the formula

$$\frac{c + bx + ay}{3y^2 + x^3}\,dx$$

for the most general holomorphic differential on this curve. The formula has three parameters a, b, c because the genus is 3. (For an easier derivation of this formula, see the examples of Essay 4.8.)

Example 8 $\chi(x, y) = y^2 - f(x)$, where $f(x)$ is a polynomial of degree $2n - 1$ or $2n$ with distinct roots (a general hyperelliptic curve).

As was seen in Example 6 of Essay 4.5, 1 and y are a normal basis for which the orders at $x = \infty$ are 0 and n. (The matrix $S(x)$ is $\left[\begin{smallmatrix} 2 & 0 \\ 0 & 2f(x) \end{smallmatrix}\right]$, whose determinant

$D(x) = 4f(x)$ has distinct roots, so 1, y is an integral basis over x. Division of $y^2 - f(x) = 0$ by x^{2n} gives $\left(\frac{y}{x^n}\right)^2 - \frac{f(x)}{x^{2n}} = 0$, which, when $v = \frac{y}{x^n}$ and $u = \frac{1}{x}$, is a curve of the same form $v^2 - F(u) = 0$ of which 1, v is an integral basis over u. It follows that 1, y is a normal basis relative to x in which the order of y at $x = \infty$ is n.) Therefore, $(h_1 + h_2 y)\,dx$ is a holomorphic differential if and only if

$$[h_1\ h_2]\begin{bmatrix} 2 & 0 \\ 0 & 2f(x) \end{bmatrix} = [0\ \ q(x)],$$

where $q(x)$ is a polynomial of degree at most $n - 2$. Thus, $h\,dx = \frac{q(x)y\,dx}{2f(x)} = \frac{q(x)\,dx}{2y}$ is the most general holomorphic differential, where $q(x)$ is a polynomial of degree at most $n - 2$. The genus is $n - 1$.

Essay 4.7 The Riemann–Roch Theorem

Dedekind and Weber say in their classic treatise that the Riemann–Roch theorem, in its usual formulation, determines the number of arbitrary constants in a function with given poles [22, §28]. Indeed, that is exactly the way Roch himself formulated the theorem [78], as his title "On the Number of Arbitrary Constants in Algebraic Functions" indicates. The answer, a formula for the dimension of the vector space of rational functions with (at most) given poles, is a corollary of the theorem of this essay, which describes the *principal parts* of rational functions on an algebraic curve.

Let $f(x, y)$ be a rational function on a curve $\chi(x, y) = 0$, say $f(x, y) = p(x, y)/q(x)$, where p and q are polynomials with integer coefficients and f is regarded as an element of the root field of $\chi(x, y)$. The **principal parts of f at finite values of** x are, by definition, the terms with negative exponents in the expansions of f in powers of $x - \alpha$ for algebraic numbers α. Such expansions are obtained by applying Newton's polygon to expand y in $n = \deg_y \chi$ ways in (possibly fractional) powers of $x - \alpha$, substituting these expansions in $p(x, y)$, and multiplying the result by the expansion of $1/q(x)$ in powers of $x - \alpha$; they can contain negative powers of $x - \alpha$ only if the expansion of $1/q(x)$ does, which is to say, only if α is a root of $q(x)$. The principal parts of $f(x, y)$ thus amount simply to a list of the roots α of $q(x)$ and, for each of them, a list of the terms, if any, with negative exponents in the n series found by substituting expansions of y in powers of $x - \alpha$ in $f(\alpha + (x - \alpha), y)$.

One can define the **principal parts of f at $x = \infty$** as the principal parts at $u = 0$ when $u = \frac{1}{x}$, but for the sake of simplicity this essay will deal only with rational functions that are finite at $x = \infty$, so that there are *no* principal parts at $x = \infty$. Specifically, the only functions considered will be those of the form $f(x, y) = p(x, y)/q(x)$, where $p(x, y)$ is in $\Theta(x^\nu)$ for $\nu = \deg q$. Expansion of numerator and denominator of $f(x, y) = \frac{p(x,y)/x^\nu}{q(x)/x^\nu}$ in powers of $1/x$ then gives a quotient of power series in $1/x$ in which neither numerator nor denominator contains terms with negative exponents (the numerator is integral over $1/x$) and the denominator is not zero when $1/x = 0$, so the expansions of $f(x, y)$ in powers of $\frac{1}{x}$ contain no terms with negative exponents.

For values α of x at which $\chi(x, y) = 0$ *ramifies*—which is to say that at least one of the expansions of y in powers of $x - \alpha$ involves fractional powers—the principal parts of a function satisfy an obvious consistency requirement, namely, since one solution $y = \beta_0 + \beta_1 s + \beta_2 s^2 + \cdots$ in which $s = \sqrt[m]{x - \alpha}$ for $m > 1$ implies $m - 1$ other solutions obtained by multiplying s by some mth root of 1 other than 1, it must be true that a term of any one of the corresponding expansions of $f(x, y)$ determines the term with the same exponent in any of the other $m - 1$ expansions: One needs merely to change s to ωs for a suitable root of unity ω.

For these reasons, a **set of proposed principal parts** of a rational function on the curve $\chi(x, y) = 0$ will be defined to consist of (1) an algebraic number field $\mathbf{A}$, (2) a finite set of elements $\alpha_1, \alpha_2, \ldots, \alpha_\mu$ of $\mathbf{A}$, and (3) for each of the α_i and for each of the n ways of expanding y in powers of $x - \alpha_i$ an expression of the form $\gamma_1(x - \alpha_i)^{-1} + \gamma_2(x - \alpha_i)^{-2} + \cdots + \gamma_l(x - \alpha_i)^{-l}$, where l is a positive integer and the γ's are in $\mathbf{A}$, except that in the case of expansions of y that are in powers of $\sqrt[m]{x - \alpha_i}$ with $m > 1$ the expressions must take the form $\gamma_1(\sqrt[m]{x - \alpha_i})^{-1} + \gamma_2(\sqrt[m]{x - \alpha_i})^{-2} + \cdots + \gamma_l(\sqrt[m]{x - \alpha_i})^{-l}$ and must satisfy the consistency requirement just described. (A natural way to handle this consistency requirement is to prescribe the expression $\gamma_1(\sqrt[m]{x - \alpha_i})^{-1} + \gamma_2(\sqrt[m]{x - \alpha_i})^{-2} + \cdots + \gamma_l(\sqrt[m]{x - \alpha_i})^{-l}$ for just *one* expansion of y in powers of $\sqrt[m]{x - \alpha_i}$ in each set of m and to derive the others from it.)

The problem is to determine whether, for a given set of proposed principal parts, there is a rational function on the curve that is finite where $x = \infty$ and that has the stated principal parts. The answer given by the theorem below is that the holomorphic differentials give simple necessary and sufficient conditions for there to be such a function.

If $h\,dx$ is a holomorphic differential and if $f(x, y) = \frac{p(x,y)}{q(x)}$ is finite when $x = \infty$, then $\mathrm{tr}_x(fh)$ on the one hand is a rational function of x whose denominator $q(x)$ has degree ν and whose numerator $\mathrm{tr}_x(ph)$ has degree at most $\nu - 2$, so the sum of the residues of $\mathrm{tr}_x(fh)$ over all finite values of the variable is zero, and on the other hand is a rational function whose residue at any finite value α of x is a linear function of the principal parts of f. In this way, a certain linear function of the principal parts of f is necessarily zero. Explicitly, the following lemma can be used to express the residue of $\mathrm{tr}_x(fh)$ at $x = \alpha$ in terms of the principal parts of f:

Lemma *Let $g(x, y)$ be a rational function on the curve $\chi(x, y) = 0$. The expansion of $\mathrm{tr}_x(g)$ in powers of $x - \alpha$ is the sum of the n expansions in (possibly fractional) powers of $x - \alpha$ obtained by substituting the n solutions of $\chi(x, y) = 0$ at $x = \alpha$ in $g(x, y)$.*

Proof Let the main theorem of Chapter 1 be used to construct a minimal splitting polynomial, call it $F(x, z)$, of $\chi(x, y)$ regarded as a polynomial in y with coefficients in $\mathbf{Z}[x]$ and let $\hat{K}$ be the root field of $F(x, z)$, which is to say that $\hat{K}$ is the splitting field of $\chi(x, y)$. Finally, let $\bar{z}$ be an expansion of a solution z of $F(x, z) = 0$ in (possibly fractional) powers of $x - \alpha$, say $\bar{z} = \delta_0 + \delta_1 s + \delta_2 s^2 + \cdots$, where $x = \alpha + s^m$. (Let m be determined by the condition that the indices i of terms in $\bar{z}$ in which $\delta_i \neq 0$

have no common divisor greater than 1.) The substitution $x = \alpha + s^m$, $z = \overline{z}$ embeds $\hat{K}$ in the field of quotients of the ring of power series in s with coefficients in $\mathbf{A}$, where $\mathbf{A}$ is an algebraic number field containing α that is constructed by the Newton polygon algorithm. Let $\mathbf{A}\langle s\rangle$ denote this field of quotients; it is, in effect, the ring of formal power series in s with coefficients in $\mathbf{A}$ enlarged to include power series with a finite number of terms with negative exponents, because the reciprocal of a power series $\gamma_i s^i + \gamma_{i+1}s^{i+1} + \cdots$ in which the term of lowest degree has degree i can be expressed as a power series $\frac{1}{\gamma_i}s^{-i} + \cdots$ in which there are i or fewer terms with negative exponents.

In short, the splitting field $\hat{K}$ of $\chi(x, y)$ can be regarded[13] as a subfield of $\mathbf{A}\langle s\rangle$ for a sufficiently large algebraic number field $\mathbf{A}$ under an embedding that carries x to $\alpha + s^m$. Let $g_1, g_2, \ldots, g_\mu$ be the distinct images of $g(x, y)$ under the Galois group of $\hat{K}$. The irreducible polynomial $\psi(X)$ with coefficients in[14] $\mathbf{Q}(x)$ of which $g(x, y)$ is a root is then $\prod_{i=1}^{\mu}(X - g_i)$, and $\mathrm{tr}_x(g)$ is $-j$ times the coefficient of $X^{\mu-1}$ in $\psi(X)$, where j is the degree of the root field K of $\chi(x, y)$ as an extension of its subfield generated by g, because the matrix that represents multiplication by g relative to a basis of K over $\mathbf{Q}(x)$ can be arranged as a $j \times j$ matrix of $\mu \times \mu$ blocks in which the blocks off the diagonal are all zero and the diagonal blocks are all the matrix whose first $\mu - 1$ rows are the last $\mu - 1$ rows of I_μ and whose last row is the negatives of the coefficients of ψ (except the leading coefficient 1) listed in reverse order. Thus, since the coefficient of $X^{\mu-1}$ is $-(g_1 + g_2 + \cdots + g_\mu)$, the expansion of $\mathrm{tr}_x(g)$ in powers of s is given by $\mathrm{tr}_x(g) = j(g_1 + g_2 + \cdots + g_\mu)$ when the g_i are represented as elements of $\mathbf{A}\langle s\rangle$. What is to be shown, then, is that substitution of the n expansions of y in powers of s, along with substitution of $\alpha + s^m$ for x, into $g(x, y)$ gives each of the μ expansions g_i for $i = 1, 2, \ldots, \mu$ exactly j times.

The Galois group of $\hat{K}$ expresses each root z of $F(x, z)$ as a polynomial in one such root with coefficients in $\mathbf{Q}(x)$. Substitution of $\overline{z}$ in these polynomials, together with substitution of $\alpha + s^m$ for x, gives $\deg_z F(x, z)$ distinct embeddings of $\hat{K}$ in $\mathbf{A}\langle s\rangle$. The possible expansions of y in powers of s and the possible expansions g_i of $g(x, y)$ all occur as images of y or $g(x, y)$, respectively, under these embeddings. The action of the Galois group on the embeddings implies the desired conclusion that each g_i occurs for the same number of different expansions of y.

(Note in particular that all fractional powers of $x - \alpha$ cancel when the sum $g_1 + g_2 + \cdots + g_\mu$ is computed. This is a clear consequence of the "consistency requirement" described above, because the sum $1 + \omega^k + \omega^{2k} + \cdots + \omega^{(m-1)k}$ is zero whenever ω is an mth root of unity and $\omega^k \neq 1$.) □

[13] Such an embedding of an algebraic field of transcendence degree 1 in $\mathbf{A}\langle s\rangle$ is analogous to an embedding of an algebraic number field in the field of complex numbers (see Essay 5.1). As in the latter case, the field $\hat{K}$ loses much of its constructive meaning when it is regarded merely as a subfield of $\mathbf{A}\langle s\rangle$, because elements of $\mathbf{A}\langle s\rangle$ are infinite series. An element of $\hat{K}$ is a root of a polynomial whose coefficients are rational functions of x, so the infinite series that represents it as an element of $\mathbf{A}\langle s\rangle$ can be specified by giving enough terms to determine an unambiguous truncated solution of the equation in question, after which all later terms are determined by Newton's polygon.

[14] As always, $\mathbf{Q}(x)$ denotes the field of quotients of $\mathbf{Z}[x]$, which is to say the field of rational functions in x.

Thus, the residue of $\mathrm{tr}_x(fh)dx$ at any α is the sum of the coefficients of $(x-\alpha)^{-1}$ over all n expansions of fh in powers of $s = \sqrt[m]{x-\alpha}$. Not only does this sum depend linearly on the principal parts of f, but the entries in the matrix that describes it are coefficients in the expansion of h in powers of s. Explicitly, if $h = h_0+h_1s+h_2s^2+\cdots$ and if $f = \gamma_1 s^{-1} + \gamma_2 s^{-2} + \cdots + \gamma_l s^{-l}$, then the coefficient of $(x-\alpha)^{-1}$ in fh is $h_0\gamma_m + h_1\gamma_{m+1} + h_2\gamma_{m+2} + \cdots + h_{l-m}\gamma_l$, a linear function of $\gamma_1, \gamma_2, \ldots, \gamma_l$. When this formula is summed over all principal parts of f it gives, for each holomorphic differential $h\,dx$, an explicit linear function of the principal parts of f that must be zero.

Theorem *If a set of proposed principal parts satisfies the condition just described for each holomorphic differential $h_i\,dx$ in a basis $h_1\,dx, h_2\,dx, \ldots, h_g\,dx$ of the holomorphic differentials on $\chi(x, y) = 0$, then it in fact gives the principal parts of some rational function on the curve. In short, these g necessary conditions for a set of proposed principal parts to be the principal parts of a function are sufficient.*

Proof Let a set of proposed principal parts be called **subordinate to** a polynomial $q(x)$ if each α for which it specifies a polynomial $\gamma_1 s^{-1} + \gamma_2 s^{-2} + \cdots + \gamma_l s^{-l}$ is a root of $q(x)$ and if, moreover, the multiplicity of α as a root of $q(x)$ is at least l/m, so that multiplication of $\gamma_1 s^{-1} + \gamma_2 s^{-2} + \cdots + \gamma_l s^{-l}$ by $q(x) = q(\alpha + s^m)$ makes all exponents nonnegative. Every set of proposed principal parts is subordinate to some $q(x)$, so it will suffice to prove that the theorem holds for all sets of proposed principal parts subordinate to a given $q(x)$. Moreover, those subordinate to $q(x)$ are also subordinate to $q(x)r(x)$ for any polynomial $r(x)$, so one can assume without loss of generality that $q(x)$ is a polynomial of high degree.

If $f(x, y)$ is finite at $x = \infty$ and its principal parts are subordinate to $q(x)$, then $q(x)f(x, y)$ has order at most $\deg q$ at $x = \infty$ and has no poles for finite x, which is to say that $f(x, y) = p(x, y)/q(x)$, where $p(x, y)$ is in $\Theta(x^\nu)$ for $\nu = \deg q$. But for large ν the set of such functions $f(x, y)$ is a vector space of dimension $n\nu - g + 1$ over the field of constants. Functions that differ by a constant have the same principal parts and conversely, so the vector space of principal parts that actually occur is seen in this way to have dimension $n\nu - g$ for large ν.

On the other hand, the dimension of the space of proposed principal parts subordinate to $q(x)$ can be found in the following way. If α is a root of $q(x)$ of multiplicity μ, and if none of the n expansions of y in powers of $x-\alpha$ involve fractional powers, the proposed principal parts subordinate to $q(x)$ contain $n\mu$ coefficients corresponding to this α, μ coefficients in each expansion (those of $(x-\alpha)^{-1}, (x-\alpha)^{-2}, \ldots, (x-\alpha)^{-\mu}$). The same formula μn holds even when some expansions involve fractional powers, because the proposed principal part corresponding to an expansion in powers of $\sqrt[m]{x-\alpha}$ contains m times as many coefficients in the required range, but by the "consistency requirement" the coefficients of just *one* determines those of a set of m of them. Therefore, a set of proposed principal parts subordinate to $q(x)$ contains $n\nu$ unknown coefficients γ, where $n = \deg_y \chi$ and $\nu = \deg q$.

In short, *the principal parts that actually occur are a subspace of codimension g in the $n\nu$-dimensional space of proposed principal parts subordinate to* $q(x)$ when $\nu = \deg q$ is sufficiently large. Since the conditions imposed by the holomorphic differentials—the sum of the residues of $\mathrm{tr}_x(fh)$ is zero for all holomorphic differentials $h\,dx$—are expressed by g homogeneous linear conditions on the coefficients of the principal parts, the actually occurring ones account for *all* of those that satisfy the necessary conditions *provided* the necessary conditions are independent (because then they determine a subspace of codimension g). In short, it will suffice to prove that every polynomial divides one for which the g necessary conditions are independent as conditions on sets of proposed principal parts subordinate to that polynomial.

The g homogeneous linear conditions imposed by the holomorphic differentials on proposed principal parts subordinate to $q(x)$ are expressed by a $g \times (n\nu)$ matrix of elements of $\mathbf{A}$, call it C_q. What is to be shown is that every polynomial divides a polynomial $q(x)$ for which the rank of C_q is g. This will be done by showing that if the rank of C_q is less than g and if β is any element of $\mathbf{A}$ that is not a root of $q(x)$, then replacing $q(x)$ with $(x-\beta)q(x)$ increases the rank of C_q, except for very extraordinary coincidences in the choice of β which can easily be avoided.

In fact, changing $q(x)$ to $(x-\beta)q(x)$ increases the degree of $q(x)$ by 1 and therefore adds n columns to C_q. Each of the new columns contains the g values of the coefficient h_i of one of the basis $h_1\,dx, h_2\,dx, \ldots, h_g\,dx$ of holomorphic differentials at one of the n points on the curve at which $x = \beta$. More precisely, let β be required to be an element of $\mathbf{A}$ (β can be taken to be a positive integer) for which $\chi(\beta, X)$ has distinct roots; then each column of the new C_q corresponds to one of the roots γ of $\chi(\beta, X)$ and it contains the values of $h_1, h_2, \ldots, h_g$ when (β, γ) is substituted for (x, y). It is to be shown that if the original C_q has rank less than g, then the extended C_q has rank greater than that of the original, except under extraordinary circumstances.

The ranks of the original and the extended C_q are unchanged by a change of basis of the holomorphic differentials. If the rank of the original C_q is less than g, then there are constants $c_1, c_2, \ldots, c_g$, not all zero, such that multiplication of the original C_q on the left by the row matrix with entries $c_1, c_2, \ldots, c_g$ gives a row of zeros. If a new basis of the holomorphic differentials is used in which the first holomorphic differential is $\sum_{i=1}^{g} c_i h_i\,dx$, the original C_q becomes a matrix whose first row is zero and the extended C_q becomes a matrix in which the first row contains $n\nu$ zeros and n new entries that are the values of the new h_1 at the n points $(x, y) = (\beta, \gamma)$. Thus, the extended C_q has greater rank unless all n of these values are zero.

Clearly, it would be an extraordinary coincidence if a value of β chosen at random were to result in even one value of h_1 that was zero, much less n of them. Since the number of zeros of the rational function h_1 (which is nonzero and does not have a pole at (β, γ) by the choice of β) is finite, one can easily find a new β for which the first row of the extended C_q contains a nonzero entry, and therefore find a β for which the rank of C_q is increased. □

This theorem determines exactly which proposed principal parts are actual principal parts and therefore solves the Riemann–Roch problem of determining the dimension of the space of rational functions with prescribed poles with, at most, prescribed multiplicities. In the notation and terminology of the proof, one can say that *the vector space of functions whose principal parts are subordinate to $q(x)$ has dimension $n\nu - \rho + 1$*, where ρ is the rank of C_q, because the proposed principal parts are a space of dimension $n\nu$, those that actually occur satisfy ρ independent conditions, and the linear function that carries functions to their principal parts has a one-dimensional kernel.

This formula $n\nu - \rho + 1$ gives the answer only in the special cases in which, roughly speaking, a pole is allowed at one point only if a pole is allowed at all other points where x has the same value (multiplicities counted). A more general case of the formula can be stated by introducing a little more terminology. Let one set of proposed principal parts be said to be **subordinate** to another if they make use of the same algebraic number field **A**, if each α of the first also occurs in the second, and if for each proposed expansion of y in powers of $x - \alpha$, the terms in the corresponding expression $\gamma_1(\sqrt[m]{x-\alpha_i})^{-1} + \gamma_2(\sqrt[m]{x-\alpha_i})^{-2} + \cdots + \gamma_l(\sqrt[m]{x-\alpha_i})^{-l}$ (where in most cases m is 1) of the first all have exponents at least as great (bearing in mind that -1 is greater than -2) as the smallest exponent of a nonzero term in the corresponding expression of the second. In short, the first set of proposed principal parts calls for no poles of greater multiplicity than are called for by the second. Let the **number of coefficients** in a set of proposed principal parts be the number of elements of **A** that need to be specified to describe a set of proposed principal parts that is subordinate to it (bearing in mind that if it calls for terms with fractional exponents, then the coefficients of *one* of the m expressions $\gamma_1(\sqrt[m]{x-\alpha_i})^{-1} + \gamma_2(\sqrt[m]{x-\alpha_i})^{-2} + \cdots + \gamma_l(\sqrt[m]{x-\alpha_i})^{-l}$ determine those in the other $m - 1$).

Corollary 3 (Riemann–Roch Theorem) *The functions whose principal parts are subordinate to a given set of proposed principal parts form a vector space of dimension* $N - \rho + 1$*, where* N *is the number of coefficients in the given set of proposed principal parts and* ρ *is the rank of the* $g \times N$ *matrix that describes the necessary and sufficient conditions of the theorem.*

Deduction The sets of principal parts subordinate to the given set form a space of dimension N; the necessary and sufficient conditions describe a subspace of codimension ρ, which is the space of possible principal parts of actual functions; and the space of functions with these principal parts has dimension one greater, because two functions with the same principal parts differ by a constant. □

The theorem also implies that the conditions (1) of Essay 4.6 satisfied by algebraic variations of N points on a curve are *sufficient*, with a few added assumptions, for a proposed variation to be algebraic:

Corollary 4 *Let* (x_i, y_i) *for* $i = 1, 2, \ldots, N$ *be pairs of algebraic numbers that satisfy* $\chi(x_i, y_i) = 0$*, and suppose that* $\chi(x_i, X)$ *has* n *distinct roots for each* x_i*,* $i = 1, 2, \ldots, N$*, so that Newton's polygon gives a unique power series solution* $y = y_i + \beta(x - x_i) + \cdots$ *of* $\chi(x_i, y)$ *for each* i*. For any list of nonzero*[15] *algebraic numbers* $\delta_1, \delta_2, \ldots, \delta_N$ *that satisfy* $\sum_{i=1}^{N} h(x_i, y_i)\delta_i = 0$ *for all holomorphic differentials, there is a rational function* f *on* $\chi(x, y) = 0$ *whose zeros are precisely at the points* (x_i, y_i) *and whose expansion* $f = \gamma_i(x - x_i) + \cdots$ *in powers of* $x - x_i$ *at each such point shows that* $\frac{dx}{df}\Big|_{(x_i,y_i)} = \delta_i$ *in the sense that* $\delta_i = \frac{1}{\gamma_i}$*.*

[15] It is natural to exclude $\delta_i = 0$, because the points (x_i, y_i) for which $\delta_i = 0$ can be omitted from the list.

More picturesquely, prescribed infinitesimal changes dx_i in the x-coordinates of the points are generated by changing f from 0 to df, which changes x_i to $x_i + \delta_i df$, provided the prescribed changes satisfy the necessary conditions $\sum h\,dx_i = 0$ for all holomorphic differentials $h\,dx$.

Deduction The required function f is found by using the theorem to construct a function θ finite at $x = \infty$ whose principal parts are $\frac{\delta_i}{x-x_i}$ for $i = 1, 2, \ldots, N$ and setting $f = \frac{1}{\theta}$. Then f is zero only at the (x_i, y_i), and at these points $\frac{1}{f} = \frac{\delta_i}{x-x_i} + \cdots$, from which $f = \frac{1}{\delta_i}(x - x_i) + \cdots$ follows. □

Corollary 5 *When (x_i, y_i) for $i = 1, 2, \ldots, N$ are points on $\chi(x, y) = 0$ as in Corollary 4, the rational functions on $\chi(x, y) = 0$ that have simple poles, at most, at these N points and no other poles form a vector space of dimension $N - g + 1 + \mu$, where g is the genus of the curve and μ is the dimension of the vector space of holomorphic differentials that are zero at all N points.*

Deduction By Corollary 4, the dimension is $N - \rho + 1$, where ρ is the rank of the $g \times N$ matrix whose columns correspond to the N given points and whose g entries in each column are the values at the corresponding point of the coefficients h_i of a basis $h_1\,dx$, $h_2\,dx, \ldots,$ $h_g\,dx$ of the holomorphic differentials. What is to be proved is that in this case $\rho = g - \mu$, which follows immediately from the observation that μ is the dimension of the kernel of the linear function from $\mathbf{A}^g$ to $\mathbf{A}^N$ given by multiplication of row matrices of length g on the right by the matrix. □

In particular, the space of functions described in Corollary 5 has dimension at least $N - g + 1$.

Essay 4.8 The Genus Is a Birational Invariant

So far in these essays, the genus of the field of rational functions on an algebraic curve $\chi(x, y) = 0$ has been described in ways that used the special element x of the field, first when the description was in terms of the dimension of the vector space $\Theta(x^\nu)$, then when it was in terms of the dimension of the space of holomorphic differentials $h\,dx$ over the field of constants. However, the geometric motivation of the concept in terms of algebraic variations leads one to believe that the genus depends only on the *curve*, not on the parameter x used to describe the curve. If z is any element of the root field of $\chi(x, y)$ that is not a constant, one can construct[16] an element w of this root field such that every element of the field can be expressed rationally in terms of

[16] See Essay 2.2. The field $\mathbf{Q}(z)$ of rational functions in z is isomorphic to a subfield of the root field, and the root field has finite degree over the subfield provided z is not a constant. Adjunction of x and y to $\mathbf{Q}(z)$ gives an explicit extension of $\mathbf{Q}(z)$ that is isomorphic to the root field of $\chi(x, y)$. By the theorem of the primitive element, such a double adjunction can be obtained by a simple adjunction, and one can describe a simple adjunction as the root field of an irreducible monic polynomial with coefficients in $\mathbf{Z}[z]$ in the usual way.

z and w and such that z and w satisfy a relation of the form $\chi_1(z, w) = 0$, in which $\chi_1(z, w)$ is an irreducible polynomial with integer coefficients that is monic in w. In short, the root field of $\chi(x, y)$ can also be described as the root field of $\chi_1(z, w)$. What is to be expected, and what will be proved in this essay, is that the genus is the same whether x or z is the special parameter used to define it. In other words, the genus depends only on the field of rational functions on the curve, which is what it means to say that the genus is a *birational invariant*.

The needed connection is obvious from the point of view of differential calculus: The rule $h\,dx \leftrightarrow \big(h\frac{dx}{dz}\big)dz$ establishes a one-to-one correspondence between differentials expressed with respect to x and differentials expressed with respect to z. The heuristic meaning of "holomorphic" is "no poles," so this correspondence between differentials with respect to x and those with respect to z should be expected to put the *holomorphic* differentials in the two cases in one-to-one, linear correspondence, implying that the dimension of these spaces of holomorphic differentials—the genus in the two cases—should be the same.

It is easy to give algebraic meaning to $\frac{dx}{dz}$. Let x and z be given elements of the root field K of $\chi(x, y)$. There is[17] an irreducible polynomial $\phi(X, Z)$ in two indeterminates with integer coefficients—it is uniquely determined, up to its sign, by x and z—with the property that $\phi(x, z) = 0$ in K. The derivative of x with respect to z is found algebraically by implicit differentiation: differentiation of $\phi(x, z) = 0$ gives $\phi_x(x, z)\,dx + \phi_z(x, z)\,dz = 0$ and therefore gives

$$\frac{dx}{dz} = -\frac{\phi_z(x, z)}{\phi_x(x, z)},$$

where the subscripts indicate partial derivatives. The theorem to be proved states that when $\frac{dx}{dz}$ is defined algebraically in this way, *$h\,dx$ is holomorphic if and only if $h\frac{dx}{dz}\,dz$ is*, when, naturally, one determines whether $h_1\,dz$ is holomorphic for $h_1 = h\frac{dx}{dz}$ by dealing with it as a differential in the root field of $\chi_1(z, w)$ rather than the root field of $\chi(x, y)$.

The notion of "principal parts" of an element of K that was defined in the preceding essay will play an important role in the proof, but now the dependence of these principal parts on x needs to be emphasized: **The principal parts of f relative to a parameter** x are the terms with negative exponents[18] in all possible expansions of f in (possibly fractional) powers of $x - \alpha$ for algebraic numbers α. (Possible

[17] Since the root field is an extension of $\mathbf{Q}(x)$ of finite degree n, the powers $1, z, z^2, \ldots, z^n$ are linearly dependent over $\mathbf{Q}(x)$, which is to say that one can find a nonzero polynomial of degree at most n with coefficients in $\mathbf{Q}(x)$ of which z is a root. The needed relation $\phi(x, z) = 0$ is found by clearing denominators and passing to an irreducible factor if necessary. Because the root field of $\chi(x, y)$ contains $\mathbf{Q}(x)$, the relation $\phi(x, z) = 0$ must involve z. To say that z is a *parameter*—that it is not a constant—means that ϕ also involves x; if this is not the case, the denominator of $\frac{dx}{dz}$ is zero and the derivative is not defined (x is not a function of z).

[18] This definition is imprecise in that it ignores the question of *multiplicities*. In Essay 4.7, $n = \deg_y \chi$ distinct embeddings of the root field of χ in $\mathbf{A}\langle s\rangle$ were constructed for each given α. A given f may have fewer than n distinct images, so the same principal parts may occur for more than one embedding. Obviously, Lemma 1 is not affected by the way in which multiplicities are treated.

principal parts at $x = \infty$ will be ignored.) As was shown in the last essay, for a given f one can find a polynomial $q(x)$ with the property that the only possible principal parts of f relative to x occur when α is a root of $q(x)$. Therefore, the determination of the principal parts relative to x is a finite—and usually quite simple—calculation. One can then use the following lemma to determine whether $f\,dx$ is holomorphic for finite x:

Lemma 1 *A differential $f\,dx$ is holomorphic for finite x if and only if all terms of all principal parts of f relative to x have exponents greater than* -1.

This criterion says roughly that $f\,dx$ has no poles, because at a point of the curve where $x = \alpha$ there is a local parameter s for which $x - \alpha = s^m$ for some m; to say that $f\,dx$ has no pole at this place where $s = 0$ is to say that $f\,d(s^m) = mfs^{m-1}\,ds$ has no pole, which is to say that multiplication of f by s^{m-1} clears its denominator, or, what is the same, that its expansion in powers of $x - \alpha = s^m$ contains no terms whose exponents are less than or equal to -1.

Proof Suppose first that all terms of all principal parts of f relative to x do have exponents greater than -1. If θ is integral over x, then its image under any embedding of K in $\mathbf{A}\langle s\rangle$ that takes x to $\alpha + s^m$ for some m is a series that contains no powers of s with negative exponents. Therefore, it contains no powers of $x - \alpha = s^m$ with negative exponents, so the image of $f\theta$ under any such embedding contains no terms in which the exponent on $x - \alpha$ is -1 or less. As was seen in Essay 4.7, the expression of $\mathrm{tr}_x(f\theta)$ as a power series in $x - \alpha$ is the sum of the n images of $f\theta$ in $\mathbf{A}\langle s\rangle$. Therefore, it is a series in which no term has an exponent less than or equal to -1. Thus, since it is a series expansion of a rational function of x, which implies that it contains no terms with fractional exponents, it is a power series in $x - \alpha$. Since this is true for every α, it follows that $\mathrm{tr}_x(f\theta)$ must in fact be a polynomial in x whenever θ is integral over x. In short, $f\,dx$ is holomorphic for finite x, as was to be shown.

Conversely, suppose that some embedding of the root field K of $\chi(x, y)$ in $\mathbf{A}\langle s\rangle$ that carries x to $\alpha + s^m$—where α is an algebraic number and m is a positive integer—carries f to a series in which the exponent on s is less than or equal to $-m$, so that the exponent on $x - \alpha$ is less than or equal to -1. Let $z_1, z_2, \ldots, z_n$ be an integral basis of the root field over x, and let $\theta = \sum c_i z_i$, where $c_1, c_2, \ldots, c_n$ are constants to be determined. Consider the terms in the n expansions of θ in (possibly fractional) powers of $x - \alpha$ in which the exponents are less than 1. When the n expansions do not involve fractional powers, the expansion of each z_i has just one term—the constant term—in which the exponent is less than 1 in each of the n embeddings, and the same is true of θ; in this case, the n constant terms in the series representations of θ are the entries of the column matrix Mc where c is the column matrix with entries $c_1, c_2, \ldots, c_n$ and M is the $n \times n$ matrix whose jth column contains the n constant terms in the images of z_j under the n embeddings.

When one or more of the n expansions do involve fractional exponents—when the curve is ramified at $x = \alpha$—a similar statement is true. An embedding involving powers of $s = \sqrt[m]{x - \alpha}$ in which $m > 1$ implies $m - 1$ others in which s is replaced by ωs for the mth roots of unity ω other than 1. There are precisely m terms (some of which may be zero) in any one of these series in which the exponent on $x - \alpha$ is less than 1, namely, the terms in $s^0, s^1, \dots, s^{m-1}$. Since the sum of the various values of m is n, it follows that all coefficients of all n expansions of θ in which the exponents on $x - \alpha$ are less than 1 are determined by just n such coefficients, namely, the coefficients of a *selection* of the expansions when just one expansion is selected from each set of m related expansions. The same is true of each z_j, and all coefficients of θ are determined by those given by a formula Mc, as before, in which the jth column of M gives the coefficients of the selected expansions of z_j.

(For example, in the case of the integral basis $1, y, y^2/x$ of the root field of $y^3 - xy + x^3$ over x, it was shown in Essay 4.4 that there are three expansions of y when $\alpha = 0$, namely, $y = \pm\sqrt{x} + \cdots$ and $y = 0 + \cdots$, where the omitted terms are divisible by x. Therefore, every θ integral over x has three expansions, but the one corresponding to $y = -\sqrt{x} + \cdots$ can be derived from the one corresponding to $y = \sqrt{x} + \cdots$ by changing $\sqrt{x}$ to $-\sqrt{x}$. When just two expansions, those corresponding to $y = \sqrt{x} + \cdots$ and $y = 0 + \cdots$, are selected, the selected expansions of $\theta = c_1 + c_2 y + c_3 y^2/x$ are given by the formula

$$Mc = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

where the first two rows contain the coefficients of 1 and $\sqrt{x}$ in the first selected expansion and the last row contains the coefficient of 1 in the second.)

That the matrix M is invertible can be proved as follows: If $Mc = 0$, then the n expansions of θ contain no terms in which the exponent on $x - \alpha$ is less than 1, so $\theta/(x - \alpha)$ is integral over x. If any c_i were nonzero, one of the coefficients $c_i/(x - \alpha)$ in the representation of $\theta/(x - \alpha)$ relative to the integral basis $z_1, z_2, \dots, z_n$ would not be a polynomial, contrary to the definition of an integral basis. Therefore, $Mc = 0$ implies $c = 0$, so the square matrix M is invertible. In other words, the coefficients in the terms of the expansion of θ in which the exponent on $x - \alpha$ is less than 1 can be given arbitrarily chosen values (subject to the relations among m such expansions when $m > 1$ that were just noted) by taking the column matrix c to be the column matrix of chosen values multiplied on the left by M^{-1}.

Because the principal parts of f are assumed to contain a nonzero term in which the exponent on $x - \alpha$ is less than or equal to -1, the least such exponent has the form $e = -i - \frac{j}{m}$, where $i \geq 1$ and $0 \leq j < m$. Let $\theta = \sum c_i z_i$ be chosen so that the coefficient of $(x - \alpha)^{j/m}$ in the expansion of θ in the embedding in $\mathbf{A}\langle s\rangle$ that gives rise to the nonzero term with exponent e in the expansion of f is nonzero, but all other expansion coefficients of terms with exponent less than 1 are zero. Then the expansion of $f\theta$ in m embeddings begins $\gamma(x - \alpha)^{-i} + \cdots$, where $\gamma \neq 0$, while in the remaining embeddings the expansion of $f\theta$ contains no terms in which the exponent on $x - \alpha$ is less than or equal to $-i$. Therefore, the sum of the n expansions of $f\theta$ is $m\gamma(x - \alpha)^{-i} + \cdots$, where $i \geq 1$. In particular, this sum is not a polynomial in x.

The bilinear form "the trace of the product" from $K \times K$ to $\mathbf{Q}(x)$ is described, relative to the integral basis $z_1, z_2, \dots, z_n$, by an $n \times n$ symmetric matrix S of polynomials in x with integer coefficients, namely, the matrix whose entry in the ith row of the jth column is $\mathrm{tr}_x(z_i z_j)$. To say that $f\,dx$ is holomorphic for finite x means simply that all entries of $[f]S$ are polynomials in x when $[f]$ denotes the row matrix whose entries are the coefficients that represent f in the integral basis $z_1, z_2, \ldots, z_n$. If this were the case, the sum of the expansions of $f\theta$, which is $[f]S[c]$ where $[c]$ contains the coefficients of $\theta = \sum c_i z_i$ as above, would also be a polynomial in x. Since it is not, $f\,dx$ must not be holomorphic for finite x, which completes the proof of Lemma 1. □

Theorem *Let z be a parameter in the root field of $\chi(x, y)$ and let $\frac{dx}{dz}$ be the element of the root field defined using implicit differentiation as above. A differential $f\,dx$ is holomorphic if and only if $f\frac{dx}{dz}\,dz$ is holomorphic.*

Proof The reciprocal of $\frac{dx}{dz}$ is $\frac{dz}{dx}$, so it will suffice to prove that "$f\,dx$ is holomorphic" implies "$f\frac{dx}{dz}\,dz$ is holomorphic."

Let δ be a given algebraic number and let an embedding of K in $\mathbf{A}\langle\sigma\rangle$ be given that carries z to $\delta + \sigma^\mu$ for some $\mu > 0$. It is to be shown that if $h\,dx$ is holomorphic, then the image of $h \cdot \frac{dx}{dz}$ under this embedding contains no terms in which the exponent on $z - \delta$ is less than or equal to -1, or, what is the same, that all exponents in the expansion of $(z - \delta) \cdot h \cdot \frac{dx}{dz}$ are positive.

Assume first that the image of x under the given embedding has no terms in which the exponent on σ is negative; say it is $\alpha + \alpha'\sigma + \alpha''\sigma^2 + \cdots$. In this case, let m be the exponent of the first nonzero term in the expansion of $x - \alpha$. (There is such a term because x is not a constant.) When an mth root ϵ_1 of the reciprocal of $\alpha^{(m)}$ is adjoined to $\mathbf{A}$, if necessary, the following lemma constructs a substitution $\sigma = \epsilon_1 s + \epsilon_2 s^2 + \epsilon_3 s^3 + \cdots$ that carries $x = \alpha + \alpha^{(m)}\sigma^m + \cdots$ to $\alpha + s^m$.

Lemma 2 *Given a nonzero power series $A_m x^m + A_{m+1}x^{m+1} + A_{m+2}x^{m+2} + \cdots$ in which the coefficients are algebraic numbers and the first nonzero term contains x to the power $m > 0$, and given an mth root C_1 of $1/A_m$, construct an infinite series $x = C_1 s + C_2 s^2 + C_3 s^3 + \cdots$ with algebraic number coefficients whose substitution in the series results in s^m.*

Proof Substitution of $x = C_1 s + C_2 s^2 + C_3 s^3 + \cdots$ in $A_m x^m + A_{m+1}x^{m+1} + A_{m+2}x^{m+2} + \cdots$ gives $B_m s^m + B_{m+1}s^{m+1} + B_{m+2}s^{m+2} + \cdots$ where $B_m = A_m C_1^m = 1$, $B_{m+1} = mA_m C_1^{m-1}C_2 + A_{m+1}C_1^{m+1}, \ldots$. The formula for B_{m+i} when $i > 0$ contains the terms $mA_m C_1^{m-1}C_{i+1}$ and $A_{m+i}C_1^{m+i}$; the remaining terms in the formula constitute a polynomial in $C_1, C_2, \ldots, C_i$ and $A_m, A_{m+1}, \ldots, A_{m+i-1}$ with integer coefficients. Thus, the requirement $B_{m+i} = 0$ for $i > 0$ is the statement that C_{i+1} is a polynomial in $C_1, C_2, \ldots, C_i$ and $A_m, A_{m+1}, \ldots, A_{m+i}$ divided by $mA_m C_1^{m-1} = m/C_1$. Since $A_m = 1/C_1^m$, it follows that each successive C_{i+1} can be expressed rationally in terms of $C_1, A_{m+1}, A_{m+2}, \ldots, A_{m+i}$. The series $C_1 s + C_2 s^2 + C_3 s^3 + \cdots$ constructed in this way has the required property. □

Because the given embedding $K \rightarrow \mathbf{A}\langle\sigma\rangle$ followed by the substitution $\sigma = \epsilon_1 s + \epsilon_2 s^2 + \cdots$ carries x to $\alpha + s^m$, and because $h\,dx$ is holomorphic, the resulting embedding $K \longrightarrow \mathbf{A}\langle s\rangle$ carries h to a series in s in which no term has an exponent less than or equal to $-m$ on s. Otherwise stated, all exponents in the expansion of $(x-\alpha)\cdot h$ in powers of s are positive. Since this expansion is found by substituting the expansion of σ in powers of s into the expansion of $(x-\alpha)h$ in powers of σ, it follows *that all exponents in the expansion of* $(x-\alpha)\cdot h$ *in powers of* σ *are positive.*

Let this expansion be multiplied by the expansion of $\frac{dx}{dz}\cdot\frac{z-\delta}{x-\alpha}$ in powers of σ. On the one hand, the result is $(z-\delta)\cdot h\cdot\frac{dx}{dz}$. On the other hand, if $\phi(x,z)=0$ is the equation satisfied by x and z, then $\phi\,(\alpha+\alpha^{(m)}\sigma^m+\cdots,\delta+\sigma^\mu)$ is identically zero, so differentiation with respect to σ gives $\phi_x(x,z)(m\alpha^{(m)}\sigma^{m-1}+\cdots)+\phi_z(x,z)\mu\sigma^{\mu-1}=0$, where x and z stand for their expansions as power series in σ and the omitted terms are divisible by σ^m. Multiplication by σ then gives $\phi_x(x,z)(m\alpha^{(m)}(x-\alpha)+\cdots)+\phi_z(x,z)\mu(z-\delta)=0$, where the omitted terms are divisible by σ^{m+1}. Division by $\phi_x(x,z)$ (which is not zero, because z is not a constant) times $x-\alpha$ gives $m\alpha^{(m)}+\cdots-\mu\frac{dx}{dz}\cdot\frac{z-\delta}{x-\alpha}=0$, where the omitted terms are all divisible by σ. This equation shows that the expansion in powers of σ of $\frac{dx}{dz}\frac{z-\delta}{x-\alpha}$ is the constant $\frac{m}{\mu}\alpha^{(m)}$ plus terms in σ. Therefore, $(z-\delta)\cdot h\cdot\frac{dx}{dz}=((x-\alpha)\cdot h)\big(\frac{m}{\mu}\alpha^{(m)}+\cdots\big)$ is a product of two series in σ, one with positive exponents and one with no negative exponents, which shows that *all terms in the expansion of* $(z-\delta)\cdot h\cdot\frac{dx}{dz}$ *in powers of* σ *have positive exponents*, a conclusion that holds for any embedding $K\longrightarrow\mathbf{A}\langle\sigma\rangle$ that carries z to $\delta+\sigma^\mu$ and carries x to a series with no negative exponents.

If an embedding that carries z to $\delta+\sigma^\mu$ carries x to a series with some negative exponents, it carries $u=\frac{1}{x}$ to a series in which all exponents are positive. Since $\frac{h}{u^2}\cdot du$ is holomorphic for finite u by virtue of the assumption that $h\,dx$ is holomorphic, it follows that all exponents in the expansion of $(z-\delta)\cdot\frac{h}{u^2}\cdot\frac{du}{dz}$ in powers of σ are positive. By the chain rule, $\frac{dx}{dz}=\frac{dx}{du}\frac{du}{dz}=\frac{-1}{u^2}\cdot\frac{du}{dz}$ when $x=\frac{1}{u}$, so it follows that all exponents in the expansion of $(z-\delta)\cdot h\cdot\frac{dx}{dz}=(z-\delta)\cdot h\cdot\frac{-1}{u^2}\cdot\frac{du}{dz}$ in powers of σ are positive in this case too.

Thus, Lemma 1 implies that $h\cdot\frac{dx}{dz}\cdot dz$ is holomorphic for finite z.

By the same token, $h\cdot\frac{dx}{dv}\cdot dv$ is holomorphic for finite v for any parameter v and in particular when $v=\frac{1}{z}$. Therefore $h\cdot\frac{dx}{dz}\cdot\frac{-1}{v^2}\cdot dv$ is holomorphic for finite $v=\frac{1}{z}$, which completes the proof that $h\cdot\frac{dx}{dz}\cdot dz$ is holomorphic. □

Corollary *The genus is a birational invariant.*

The determination of the genus can be accomplished by finding holomorphic differentials, for which the following proposition is useful.

An algebraic curve $\chi(x,y)=0$ is **nonsingular for finite** x if no pair (α,β) of algebraic numbers satisfies all three conditions $\chi(\alpha,\beta)=0$, $\chi_x(\alpha,\beta)=0$, and $\chi_y(\alpha,\beta)=0$.

Proposition *If* $\chi(x,y)=0$ *is nonsingular for finite* x*, then* $h\,dx$ *is holomorphic for finite* x *if and only if* $h\cdot\chi_y$ *is integral over* x.

In other words, when $\chi(x,y)=0$ is nonsingular for finite x, the differentials holomorphic for finite x are those of the form $\frac{\phi(x,y)\,dx}{\chi_y(x,y)}$, where $\phi(x,y)$ is integral over x.

Proof First assume that $h\,dx$ is holomorphic for finite x. By the proposition of Essay 4.5, it will suffice to prove that the image of $h \cdot \chi_y$ in each embedding of K in $\mathbf{A}\langle s\rangle$ that carries x to $\alpha + s^m$, and carries rational numbers to themselves, is without negative exponents.

When $\chi_y(\alpha,\beta) \neq 0$, β is a simple root of $\chi(\alpha, y)$, which implies, as was shown in Essay 4.4, that $x = \alpha + s$, $y = \beta$ is an unambiguous truncated solution of $\chi(x, y) = 0$. Such a truncated solution implies an infinite series solution $y = \beta + \beta'(x-\alpha) + \beta''(x-\alpha)^2 + \cdots$. The corresponding embedding $K \longrightarrow \mathbf{A}\langle s\rangle$ does not involve fractional powers of $x-\alpha$. The assumption that $h\,dx$ is holomorphic for finite x implies that the image of h in $\mathbf{A}\langle s\rangle$ contains no exponents less than or equal to -1, so all exponents are greater than or equal to zero. The same is true of the image of χ_y—it is a polynomial in x and y and is therefore integral over x—so the image of $h \cdot \chi_y$ under the embedding has no negative exponents, as was to be shown.

Otherwise, $\chi_x(\alpha,\beta) \neq 0$, because the curve is nonsingular for finite x. In this case, the polynomial $\Phi_0(t)$ in $\chi(\alpha+s, \beta+t) = \Phi_0(s) + \Phi_1(s)t + \cdots + t^n$ is divisible by s but not s^2, so the Newton polygon algorithm leads to a "polygon" with one segment from $(0, 1)$ to a point where $j_i = 0$; call it $(\tau, 0)$. The ambiguity of the truncated solution $x = \alpha + s$, $y = \beta$ is then τ, and the output of Newton's polygon is τ unambiguous truncated solutions $x = \alpha + s_1^\tau$, $y = \beta + \sqrt[\tau]{\zeta_0} \cdot s_1$ (one solution for each of the τ possible values of $\sqrt[\tau]{\zeta_0}$). By Lemma 2, the infinite series expansion $y - \beta = \sqrt[\tau]{\zeta_0} \cdot s_1 + \beta'' s_1 + \cdots$ implies an infinite series expansion $s_1 = \epsilon_1(y-\beta) + \epsilon_2(y-\beta)^2 + \cdots$, whose substitution in $\sqrt[\tau]{\zeta_0} \cdot s_1 + \beta'' s_1 + \cdots$ gives $y - \beta$ and whose substitution in the embedding $K \longrightarrow \mathbf{A}\langle s_1\rangle$ therefore gives an embedding $K \longrightarrow \mathbf{A}\langle y-\beta\rangle$ that carries y to $\beta + (y-\beta)$. Because $h \cdot \frac{dx}{dy} \cdot dy$ is holomorphic, it follows that the image of $h \cdot \frac{dx}{dy} = -h \cdot \frac{\chi_y}{\chi_x}$ under this embedding has no exponents less than or equal to -1. It has no fractional exponents, so all exponents in the expansion of $h \cdot \frac{\chi_y}{\chi_x}$ in powers of $y - \beta$ are at least zero. Therefore, the same is true of its expansion in powers of s_1. Since χ_x is a polynomial in $x = \alpha + s_1^\tau$ and $y = \beta + \sqrt[\tau]{\zeta_0} \cdot s_1 + \cdots$, the expansion of $h \cdot \chi_y = h \cdot \frac{\chi_y}{\chi_x} \cdot \chi_x$ in powers of s_1 has no terms with negative exponents, as was to be shown. Thus, the proof that "$h\,dx$ is holomorphic for finite x" implies "$h \cdot \chi_y$ is integral over x" is complete.

To prove, conversely, that all differentials of the form $\frac{\phi}{\chi_y}\,dx$ in which ϕ is integral over x are holomorphic for finite x it will suffice to prove that $\frac{1}{\chi_y}\,dx$ is holomorphic for finite x. Certainly for any embedding $x = \alpha + s^m$, $y = \beta + \beta' s + \beta'' s^2 + \cdots$ for which $\chi_y(\alpha,\beta) \neq 0$ the expansion of $\frac{1}{\chi_y}$ in powers of s contains no negative exponents. All other embeddings have the form $x = \alpha + s_1^\tau$, $y = \beta + \beta' s_1 + \beta'' s_1^2 + \cdots$, where $\beta' \neq 0$—as was just seen—by virtue of the assumption that $\chi(x, y)$ has no singularities for finite x. Differentiation of $\chi(\alpha + s_1^\tau, \beta + \beta' s_1 + \cdots) = 0$ gives $\chi_x(\alpha + s_1^\tau, \beta + \beta' s_1 + \cdots) \cdot \tau s_1^{\tau-1} + \chi_y(\alpha + s_1^\tau, \beta + \beta' s_1 + \cdots) \cdot (\beta' + \cdots) = 0$. Thus, $\frac{1}{\chi_y} = -\frac{\beta' + \cdots}{\chi_x \cdot \tau \cdot s_1^{\tau-1}}$. Since $\chi_x(\alpha,\beta) \neq 0$, it follows that $-\tau + 1$ is the least exponent in the expansion of $\frac{1}{\chi_y}$ in powers of s_1. Since $s_1^{-\tau+1} = (y-\beta)^{-1+\frac{1}{\tau}}$, the desired conclusion that the principal parts of $\frac{1}{\chi_y}$ relative to this embedding $K \longrightarrow \mathbf{A}\langle s\rangle$ contain no exponents that are -1 or less follows. □

Example 9 $\chi(x, y) = y^2 + x^4 - 1$ (the elliptic curve mentioned in Essay 4.2)

Since $\chi_y = 0$ implies $y = 0$ and $\chi_x = 0$ implies $x = 0$, this curve is nonsingular for finite x, because $(\alpha, \beta) = (0, 0)$ does not satisfy $\beta^2 + \alpha^4 - 1 = 0$. Therefore, the differentials holomorphic for finite x are those of the form $\frac{\phi\, dx}{2y}$, where ϕ is integral over x. The substitution $x = \frac{1}{u}$, $y = \frac{v}{u^2}$ puts this curve in the form $v^2 + 1 - u^4 = 0$ and puts $dx/2y$ in the form $-(\frac{1}{u^2})du/2(\frac{v}{u^2}) = -du/2v$, so $\phi\, dx/2y$ is holomorphic if and only if ϕ is integral over x *and* over $u = 1/x$, which is to say, if and only if ϕ is constant.

Example 10 $\chi(x, y) = y^3 + x^3 y + x$ (the Klein curve)

This curve is nonsingular for finite x, because $3\beta^2 + \alpha^3 = 0$ and $3\alpha^2\beta + 1 = 0$ imply that $\beta = -\frac{1}{3\alpha^2}$ and $\alpha^3 = -3\beta^2 = -\frac{1}{3\alpha^4}$, so that $\alpha^7 = -\frac{1}{3}$; therefore $\beta^3 + \alpha^3\beta + \alpha \neq 0$, because $\alpha^3\beta + \alpha = \alpha^3 \cdot \frac{-1}{3\alpha^2} + \alpha = \frac{2\alpha}{3}$ is not the negative of $\beta^3 = -\frac{1}{27\alpha^6} = -\frac{\alpha}{27\cdot(-1/3)} = \frac{\alpha}{9}$.

Since $\chi_y(x, y) = 3y^2 + x^3$, every holomorphic differential can be written $\frac{\phi\, dx}{3y^2 + x^3}$ for some ϕ integral over x. Because $1, y, y^2$ is an integral basis over x (Example 2 of Essay 4.5), ϕ must be a polynomial in x and y. Determining the holomorphic differentials on the Klein curve therefore amounts to determining the polynomials $\phi(x, y)$ in x and y for which $\frac{\phi\, dx}{3y^2 + x^3}$ is holomorphic. Such a differential is holomorphic for finite x, and the problem is to determine the conditions under which it is without poles at $x = \infty$.

The substitution $x = \frac{1}{u}$, $y = \frac{v}{u^3}$ transforms the curve into $v^3 + u^3 v + u^8 = 0$, a curve with a singularity at $(u, v) = (0, 0)$. The first step of the Newton polygon algorithm in the case in which the value of u is 0 calls for setting $u = s$ and $v = 0 + t$, which leads to $s^8 + s^3 t + t^3$. The polygon is based on the points $(0, 8)$, $(1, 3)$, and $(3, 0)$, so it consists of two segments $5i + j = 8$ and $3i + 2j = 9$. The first segment furnishes an unambiguous truncated solution $u = s$, $v = -s^5$, which implies an infinite series solution, and the second furnishes the remaining two solutions in the form of the unambiguous truncated solutions $u = \sigma^2$, $v = \pm i\sigma^3$.

The expression of $\frac{dx}{3y^2 + x^3}$ relative to the first of these is

$$\frac{dx}{3y^2 + x^3} = \left(-\frac{du}{u^2}\right) \cdot \frac{1}{3 \cdot \frac{v^2}{u^6} + \frac{1}{u^3}} = \frac{-s^4 ds}{(3s^{10} + \cdots)^2 + s^3} = (-s + \cdots)ds.$$

Therefore, there is no pole for this embedding if this differential is multiplied by $x = \frac{1}{s} + \cdots$ or by any power of $y = \frac{v}{u^3} = -s^2 + \cdots$. The expression of $\frac{dx}{3y^2 + x^3}$ relative to the second is

$$\frac{dx}{3y^2 + x^3} = \frac{-2d\sigma}{\sigma^3} \cdot \frac{1}{3 \cdot \frac{(-\sigma^6 + \cdots)}{\sigma^{12}} + \frac{1}{\sigma^6}} = \frac{2d\sigma}{\sigma^3} \cdot \frac{\sigma^6}{2 + \cdots} = (\sigma^3 + \cdots)d\sigma.$$

Therefore, it remains finite for this embedding if it is multiplied by $x = \frac{1}{\sigma^2}$ or by $y = \pm\frac{i}{\sigma^3} + \cdots$ but not if it is multiplied by any polynomial of higher degree in x or y.

In conclusion, the holomorphic differentials in this case are

$$\frac{(a + bx + cy)\,dx}{3y^2 + x^3},$$

as was already found in Example 7 of Essay 4.6.

Chapter 9
Abel's Theorem

> . . . every theorem of algebra or higher analysis, no matter how remote it may seem, can be expressed as a statement about natural numbers . . . [This is] something I often heard Dirichlet say.[1]—R. Dedekind [23, p. 338, vol. 3]

Essay 9.1 What Was Abel's Theorem?

Abel's 1826 Paris Memoir *Mêmoire sur une propriété générale d'une classe très étendue de fonctions transcendentes* [2] deals with a certain type of transcendental function, namely, indefinite integrals of the form $\int f(x, y)\,dx$, where $f(x, y)$ is a rational function of two variables and y is an algebraic function of x. Such integrals are now called **Abelian integrals**. They generalize **elliptic integrals**,[2] which are Abelian integrals in which the definition of y as an algebraic function of x has the particular form $y^2 = f(x)$, where $f(x)$ is a polynomial of degree 3 or 4 with rational coefficients and distinct roots. Abel's intention in the Paris Memoir, as he described it in the introduction, was to generalize a known property of elliptic integrals to all Abelian integrals.

For Abel, an **algebraic function** y of x is defined by an equation of the form

$$\chi(x, y) = p_0(x) + p_1(x)y + p_2(x)y^2 + \cdots + p_{n-1}(x)y^{n-1} + y^n = 0,$$

where $p_0(x), \ldots, p_{n-1}(x)$ are polynomials. According to Abel, this "gives for the function y a number n of different forms." Algebraic functions are not functions in the modern sense, because they can take two different values for the same x. Over the complex numbers, the implicit function theorem gives one way to make this

[1] . . . *daß jeder auch noch so fern liegende Satz der Algebra und höheren Analysis sich als ein Satz über die natürlichen Zahlen aussprechen läßt, eine Behauptung, die ich auch wiederholt aus dem Munde von Dirichlet gehört habe.*

[2] Abel calls these "elliptic functions," but today the term "elliptic function" refers to a function whose inverse is an elliptic integral.

H. M. Edwards, *Essays in Constructive Mathematics*,
https://doi.org/10.1007/978-3-030-98558-5_9

precise.[3] However, analytic continuation can change one of these locally defined functions into another, which is part of what led Riemann to introduce the concept of a Riemann surface. These approaches involve complex numbers and hence are not constructive. Instead, observe that the splitting field of $\chi(x, y)$ over $\mathbf{Q}(x)$ contains n roots of the polynomial, which are algebraic versions of the "different forms" mentioned by Abel. Adjoining one the roots, call it y, to $\mathbf{Q}(x)$ gives the curve field $\mathcal{K} = \mathbf{Q}(x, y)$ as studied in Chapter 8.

One expression of the property of elliptic integrals that Abel wanted to generalize is that the sum of two elliptic integrals $\int f(x_1, y_1)\,dx_1 + \int f(x_2, y_2)\,dx_2$ is an elementary function when the points (x_1, y_1) and (x_2, y_2) satisfy a suitable condition. The theorem sketched by Abel in the introduction to the Paris Memoir generalizes this to a statement about arbitrary Abelian integrals. Specifically, it states that, for an Abelian integral $\int f(x, y)\,dx$, the sum

$$\int f(x_1, y_1)\,dx_1 + \cdots + \int f(x_\mu, y_\mu)\,dx_\mu = v,$$

where the x_i and y_i are algebraic functions of parameters $a, a', a'', \ldots$, is an "algebraic and logarithmic function" when x_i and y_i satisfy certain algebraic conditions. Unlike the special case of elliptic integrals, however, in the general case it is normally necessary to impose more than one condition. The number of conditions is in fact the genus of the curve field $\mathcal{K} = \mathbf{Q}(x, y)$.

Abel explains "algebraic and logarithmic function" as follows:

> If now dv is a rational differential function of the quantities $a, a', a'', \ldots$, its integral or the quantity v will be an algebraic and logarithmic function of the $a, a', a'', \ldots$ [2, p. 149][4]

In other words, Abel wants to study when the sum

$$f(x_1, y_1)\,dx_1 + \cdots + f(x_\mu, y_\mu)\,dx_\mu = dv \tag{1}$$

is a rational differential in the parameters $a, a', a'', \ldots$ This is a purely algebraic question.

The memoir [2] has different versions of what might be called "Abel's theorem." After developing background material on normal bases and holomorphic differentials in Essays 9.2–9.5, four "Abel theorems" inspired by [2] will be discussed. The theorem of Essay 9.7 shows that (1) is a rational differential in the parameters when the (x_i, y_i) are a full set of conjugate solutions (this will be made precise), and the theorem of Essay 9.8 proves that in the situation of Essay 9.7, the sum in (1) equals zero when $f(x, y)$ is a holomorphic differential. This is where the genus makes its

[3] The equation $\chi(x, y) = 0$ defines y locally as a function of x in the sense that, by the implicit function theorem (see, for example, [27, pp. 133–134]), the relation $\chi(x, y) = 0$ defines y as a function of x for all values of x that are sufficiently near a value x_1 of x for which there is a y_1 that satisfies $\chi(x_1, y_1) = 0$ and $\frac{\partial\chi}{\partial y}(x_1, y_1) \neq 0$. The resulting function satisfies $y(x_1) = y_1$.

[4] Loosely speaking, this claim follows from the method of partial fractions explained in Endnote 9.1. See Essays 9.7 and 9.9 for the precise relation between dv and "algebraic and logarithmic" functions.

first appearance. Finally, the two theorems of Essay 9.10 are versions of an addition theorem that makes essential use of the genus.

Steven Kleiman's paper *What is Abel's Theorem Anyway?* [50] is an important source for the history of Abel's theorem. He discusses four "Abel theorems" implicit in [2] that are related to the theorems stated in Essays 9.7, 9.8, and 9.10. However, unlike [50], the treatment here is constructive and purely algebraic.

The final essay of Chapter 9 discusses the relation between Abel's Theorem and the well-known "addition" defined for points on elliptic curves.

Essay 9.2 Normal Bases

As will be seen, the idea of a **normal basis** of the field of rational functions on the algebraic curve $\chi(x, y) = 0$ as an extension of $\mathbf{Q}(x)$ plays a major role in the statements and proofs of the theorems to be proved in this chapter. The notion of a normal basis was defined in the paper of Dedekind and Weber [22, §22], and, as far as I know, this was its first appearance. It leads to what is now called the **genus** of the curve in question for any curve $\chi(x, y) = 0$, and, as far as I can see, Abel did not determine the genus in all cases, only large classes of special curves. In this respect, then, Dedekind and Weber had made an important advance with the idea of a normal basis (although their definition was not constructive).

As in Essay 8.11, a quantity z in a curve field $\mathcal{K}$ defined by $\chi(x, y) = 0$ is **integral over** x if it has poles only at places where x, regarded as a quantity in $\mathcal{K}$, has poles or, what is the same, if some power of it can be written as a linear combination of lower powers

$$(1) \qquad z^k = \phi_1(x)z^{k-1} + \phi_2(x)z^{k-2} + \cdots + \phi_k(x)$$

with coefficients $\phi_i(x)$ that are polynomials in x with rational coefficients.

If z is integral over x, then $\frac{z}{x^\nu}$ is integral over $\frac{1}{x}$ for all sufficiently large integers ν, because division of (1) by $x^{k\nu}$ gives $(\frac{z}{x^\nu})^k = \frac{\phi_1(x)}{x^\nu}\cdot(\frac{z}{x^\nu})^{k-1} + \frac{\phi_2(x)}{x^{2\nu}}\cdot(\frac{z}{x^\nu})^{k-2} + \cdots + \frac{\phi_k(x)}{x^{\nu k}}$, which shows that $\frac{z}{x^\nu}$ is integral over $\frac{1}{x}$ whenever ν is large enough that the degree of $\phi_i(x)$ is at most νi for all i. The **order of** z **at** $x = \infty$ of a quantity z that is integral over x is by definition the least integer ν for which $\frac{z}{x^\nu}$ is integral over $\frac{1}{x}$.

A **basis** of $\mathcal{K}$ over $\mathbf{Q}(x)$ is of course a subset $y_1, y_2, \ldots, y_n$ of $\mathcal{K}$ for which each z in $\mathcal{K}$ has one and only one representation in the form

$$z = \phi_1(x)y_1 + \phi_2(x)y_2 + \cdots + \phi_n(x)y_n,$$

where the coefficients $\phi_i(x)$ are in $\mathbf{Q}(x)$. Loosely speaking, a *normal* basis of $\mathcal{K}$ with respect to x is one for which the representation of z in the basis reveals whether z is integral over x and, if so, what its order is at $x = \infty$.

Specifically, a basis $y_1, y_2, \ldots, y_n$ of $\mathcal{K}$ over $\mathbf{Q}(x)$ is a **normal basis** of $\mathcal{K}$ over $\mathbf{Q}(x)$ if (1) the quantities z in $\mathcal{K}$ that are integral over x are those whose representations in the basis have coefficients $\phi_i(x)$ that are *polynomials* with rational coefficients, and if (2) a quantity z in $\mathcal{K}$ that is integral over x has order at most ν at $x = \infty$ if and only if each term $\phi_i(x)y_i$ in its representation relative to the basis has order at most ν at $x = \infty$.

Since the order of $\phi_i(x)y_i$ at $x = \infty$ for a polynomial $\phi_i(x)$ is $\deg \phi_i(x)$ plus the order of y_i at $x = \infty$, condition (2) states that, for large integers ν, a quantity z in $\mathcal{K}$ that is integral over x has order at most ν at $x = \infty$ if and only if, in its representation relative to the basis, $\deg \phi_i(x) \le \nu - \mu_i$ for all i, where μ_i is the order at $x = \infty$ of y_i.

A basis with just property (1) is called an **integral basis**. A two-part algorithm can be given (see **Construction of an Integral Basis** in Essay 4.5) for constructing an integral basis for a given $\mathcal{K} \supset \mathbf{Q}(x)$. Briefly, the first part supplements the basis $1, y, y^2, \ldots, y^{n-1}$ of $\mathcal{K}$ over $\mathbf{Q}(x)$ by finding enough quantities in $\mathcal{K}$ that are integral over x, if more are needed, to span, over $\mathbf{Q}[x]$, the set of *all* quantities of $\mathcal{K}$ that are integral over x. It makes use of the fact that if $\frac{p(x,y)}{q(x)}$ is integral over x, then the square of $q(x)$ must divide the determinant of the matrix S whose entry in the ith row of the jth column is $\mathrm{tr}_x(y^{i+j-2})$. This observation provides a common denominator $d(x)$ for the quantities that are integral over x. Then the condition that $p(x, y)$ must, in order for a proper fraction $\frac{p(x,y)}{q(x)}$ to be integral over x, have a zero of order at least ν at every point P where $d(x)$ has a zero of order ν provides a set of homogeneous linear equations that the finite number of coefficients of $p(x, y)$ must satisfy. A basis of the solution space of these equations gives the needed spanning set.

The second part of the construction of an integral basis is an algorithm that takes as input a set of m quantities in $\mathcal{K}$ that span, over $\mathbf{Q}[x]$, the quantities that are integral over x and gives as output a set of $m - 1$ quantities with the same property whenever m is greater than the degree n of the extension $\mathcal{K} \supset \mathbf{Q}(x)$. Iteration of this algorithm terminates with a set of just n quantities integral over x that span, over $\mathbf{Q}[x]$, all quantities integral over x. Such a set is an integral basis of $\mathcal{K}$ over x.

As follows directly from the definitions, an integral basis $y_1, y_2, \ldots, y_n$ of $\mathcal{K}$ over $\mathbf{Q}(x)$ is a normal basis of $\mathcal{K}$ over $\mathbf{Q}(x)$ if and only if $\frac{y_1}{x^{\mu_1}}, \frac{y_2}{x^{\mu_2}}, \ldots, \frac{y_n}{x^{\mu_n}}$ is an integral basis of $\mathcal{K}$ over $\mathbf{Q}(\frac{1}{x})$, where, as above, μ_i is the order of y_i at $x = \infty$. Given an integral basis that is not a normal basis, a normal basis can be constructed using the following algorithm to replace one y_i with another for which μ_i is reduced but the new set of y's is still an integral basis.

Let an integral basis $z_1, z_2, \ldots, z_n$ of $\mathcal{K}$ over $\mathbf{Q}(\frac{1}{x})$ be constructed. If the representation of each z_i relative to the basis $\frac{y_1}{x^{\mu_1}}, \frac{y_2}{x^{\mu_2}}, \ldots, \frac{y_n}{x^{\mu_n}}$ of the curve field over $\mathbf{Q}(x)$ has coefficients that are polynomials in $\frac{1}{x}$ with rational coefficients, then both bases are integral bases over $\mathbf{Q}(\frac{1}{x})$, and $y_1, y_2, \ldots, y_n$ is a normal basis over $\mathbf{Q}(x)$. Otherwise, at least one of $z_1, z_2, \ldots, z_n$ is a quantity z that is integral over $\mathbf{Q}(\frac{1}{x})$ but its representation $z = \psi_1(x) \cdot \frac{y_1}{x^{\mu_1}} + \psi_2(x) \cdot \frac{y_2}{x^{\mu_2}} + \cdots + \psi_n(x) \cdot \frac{y_n}{x^{\mu_n}}$ has at least one coefficient $\psi_j(x)$ that is not a polynomial in $\frac{1}{x}$. Therefore, when each coefficient is written in the form $\psi_j(x) = x\xi_i(x) + \theta_j(\frac{1}{x})$ where $\xi_j(x)$ is a polynomial in x and

$\theta_j(\frac{1}{x})$ is a polynomial in $\frac{1}{x}$,[5] both with rational coefficients, and then at least one $\xi_j(x)$ is nonzero. Let k be an index for which $\xi_k(x) \neq 0$ has maximum degree, say σ is this maximum degree, and, if $\xi_k(x)$ has degree σ for more than one index, choose k to be an index for which $\deg \xi_k(x) = \sigma$ and μ_k is as large as possible. Define y_k' to be $\sum_j c_j x^{\mu_k - \mu_j} y_j$, where c_j is the coefficient of x^σ in $\xi_j(x)$ (so $c_j = 0$ when $\deg \xi_j(x) \neq \sigma$). Replacement of y_k with y_k' gives a new integral basis over $\mathbf{Q}(x)$ in which $\mu_1 + \mu_2 + \cdots + \mu_n$ is reduced.

(It is still an integral basis over $\mathbf{Q}(x)$ because the coefficient of y_k in y_k' is $c_k \neq 0$ and the exponent $\mu_k - \mu_j$ on x in the terms of y_k' is never negative. Moreover, y_i is unchanged when $i \neq k$, so μ_i is also unchanged, and what is to be shown is that μ_k is reduced. Note that $\frac{z}{x^{\sigma+1}} = \sum_j (c_j + \cdots) \cdot \frac{y_j}{x^{\mu_j}}$, where the omitted terms in the parentheses all contain positive powers of $\frac{1}{x}$. Thus, $\frac{z}{x^{\sigma+1}} = \frac{y_k'}{x^{\mu_k}} + \cdots$, where the omitted terms are not only integral over $\frac{1}{x}$ but remain integral over $\frac{1}{x}$ when they are multiplied by x. Since $\sigma \geq 0$, $\frac{z}{x^{\sigma+1}}$ is also integral over $\frac{1}{x}$ and remains integral over $\frac{1}{x}$ when it is multiplied by x. Therefore, $\frac{y_k'}{x^{\mu_k}}$, as the difference of two quantities with this property, also has this property, which implies that $\frac{x y_k'}{x^{\mu_k}} = \frac{y_k'}{x^{\mu_k - 1}}$ is integral over $\frac{1}{x}$, as was to be shown.)

Since one μ_i can be reduced without changing μ_j for $j \neq i$ whenever the given integral basis is not a normal basis, and since $\mu_1 + \mu_2 + \cdots + \mu_n$ can be reduced at most a finite number of times, the algorithm must terminate with a basis that is normal over $\mathbf{Q}(x)$, as was to be constructed.

In conclusion, then:

Theorem *Let $\mathcal{K}$ be a curve field and let x be a parameter in $\mathcal{K}$. A normal basis of the field extension $\mathcal{K} \supset \mathbf{Q}(x)$ can be constructed by means of rational arithmetic.*

Example 1 Given $\chi(x, y) = p_0(x) + p_1(x)y + \cdots + p_{n-1}(x)y^{n-1} + y^n$, Abel constructs an auxiliary polynomial $\theta(x, y, a', a'', \dots)$ using the basis $1, y, \dots, y^{n-1}$ of $\mathcal{K}$ over $\mathbf{Q}(x)$ (see [2, p. 147]). Although $1, \dots, y^{n-1}$ are clearly integral over $\mathbf{Q}(x)$, they need not form a normal basis. Consider $\chi(x, y) = y^3 - xy + x^3$ (folium of Descartes). One can check that $\alpha = y^2/x$ satisfies

$$\alpha^3 = 2\alpha^2 - \alpha + x^3.$$

Thus α in integral over $\mathbf{Q}(x)$. But in the basis $1, y, y^2$, α has the representation

$$\alpha = \frac{y^2}{x} = 0 \cdot 1 + 0 \cdot y + \frac{1}{x} \cdot y^2.$$

The coefficient of y^2 is not a polynomial in x, so $1, y, y^2$ is not an integral basis, hence not normal. Example 1 of Essay 4.5 shows that $1, y, y^2/x$ is a normal basis. Essay 9.6 will construct a version of $\theta(x, y, a', a'', \dots)$ that uses a normal basis.

[5] That $\psi_j(x)$ can be written in this form is proved as follows. The quantity z integral over $\mathbf{Q}(\frac{1}{x})$, so that $x^\nu z$ is integral over $\mathbf{Q}(x)$ when ν is sufficiently large. Therefore, $x^{\nu - \mu_j} \psi_j(x)$ is a polynomial in x because $y_1, \dots, y_n$ is an integral basis over $\mathbf{Q}(x)$.

Essay 9.3 The Field of Constants

The constants in a curve field—those quantities that are roots of polynomials with rational coefficients—constitute a subfield of $\mathcal{K}$. This subfield, the **field of constants** of $\mathcal{K}$, will be denoted by $\mathcal{K}_0$. A quantity in $\mathcal{K}$ is in $\mathcal{K}_0$ if and only if it is integral over x and has order 0 at $x = \infty$. Thus, the construction of a normal basis of $\mathcal{K}$ implies a construction of its field of constants, because $\mathcal{K}_0$ *is the span over* $\mathbf{Q}$ *of those elements in a normal basis that are constants.*

The extension $\mathcal{K}_0 \supset \mathbf{Q}$ is algebraic, and the degree of this extension is simply the number of constants in a normal basis of $\mathcal{K}$ over x. This degree is also the degree of the extension $\mathcal{K}_0(x) \supset \mathbf{Q}(x)$, because the constants in a normal basis are also a basis of $\mathcal{K}_0(x) \supset \mathbf{Q}(x)$.

The construction of the preceding essay can easily be modified to construct a **normal basis of $\mathcal{K}$ over $\mathcal{K}_0(x)$**. Such a normal basis contains just one basis element that has order 0 at $x = \infty$—that is, just one basis element that is a constant.

By the theorem of the primitive element (see Endnote 7.2) every curve field $\mathcal{K}$ can be written as an algebraic extension of $\mathbf{Q}(x)$ obtained by adjoining a single root y of an irreducible polynomial in x and y with integer coefficients that is monic in y. However, intrinsic properties of $\mathcal{K}$ are more likely to be apparent when $\mathcal{K}$ is represented as an algebraic extension of $\mathcal{K}_0(x)$. For example, $y^4 + 2x^2y^2 + x^4 - 2 = 0$ describes a curve field that is more clearly described as the field constructed by adjoining a root y of $y^2 + x^2 - \sqrt{2}$ to the field $\mathcal{K}_0(x)$, where $\mathcal{K}_0$ is the field of constants $\mathbf{Q}(\sqrt{2})$ of this curve field.

For this reason, it is natural to break the extension $\mathcal{K} \supset \mathbf{Q}(x)$ into two steps, the first being the extension $\mathcal{K}_0(x) \supset \mathbf{Q}(x)$, that adjoins the constants in $\mathcal{K}$, and the second being the extension $\mathcal{K} \supset \mathcal{K}_0(x)$. Otherwise stated, it is natural to make two uses of the theorem of the primitive element to find first a monic polynomial $f(x)$ with integer coefficients that is irreducible over $\mathbf{Q}$ and describes $\mathcal{K}_0 \supset \mathbf{Q}$ (as the extension that adjoins one root of $f(x)$ to $\mathbf{Q}$), and then a monic polynomial $\chi(x, y)$ in y with coefficients in $\mathcal{K}_0[x]$ that is irreducible over $\mathcal{K}_0(x)$, to describe $\mathcal{K} \supset \mathcal{K}_0(x)$ (as the extension which adjoins one root y of $\chi(x, y)$ to $\mathcal{K}_0(x)$). The remainder of Chapter 9 will use this $\chi(x, y)$.

An important property of $\chi(x, y)$ is that it remains irreducible when the field of constants is enlarged:

Proposition 1 *Let $\mathcal{K}$ be a curve field with field of constants $\mathcal{K}_0$, and let $\chi(x, y)$ be as above. If an extension of $\mathcal{K}$ contains a finite extension $\mathcal{K}_0'$ of $\mathcal{K}_0$, then $\chi(x, y)$ is irreducible over $\mathcal{K}_0'(x)$ and $\mathcal{K}_0'$ is the field of constants of the curve field $\mathcal{K}' = \mathcal{K}_0'(x, y)$.*

Proof By the theorem of the primitive element, $\mathcal{K}_0' = \mathcal{K}_0(v)$ for some root v of an irreducible polynomial $\phi(z)$ with coefficients in $\mathcal{K}_0$. That $\phi(z)$ is irreducible over $\mathcal{K}$ is shown as follows. Let $g(z)$ in $\mathcal{K}[z]$ be an irreducible factor of $\phi(z)$. The coefficients of $g(z)$ are (up to sign) the elementary symmetric polynomials of its roots, which are also roots of $\phi(z)$. It follows that the coefficients of $g(z)$ are algebraic over $\mathcal{K}_0$. But they also lie in $\mathcal{K}$ and hence lie in the field of constants $\mathcal{K}_0$. So $g(z)$ has coefficients in $\mathcal{K}_0$, which implies that $g(z) = \phi(z)$ up to a constant. Thus $\phi(z)$ is irreducible over $\mathcal{K}$ and hence also irreducible over $\mathcal{K}_0(x)$.

Let m be the degree of $\phi(z)$. The previous paragraph implies that the extensions $\mathcal{K} \subset \mathcal{K}' = \mathcal{K}(v)$ and $\mathcal{K}_0(x) \subset \mathcal{K}_0'(x) = \mathcal{K}(x, v)$ have degree m. Furthermore, $\mathcal{K}_0(x) \subset \mathcal{K} = \mathcal{K}_0(x, y)$ has degree n since $\chi(x, y)$ is irreducible of degree n over $\mathcal{K}_0(x)$. Therefore,

$$\mathcal{K}_0(x) \subset \mathcal{K} \subset \mathcal{K}'$$

has degree nm. Now consider

$$\mathcal{K}_0(x) \subset \mathcal{K}_0'(x) \subset \mathcal{K}'.$$

The whole extension has degree nm, and the extension on the left has degree m. Thus, $\mathcal{K}_0'(x) \subset \mathcal{K}' = \mathcal{K}_0'(x, y)$ has degree n, which proves that $\chi(x, y)$ is irreducible over $\mathcal{K}_0'(x)$.

Finally, let $\mathcal{K}_0''$ be the field of constants of the curve field $\mathcal{K}' = \mathcal{K}_0'(x, y)$. Since $\mathcal{K}_0''$ is algebraic over $\mathcal{K}_0$, the previous paragraph implies that $\chi(x, y)$ is irreducible over both $\mathcal{K}_0'(x)$ and $\mathcal{K}_0''(x)$, which implies

$$\text{degree of } \mathcal{K}_0'(x) \subset \mathcal{K}' \quad = \quad \text{degree of } \mathcal{K}_0''(x) \subset \mathcal{K}'.$$

Since $\mathcal{K}_0'(x) \subset \mathcal{K}_0''(x)$, this forces $\mathcal{K}_0'(x) = \mathcal{K}_0''(x)$, and $\mathcal{K}_0' = \mathcal{K}_0''$ follows because x is transcendental over $\mathcal{K}_0'$ and $\mathcal{K}_0''$. Therefore, $\mathcal{K}_0'$ is the field of constants of $\mathcal{K}'$, as was to be shown. □

A final observation is that if $\mathbf{Q}$ is replaced by the field $\hat{\mathbf{Q}} = \mathbf{Q}(\alpha_1, \dots, \alpha_h)$ for indeterminates $\alpha_1, \dots, \alpha_h$ as in Essay 8.2, then everything done so far in Chapter 9 remains true for a curve field $\mathcal{K} = \hat{\mathbf{Q}}(x, y)$. In particular, the field of constants $\mathcal{K}_0$ is now a finite extension of $\hat{\mathbf{Q}}$.

Essay 9.4 Differentials and Holomorphic Differentials

If x and y are parameters in a curve field $\mathcal{K}$, the **derivative** $\frac{dy}{dx}$ of y with respect to x is the unique quantity in $\mathcal{K}$ whose value at every point $(x, y) = (a, b)$ of $\mathcal{K}$ at which x is a local parameter as in Essay 8.4 and x and y have the finite values a and b respectively is the value of $\frac{y-b}{x-a}$ at the point. The existence of such a quantity in $\mathcal{K}$ follows from the fact that, when x and y are parameters in $\mathcal{K}$, there is a polynomial $\phi(X, Y)$ with integer coefficients for which the quantity $\phi(x, y)$ of $\mathcal{K}$ is zero. Let $t = x - a$ at a point $(x, y) = (a, b)$ of $\mathcal{K}$ where x is a local parameter, and let $y = b + ct + \cdots$ be the expansion of y in powers of t. All coefficients in the expansion of $\phi(a + t, b + ct + \cdots)$ in powers of t must be zero, because $\phi(x, y) = 0$ in $\mathcal{K}$. On the other hand, the coefficient of t in this expansion can be found by differentiation (an algebraic operation) of $\phi(a + t, b + ct + \cdots)$ at $t = 0$ to find that it is the rational function $\frac{\partial\phi(x,y)}{\partial x} \cdot 1 + \frac{\partial\phi(x,y)}{\partial y} \cdot c$ of x and y evaluated at the point. Thus, on the one hand, the coefficient c of $t = x - a$ in the expansion $ct + \cdots$ of $y - b$ is the value of $\frac{y-b}{x-a}$ at the point, while on the other hand, it is the value of $-\frac{\partial\phi(x,y)}{\partial x} / \frac{\partial\phi(x,y)}{\partial y}$ at the same point. Since a nonzero quantity in $\mathcal{K}$ is zero at at most a finite number of points, and since $-\frac{\partial\phi(x,y)}{\partial x} / \frac{\partial\phi(x,y)}{\partial y}$ has the desired value at infinitely many points, it is the *only* quantity in $\mathcal{K}$ that fulfills the defining properties of $\frac{dy}{dx}$. (Note that the limit concept plays no role in the definition of $\frac{dy}{dx}$.)

A **differential of $\mathcal{K}$ with respect to** x, where x is a parameter of $\mathcal{K}$, is an expression of the form $F\,dx$ in which F is a quantity in $\mathcal{K}$. A differential $F\,dx$ of $\mathcal{K}$ with respect to x is said to be **equal** to a differential $G\,du$ with respect to another parameter u of $\mathcal{K}$ if $\frac{F}{G} = \frac{du}{dx}$. The proof that this definition of "equality" defines an equivalence relation depends on the chain rule of differentiation, which has a simple algebraic proof.

A differential **has a pole** at a point if it is equal to a differential $F\,dt$ where F has value ∞ at the point and t is a local parameter as defined in Essay 8.10. This is easily seen to be independent of the local parameter. It follows immediately that equal differentials have the same poles.

Finally, a differential is **holomorphic** if it has no poles.

Proposition *The differential $F\,dx$ is holomorphic if and only if* (1) *xF is zero at every point where $x = \infty$ and* (2) *$(x - a)F$ is zero at each point where x has the finite value a.*

Proof First suppose that x has the finite value a at a point, and take an expansion rule with $x = a + \tau^\delta$. Let t be a local parameter, so that $t = \tau + \cdots$ after multiplying by a suitable nonzero constant (omitted terms have higher degree in τ). If $\Phi(x, t) = 0$ is the irreducible algebraic relation between x and t, then $\Phi(a + \tau^\delta, \tau + \cdots) = 0$. Differentiating formally with respect to τ gives

$$\Phi_x(a + \tau^\delta, \tau + \cdots) \cdot \delta\tau^{\delta-1} + \Phi_t(a + \tau^\delta, \tau + \cdots) \cdot (1 + \cdots) = 0,$$

where subscripts indicate partial derivatives. Therefore, the expansion of $\frac{dx}{dt} = -\frac{\Phi_t}{\Phi_x}$ at the point is given by

$$\frac{dx}{dt} = -\frac{\Phi_t(a + \tau^\delta, \tau + \cdots)}{\Phi_x(a + \tau^\delta, \tau + \cdots)} = \frac{\delta\tau^{\delta-1}}{1 + \cdots} = \delta\tau^{\delta-1} + \cdots .$$

If the expansion of F is $c\tau^k + \cdots$, where c is nonzero and k is an integer, then

$$(x - a)F \text{ is zero at the point } \Leftrightarrow \delta + k > 0$$

since the expansion of $(x - a)F$ is $\tau^\delta(c\tau^k + \cdots) = c\tau^{\delta+k} + \cdots$. On the other hand, $F\,dx = F\frac{dx}{dt}\,dt$, and $F\frac{dx}{dt} = (c\tau^k + \cdots)(\delta\tau^{\delta-1} + \cdots) = c\delta\tau^{\delta+k-1} + \cdots$. Since t is a local parameter,

$$F\frac{dx}{dt}\,dt \text{ has no pole at the point } \Leftrightarrow F\frac{dx}{dt} \text{ is finite } \Leftrightarrow \delta + k - 1 \geq 0 \Leftrightarrow \delta + k > 0.$$

Therefore, $F\,dx$ has no poles where x is finite if and only if (2) holds.

Where $x = \infty$, $u = 1/x$ has the value zero. Then $F\,dx = F \cdot \frac{dx}{du} \cdot du = F \cdot \frac{-1}{u^2} \cdot du = -\frac{F}{u^2}\,du$, so that by (2), $-\frac{F}{u^2}\,du$ has no pole where $x = \infty$ if and only if

$$(u-0)\Big(-\frac{F}{u^2}\Big) = -\frac{F}{u} = -xF$$

is zero where $x = \infty$. Thus $F\,dx$ has no poles where $x = \infty$ if and only if (1) holds.

□

Later essays will consider differentials in several independent variables, say $a_1, a_2, \ldots, a_N$. In this setting, a **rational differential** has the form

$$\sum_{j=1}^{N} R_j(a_1, a_2, \ldots, a_N)\,da_j.$$

where $R_j(a_1, a_2, \ldots, a_N)$ is a rational function of $a_1, a_2, \ldots, a_N$ with coefficients in the field of constants $\mathcal{K}_0$. Such a form is **closed** if in addition

$$\frac{\partial R_j}{\partial a_i} = \frac{\partial R_i}{\partial a_j}$$

for all $i, j = 1, \ldots, N$.

Essay 9.5 The Construction of Holomorphic Differentials

Theorem *The holomorphic differentials of a curve field $\mathcal{K}$ can be constructed using rational arithmetic.*

Specifically, if $y_1, y_2, \ldots, y_n$ is a normal basis of $\mathcal{K}$ over $\mathcal{K}_0(x)$ for some parameter x of $\mathcal{K}$, where $\mathcal{K}_0$ is the field of constants of the curve field, then $F\,dx$ is holomorphic if and only if the coefficients $\xi_i(x)$ of the representation $F = \sum_{i=1}^{n} \xi_i(x) y_i$ of F with respect to the given normal basis have the property that the kth entry of the product matrix $S[\xi_i(x)]$, where $[\xi_i(x)]$ stands for the column matrix of height n whose ith entry is $\xi_i(x)$ and S is the $n \times n$ matrix whose entry in the ith row of the jth column is $\mathrm{tr}_x(y_i y_j)$, is a polynomial of degree at most $\mu_k - 2$, where μ_k is the order of y_k at $x = \infty$. (In particular, the kth entry must be zero when μ_k is 0 or 1.)

The matrix S is invertible, so this observation implies a formula for the most general holomorphic differential $F\,dx$.

Corollary 1 *The holomorphic differentials of $\mathcal{K}$ form a vector space over $\mathcal{K}_0$ of dimension $1 - n + \sum_{i=1}^{n} \mu_i$.*

Deduction The formula implied by the theorem is $[\xi_i(x)] = S^{-1}[\pi_i(x)]$, where $[\pi_i(x)]$ is a column matrix of polynomials in x with coefficients in $\mathcal{K}_0$ in which the polynomial in the ith row has degree at most $\mu_i - 2$. Therefore, the holomorphic differentials form a vector space over $\mathcal{K}_0$ whose dimension is the number of arbitrary constants in the column matrix $[\pi_i(x)]$. Because the number of arbitrary constants in the ith entry is $(\mu_i - 2) + 1$, except that it is zero for the one value of i for which $\mu_i = 0$, this number is the sum of the $n - 1$ numbers $\mu_i - 1$ in which $\mu_i \neq 0$, which is the number given in the statement of the corollary. □

Since the holomorphic differentials are intrinsic to $\mathcal{K}$, this corollary implies that the number $1 - n + \sum_{i=1}^{n} \mu_i$ is independent of the choice of both the parameter x and the normal basis of $\mathcal{K} \supset \mathcal{K}_0(x)$ that are used. It is the **genus** of $\mathcal{K}$, denoted by g.

Proof of the Theorem It will first be shown that *if $F\,dx$ is holomorphic then* $\mathrm{tr}_x(\theta F)$ *is a polynomial in x whenever θ is integral over x.*

Let $\mathcal{K}$ be the curve field determined by $\chi(x, y) = 0$, where $\chi(x, y)$ is a polynomial with coefficients in the field of constants $\mathcal{K}_0$ of $\mathcal{K}$ that is irreducible over $\mathcal{K}_0$ and monic in y. When the Newton diagram algorithm is adapted to the case in which the initial field of constants is $\mathcal{K}_0$ rather than $\mathbf{Q}$, one finds that for every algebraic number a there are $n = \deg_y \chi(x, y)$ points (counted with multiplicity) of $\mathcal{K}$ at which x has the value a, and they are determined by the infinite series solutions $\tilde{y}$ of $\chi(a + t, y) = 0$ in powers (nonnegative but possibly fractional) of t.

Let z be a quantity in $\mathcal{K}$ that is integral over x, say $z = \frac{p(x,y)}{q(x)}$, where numerator and denominator are polynomials in the indicated quantities with coefficients in $\mathcal{K}_0$. The expansion of z at the point that is described by $\tilde{y}$ is the infinite series in powers of $t = x - a$ described by the formula $\frac{p(a+t,\tilde{y})}{q(a+t)}$. Let $\tilde{z}$ denote this infinite series in (possibly fractional) powers of t with algebraic number coefficients. Because z is integral over x, $\tilde{z}$ contains no terms with negative coefficients, so the value of z at the point in question is the constant term of $\tilde{z}$.

If the field $\mathcal{K}_0(z, x)$ that adjoins z to $\mathcal{K}_0(x)$ is an extension of degree k of $\mathcal{K}_0(x)$, then, because z is integral over x, an equation of the form $z^k + \phi_1(x)z^{k-1} + \cdots + \phi_k(x) = 0$ holds in which the coefficients $\phi_j(x)$ are polynomials in x with coefficients in $\mathcal{K}_0$ and k divides n—specifically, $n = kl$, where l is the degree of the extension $\mathcal{K} \supset \mathcal{K}_0(z, x)$. The n expansions $\tilde{z}_i$, corresponding to the n expansions $\tilde{y}_i$ that determine the points where $x = a$, are roots of $(Z^k + \phi_1(a+t)Z^{k-1} + \cdots + \phi_k(a+t))^l = Z^n + l\phi_1(a+t)Z^{n-1} + \cdots$, which means that the sum[6] of the expansions $\tilde{z}_i$ in fractional powers of t determined by the algorithm is a polynomial in t with algebraic number coefficients, namely, $-l\phi_1(a + t)$. (The sum of the roots of a monic polynomial of degree n is minus the coefficient of the term of degree $n - 1$.) Therefore, the sum of the values of z at points where $x = a$ is the constant term of $-l\phi_1(a + t)$, which is the value at a of the trace of z relative to the extension $\mathcal{K} \supset \mathcal{K}_0(x)$.[7]

When θ is integral over x and $F\,dx$ is a holomorphic differential of $\mathcal{K}$, θ is finite and $(x - a)F$ is zero at every point where x has a finite value a, so $\theta \cdot (x - a)F$ is zero at all points where $x = a$, which implies, on the one hand, that the sum of these values over the points where $x = a$ is zero, and, on the other hand, that

[6] This sum is a quantity in $\mathbf{A}\langle\tau\rangle$, $\tau = \sqrt[\delta]{x - a}$ for suitable δ (see Essay 8.8), of which any number of initial terms can be computed exactly and which then, because it satisfies an explicit algebraic relation, can be determined as an infinite series.

[7] By definition, $\mathrm{tr}_x(z)$ is the trace of the matrix that represents multiplication by z relative to a basis of the extension $\mathcal{K} \supset \mathcal{K}_0(x)$. As is easily shown, $\mathrm{tr}_x(z)$ is $-l\phi_1(x)$.

this sum is the value of $\mathrm{tr}_x(\theta \cdot (x-a)F)$ at $x = a$. As follows from the definition, $\mathrm{tr}_x(\theta \cdot (x-a)F) = (x-a) \cdot \mathrm{tr}_x(\theta F)$. Thus, $\mathrm{tr}_x(\theta F)$ is a rational function $\rho(x)$ of x with the property that $(x-a)\rho(x)$ is zero at all points where x has a finite value a. Such a rational function is a polynomial, as was to be shown.

It will next be shown that S is invertible. A nonzero column vector of rational functions $[u_j(x)]$ gives the nonzero element $z = \sum_{j=1}^{n} u_j(x)y_j$ of the curve field $\mathcal{K}$. If $z^{-1} = \sum_{i=1}^{n} v_i(x)y_i$ for rational functions $v_i(x)$, then

$$\begin{aligned} n = \mathrm{tr}_x(1) = \mathrm{tr}_x(z^{-1} \cdot z) = \mathrm{tr}_x\Big(\sum_{ij} v_i(x)u_j(x)y_iy_j\Big) = \sum_{ij} v_i(x)u_j(x)\,\mathrm{tr}_x(y_iy_j) \\ = [v_i(x)]S[u_j(x)], \end{aligned}$$

where $[v_i(x)]$ is the row vector of the $v_i(x)$'s. Therefore, $S[u_j(x)]$ is nonzero whenever $[u_j(x)]$ is nonzero. Such a matrix is invertible, as was to be shown.

Now consider $S[\xi_i(x)]$ when $F = \sum_{i=1}^{n} \xi_i(x)y_i$. Since the ith entry of $S[\xi_i(x)]$ is $\sum_{j=1}^{n} \mathrm{tr}_x(y_iy_j) \cdot \xi_j(x) = \mathrm{tr}_x(y_i \cdot \sum_{j=1}^{n} \xi_j(x)y_j) = \mathrm{tr}_x(y_iF)$, it follows that this entry is a polynomial in x when $F\,dx$ is holomorphic. But if $F\,dx$ is holomorphic, then $\mathrm{tr}_x(\theta \cdot (-x^2F))$ must be a polynomial in $\frac{1}{x}$ whenever θ is integral over $\frac{1}{x}$. Since $\frac{y_i}{x^{\mu_i}}$ is integral over $\frac{1}{x}$, it follows that $\mathrm{tr}_x(y_ix^{2-\mu_i}F) = \frac{1}{x^{\mu_i-2}} \cdot \mathrm{tr}_x(y_iF)$ is a polynomial in $\frac{1}{x}$. Therefore, the polynomial $\mathrm{tr}_x(y_iF)$ in the ith row of $S[\xi_i(x)]$ has degree at most $\mu_i - 2$. Thus, if $F\,dx$ is holomorphic then $S[\xi_i(x)]$ has the stated form.

It remains only to show that if $F\,dx$ is not holomorphic then $S[\xi_i(x)]$ does not have the stated form. If $F\,dx$ is not holomorphic, then it must have a pole, and, by changing x to $\frac{1}{x}$ if necessary, one can assume without loss of generality that it has a pole at a point where x is finite—that is, $(x-a)F$ is not zero at some point where $x = a$. Since the ith entry of $S[\xi_i(x)]$ is $\mathrm{tr}_x(y_iF)$, it will suffice to show that if $(x-a)F$ is not zero at some point where $x = a$, then there is a θ that is integral over x for which $\mathrm{tr}_x(\theta F)$ is not a polynomial.

If $(x-a)F$ is not zero at a point where $x = a$, then the expansion of F in powers of $t = x - a$ at the point must contain a term with a (possibly fractional) exponent less than or equal to -1. Consider an integral quantity θ with the property that at each point where $x = a$, the expansion of θF in powers of t is $\theta F = t^m + \cdots$, where m is an integer that is negative for at least one point. As was seen above, the expansion of $\mathrm{tr}_x(\theta F)$ is the sum of the n expansions of θF at the points where $x = a$.

To describe these expansions, suppose that $x - a$ has a zero of multiplicity δ_i at a point where $x = a$. Then there are $\delta_i - 1$ other expansion rules at the point, namely, one that replaces $\tau_i = \sqrt[\delta_i]{x-a}$ with $\alpha\tau_i$ for each δ_ith root of unity α other than 1. In this way, the n expansion rules determined by infinite series solutions $\tilde{y}$ of $\chi(a+t, y) = 0$ are partitioned by the points to which they correspond—say $n = \delta_1 + \delta_2 + \cdots + \delta_l$, where there are l distinct points, at the ith one of which there are δ_i distinct expansion rules.

In this notation, the expansion of θF at the ith point is $\theta F = t^{m_i} + \cdots$. Because m_i is an integer, all δ_i expansion rules at the ith point begin this way. If m is the minimum of the m_i, it follows that

$$\mathrm{tr}_x(\theta F) = \Big(\sum_i \delta_i\Big) t^m + \cdots,$$

where the sum is over all i such that $m_i = m$. Since m is negative, $\mathrm{tr}_x(\theta F)$ is *not* a polynomial in x, as was to be shown.

It remains to construct θ. Let an expansion rule with $x - a = t = \tau_i^{\delta_i}$ be chosen for ith point, and for this rule, suppose that $F = c_i\tau_i^{k_i} + \cdots$ with $c_i \neq 0$. Pick an integer ℓ_i between 0 and $\delta_i - 1$ such that $\ell_i + k_i$ is an integer multiple of δ_i, call it $\delta_i m_i$. If θ can be chosen so that at the ith point it has the expansion $\theta = \frac{1}{c_i}\tau_i^{\ell_i} + \cdots$, then $\theta F = \tau_i^{\ell_i + k_i} + \cdots = \tau_i^{\delta_i m_i} + \cdots = t^{m_i} + \cdots$, as desired. Therefore, the proof of the theorem is completed by:

Lemma *Let expansion rules for the l points where $x = a$ be chosen at above, where $x - a = \tau_i^{\delta_i}$ at the ith point and $n = \delta_1 + \cdots + \delta_l$. A quantity θ that is integral over x can be constructed for which the n coefficients b_j that occur in the first δ_i terms $b_1 + b_2\tau_i + b_2\tau_i^2 + \cdots + b_{\delta_i}\tau_i^{\delta_i - 1}$ in each of the l chosen expansions of θ assume arbitrarily chosen values in $\mathcal{K}_0$ in all l cases.*

Proof Consider the mapping $\mathcal{L}$ that sends quantities θ of $\mathcal{K}$ that are integral over x to the n coefficients b_j in their l expansions. What is to be shown is that $\mathcal{L}$ (a linear mapping of vector spaces over $\mathcal{K}_0$) is onto.

Since the expansion of $\frac{1}{x-a}$ has no terms with negative coefficients at points where $x \neq a$, $\frac{\theta}{x-a}$ is integral over x (its expansions at points where x is finite contain no terms with negative exponents) if and only if the expansions of $\frac{\theta}{x-a}$ at points where $x = a$ contain no terms with negative exponents. Since, for each i, the expansion of $x - a$ in powers of τ_i at the corresponding point where $x = a$ is $\tau_i^{\delta_i}$, $\frac{\theta}{x-a}$ is integral over x if and only θ is in the kernel of $\mathcal{L}$.

By the definition of a normal basis, $\frac{\theta}{x-a}$ is integral over x if and only if each of the coefficients $\theta_i(x)$ in the expansion $\theta = \sum_{i=1}^n \theta_i(x)y_i$ is divisible by $x - a$ as a polynomial with coefficients in $\mathcal{K}_0$. Thus θ is in the kernel of $\mathcal{L}$ if and only if each $\theta_i(x)$ in $\theta = \sum_{i=1}^n \theta_i(x)y_i$ is divisible by $x - a$.

Consider the n-dimensional subspace consisting of $\theta = \sum_{i=1}^n c_i y_i$ for c_i in $\mathcal{K}_0$. For such a θ, the previous paragraph implies that θ is in the kernel of $\mathcal{L}$ only when $\theta = 0$. In other words, $\mathcal{L}$ is one-to-one on this subspace, and therefore, because domain and range have the same dimension n, onto, as was to be shown. □

The lemma just proved implies the existence of local parameters:

Corollary *Construct a local parameter at a given a point of a curve field $\mathcal{K}$.*

Deduction For any parameter x in $\mathcal{K}$, x or $\frac{1}{x}$ is finite at the point. Without loss of generality assume that x has a finite value a at the point. Then the point is given by one of the l chosen expansion rules in the statement of the lemma. If $x - a$ has multiplicity $\delta_i = 1$ at the point, then $x - a$ is a local parameter. If $\delta_i > 1$, the integral quantity θ in the lemma has an expansion that begins $b_1 + b_2\tau_i + b_2\tau_i^2 + \cdots + b_{\delta_i}\tau_i^{\delta_i - 1}$, and the lemma shows that θ can be chosen so that $b_1 = 0$ and $b_2 = 1$. Therefore, θ vanishes at the point with multiplicity one and is the desired local parameter. □

Here is another useful corollary of the main theorem of the essay:

Corollary 2 *If $F\,dx$ is holomorphic then* $\mathrm{tr}_x(F) = 0$.

Deduction Suppose that $F\,dx = -u^{-2}F\,du$ is holomorphic. Since 1 is integral over both x and u, $\mathrm{tr}_x(F)$ must be a polynomial in x, and $\mathrm{tr}_u(u^{-2}F) = x^2\mathrm{tr}_x(F)$ must be a polynomial in $u = \frac{1}{x}$, which implies $\mathrm{tr}_x(F) = 0$ because zero is the only polynomial in x that becomes a polynomial in $\frac{1}{x}$ when it is multiplied by x^2. □

Since the property of being holomorphic does not depend on the parameter x that is used in the presentation of the field, the corollary is true for every parameter x of the curve field in question.

Example 2 When $\mathcal{K}$ is defined by $y^2 = x^3 + 1$, a normal basis is given by $y_1 = 1$, $y_2 = y$, from which

$$S = \begin{bmatrix} 2 & 0 \\ 0 & 2(x^3+1) \end{bmatrix}.$$

Since the order of y at $x = \infty$ is 2, the genus of this curve field is $1 - 2 + (0 + 2) = 1$ and the most general holomorphic differential has the form $(\xi_1(x) \cdot 1 + \xi_2(x) \cdot y)\,dx$ where

$$\begin{bmatrix} 2 & 0 \\ 0 & 2(x^3+1) \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \pi(x) \end{bmatrix} = \begin{bmatrix} \xi_1(x) \\ \xi_2(x) \end{bmatrix}$$

holds for a polynomial $\pi(x)$ of degree zero. Therefore, the most general holomorphic differential is a rational number times $\frac{y}{2(x^3+1)} \cdot dx = \frac{y\,dx}{2y^2} = \frac{dx}{2y}$. That is, $\frac{dx}{y}$ is a basis of the vector space of holomorphic functions over the field of constants $\mathbf{Q}$ of this curve field. (See also Example 6 in Essay 4.5 and Example 8 in Essay 4.6.)

Essay 9.6 Parametric Points Constructed Using a Normal Basis

In his Paris Memoir [2, p. 147], Abel defines

$$\theta(x, y) = q_0 + q_1 y + q_2 y^2 + \cdots + q_{n-1} y^{n-1},$$

where q_i is a polynomial in x, and “a certain number of coefficients of various powers of x in these functions will be assumed to be undetermined; we designate them by a, a', a'', etc.” The solutions $(x_1, y_1), \ldots, (x_\mu, y_\mu)$ of the simultaneous equations $\chi(x, y) = 0$ and $\theta(x, y) = 0$ are algebraic functions of $a, a', a'', \ldots$. These are the variables x_i and y_i that appear in equation (1) in Essay 9.1.

Let $\mathcal{K}$ be the curve field determined by $\chi(x, y) = 0$. The polynomial $\theta(x, y)$ defined by Abel uses the basis $1, y, \ldots, y^{n-1}$ of $\mathcal{K}$ over $\mathcal{K}_0(x)$. A better choice is a normal basis $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ of $\mathcal{K}$ over $\mathcal{K}_0(x)$. Let μ_i be the order of $\mathbf{y}_i$ at $x = \infty$. For any integer ν that is at least μ_i for each i, set

$$\theta_\nu(x, y) = \sum_{i=1}^{n} \theta_i(x)\mathbf{y}_i,$$

where $\theta_i(x)$ is the polynomial in x of degree $\nu - \mu_i$ whose $\nu - \mu_i + 1$ coefficients are indeterminates. Then $\theta_\nu(x, y)$ contains a total of

$$\sum_{i=1}^{n}(\nu - \mu_i + 1) = n\nu + n - \sum_{i=1}^{n} \mu_i = n\nu - g + 1$$

indeterminate coefficients, where $g = 1 - n + \sum_{i=1}^{n} \mu_i$ is the genus of $\mathcal{K}$ as defined in the previous essay.[8]

Heuristically, this choice of $\theta_\nu(x, y)$ represents—when numerical values are given to the $n\nu - g + 1$ indeterminates in $\theta_\nu(x, y)$—a quantity in $\mathcal{K}$ that is integral over x and has order at most ν at $x = \infty$. In the generic case, it has $n\nu$ zeros, as will be shown below. Thus, there are $n\nu$ zeros, which are the solutions of the simultaneous equations $\chi(x, y) = 0$ and $\theta_\nu(x, y) = 0$. They are algebraic functions of the $n\nu - g + 1$ indeterminates, which can be regarded as parameters. From a geometric point of view, these $n\nu$ parametric points (x_i, y_i) in which $\chi(x, y) = 0$ and $\theta_\nu(x, y) = 0$ intersect vary as the $n\nu - g + 1$ coefficients of $\theta_\nu(x, y)$ vary.

The solutions (x_i, y_i) depend on just $n\nu - g$ parameters, not $n\nu - g + 1$. This conclusion follows from the observation that multiplying $\theta_\nu(x, y)$ by a nonzero constant does not change its zeros. This redundancy in the coefficients of $\theta_\nu(x, y)$ can be eliminated by setting one of the coefficients equal to one. As noted at the end of Essay 9.3, just one element, say $\mathbf{y}_1$, of the normal basis $\mathbf{y}_1, \ldots, \mathbf{y}_n$ is a constant, which can be assumed to be 1. Then the degree of $\theta_1(x)$ is ν. Setting the coefficient of x^ν in $\theta_1(x)$ to be 1 gives a rational function $\Theta_\nu(x, y)$ that depends on just $n\nu - g$ parameters.

For the moment, regard $\Theta_\nu(x, y)$ as a quantity in the curve field $\hat{\mathcal{K}}$ defined by $\chi(x, y) = 0$ over $\hat{\mathcal{K}}_0(x)$, where $\hat{\mathcal{K}}_0(x)$ is the field $\mathcal{K}_0(a_1, a_2, \ldots, a_{n\nu-g})$ of rational functions in the $n\nu - g$ indeterminate coefficients $a_1, a_2, \ldots, a_{n\nu-g}$ of $\Theta_\nu(x, y)$ with coefficients in $\mathcal{K}_0$. Then the solutions (x_i, y_i) of the simultaneous equations

$$\chi(x, y) = 0 \ \text{ and } \ \Theta_\nu(x, y) = 0 \tag{1}$$

are then the zeros of $\Theta_\nu(x, y) = 0$ as an element of this curve field, hence finite in number. Furthermore, x_i and y_i are quantities in an algebraic extension of $\hat{\mathcal{K}}_0$—they are "algebraic functions" of $a_1, a_2, \ldots, a_{n\nu-g}$ that can be found by algebraic methods.

In what follows, ν is an integer such that $\nu \geq \mu_i$ for all i and also $\nu \geq \mu$, where μ is the order of y at $x = \infty$.

Lemma 1 *The equations (1) have $n\nu$ solutions (x_i, y_i), all of multiplicity one. Furthermore, the x-coordinates x_i are distinct and transcendental over $\mathcal{K}_0$.*

[8] $\theta_\nu(x, y)$ need not be a polynomial in x and y since, as shown by Example 1 in Essay 9.2, a normal basis might have denominators when expressed in terms of x and y.

Proof As just noted, the solutions of (1) are the zeros of $\Theta_\nu(x,y)$ when regarded as an element of the curve field $\hat{\mathcal{K}}$. Proposition 1 of Essay 8.10 implies that, counted with multiplicity, the number of zeros equals the number poles. However, $\Theta_\nu(x,y)$ is integral over x and hence has no poles at finite values of x. For poles where $x = \infty$, note that

$$\begin{aligned}\frac{\Theta_\nu(x,y)}{x^\nu} &= \frac{\theta_1(x) + \theta_2(x)\mathbf{y}_2(x,y) + \cdots + \theta_n(x)\mathbf{y}_n(x,y)}{x^\nu}\\ &= 1 + \frac{a_1}{x} + \cdots + \frac{a_\nu}{x^\nu} + \frac{\theta_2(x)}{x^{\nu-\mu_2}}\frac{\mathbf{y}_2}{x^{\mu_2}} + \cdots + \frac{\theta_n(x)}{x^{\nu-\mu_n}}\frac{\mathbf{y}_n}{x^{\mu_n}},\end{aligned}$$

where μ_j is the order of $\mathbf{y}_j$ at $x = \infty$ and $\theta_j(x)$ has degree $\nu - \mu_j$ with indeterminate coefficients. Because of these coefficients, nothing can cancel the 1 at the beginning of the right side of the second line. Thus $\Theta_\nu(x,y)/x^\nu$ has a finite nonzero value at any point where $x = \infty$, so that $\Theta_\nu(x,y)$ has a pole of order ν at such a point. Since there are n points (counted with multiplicity) where $x = \infty$ by Proposition 1 of Essay 8.10, it follows that $\Theta_\nu(x,y)$ has $n\nu$ poles. Consequently, $\Theta_\nu(x,y)$ also has $n\nu$ zeros, so that the system (1) has $n\nu$ solutions, counting multiplicities.

The number of actual solutions cannot increase when the indeterminate coefficients of $\Theta_\nu(x,y)$ are allowed to take values in $\hat{\mathcal{K}}_0$. With this in mind, consider the case where the coefficients of $\theta_2(x),\ldots,\theta_n(x)$ are set to zero and the coefficients of $\theta_1(x)$ remain indeterminate (except for the leading term 1). Then (1) becomes

$$\chi(x,y) = 0 \ \text{ and } \ \theta_1(x) = 0.$$

The second equation has ν solutions x_i, $i = 1,\ldots,\nu$, all of which are transcendental over $\mathcal{K}_0$. Then $\chi(x_i,y)$ has n distinct solutions for each i, so that the above system has exactly $n\nu$ solutions. Hence (1) must have at least $n\nu$ distinct solutions. Since it has at most this many, there must be exactly $n\nu$ solutions, all of multiplicity 1.

Consider the x-coordinates of the solutions. The number of distinct x-coordinates cannot increase when the coefficients of $\Theta_\nu(x,y)$ are allowed to take values in $\hat{\mathcal{K}}_0$. Since y is integral over x, it can be expressed as a sum $y = \phi_1(x)\mathbf{y}_1 + \cdots + \phi_n(x)\mathbf{y}_n$ with $\deg \phi_i(x) \le \mu - \mu_i$, where μ is the order of y at $x = \infty$. Let $\theta_1(x) = x^\nu + a_1x^{\nu-1} + \cdots + a_\nu$ and assign values to $a_{\nu+1},\ldots,a_{n\nu-g}$ as follows: If a_i is a coefficient of θ_{j_i}, then replace a_i with $a_{\nu+1}$ times the corresponding coefficient of ϕ_{j_i}. Since $\nu \ge \mu$ and $\mathbf{y}_1 = 1$, the system (1) transforms into

$$\chi(x,y) = 0 \ \text{ and } \ \theta_1(x) + a_{\nu+1}\Big(\sum_{i=2}^{n}\phi_i(x)\mathbf{y}_i\Big) = \theta_1(x) + a_{\nu+1}\big(y - \phi_1(x)\big) = 0.$$

This system has $n\nu$ solutions of multiplicity 1 since setting $a_{\nu+1} = 0$ gives the system of the previous paragraph. But any solution of the above system satisfies

$$y = \phi_1(x) - \frac{1}{a_{\nu+1}}\theta_1(x),$$

so that the y-coordinate is uniquely determined by the x-coordinate. Therefore, the $n\nu$ distinct solutions of the transformed system have $n\nu$ distinct x-coordinates. It follows that solutions of the original system (1) also have distinct x-coordinates.

It remains to show that the x_i are transcendental over $\mathcal{K}_0$. Suppose there is a solution (x_i, y_i) with x_i algebraic over $\mathcal{K}_0$. Then the same is true for y_i since $\chi(x, y)$ has coefficients in $\mathcal{K}_0$. In the normal basis $\mathbf{y}_1 = 1, \mathbf{y}_2, \dots, \mathbf{y}_n$, each $\mathbf{y}_j$ is integral over x and hence has a finite value at (x_i, y_i). Therefore,

$$0 = \Theta_\nu(x_i, y_i) = \theta_1(x_i) + \theta_2(x_i)\mathbf{y}_2(x_i, y_i) + \cdots + \theta_n(x_i)\mathbf{y}_n(x_i, y_i)$$

is a nontrivial linear relation in $a_1, a_2, \dots, a_{n\nu-g}$ whose coefficients are algebraic over $\mathcal{K}_0$, contrary to the algebraic independence of the a_i. □

The solutions (x_i, y_i) can be constructed as follows. The first step is the algebraic elimination of y from the simultaneous equations (1) to express x as an algebraic function of the indeterminate coefficients of $\Theta_\nu(x, y)$, and then to express y as a *rational* function of x whose coefficients involve the indeterminate coefficients of $\Theta_\nu(x, y)$ to make both coordinates of the solutions (x_i, y_i) algebraic functions of those indeterminate coefficients.

The elimination of y from the simultaneous equations (1) can be accomplished in the following way. Let $q(x)$ be a common denominator of the normal basis elements $\mathbf{y}_i$ in the sense that $q(x) \cdot \mathbf{y}_i$ can be expressed as a polynomial in x and y with coefficients in $\mathcal{K}_0$ for all i. Then $\Theta_\nu(x, y)$ has the form

$$\frac{p(x, y, a_1, a_2, \dots, a_{n\nu-g})}{q(x)},$$

where numerator and denominator are polynomials in the indicated indeterminates with coefficients in $\mathcal{K}_0$. Moreover, the degree of the numerator in y can be taken to be less than n, because otherwise $\chi(x, y) = 0$ could be used to reduce that degree. Finally, common factors can be canceled in numerator and denominator to express $\Theta_\nu(x, y)$ as a quotient of relatively prime polynomials with coefficients in $\mathcal{K}_0$.

The main step in the solution is to regard $\chi(x, y)$ and $p(x, y, a_1, a_2, \dots, a_{n\nu-g})$ as polynomials in y with coefficients in the field $\hat{\mathcal{K}}_0(x)$, where $\hat{\mathcal{K}}_0$ is the field $\mathcal{K}_0(a_1, a_2, \dots, a_{n\nu-g})$, and to find their greatest common divisor using the Euclidean algorithm for polynomials to write that common divisor in the form $A(y)\chi(x, y) + B(y)p(x, y, a_1, a_2, \dots, a_{n\nu-g}) = C(y)$ where $A(y)$, $B(y)$, and $C(y)$ are polynomials in y with coefficients in $\hat{\mathcal{K}}_0(x)$. Since $\chi(x, y)$ is irreducible and $p(x, y, a_1, a_2, \dots, a_{n\nu-g})$ has lower degree in y than $\chi(x, y)$ does, the common divisor C must have degree 0 in y, so that

$$A(y)\chi(x, y) + B(y)p(x, y, a_1, a_2, \dots, a_{n\nu-g}) = C, \tag{2}$$

where C is in $\hat{\mathcal{K}}_0(x)$.

Recall that $\deg_y \chi(x, y) = n$, and set $\deg_y p(x, y, a_1, \dots, a_{n\nu-g}) = m$. Then one can assume without loss of generality that $\deg A(y) < m$ and $\deg B(y) < n$ in (2), as can be seen as follows. Division defines polynomials $q_1(y)$, $q_2(y)$, $r_1(y)$, and $r_2(y)$ for which $A(y) = q_1(y)p(x, y, a_1, \ldots, a_{n\nu-g}) + r_1(y)$ and $B(y) = q_2(y)\chi(x, y) + r_2(y)$, where $\deg r_1(y) < m$ and $\deg r_2(y) < n$. Then

$$\begin{aligned}(q_1(y) + q_2(y))\chi(x, y)p(x, y, a_1, \ldots, a_{n\nu-g})&\\ +r_1(y)\chi(x, y) + r_2(y)p(x, y, a_1, \ldots, a_{n\nu-g}) &= C.\end{aligned}$$

In this equation, $\chi(x, y)p(x, y, a_1, \ldots, a_{n\nu-g})$ has degree $n+m$ while the sum involving $r_2(y)$ and $r_1(y)$ has degree strictly less than $n+m$. Thus $q_1(y)+q_2(y) = 0$ because C contains no terms in y. Setting $A(y) = r_1(y)$ and $B(y) = r_2(y)$, and $\deg A(y) < m$, $\deg B(y) < n$ follows. In particular, the number of unknown coefficients of $A(y)$ and $B(y)$ is $m + n$.

Let C be written as a quotient of relatively prime polynomials in $x, a_1, a_2, \ldots, a_{n\nu-g}$ with coefficients in $\mathcal{K}_0$ and let $\varrho(x, a_1, a_2, \ldots, a_{n\nu-g})$ be the numerator of this quotient (determined up to multiplication by a nonzero quantity in $\mathcal{K}_0$). One obtains

$$A(y)\chi(x, y) + B(y)p(x, y, a_1, a_2, \ldots, a_{n\nu-g}) = \varrho(x, a_1, a_2, \ldots, a_{n\nu-g}).$$

where $A(y)$ and $B(y)$ have coefficients in $\mathcal{K}_0[x, a_1, a_2, \ldots, a_{n\nu-g}]$. To highlight the dependence on x and y, this equation will be written as

$$A(x, y)\chi(x, y) + B(x, y)p(x, y) = \varrho(x),$$

where $\chi(x, y)$ has coefficients in $\mathcal{K}_0$, and the other polynomials have coefficients in $\mathcal{K}_0[a_1, a_2, \ldots, a_{n\nu-g}]$.

Roots of $q(x)$ create extraneous roots of $\varrho(x)$. Given a root $q(x_0) = 0$, construct y_0 such that $\chi(x_0, y_0) = 0$ in some splitting field. Note that $\Theta_\nu(x_0, y_0)$ is finite since $\Theta_\nu(x, y)$ is integral over x. Therefore, $p(x_0, y_0) = 0$ since $\Theta_\nu(x, y) = \frac{p(x,y)}{q(x)}$ and $q(x_0) = 0$. It follows that $\varrho(x_0) = 0$. Let $h(x)$ be the greatest common divisor of $\varrho(x)$ and $q(x)^{\deg \varrho(x)}$. Then

$$\varrho(x) = h(x)\varrho_1(x)$$

where $h(x)$ has coefficients in $\mathcal{K}_0$ and $\varrho_1(x)$ has coefficients in $\mathcal{K}_0[a_1, a_2, \ldots, a_{n\nu-g}]$. The exponent $\deg \varrho(x)$ in the definition of $h(x)$ guarantees that $h(x)$ and $\varrho_1(x)$ are relatively prime. In this notation,

$$A(x, y)\chi(x, y) + B(x, y)p(x, y) = h(x)\varrho_1(x). \tag{3}$$

One can assume that $A(x, y)$, $B(x, y)$, and $h(x)\varrho_1(x)$ are relatively prime.

If (x_i, y_i) is a zero of $\Theta_\nu(x, y)$ as a quantity in the curve field defined by $\chi(x, y) = 0$ over $\hat{\mathcal{K}}_0$, then substitution of (x_i, y_i) in (3) makes the left side zero, so x_i must be a root of $h(x)\varrho_1(x)$ and thus a root of $\varrho_1(x)$ since x_i is transcendental over $\mathcal{K}_0$ by Lemma 1 (roots of $h(x)$ are roots of $q(x)$, which are algebraic over $\mathcal{K}_0$). The splitting field of $\varrho_1(x) = \varrho_1(x, a_1, a_2, \ldots, a_{n\nu-g})$ as a polynomial in x with coefficients in $\hat{\mathcal{K}}_0$ is therefore a field that contains all x-coordinates x_i of zeros of $\Theta_\nu(x, y)$ on $\chi(x, y) = 0$.

Lemma 2 *There is a rational function $Y(x)$ of x with coefficients in $\hat{\mathcal{K}}_0$ with the property that if x_i is a root of $\varrho_1(x)$ in a splitting field of this polynomial in x with coefficients in $\hat{\mathcal{K}}_0$, then $(x_i, Y(x_i))$ is a solution of $\chi(x, y) = \Theta_\nu(x, y) = 0$. Furthermore, all solutions arise in this way.*

Proof Let x_i be a root of $\varrho_1(x)$. Note that $q(x_i) \neq 0$ by the construction of $\varrho_1(x)$. Substitution of $x = x_i$ in (3) shows that

$$A(x_i, y)\chi(x_i, y) + B(x_i, y)p(x_i, y) = 0.$$

Since $\chi(x_i, y)$ is monic of degree n, $B(x_i, y) = 0$ implies $A(x_i, y) = 0$, contrary to $A(x, y)$, $B(x, y)$, and $h(x)\varrho_1(x)$ being relatively prime. Thus $B(x_i, y)$ is a nonzero polynomial of degree at most $n - 1$. It therefore cannot be divisible by $\chi(x_i, y)$, which proves that $\chi(x_i, y)$ and $p(x_i, y)$ have a common factor of positive degree. Any root y_i in a splitting field of this common factor gives a solution (x_i, y_i) of $\chi(x, y) = p(x, y) = 0$. But $p(x, y) = q(x)\Theta_\nu(x, y)$ and $q(x_i) \neq 0$, so that (x_i, y_i) is a solution of $\chi(x, y) = \Theta_\nu(x, y) = 0$. The comments made before Lemma 2 show that this gives all solutions of $\chi(x, y) = \Theta_\nu(x, y) = 0$.

It remains to construct a rational function Y such that $y_i = Y(x_i)$ for all solutions. Equation (2) represents a solution of the problem

$$\text{“}A(y)\chi(x, y) + B(y)p(x, y, a_1, \ldots, a_{n\nu-g}) \text{ has degree zero in } y,\text{”}$$

which can be seen as the solution of the system of $m + n - 1$ homogeneous linear equations in the $m + n$ coefficients of $A(y)$ and $B(y)$, because all of the $m + n$ terms of $A(y)\chi(x, y) + B(y)p(x, y, a_1, \ldots, a_{n\nu-g})$, except the constant term, must be zero.[9] Since $\chi(x, y)$ and $p(x, y, a_1, \ldots, a_{n\nu-g})$ determine their greatest common divisor as polynomials in y up to multiplication by a nonzero quantity in $\hat{\mathcal{K}}_0(x)$, the solution space of this $(m + n - 1) \times (m + n)$ system of homogeneous linear equations is 1-dimensional, which means that the matrix of coefficients has rank $m + n - 1$, the largest rank that a matrix of this size can have. Therefore, the matrix that results when the bottom row is omitted has the full rank $m + n - 2$, which means that the solution space of the problem

$$\tilde{A}(y)\chi(x, y) + \tilde{B}(y)p(x, y, a_1, \ldots, a_{n\nu-g}) = \alpha(x)y + \beta(x)$$

has dimension 2. In particular, there is a solution[10] of this problem in which $\alpha(x) \neq 0$ with $\deg \tilde{A}(y) < m$ and $\deg \tilde{B}(y) < n$. Similar to (3), write the above equation as

$$\tilde{A}(x, y)\chi(x, y) + \tilde{B}(x, y)p(x, y) = \alpha(x)y + \beta(x), \tag{4}$$

[9] If, instead, one considers (2) to represent the $(m + n) \times (m + n)$ inhomogeneous system of linear equations in which the right side is given, the determinant of the matrix of coefficients is the **resultant** of $\chi(x, y)$ and $p(x, y, a_1, \ldots, a_{n\nu-g})$ with respect to y as it is normally defined.

[10] An *explicit* solution can of course be expressed in terms of the entries of the matrix of coefficients using Cramer's rule.

where $\alpha(x) \neq 0$ and $\tilde{A}(x, y)$, $\tilde{B}(x, y)$, $\alpha(x)$, and $\beta(x)$ are relatively prime polynomials with $\deg_y \tilde{A}(x, y) < m$ and $\deg_y \tilde{B}(x, y) < n$.

Substitution in (3) and (4) of a solution (x_i, y_i) of $\chi(x, y) = \Theta_\nu(x, y) = 0$ shows that $\varrho_1(x_i) = \alpha(x_i)y_i + \beta(x_i) = 0$. If $\alpha(x_i) = 0$, then $\beta(x_i) = 0$, so that (4) becomes

$$\tilde{A}(x_i, y)\chi(x_i, y) + \tilde{B}(x_i, y)p(x_i, y) = 0. \tag{5}$$

By Lemma 1, only one solution has x-coordinate x_i. Therefore, the monic greatest common divisor of $\chi(x_i, y)$ and $p(x_i, y)$ is $y - y_i$. Then (5) implies that there is a polynomial $C(y)$ such that

$$\tilde{A}(x_i, y) = C(y)\frac{p(x_i, y)}{y - y_i} \quad \text{and} \quad \tilde{B}(x_i, y) = -C(y)\frac{\chi(x_i, y)}{y - y_i}.$$

However, $\deg \chi(x_i, y) = n$ since $\chi(x, y)$ is monic in y. Thus $\deg \tilde{B}(x_i, y) = \deg C(y) + n - 1$. But $\deg \tilde{B}(x_i, y) \leq \deg_y \tilde{B}(x, y) \leq n - 1$. Therefore, $\deg C(y) = 0$, so that $C(y)$ is a constant, call it c. Therefore,

$$\tilde{A}(x_i, y) = c\,\frac{p(x_i, y)}{y - y_i} \quad \text{and} \quad \tilde{B}(x_i, y) = -c\,\frac{\chi(x_i, y)}{y - y_i}. \tag{6}$$

If $c = 0$, then $\tilde{A}(x_i, y) = \tilde{B}(x_i, y) = 0$, so $\tilde{A}(x, y)$ and $\tilde{B}(x, y)$ in (4) would be divisible by $x - x_i$. The same is true for $\alpha(x)$ and $\beta(x)$, yet $\tilde{A}(x, y)$, $\tilde{B}(x, y)$, $\alpha(x)$, and $\beta(x)$ are relatively prime. Thus, $c \neq 0$ in (6)

Since $\varrho_1(x_i) = 0$, equation (3) implies that

$$A(x_i, y)\chi(x_i, y) + B(x_i, y)p(x_i, y) = 0.$$

Comparing this to (5), it follows that $A(x_i, y)$ and $B(x_i, y)$ are given by formulas similar to (6), with a possibly different constant c', and arguing as above shows that $c' \neq 0$. It follows easily that

$$cA(x_i, y) = c'\tilde{A}(x_i, y) \quad \text{and} \quad cB(x_i, y) = c'\tilde{B}(x_i, y).$$

Divide $A(x, y)$, $B(x, y)$, $\tilde{A}(x, y)$, $\tilde{B}(x, y)$ by $x - x_i$ with respective remainders $A(x_i, y)$, $B(x_i, y)$, $\tilde{A}(x_i, y)$, $\tilde{B}(x_i, y)$, and then multiply (4) by c' and (3) by c and subtract. The remainder terms cancel, leaving a solution of (4) (with a different $\beta(x)$) whose coefficients are not relatively prime. Dividing out a common factor gives an equation

$$\hat{A}(x, y)\chi(x, y) + \hat{B}(x, y)p(x, y) = \hat{\alpha}(x)y + \hat{\beta}(x)$$

where $\deg \hat{\alpha}(x) < \deg \alpha(x)$. Therefore, $\alpha(x_i) = 0$ leads to a solution of (4) where $\alpha(x)$ has strictly smaller degree. By infinite descent, there is a solution with $\alpha(x_i) \neq 0$. Repeating this for the other indices gives a solution of (4) where $\alpha(x_i) \neq 0$ for all i.

For such a solution, the rational function $Y(x) = -\frac{\beta(x)}{\alpha(x)}$ has the required property that $y_i = Y(x_i)$ for all (x_i, y_i) that satisfy $\chi(x, y) = \Theta_\nu(x, y) = 0$. □

Thus, the $2n\nu$ coordinates of solutions (x_i, y_i) of $\Theta_\nu(x, y) = 0$ and $\chi(x, y) = 0$ are algebraic functions of the $n\nu - g$ indeterminates $a_1, a_2, \ldots, a_{n\nu-g}$ of $\hat{\mathcal{K}}_0$. Specifically, the roots of $\varrho_1(x) = 0$ determine the x-coordinates x_i of the zeros of $\Theta_\nu(x, y)$ on $\chi(x, y) = 0$ as algebraic functions of $a_1, a_2, \ldots, a_{n\nu-g}$, after which the formula $y_i = Y(x_i)$ determines the corresponding y_i as an explicit rational function of x_i with coefficients in $\hat{\mathcal{K}}_0$ for each i and therefore as an algebraic function of $a_1, a_2, \ldots, a_{n\nu-g}$ for each i. In this sense, *the $n\nu$ zeros $(x_i, Y(x_i))$ of $\Theta_\nu(x, y)$ on $\chi(x, y) = 0$ are parameterized by the $n\nu - g$ indeterminates* $a_1, a_2, \ldots, a_{n\nu-g}$ that define $\hat{\mathcal{K}}_0$.

In particular, the partial derivatives of the x-coordinates x_i with respect to the parameters a_j have algebraic meaning and can be found by implicit differentiation. In short, the *differentials* of the points (x_i, y_i) along $\chi(x, y) = 0$ that are described by $\Theta_\nu(x, y)$ can be expressed as linear combinations of the $n\nu - g$ differentials da_j in which the coefficients are quantities in the splitting field of $\varrho_1(x) = \varrho_1(x, a_1, a_2, \ldots, a_{n\nu-g})$. The resulting algebraic expressions of the differentials dx_i of the x-coordinates of the solutions in terms of the differentials da_j of the indeterminates of $\Theta_\nu(x, y)$ can be regarded, heuristically, as describing the infinitesimal motion of the zeros of $\Theta_\nu(x, y)$ along the curve $\chi(x, y) = 0$ that result from changes in the variable coefficients of $\Theta_\nu(x, y)$.

Example 3 When the curve field is defined by $y^2 = x^3 + 1$, a normal basis is given by 1 and y, and $\Theta_2(x, y)$ becomes $x^2 + ax + b + cy$. Then the intersection points are roots of $(\frac{-x^2-ax-b}{c})^2 - (x^3 + 1)$, which means that $\varrho_1(x, a, b, c) = \varrho(x, a, b, c)$ (since $q(x) = 1$) can be taken to be

$$\begin{aligned}\varrho_1(x, a, b, c) &= (x^2 + ax + b)^2 - c^2(x^3 + 1)\\ &= x^4 + (2a - c^2)x^3 + (2b + a^2)x^2 + 2abx + (b^2 - c^2).\end{aligned}$$

The rational function $Y(x)$ in Lemma 2 is clearly $Y(x) = \frac{-x^2-ax-b}{c}$.

Implicit differentiation of $\varrho_1(x, a, b, c) = 0$ with respect to a gives

$$\frac{dx}{da} = -\frac{\dfrac{\partial \varrho_1}{\partial a}}{\dfrac{\partial \varrho_1}{\partial x}} = -\frac{2x^3 + 2ax^2 + 2bx}{D} = -\frac{2x(x^2 + ax + b)}{D} = \frac{2cxy}{D},$$

where $D = \frac{\partial \varrho_1}{\partial x} = 4x^3 + 3(2a - c^2)x^2 + 2(2b + a^2)x + 2ab$. Similarly,

$$\frac{dx}{db} = \frac{2cy}{D} \text{ and } \frac{dx}{dc} = \frac{cy^2}{D}.$$

Thus, the parametric points (x_i, y_i) give differentials

$$\begin{aligned}dx_i &= \frac{2cx_iy_i}{D_i} \cdot da + \frac{2cy_i}{D_i} \cdot db + \frac{cy_i^2}{D_i} \cdot dc\\ &= \frac{2x_i(-x_i^2 - ax_i - b)}{D_i} \cdot da + \frac{2(-x_i^2 - ax_i - b)}{D_i} \cdot db + \frac{c(x_i^3 + 1)}{D_i} \cdot dc,\end{aligned} \tag{7}$$

where $D_i = 4x_i^3 + 3(2a - c^2)x_i^2 + 2(2b + a^2)x_i + 2ab$. These computations will be used in later examples.

Essay 9.7 The Theorem of Abel's Last Paper

> [Abel's paper [1]] *still stands for me as pure magic. Neither with Gauss nor Riemann, nor with anybody else, have I found anything that really measures up to this.*—Alte Selberg [9, p. 648].

The themes of Abel's 1826 Paris Memoir [2] were developed in two subsequent papers [1, 5] published in 1828 and 1829 respectively. The Paris Memoir did not appear in print until 1841, so these two papers represent how mathematicians first learned of Abel's work on algebraic curves. The hyperelliptic case is covered in detail in [5], while [1], Abel's last paper, gives an elegant proof of a result from [2].

Here is a version of Abel's result adapted to the framework of the previous essay:

Abel's Theorem 1 *Let (x_i, y_i) be the $n\nu$ solutions of $\chi(x, y) = 0$ and $\Theta_\nu(x, y) = 0$ constructed in* Essay 9.6. *Then the x_i and y_i are algebraic functions of the parameters $a_1, a_2, \ldots, a_{n\nu-g}$, and for any rational function $f(x, y)$ with coefficients in $\mathcal{K}_0$, the sum*

$$dv = f(x_1, y_1)\,dx_1 + \cdots + f(x_{n\nu}, y_{n\nu})\,dx_{n\nu}$$

is a rational differential in $a_1, a_2, \ldots, a_{n\nu-g}$, i.e.,

$$dv = R_1(a_1, \ldots, a_{n\nu-g})\,da_1 + \cdots + R_{n\nu-g}(a_1, \ldots, a_{n\nu-g})\,da_{n\nu-g}, \tag{1}$$

where the $R_j(a_1, \ldots, a_{n\nu-g})$ are rational functions with coefficients in $\mathcal{K}_0$.

Furthermore, the differential (1) is ***closed****, meaning that the compatibility condition*

$$\frac{\partial R_j}{\partial a_k} = \frac{\partial R_k}{\partial a_j}$$

is satisfied for all $j < k$.

Proof The $x_1, \ldots, x_{n\nu}$ are roots of $\varrho_1(x) = \varrho_1(x, a_1, a_2, \ldots, a_{n\nu-g})$, which has coefficients in $\hat{\mathcal{K}}_0 = \mathcal{K}_0(a_1, a_2, \ldots, a_{n\nu-g})$. For each $j = 1, \ldots, n\nu - g$, define the rational function

$$S_j(x) = -\frac{\dfrac{\partial \varrho_1}{\partial a_j}\big(x, a_1, \ldots, a_{n\nu-g}\big)}{\dfrac{\partial \varrho_1}{\partial x}\big(x, a_1, \ldots, a_{n\nu-g}\big)}$$

with coefficients in $\hat{\mathcal{K}}_0$. Since $\varrho_1(x_i, a_1, \ldots, a_{n\nu-g}) = 0$, implicit differentiation implies that

$$dx_i = \sum_{j=1}^{n\nu-g} \frac{dx_i}{da_j} \cdot da_j,$$

where

$$\text{(2)} \qquad \frac{dx_i}{da_j} = -\frac{\dfrac{\partial \varrho_1}{\partial a_j}\big(x_i, a_1, \ldots, a_{n\nu-g}\big)}{\dfrac{\partial \varrho_1}{\partial x}\big(x_i, a_1, \ldots, a_{n\nu-g}\big)} = S_j(x_i).$$

By Lemma 2 of Essay 9.6, there is a rational function $Y(x)$ with coefficients in $\hat{\mathcal{K}}_0$ such that $y_i = Y(x_i)$ for all i. It follows that

$$\sum_{i=1}^{n\nu} f(x_i, y_i)\, dx_i = \sum_{j=1}^{n\nu-g} \sum_{i=1}^{n\nu} f(x_i, y_i) \frac{dx_i}{da_j}\, da_j = \sum_{j=1}^{n\nu-g} \underbrace{\Bigg(\sum_{i=1}^{n\nu} f(x_i, Y(x_i)) S_j(x_i) \Bigg)}_{R_j} da_j.$$

The rational function R_j inside the large parentheses is symmetric in $x_1, \ldots, x_{n\nu}$ with coefficients in $\hat{\mathcal{K}}_0$. By the theory of symmetric functions, R_j is a rational function in the elementary symmetric polynomials of the x_i with coefficients in $\hat{\mathcal{K}}_0$. The x_i are the roots of $\varrho_1(x, a_1, a_2, \ldots, a_{n\nu-g})$, which as a polynomial in x also has coefficients in $\hat{\mathcal{K}}_0$. It follows that R_j is in $\hat{\mathcal{K}}_0$, as was to be shown.[11]

Finally, $R_j = \sum_{i=1}^{n\nu} f(x_i, y_i) \frac{dx_i}{da_j}$, where y_i is an algebraic function of x_i. The compatibility conditions for all $j < k$ follow easily. Hence (1) is closed. □

As noted in Essay 9.1, Abel's version of the theorem used integrals rather than differentials. Deducing his result from Abel's Theorem 1 requires more than just the method of partial fractions, which applies to differential $R(x)\,dx$ of a single variable (see Endnote 9.1). This is because the differential (1) of Abel's Theorem 1 is rational in $a_1, \ldots, a_{n\nu-g}$. However, since (1) is a *closed* differential, a result proved by Chen and Koutschan [17, Theorem 8] implies that there is finite extension $\mathcal{K}_0'$ of $\mathcal{K}_0$ and a function

$$\text{(3)} \qquad v = g_0(a_1, \ldots, a_{n\nu-g}) + \sum_{k=1}^{L} \gamma_k \log\big(g_k(a_1, \ldots, a_{n\nu-g})\big),$$

where $\gamma_1, \ldots, \gamma_L$ are constants in $\mathcal{K}_0'$, g_0 is a rational function with coefficients in $\mathcal{K}_0$, and $g_1, \ldots, g_L$ are rational functions with coefficients in $\mathcal{K}_0'$, such that

$$dv = R_1(a_1, \ldots, a_{n\nu-g})\, da_1 + \cdots + R_{n\nu-g}(a_1, \ldots, a_{n\nu-g})\, da_{n\nu-g}$$

(see Endnote 9.2 for more details). It follows that

$$\int R_1(a_1, \ldots, a_{n\nu-g})\, da_1 + \cdots + \int R_{n\nu-g}(a_1, \ldots, a_{n\nu-g})\, da_{n\nu-g} = v + C,$$

[11] This argument is similar to what Abel did in [1], though his exposition is terse. Earlier, in his 1826 memoir, Abel hints at the above proof [2, p. 149] and then gives another argument that leads to an explicit formula for the R_j [2, pp. 150–159]. A detailed derivation of his formula can be found in [41, pp. 579–582].

where v is a "rational and logarithmic" function and C is an arbitrary constant. This differs from the "algebraic and logarithmic" terminology used by Abel. His more general functions will appear in Essay 9.9.

Example 4 Consider $y^2 = x^3 + 1$ as in Example 3 of Essay 9.6. When $\nu = 2$, recall that

$$\varrho_1(x, a, b, c) = x^4 + (2a - c^2)x^3 + (2b + a^2)x^2 + 2abx + (b^2 - c^2)$$

and $Y(x) = \frac{-x^2-ax-b}{c}$. For the rational function $f(x, y) = \frac{x}{y}$, the proof of Abel's Theorem 1 and the formulas from Example 3 in Essay 9.6 imply that

$$\text{(4)} \qquad \sum_{i=1}^{4} \frac{x_i}{y_i}\, dx_i = \left(\sum_{i=1}^{4} \frac{2cx_i^2}{D_i}\right) da + \left(\sum_{i=1}^{4} \frac{2cx_i}{D_i}\right) db + \left(\sum_{i=1}^{4} \frac{cx_i y_i}{D_i}\right) dc,$$

where $D_i = 4x_i^3 + 3(2a - c^2)x_i^2 + 2(2b + a^2)x_i + 2ab$. The expressions inside the large parentheses can be computed by standard algorithms for symmetric functions (for the third sum on the right, use $cy_i = -x_i^2 - ax_i - b$).

These computations can be simplified by using Abel's Theorem 2 of Essay 9.8, which says that

$$\frac{dx_1}{y_1} + \frac{dx_2}{y_2} + \frac{dx_3}{y_3} + \frac{dx_4}{y_4} = 0$$

since $\frac{dx}{y}$ is holomorphic. Writing the left side in terms of da, db, dc gives

$$\left(\sum_{i=1}^{4} \frac{2cx_i}{D_i}\right) da + \left(\sum_{i=1}^{4} \frac{2c}{D_i}\right) db + \left(\sum_{i=1}^{4} \frac{cy_i}{D_i}\right) dc = 0,$$

so that

$$\text{(5)} \qquad \sum_{i=1}^{4} \frac{2cx_i}{D_i} = \sum_{i=1}^{4} \frac{2c}{D_i} = \sum_{i=1}^{4} \frac{cy_i}{D_i} = 0.$$

The vanishing of the third sum of (5) and $cy_i = -x_i^2 - ax_i - b$ then imply that

$$0 = \sum_{i=1}^{4} \frac{x_i^2}{D_i} + a \sum_{i=1}^{4} \frac{x_i}{D_i} + b \sum_{i=1}^{4} \frac{1}{D_i}.$$

This and the vanishing of the first two sums of (5) yield $\sum_{i=1}^{4} \frac{x_i^2}{D_i} = 0$.

It follows that in (4), the first two sums on the right vanish. Furthermore, again using $cy_i = -x_i^2 - ax_i - b$ and the sums already known to vanish, one obtains

$$\sum_{i=1}^{4} \frac{x_i}{y_i}\, dx_i = \left(\sum_{i=1}^{4} \frac{cx_i y_i}{D_i}\right) dc = -\left(\sum_{i=1}^{4} \frac{x_i^3}{D_i}\right) dc.$$

A calculation using the *Magma Computational Algebra System* reveals that the sum $\sum_{i=1}^{4} \frac{x_i^3}{D_i}$ reduces to 1, with the result that

$$\frac{x_1\,dx_1}{y_1} + \frac{x_2\,dx_1}{y_2} + \frac{x_3\,dx_1}{y_3} + \frac{x_4\,dx_1}{y_4} = -dc,$$

which is a closed rational differential in a, b, c.

Essay 9.8 A Theorem About Holomorphic Differentials

Holomorphic differentials appear indirectly in Abel's Paris Memoir [2] when he asks when a sum of integrals $\int f(x, y)\,dx$ can equal a constant, or equivalently, when their differentials $f(x, y)\,dx$ sum to zero. He realizes that this imposes conditions on $f(x, y)$. His analysis leads to differentials that are (almost) holomorphic, and the "number of arbitrary constants" that describe these differentials (see [2, p. 167]) is closely related to the genus g. Although the relation is imperfect,[12] Abel was asking the right question, and the answer definitely involves holomorphic differentials.

The following theorem shows how this works in the setting of the previous essay.

Abel's Theorem 2 *Let* (x_i, y_i) *be the* $n\nu$ *solutions of* $\chi(x, y) = 0$ *and* $\Theta_\nu(x, y) = 0$ *constructed in* Essay 9.6. *If* $h(x, y)$ *is a rational function with coefficients in* $\mathcal{K}_0$, *then*

$$h(x_1, y_1)\,dx_1 + \cdots + h(x_{n\nu}, y_{n\nu})\,dx_{n\nu} = 0$$

if and only if $h(x, y)\,dx$ *is a holomorphic differential.*

Proof Given $h(x, y)$, the proof of Abel's Theorem 1 (especially equation (2) from Essay 9.7) implies that

$$(1) \qquad \sum_{i=1}^{n\nu} h(x_i, y_i)\,dx_i = \sum_{j=1}^{n\nu-g} \left(\sum_{i=1}^{n\nu} h(x_i, y_i) \frac{dx_i}{da_j} \right) da_j,$$

where $\frac{dx_i}{da_j}$ is a rational function of $x, a_1, a_2, \ldots, a_{n\nu-g}$ with coefficients in $\mathcal{K}_0$.

If $h(x, y)\,dx$ is holomorphic, then so is $h(x, y)\frac{dx}{da_j}\,da_j$, because these are by definition two descriptions of the same differential. Because y_i can be expressed as a rational function $Y(x_i)$ of x_i with coefficients in $\hat{\mathcal{K}}_0$, it follows that $h(x_i, y_i)\frac{dx_i}{da_j}$ can be expressed as a quantity in the splitting field of $\varrho_1(x, a_1, a_2, \ldots, a_{n\nu-g})$. The coefficient of da_j in the right side of (1) is then the trace of $h(x, y)\frac{dx}{da_j}$. Since $h(x, y)\frac{dx}{da_j}\,da_j$ is a holomorphic differential, this trace equals zero by Corollary 2 of Essay 9.5. Therefore, the sum on the left side of (1) equals zero, as was to be shown.

For the converse, let $\mathcal{L}$ be the splitting field of $\varrho_1(x, a_1, a_2, \ldots, a_{n\nu-g})$ over $\hat{\mathcal{K}}_0 = \mathcal{K}_0(a_1, a_2, \ldots, a_{n\nu-g})$. This field contains $x_1, \ldots, x_{n\nu}$, and differentials in dx_i with

[12] See [12, pp. 215–222] and [50] for a discussion of Abel's work on holomorphic differentials.

coefficients in $\mathcal{L}$ form a vector space over $\mathcal{L}$. Proposition 2, to be proved below, shows that exactly $n\nu - g$ of the x_i are algebraically independent over $\mathcal{K}_0$. It follows that this vector space of differentials has dimension $n\nu - g$ over $\mathcal{L}$.

Suppose there are differentials $h_1(x, y)\,dx, \ldots, h_{g+1}(x, y)\,dx$ with coefficients in $\mathcal{K}_0$ that are linearly independent over $\mathcal{K}_0$ and satisfy

$$(2) \qquad \sum_{i=1}^{n\nu} h_j(x_i, Y(x_i))\,dx_i = 0, \quad j = 1, \ldots, g+1.$$

Since the h_j are linearly independent over $\mathcal{K}_0$, it follows that (2) imposes $g + 1$ independent conditions on $dx_1, \ldots, dx_{n\nu}$. Thus the vector space of differentials has dimension at most $n\nu - (g + 1) = n\nu - g - 1$. Yet its dimension is $n\nu - g$ by the previous paragraph. This proves that over $\mathcal{K}_0$, the vector space of differentials for which

$$\sum_{i=1}^{n\nu} h(x_i, y_i)\,dx_i = \sum_{i=1}^{n\nu} h(x_i, Y(x_i))\,dx_i = 0$$

has dimension at most g over $\mathcal{K}_0$. This vector space contains the subspace of holomorphic differentials, which has dimension g, so the two are equal. Thus every $h(x, y)\,dx$ satisfying the above equation is holomorphic. □

This proof works because g arises in two seemingly different contexts: g is the number of linearly independent holomorphic differentials, and $n\nu - g + 1$ is the number of indeterminates in the polynomial $\theta_\nu(x, y)$ of Essay 9.6.

Example 5 For the elliptic curve $y^2 = x^3 + 1$ from the example of the preceding essay, Abel's Theorem 2 implies that

$$\frac{dx_1}{y_1} + \frac{dx_2}{y_2} + \frac{dx_3}{y_3} + \frac{dx_4}{y_4} = 0$$

since $\frac{dx}{y}$ is a holomorphic differential. Writing this in terms of da, db, dc, one obtains

$$\sum_{i=1}^{4} \frac{2cx_i}{D_i} = 0, \quad \sum_{i=1}^{4} \frac{2c}{D_i} = 0, \quad \text{and} \quad \sum_{i=1}^{4} \frac{cy_i}{D_i} = -\sum_{i=1}^{4} \frac{x_i^2 + ax_i + b}{D_i} = 0,$$

where $D_i = 4x_i^3 + 3(2a - c^2)x_i^2 + 2(2b + a^2)x_i + 2ab$. Direct verification of these equations involves expressing these symmetric polynomials in x_1, x_2, x_3, and x_4 in terms of a, b, and c and leads to a long calculation. This can done using a computer as in Example 4 from Essay 9.7.

A partial, but convincing, check can be made by approximating the roots x_i of $\varrho_1(x, a, b, c)$ numerically in the case $a = 1, b = 2, c = 3$, in which case $\varrho_1(x, a, b, c) = x^4 - 7x^3 + 5x^2 + 4x - 5$, and using these numerical approximations and $D_i = 4x_i^3 - 21x_i^2 + 10x_i + 4$, to approximate the three sums numerically. The results will be found to differ from 0 by an amount that can be accounted for by round-off error.

The final task of this essay is to prove the following result used in the proof of Abel's Theorem 2.

Proposition 2 *Let (x_i, y_i) be the $n\nu$ solutions of $\chi(x, y) = 0$ and $\Theta_\nu(x, y) = 0$ constructed in* Essay 9.6, *and let $a_1, a_2, \ldots, a_{n\nu-g}$ be the indeterminates appearing in $\Theta_\nu(x, y)$. Then the (x_i, y_i) can be renumbered so that the following hold*:

1. *Each a_j is a rational function of x_i, y_i, $i = 1, \ldots, n\nu - g$, with coefficients in $\mathcal{K}_0$.*
2. *There is a rational function $Y(x, x_1, y_1, \ldots, x_{n\nu-g}, y_{n\nu-g})$ with coefficients in $\mathcal{K}_0$ such that $y_i = Y(x_i, x_1, y_1, \ldots, x_{n\nu-g}, y_{n\nu-g})$ for $i = 1, \ldots, n\nu$.*
3. *$x_1, \ldots, x_{n\nu-g}$ are algebraically independent over $\mathcal{K}_0$ and $x_{n\nu-g+1}, \ldots, x_{n\nu}$ are algebraic over $\mathcal{K}_0(x_1, \ldots, x_{n\nu-g})$.*

Proof The construction of Essay 9.6 implies that $\Theta_\nu(x, y)$ can be written in the form

$$\Theta_\nu(x, y) = x^\nu + \sum_{j=1}^{n\nu-g} a_j \Phi_j(x, y),$$

where the rational functions $\Phi_j(x, y)$ have coefficients in $\mathcal{K}_0$. The vanishing of $\Theta_\nu(x, y)$ at (x_i, y_i) shows that the a_j are solutions of the linear equations

$$\sum_{j=1}^{n\nu-g} a_j \Phi_j(x_i, y_i) = -x_i^\nu, \quad i = 1, \ldots, n\nu. \tag{3}$$

Assume there is another solution $b_1, b_2, \ldots, b_{n\nu-g}$ in $\hat{\mathcal{K}}_0 = \mathcal{K}_0(a_1, a_2, \ldots, a_{n\nu-g})$. Then $\Theta_\nu(x, y)$ and

$$\hat{\Theta}_\nu(x, y) = x^\nu + \sum_{j=1}^{n\nu-g} b_j \Phi_j(x, y)$$

vanish at the $n\nu$ points (x_i, y_i). Thus, when regarded as a quantity in the curve field of $\chi(x, y)$, $\hat{\Theta}_\nu(x, y)$ has **at least** $n\nu$ zeros, counted with multiplicity. But the above formula for $\hat{\Theta}_\nu(x, y)$ implies that it is integral over x with order at most ν at $x = \infty$. Hence $\hat{\Theta}_\nu(x, y)$ has **at most** $n\nu$ poles, counted with multiplicity. Since the number of zeros equals the number of poles, $\hat{\Theta}_\nu(x, y)$ has exactly $n\nu$ zeros, all of multiplicity one, and its order at $x = \infty$ is exactly ν, so that it has poles of order ν at all n poles of x. The same is true for $\Theta_\nu(x, y)$, so that their quotient has no poles. Therefore, the quotient is a constant, i.e., $\hat{\Theta}_\nu(x, y)$ is a constant multiple of $\Theta_\nu(x, y)$. Then $\hat{\Theta}_\nu(x, y) = \Theta_\nu(x, y)$ since both come from solutions of (3). It follows that the a_j are the unique solution of (3).

The inhomogeneous system (3) consists of $n\nu$ equations (one for each (x_i, y_i)) in $n\nu - g$ unknowns $a_1, a_2, \ldots, a_{n\nu-g}$. Hence a subset of the equations gives a square system with a unique solution. Renumbering if necessary, the square system

$$\sum_{j=1}^{n\nu-g} a_j \Phi_j(x_i, y_i) = -x_i^\nu, \quad i = 1, \ldots, n\nu - g$$

has a unique solution. Cramer's rule expresses each a_j as a rational function of (x_i, y_i) for $i = 1, \dots, n\nu - g$ with coefficients in $\mathcal{K}_0$. This proves the first assertion of the proposition.

The second assertion follows without difficulty. By Lemma 2 in Essay 9.6, there is a rational function $Y(x)$ with coefficients in $\hat{\mathcal{K}}_0 = \mathcal{K}_0(a_1, \dots, a_{n\nu-g})$ such that $y_i = Y(x_i)$ for all i. Replacing each a_j appearing in $Y(x)$ with the rational function constructed in the first assertion gives the desired $Y(x, x_1, y_1, \dots, x_{n\nu-g}, y_{n\nu-g})$.

For the third assertion, observe that y_i algebraic over x_i for all i. Thus the first assertion implies that $a_1, \dots, a_{n\nu-g}$ are algebraic over $x_1, \dots, x_{n\nu-g}$. Then the algebraic independence of $a_1, \dots, a_{n\nu-g}$ immediately implies that $x_1, \dots, x_{n\nu-g}$ are algebraically independent.

Finally, $x_{n\nu-g+1}, \dots, x_{n\nu}$ are algebraic over $a_1, \dots, a_{n\nu-g}$ by construction, and hence algebraic over $x_1, y_1, \dots, x_{n\nu-g}, y_{n\nu-g}$ by the first assertion. Each y_i is algebraic over x_i, so $x_{n\nu-g+1}, \dots, x_{n\nu}$ are algebraic over $x_1, \dots, x_{n\nu-g}$, as was to be shown. □

Example 6 For the elliptic curve $y^2 = x^3 + 1$, the smallest value of ν is 2 since y has order 2 at $x = \infty$. For $\Theta_2(x, y, a, b, c) = x^2 + ax + b + cy$, the four solutions (x_i, y_i) satisfy

$$ax_i + b + cy_i = -x_i^2, \quad i = 1, 2, 3, 4.$$

Using $i = 1, 2, 3$ and Cramer's rule, one obtains

$$a = \frac{\det\begin{bmatrix} -x_1^2 & 1 & y_1 \\ -x_2^2 & 1 & y_2 \\ -x_3^2 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} x_1 & 1 & y_1 \\ x_2 & 1 & y_2 \\ x_3 & 1 & y_3 \end{bmatrix}} = \frac{x_1^2 y_2 + x_2^2 y_3 + x_3^2 y_1 - x_1^2 y_3 - x_2^2 y_1 - x_3^2 y_2}{x_1 y_3 + x_2 y_1 + x_3 y_2 - x_1 y_2 - x_2 y_3 - x_3 y_1},$$

with similar formulas for b and c. This is the first assertion of the proposition. Then $Y(x) = \frac{-x^2 - ax - b}{c}$ leads to an explicit formula for $Y(x, x_1, y_1, x_2, y_2, x_3, y_3)$ as in the second assertion. Finally, since

$$\varrho_1(x, a, b, c) = x^4 + (2a - c^2)x^3 + (2b + a^2)x^2 + 2abx + (b^2 - c^2)$$

(see Example 3 in Essay 9.6), $x_1 + x_2 + x_3 + x_4 = -(2a - c^2)$. This gives an explicit formula for x_4 in terms of $x_1, y_1, x_2, y_2, x_3, y_3$, as claimed in the third assertion.

This formula for x_4 and $y_4 = Y(x_4, x_1, y_1, x_2, y_2, x_3, y_3)$ imply that (x_4, y_4) is uniquely determined by the points (x_1, y_1), (x_2, y_2), (x_3, y_3). In fact, Essay 9.11 will show that $(x_4, y_4) = -(x_1, y_1) - (x_2, y_2) - (x_3, y_3)$ with respect to the usual addition law on the elliptic curve $y^2 = x^3 + 1$.

Something remarkable has happened here—thinking about how the points (x_i, y_i) relate to each other (as codified in the proposition) leads to formulas related to the addition law on $y^2 = x^3 + 1$. The framework developed in Essay 9.6 always involves $n\nu$ points on the curve, which for $y^2 = x^3 + 1$ is always at least four. Thus further thought is needed to study what happens for an arbitrary number of points on the curve. This is the subject of the next essay.

Essay 9.9 A Change of Parameters

The indeterminates $a_1, a_2, \ldots, a_{n\nu-g}$ introduced in Essay 9.6 have played a prominent role so far. Abel used similar quantities $a, a', a'', \ldots$ in the first five sections of his Paris Memoir [2]. The situation changed in Section 6, where he let α denote their number and asserted that their values are "functions of a number α of the quantities $x_1, \ldots, x_\mu$; for example, as functions of $x_1, \ldots, x_\alpha$" [2, p. 170]. Abel also stated that the difference $\mu - \alpha$ is "very remarkable" [2, p. 172].

In the situation of these essays, $n\nu - g$ plays that role of α and $n\nu$ plays the role of μ, so that

$$\mu - \alpha = n\nu - (n\nu - g) = g.$$

Very remarkable indeed![13] The two versions of Abel's theorem stated and proved in the next essay will give insight into why Abel thought the genus was so important. The preliminary lemmas proved here lay the groundwork for the theorems. In these lemmas, the parameters $a_1, a_2, \ldots, a_{n\nu-g}$ will be replaced with an equal number of the x_i's. This is the "change of parameters" in the title of the essay.

In what follows, set $N = n\nu - g$ and $\hat{\mathcal{K}}_0 = \mathcal{K}_0(a_1, a_2, \ldots, a_N)$. Then, as in Abel's Theorems 1 and 2, let $\mathcal{L}$ be the splitting field of $\varrho_1(x, a_1, a_2, \ldots, a_N)$ over $\hat{\mathcal{K}}_0$. Also set $y_i = Y(x_i)$ for $i = 1, \ldots, n\nu$, where $Y(x)$ is the rational function with coefficients in $\hat{\mathcal{K}}_0$ from Lemma 2 of Essay 9.6. The y_i lie in $\mathcal{L}$ and satisfy $\chi(x_i, y_i) = 0$ for all i.

By Proposition 2 of the previous essay, the x_i and y_i can be renumbered so that the a_j are rational functions of $x_1, \ldots, x_N, y_1, \ldots, y_N$ with coefficients in $\mathcal{K}_0$. It follows that

$$\mathcal{L} = \mathcal{K}_0(x_1, \ldots, x_{n\nu}, y_1, \ldots, y_{n\nu}).$$

The proposition also implies that $x_1, \ldots, x_N$ are algebraically independent over $\mathcal{K}_0$ and $x_{N+1}, \ldots, x_{N+g} = x_{n\nu}$ are algebraic over $x_1, \ldots, x_N$. Consider the subfield

$$\mathcal{L}_N = \mathcal{K}_0(x_1, \ldots, x_N, y_1, \ldots, y_N)$$

of $\mathcal{L}$. The field $\mathcal{L}_N$ has a simple structure since $x_1, \ldots, x_N$ are algebraically independent over $\mathcal{K}_0$ and each y_i is a root of $\chi(x_i, y) = 0$. Heuristically, $\mathcal{L}_N$ is the field associated with N general points on the curve $\chi(x, y) = 0$.

Given a rational function $f(x, y)$, Abel's Theorem 1 from Essay 9.7 constructs a closed differential $dv = \sum_{j=1}^{N} R_j(a_1, \ldots, a_N)\,da_j$ that satisfies the theorem. Replacing $a_1, \ldots, a_N$ with the rational functions of $x_1, \ldots, x_N, y_1, \ldots, y_N$ constructed in the previous essay gives

[13] Similar to footnote 12, Abel's number $\mu - \alpha$ has a imperfect relation to the genus g.

$$dv = \sum_{j=1}^{N} R_j(a_1, \dots, a_N)\,da_j = \sum_{j=1}^{N} R_j(a_1, \dots, a_N)\Big(\sum_{k=1}^{N} \frac{\partial a_j}{\partial x_k}\,dx_k + \sum_{j=1}^{N} \frac{\partial a_j}{\partial y_k}\,dy_k\Big).$$

This is rational in $x_1, \dots, x_N, y_1, \dots, y_N$ and hence algebraic in $x_1, \dots, x_N$ since each y_i is algebraic over x_i. It follows that dv is constructed via substitution of algebraic functions into a closed rational differential in algebraically independent variables. Such a differential is said to be **rationally closed**.

Preliminary Lemma 1 *The algebraic extension $\mathcal{L}_N \subset \mathcal{L}$ contains solutions (x_j', y_j') of $\chi(x, y) = 0$ for $j = 1, \dots, g$ such that for any rational function $f(x, y)$ over $\mathcal{K}_0$, there is a differential dv satisfying*

$$f(x_1, y_1)\,dx_1 \ + \cdots + f(x_N, y_N)\,dx_N = dv - \big(f(x_1', y_1')\,dx_1' + \cdots + f(x_g', y_g')\,dx_g'\big).$$

Furthermore, the differential dv satisfies

1. *dv is rational in $x_1, \dots, x_N, y_1, \dots, y_N$ and rationally closed.*
2. *$dv = 0$ when $f(x, y)\,dx$ is a holomorphic differential.*

Proof Abel's Theorem 1 implies that $f(x_1, y_1)\,dx_1 \ + \cdots + f(x_N, y_N)\,dx_N$ equals

$$dv - \big(f(x_{N+1}, y_{N+1})\,dx_{N+1} + \cdots + f(x_{n\nu}, y_{n\nu})\,dx_{n\nu}\big)$$

over $\hat{\mathcal{K}}_0 = \mathcal{K}_0(a_1, \dots, a_N)$. The equation of the lemma follows by setting $(x_j', y_j') = (x_{N+j}, y_{N+j})$ for $j = 1, \dots, g$ (recall that $N = n\nu - g$).

This equation continues to hold when $a_1, \dots, a_N$ are replaced by rational functions in $x_1, \dots, x_N, y_1, \dots, y_N$, and the resulting dv is rationally closed by the discussion preceding the lemma.

The final assertion of the lemma ($dv = 0$ when $f(x, y)\,dx$ is holomorphic) follows from Abel's Theorem 2. □

The reason for being careful about dv becomes apparent when one considers the integral version of the equation in Preliminary Lemma 1:

$$\int f(x_1, y_1)\,dx_1 + \cdots + \int f(x_N, y_N)\,dx_N$$
$$= v - \Big(\int f(x_1', y_1')\,dx_1' + \cdots + \int f(x_g', y_g')\,dx_g'\Big).$$

In equation (3) of Essay 9.7, $v = g_0(a_1, \dots, a_{n\nu-g}) + \sum_{k=1}^{L} \gamma_k \log\big(g_k(a_1, \dots, a_{n\nu-g})\big)$, while here, the function v arises by replacing a_j with the appropriate rational function of $x_1, \dots, x_N, y_1, \dots, y_N$. Since y_i is an algebraic function of x_i, it follows that v is now an "algebraic and logarithmic" function of $x_1, \dots, x_N$. This is the terminology used by Abel in [2].

The second preliminary lemma generalizes Preliminary Lemma 1 by replacing the indeterminates $x_1, \dots, x_N$, $N = n\nu - g$, with $x_1, \dots, x_\alpha$, where α is now allowed to be arbitrary. This is done as follows.

Fix a positive integer α and let $x_1, \ldots, x_\alpha$ be indeterminates over the field of constants $\mathcal{K}_0$. For each $i = 1, \ldots, \alpha$, adjoin a root y_i of $\chi(x_i, y)$ to get a c-field

$$(1) \qquad \mathcal{L}_\alpha = \mathcal{K}_0(x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha).$$

Heuristically, one can think of the (x_i, y_i) as giving α general points on the curve defined by $\chi(x, y) = 0$.

Preliminary Lemma 2 *Given a positive integer α, construct an algebraic extension $\mathcal{L}_\alpha \subset \mathcal{L}'_\alpha$ that contains solutions (x'_j, y'_j) of $\chi(x, y) = 0$ for $j = 1, \ldots, g$ and a finite extension $\mathcal{K}'_0$ of $\mathcal{K}_0$ such that for any rational function $f(x, y)$ over $\mathcal{K}_0$, there is a differential dv satisfying*

$$f(x_1, y_1)\,dx_1 \ + \cdots + f(x_\alpha, y_\alpha)\,dx_\alpha = dv - \big(f(x'_1, y'_1)\,dx'_1 + \cdots + f(x'_g, y'_g)\,dx'_g\big).$$

Furthermore, the differential dv satisfies
1. *dv is rational in $x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha$ over $\mathcal{K}'_0$ and rationally closed.*
2. *$dv = 0$ when $f(x, y)\,dx$ is a holomorphic differential.*

Proof Pick an integer ν as in Essay 9.6 such that $n\nu - g \geq \alpha$ and set $N = n\nu - g$. When $\alpha = N$, the lemma holds with $\mathcal{L}'_N = \mathcal{L}$ and $\mathcal{K}'_0 = \mathcal{K}_0$ by Preliminary Lemma 1.

To prove the lemma for $\alpha = N-1$, the strategy is to replace (x_1, y_1) with a constant solution $(x_1^{\circ}, y_1^{\circ})$ that lies in a finite extension $\mathcal{K}'_0$ of $\mathcal{K}_0$. To evaluate a quantity z in $\mathcal{L}$ at $(x_1^{\circ}, y_1^{\circ})$, the first step is to isolate x_1 and y_1 in z and write $z = \frac{p(x_1,y_1)}{q(x_1)}$, where $p(x, y)$ and $q(x)$ have coefficients in $\mathcal{K}_0(x_2, \ldots, x_{n\nu}, y_2, \ldots, y_{n\nu})$. This is possible because y_1 is algebraic over $\mathcal{K}_0(x_1)$.

As explained in Essay 8.1, the evaluation of z at $(x_1^{\circ}, y_1^{\circ})$ is equal to $\frac{p(x_1^{\circ},y_1^{\circ})}{q(x_1^{\circ})}$ whenever $q(x_1^{\circ}) \neq 0$. So evaluating finitely many known quantities in $\mathcal{L}$ at $(x_1^{\circ}, y_1^{\circ})$ works provided that x_1° is not a root of any of the denominators involved. If the product of denominators has degree m, this can be accomplished constructively by trying $m + 1$ distinct integers—let x_1° be the first where none of them vanish.

The quantities to evaluate are (x'_j, y'_j), $j = 1, \ldots, g$, and the rational differential dv. For each j, there is also the monic algebraic relation of x'_j on $x_1, \ldots, x_N$ to consider, along with the rational function $Y(x)$ that satisfies $y'_j = Y(x'_j)$ for $j = 1, \ldots, g$. (The relation $\chi(x'_j, y'_j) = 0$ is unaffected by the evaluation since $\chi(x, y)$ has coefficients in $\mathcal{K}_0$.). As indicated above, x_1° can be chosen so that none of the denominators vanish.

Given x_1°, adjoining a root y_1° of $\chi(x_1^{\circ}, y)$ to $\mathcal{K}_0$ gives a finite extension $\mathcal{K}'_0$ of $\mathcal{K}_0$. Proposition 1 of Essay 9.3 implies that $\chi(x, y)$ remains irreducible when $\mathcal{K}_0$ is replaced with $\mathcal{K}'_0$, which gives a new field $\mathcal{L}'$. Then evaluating (x'_j, y'_j), $j = 1, \ldots, g$, $Y(x)$, and dv at $(x_1^{\circ}, y_1^{\circ})$ gives (x''_j, y''_j), $j = 1, \ldots, g$, $Y'(x)$, and dv' with $y''_i = Y'(x''_i)$. Since $dx_1^{\circ} = 0$, this evaluation transforms

$$f(x_1, y_1)\,dx_1 \ + \cdots + f(x_N, y_N)\,dx_N = dv - \big(f(x'_1, y'_1)\,dx'_1 + \cdots + f(x'_g, y'_g)\,dx'_g\big)$$

into

$$f(x_2, y_2)\,dx_2 \ + \cdots + f(x_N, y_N)\,dx_N = dv' - \big(f(x''_1, y''_1)\,dx''_1 + \cdots + f(x''_g, y''_g)\,dx''_g\big),$$

which has the desired form with $\mathcal{L}'_{N-1} \simeq \mathcal{K}'_0(x_2, \dots, x_N, y_2, \dots, y_N)$. Continuing in this way, equation of the lemma follows for any positive $\alpha \le N$.

For the final assertions of the lemma, first note that dv from Preliminary Lemma 1 remains rationally closed when evaluated at (x_1^o, y_1^o). Furthermore, when $N = nv - g$ and $f(x, y)\,dx$ is a holomorphic differential, Abel's Theorem 2 implies that $dv = 0$. This continues to hold when (x_1, y_1) is evaluated at (x_1^o, y_1^o). Thus the final assertions of the lemma hold for $\alpha = N - 1$. As above, they hold for any positive $\alpha \le N$. □

Example 7 For $y^2 = x^3 + 1$, Example 6 from Essay 9.8 showed that $(x_4.y_4)$ can be expressed in terms of (x_1, y_1), (x_2, y_3) and (x_3, y_3). Evaluating this formula at a careful choice for (x_1, y_1) will yield something familiar.

As noted in Chapter 8, a "point" is a way of assigning a value (or the symbol ∞) to each quantity in the curve field $\mathcal{K} = \mathbf{Q}(x_1, y_1)$, where $y_1^2 = x_1^3 + 1$. By Example 4 from Essay 8.9, $z = \frac{1}{x_1}$ and $w = \frac{y_1}{x_1^2}$ give the equation $w^2 = z + z^4$ with the same curve field. Consider the point $P = (0, 0)$ of $w^2 = z + z^4$. This corresponds to the unique point of $y^2 = x^3 + 1$ where $x = \infty$.

Recall that $Y = \frac{-x^2 - ax - b}{c} = -\frac{1}{c}x^2 - \frac{a}{c}x - \frac{b}{c}$. To evaluate these coefficients when $z, w \longrightarrow 0$, first observe that $x_1 = \frac{1}{z}$ and $y_1 = \frac{w}{z^2}$. Then solve for a, b, c using Cramer's rule as in Example 6 from Essay 9.8 to obtain:

$$\frac{1}{c} = \frac{\det\begin{bmatrix} x_1 & 1 & y_1 \\ x_2 & 1 & y_2 \\ x_3 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} x_1 & 1 & -x_1^2 \\ x_2 & 1 & -x_2^2 \\ x_3 & 1 & -x_3^2 \end{bmatrix}} = \frac{\det\begin{bmatrix} z & z^2 & w \\ x_2 & 1 & y_2 \\ x_3 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} z & z^2 & -1 \\ x_2 & 1 & -x_2^2 \\ x_3 & 1 & -x_3^2 \end{bmatrix}} \xrightarrow{z,w \to 0} \frac{\det\begin{bmatrix} 0 & 0 & 0 \\ x_2 & 1 & y_2 \\ x_3 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} 0 & 0 & -1 \\ x_2 & 1 & -x_2^2 \\ x_3 & 1 & -x_3^2 \end{bmatrix}} = 0,$$

where the second equality expresses x_1, y_1 in terms of z, w and multiplies the top and bottom determinants by z^2 to clear denominators. Similarly,

$$\frac{a}{c} = \frac{\det\begin{bmatrix} -x_1^2 & 1 & y_1 \\ -x_2^2 & 1 & y_2 \\ -x_3^2 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} x_1 & 1 & -x_1^2 \\ x_2 & 1 & -x_2^2 \\ x_3 & 1 & -x_3^2 \end{bmatrix}} = \frac{\det\begin{bmatrix} -1 & z^2 & w \\ x_2 & 1 & y_2 \\ x_3 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} z & z^2 & -1 \\ x_2 & 1 & -x_2^2 \\ x_3 & 1 & -x_3^2 \end{bmatrix}} \xrightarrow{z,w \to 0} \frac{\det\begin{bmatrix} -1 & 0 & 0 \\ x_2 & 1 & y_2 \\ x_3 & 1 & y_3 \end{bmatrix}}{\det\begin{bmatrix} 0 & 0 & -1 \\ x_2 & 1 & -x_2^2 \\ x_3 & 1 & -x_3^2 \end{bmatrix}} = \frac{y_2 - y_3}{x_3 - x_2},$$

and

$$\frac{b}{c} \xrightarrow{z,w \to 0} \frac{x_2 y_3 - x_3 y_2}{x_3 - x_2}.$$

It follows that

$$\begin{aligned} Y(x) = -\frac{1}{c}x^2 - \frac{a}{c}x - \frac{b}{c} &\xrightarrow{z,w\to 0} 0\cdot x^2 - \frac{y_2-y_3}{x_3-x_2}x - \frac{x_2y_3-x_3y_2}{x_3-x_2} \\ &= \frac{y_2-y_3}{x_2-x_3}x + \frac{x_2y_3-x_3y_2}{x_2-x_3}. \end{aligned}$$

Denote this linear function by $Y'(x)$.

As in the proof of Preliminary Lemma 2, $(x_4, y_4) \xrightarrow{z,w\to 0} (x_1', y_1')$, which satisfies $y_1' = Y'(x_1')$. An explicit formula for this point is obtained as follows. In Example 3 of Essay 9.6, $\varrho_1(x)$ was constructed by clearing the denominator of $Y(x)^2 - (x^3+1)$. Here, it is more useful to let $\varrho_1(x) = Y(x)^2 - (x^3+1)$. Since $Y(x) \xrightarrow{z,w\to 0} Y'(x)$, $\varrho_1(x)$ becomes the cubic

$$Y'(x)^2 - (x^3+1) = \left(\frac{y_2-y_3}{x_2-x_3}x + \frac{x_2y_3-x_3y_2}{x_2-x_3}\right)^2 - (x^3+1).$$

Since its roots satisfy $x_2+x_3+x_1' = -$coefficient of x^2/coefficient of x^3, one obtains

$$\begin{aligned} x_1' &= -x_2 - x_3 + \left(\frac{y_2-y_3}{x_2-x_3}\right)^2 \\ y_1' &= Y'(x_1') = \frac{y_2-y_3}{x_2-x_3}x_1' + \frac{x_2y_3-x_3y_2}{x_2-x_3}. \end{aligned} \tag{2}$$

These formulas, interpreted geometrically, say that $Y'(x) = 0$ is the equation of the line determined by $P = (x_2, y_2)$ and $Q = (x_3, y_3)$, and then $S = (x_1', y_1')$ is where this line meets the curve $y^2 = x^3 + 1$. Figure 4.3 in Essay 4.2 shows what this looks like for the elliptic curve $y^2 = \frac{3}{4}x^3 - \frac{1}{2}x + \frac{1}{4}$.

For the holomorphic differential $\frac{dx}{y}$, Preliminary Lemma 2 says

$$\frac{dx_2}{y_2} + \frac{dx_3}{y_3} = -\frac{dx_1'}{y_1'}, \quad \text{i.e.,} \quad \frac{dx_2}{y_2} + \frac{dx_3}{y_3} + \frac{dx_1'}{y_1'} = 0$$

when $x_2, y_2, x_3, y_3, x_1', y_1'$ are related by (2). Thus the equations (2) represent an "algebraic integration" of the above differential equation.

Essay 9.10 Abel's Addition Theorems

Given $\alpha > 0$, Preliminary Lemma 2 from the previous essay gives an equation

$$f(x_1,y_1)dx_1 + \cdots + f(x_\alpha,y_\alpha)dx_\alpha = dv - \big(f(x_1',y_1')dx_1' + \cdots + f(x_g',y_g')dx_g'\big). \tag{1}$$

Setting $\psi_i x_i = \int f(x_i, y_i)\,dx_i$, Abel [2, p. 172] wrote the integral version of (1) as

$$\psi_1 x_1 + \cdots + \psi_\alpha x_\alpha = v - \big(\psi_{\alpha+1}x_{\alpha+1} + \cdots + \psi_\mu x_\mu\big).$$

The number of integrals on the right is $\mu - \alpha$, which in Preliminary Lemma 2 is the genus g, and v is an algebraic and logarithmic function as explained in Essay 9.9.

Abel had two ways of dealing the minus sign in the right side of these equations. This will lead to Abel's Theorems 3 and 4. In the first of these theorems, he used a clever trick [2, p. 186] to change the minus sign in (1) into a plus sign:

Abel's Theorem 3 *Given a positive integer α, construct an algebraic extension $\mathcal{L}_\alpha \subset \mathcal{L}'_\alpha$ that contains solutions (x''_j, y''_j) of $\chi(x, y) = 0$ for $j = 1, \dots, g$, and a finite extension $\mathcal{K}'_0$ of $\mathcal{K}_0$ such that for any rational function $f(x, y)$ over $\mathcal{K}_0$, there is a differential dv_1 satisfying*

$$f(x_1, y_1)\,dx_1 \; + \cdots + f(x_\alpha, y_\alpha)\,dx_\alpha = dv_1 + f(x''_1, y''_1)\,dx''_1 + \cdots + f(x''_g, y''_g)\,dx''_g.$$

Furthermore, dv_1 is rationally closed over $\mathcal{K}'_0$ and equals 0 when $f(x, y)\,dx$ is a holomorphic differential.

Proof For independent variables $z_1, \dots, z_g$, Preliminary Lemma 2 applied with $\alpha = g$ gives an equation

$$f(z_1, w_1)\,dz_1 \; + \cdots + f(z_g, w_g)\,dz_g = dv' - \big(f(z'_1, w'_1)\,dz'_1 + \cdots + f(z'_g, w'_g)\,dz'_g\big).$$

Evaluating (z_i, w_i) at (x'_i, y'_i) for $i = 1, \dots, g$ gives

$$f(x'_1, y'_1)\,dx'_1 \; + \cdots + f(x'_g, y'_g)\,dx'_g = dv'' - \big(f(x''_1, y''_1)\,dx''_1 + \cdots + f(x''_g, y''_g)\,dx''_g\big).$$

Adding $0 = \sum_{i=1}^{\alpha} f(x'_i, y'_i)\,dx'_i - dv'' + \sum_{j=1}^{g} f(x''_i, y''_i)\,dx''_i$ to the right side of (1) yields the desired equation with $dv_1 = dv - dv''$, which is a differential of the required form.

Finally, when $f(x, y)\,dx$ is holomorphic, $dv = dv' = 0$ by Preliminary Lemma 2, so that $dv_1 = 0$ as well. □

Example 8 Continuing with the elliptic curve $y^2 = x^3 + 1$ from Example 7 of the previous essay, Preliminary Lemma 2 can be stated as

$$f(x_1, y_1)\,dx_1 + f(x_2, y_2)\,dx_2 = dv - f(x'_1, y'_1)\,dx'_1 \tag{2}$$

after replacing x_2, x_3 with x_1, x_2, and similarly for y_2, y_3. With this change of notation, formula (2) from Example 7 becomes

$$\begin{aligned} x'_1 &= -x_1 - x_2 + \left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2 \\ y'_1 &= Y'(x'_4) = \frac{y_1 - y_2}{x_1 - x_2}x'_1 + \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}. \end{aligned} \tag{3}$$

Both (2) and (3) have interesting consequences. First note that since (x'_1, y'_1) is a solution of $y^2 = x^3 + 1$, so is $(x_3, y_3) = (x'_1, -y'_1)$. Writing $f(x, y) = g(x) + h(x)y$ for rational functions $g(x)$ and $h(x)$, one easily sees that

$$f(x_1', y_1')\,dx_1' = dv'' - f(x_3, y_3)\,dx_3, \quad dv'' = 2g(x_3)\,dx_3.$$

Combining this with (2) gives

$$f(x_1, y_1)\,dx_1 + f(x_2, y_2)\,dx_2 = dv_1 + f(x_3, y_3)\,dx_3, \quad dv_1 = dv - dv'',$$

exactly as in the proof of Abel's Theorem 3. Furthermore, $(x_3, y_3) = (x_1', -y_1')$ transforms (3) into

$$\begin{aligned} x_3 &= -x_1 - x_2 + \left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2 \\ y_3 &= -\frac{y_1 - y_2}{x_1 - x_2} x_3 - \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}, \end{aligned}$$

which is the usual addition law on the elliptic curve $y^2 = x^3 + 1$.

Abel's second way of dealing with the minus signs in (1) was to make them go away by letting the (x_j', y_j') be constant, so that $dx_1' = \cdots = dx_g' = 0$ in (1). This is accomplished by imposing g conditions on (x_i, y_i) for $i = 1, \ldots, \alpha$.

Some preparation is needed before stating the result. Fix ν and set $N = n\nu - g$. Lemma 2 of Essay 9.6 shows that $x_1, \ldots, x_{n\nu}$ are the roots of $\varrho_1(x, a_1, a_2, \ldots, a_N)$, which here will be assumed to be monic in x. By Proposition 2 in Essay 9.8, the a_j are rational functions of $x_1, \ldots, x_N, y_1, \ldots, y_N$ after relabeling. The resulting $\varrho_1(x, x_1, \ldots, x_N, y_1, \ldots, y_N)$ is a polynomial in x that factors as

$$\varrho_1(x, x_1, \ldots, x_N, y_1, \ldots, y_N) = \prod_{i=1}^{N}(x - x_i) \cdot F(x, x_1, \ldots, x_N, y_1, \ldots, y_N),$$

where F has degree g in x. Its roots $x_1', \ldots, x_g'$ appear in the preliminary lemmas of the previous essay when $\alpha = N$.

For $\alpha < N$, Preliminary Lemma 2 proved (1) by replacing the first $N - \alpha$ of the (x_i, y_i) with carefully chosen constant solutions $(x_i^{\mathrm{o}}, y_i^{\mathrm{o}})$ and relabeling $(x_{N-\alpha+1}, y_{N-\alpha+1}), \ldots, (x_N, y_N)$ as $(x_1, y_1), \ldots, (x_\alpha, y_\alpha)$. The above factorization of $\varrho_1(x, x_1, \ldots, x_N, y_1, \ldots, y_N)$ then transforms into

$$\varrho_1^{\mathrm{o}}(x, x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha) = \prod_{i=1}^{N-\alpha}(x - x_i^{\mathrm{o}}) \cdot \prod_{i=1}^{\alpha}(x - x_i) \cdot F_1(x, x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha),$$

where

$$F_1(x, x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha) = F(x, x_1^{\mathrm{o}}, \ldots, x_{N-\alpha}^{\mathrm{o}}, x_1, \ldots, x_\alpha, y_1^{\mathrm{o}}, \ldots, y_{N-\alpha}^{\mathrm{o}}, y_1, \ldots, y_\alpha).$$

The quantities $x_1', \ldots, x_g'$ that appear in (1) are the roots of F_1 in some splitting field, and $y_i' = Y'(x_i')$ as in the proof of Preliminary Lemma 2.

The basic idea of Abel's Theorem 4 is that requiring the points (x_j', y_j') to be constant imposes g conditions on $(x_1, y_1), \ldots, (x_\alpha, y_\alpha)$, and when these conditions are satisfied, the right side of (1) reduces to dv.

Algebraically, this is done using the parametric setting introduced in Essay 8.2, where one adds algebraically independent parameters to the field of constants. Suppose h of $x_1', \ldots, x_g'$, say $x_1', \ldots, x_h'$, are algebraically independent over $\mathcal{K}_0'$ and the remaining x_i' are algebraic over these.[14] As in Essay 8.2, this gives a curve field over $\hat{\mathbf{Q}} = \mathbf{Q}(x_1', \ldots, x_h')$.

Originally, $\chi(x, y)$ was irreducible over $\mathcal{K}_0$, and when this field was enlarged to $\mathcal{K}_0'$ in Preliminary Lemma 2 of Essay 9.9, it remained irreducible by Proposition 1 of Essay 9.3. Since $\hat{\mathcal{K}}_0' = \mathcal{K}_0'(x_1', \ldots, x_h')$ simply adds new indeterminates, $\chi(x, y)$ is irreducible over $\hat{\mathcal{K}}_0'$. The choice of $x_1', \ldots, x_h'$ guarantees that the new field of constants $\hat{\mathcal{K}}_0''$ contains $x_1', \ldots, x_g', y_1', \ldots, y_g'$.

In this setting, $dx_i' = 0$ for $i = 1, \ldots, h$, and then $dx_i' = 0$ for $i = 1, \ldots, g$ since the others are algebraic over $x_1', \ldots, x_h'$. It follows that $dy_i' = 0$ since y_i' is algebraic over x_i'. In this way, (x_i', y_i') becomes a constant solution of $\chi(x, y) = 0$ for $i = 1, \ldots, g$. Also note that $x_1, \ldots, x_\alpha$ are no longer algebraically independent since the equations

$$F_1(x_i', x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha) = 0, \quad i = 1, \ldots, g \tag{4}$$

lead to algebraic relations among $x_1, \ldots, x_\alpha$ over $\hat{\mathcal{K}}_0'$ because y_i is algebraic over x_i for $i = 1, \ldots, \alpha$.

Since $dx_j' = 0$ for all j, equation (1) has the following immediate consequence.

Abel's Theorem 4 *Let the field of constants be $\hat{\mathcal{K}}_0''$ containing $x_1', \ldots, x_g', y_1', \ldots, y_g'$ as above. Then for any rational function $f(x, y)$ with coefficients in $\mathcal{K}_0$, there is a rationally closed differential dv, rational in $x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha$, such that*

$$f(x_1, y_1)\, dx_1 \ + \cdots + f(x_\alpha, y_\alpha)\, dx_\alpha = dv.$$

Furthermore, $x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha$ satisfy the g relations given by (4).

This theorem is remarkable because, first, the relations (4) are independent of the differential $f(x, y)\, dx$, and second, the number of relations is the genus of the curve. Since $dv = 0$ when the differential is holomorphic, Abel's Theorem 4 has the following immediate corollary:

Corollary 3 *Let the field of constants be $\hat{\mathcal{K}}_0' = \mathcal{K}_0'(x_1', \ldots, x_h')$ as above. Then for any holomorphic differential $h(x, y)\, dx$ with coefficients in $\mathcal{K}_0$,*

$$h(x_1, y_1)\, dx_1 \ + \cdots + h(x_\alpha, y_\alpha)\, dx_\alpha = 0.$$

Furthermore, $x_1, \ldots, x_\alpha, y_1, \ldots, y_\alpha$ satisfy the g relations given by (4).

In other words, an "algebraic integral" of the differential equations in the corollary is given by the g equations (4).

[14] Intuitively, h is the number of independent points among $(x_1', y_1'), \ldots, (x_g', y_g')$. The relation between g and h is subtle. The results of [50, Section 4] (proved non-constructively) imply that $h = g$ when $\alpha \geq 2g - 1$.

Example 9 Continuing with $y^2 = x^3 + 1$ from Example 8, the points (x_1, y_1), (x_2, y_2), and (x_1', y_1') on the curve satisfy

$$x_1' = -x_1 - x_2 + \left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2$$
$$y_1' = Y'(x_1') = \frac{y_1 - y_2}{x_1 - x_2} x_1' + \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}.$$

When (x_1', y_1') is constant as in Abel's Theorem 4, the first line expresses an algebraic relation between x_1 and x_2 (there is only one since $g = 1$ in this case). Then, for the holomorphic differential $\frac{dx}{y}$, the corollary says that

$$\frac{dx_1}{y_1} + \frac{dx_2}{y_2} = 0$$

when (x_1, y_1) and (x_2, y_2) satisfy the algebraic condition that $-x_1 - x_2 + \left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2$ is constant. This is the "algebraic integration" mentioned in the Synopsis to Part II. Geometrically, this condition says that (x_1, y_1) and (x_2, y_2) vary in such a way that the line connecting them always goes through the fixed point (x_1', y_1').

Essay 9.11 Addition on an Elliptic Curve

The well-known operation of "addition" on an elliptic curve can be understood in the following way.

An elliptic curve can be defined as one with a defining equation of the form $y^2 = f(x)$ in which $f(x)$ is a polynomial of degree 3 or 4 with algebraic number coefficients and distinct roots. If $f(x)$ has degree 4, the defining equation can be recast as one in which $f(x)$ has degree 3 in as follows. Let x_0 be one of the four roots of $f(x)$ in a splitting field and let $x' = x - x_0$. Then $f(x' + x_0)$ has degree 4 in x', and zero is one of its roots. Therefore, there is no loss of generality in assuming that 0 is a root of $f(x)$. Division of $y^2 = f(x)$ by x^4 then gives $v^2 = g(u)$ where $u = \frac{1}{x}$, $v = \frac{y}{x^2}$, and $g(u)$ has degree 3. Therefore, there is no loss of generality in assuming that a given elliptic curve is presented in the form $y^2 = f(x)$ in which $f(x)$ has degree 3 and algebraic number coefficients.[15]

Given two points P and Q in the xy-plane that lie on $y^2 = f(x)$, the line through them intersects $y^2 = f(x)$ in a third point S on $y^2 = f(x)$ (unless the two points are equal, in which case the line is the tangent line at the point, or are unequal with the same x-coordinate, in which case the line is vertical and, by convention,

[15] There is a similar simplification for a hyperelliptic curve—that is, of a curve of the form $y^2 = f(x)$, where $f(x)$ is a polynomial with algebraic number coefficients whose roots are distinct. In this case, too, if the degree of $f(x)$ is $2n$, the curve can also be presented—in an analogous way—as the curve field determined by $y^2 = f(x)$, where the degree of $f(x)$ is $2n - 1$. The genus is $n - 1$ and a basis of the holomorphic differentials is given by $\frac{x^i dx}{y}$ for $i = 0, 1, \ldots, n - 2$ (see Example 6 in Essay 4.5 and Example 8 in Essay 4.6).

the third point of intersection is considered to be the unique point at infinity, where $x = \infty$). Then the second point in which the vertical line through S intersects the curve $y^2 = f(x)$ is the sum of P and Q. The point at infinity is denoted O and the sum of P and Q is denoted $P + Q$.[16] In Figure 4.3 of Essay 4.2, $P + Q$ is the point R.

This description of the addition operation gives no insight into the remarkable and useful fact that this binary operation on the points of $y^2 = f(x)$ makes them an Abelian group with O as the additive identity. Most significantly, it leads to no obvious proof that the associative law $(P + Q) + R = P + (Q + R)$ holds. But the best definition of the addition operation and the best proof of its associativity relies on properties of the *curve field* of the elliptic curve and its *divisors* (as defined Essay 8.12). A crucial fact is that the curve field has genus 1 in this case.

Recall from Essay 8.12 that the divisor of a nonzero quantity z in a curve field is the formal quotient $\frac{P_1 \cdots P_m}{Q_1 \cdots Q_m}$ where $P_1, \ldots, P_m$ and $Q_1, \ldots, Q_m$ are the zeros and poles respectively of z, with repetitions determined by their multiplicities.

Lemma 3 *If the divisors of two quantities in the curve field $\mathcal{K}$ of an elliptic curve share the same points with at most one exception, then the divisors are equal.*

Proof Let z and z' be quantities in $\mathcal{K}$ with divisors $\frac{P_1 \cdots P_{m-1} P}{Q_1 \cdots Q_m}$ and $\frac{P_1 \cdots P_{m-1} P'}{Q_1 \cdots Q_m}$. Then the divisor of $w = z/z'$ is

$$\frac{P_1 \cdots P_{m-1} P}{Q_1 \cdots Q_m} \cdot \frac{Q_1 \cdots Q_m}{P_1 \cdots P_{m-1} P'} = \frac{P}{P'}.$$

When w is a parameter, the number of zeros (counted with multiplicity) is the degree of $\mathcal{K}_0(w) \subset \mathcal{K}$ by Proposition 1 of Essay 8.10. However, $\mathcal{K}_0(w) \neq \mathcal{K}$ since $\mathcal{K}$ has genus 1 and $\mathcal{K}_0(w)$ has genus 0.[17] Therefore, the degree (and hence the number of zeros) is greater than 1. Since the divisor of w is $\frac{P}{P'}$, it follows that w is constant and $P = P'$, as was to be shown.

When the possible exception occurs in the denominator, replace the quantities with their inverses. The desired conclusion follows from the previous paragraph. □

The geometric description of the sum $P + Q$ can be recast in terms of divisors:

Proposition *Given points P and Q lying on an elliptic curve, the point $P + Q$ is the unique point such that $\frac{PQ}{O(P+Q)}$ is the divisor of a quantity in the curve field.*

Proof First suppose that P and Q are distinct points lying on the non-vertical line $y = ax + b$. Then $y - ax - b$, regarded as a quantity in the curve field, has divisor $\frac{PQS}{O^3}$ because $y - ax - b$ has zeros at P, Q, and S, and its only possible pole is the point O where $x = \infty$, necessarily a triple pole since there are three zeros. Also, if c is the x-coordinate of S, then $x - c$ has zeros at S and $P + Q$ and a pole at O, so its divisor is $\frac{S(P+Q)}{O^2}$. Therefore, the divisor of $\frac{y-ax-b}{x-c}$ is

[16] The field of constants of the curve field will always be extended to include the coordinates of all points under consideration. Thus, for P and Q as above, the line $y = ax + b$ through them gives the quantity $y - ax - b$ in the curve field whose zeros are P, Q, and S.

[17] $\mathcal{K}_0(w)$ is the curve field defined by $\phi(w, y) = y$, with normal basis $y_1 = 1$ and $\mu_1 = 0$. Thus, the genus is 0 by the genus formula in Corollary 1 of Essay 9.5.

$$\frac{PQS}{O^3} \cdot \frac{O^2}{S(P+Q)} = \frac{PQ}{O(P+Q)}.$$

The uniqueness of $P + Q$ follows from the lemma just proved.

The other cases ($P = Q$ with nonvertical tangent line, $P \neq Q$ but on the same vertical line, $P = O$, etc.) are handled similarly. □

For points P, Q, and R on the curve, by the proposition, there are quantities in the curve field whose divisors are $\frac{PQ}{O(P+Q)}$ and $\frac{(P+Q)R}{O((P+Q)+R)}$. The product of the two has divisor $\frac{PQ(P+Q)R}{O^2(P+Q)((P+Q)+R)} = \frac{PQR}{O^2((P+Q)+R)}$. But there are also quantities with divisors $\frac{QR}{O(Q+R)}$ and $\frac{P(Q+R)}{O(P+(Q+R))}$, so that $\frac{PQR}{O^2(P+(Q+R))}$ is the divisor of a quantity in the curve field. Lemma 3 implies that $(P+Q)+R = P+(Q+R)$. Thus, addition is associative, proving that $P + Q + R$ does not rely on how the points are grouped.

Corollary *When a line $y = ax + b$ intersects the elliptic curve $y^2 = f(x)$ in points P, Q, S, these points satisfy $P + Q + S = O$.*

Proof Associativity and the proposition imply that $\frac{(P+Q)S}{O(P+Q+S)}$ is the divisor of a quantity in the curve field, and the proof of the proposition shows that the same is true for $\frac{S(P+Q)}{O^2} = \frac{(P+Q)S}{O \cdot O}$. Thus, $P + Q + S = O$ by Lemma 3. □

Given points $P_1, \ldots, P_m$ (with repetitions allowed) that lie on the elliptic curve, the proposition constructs quantities in the curve field with divisors

$$\frac{P_1P_2}{O(P_1+P_2)}, \frac{(P_1+P_2)P_3}{O(P_1+P_2+P_3)}, \frac{(P_1+P_2+P_3)P_4}{O(P_1+P_2+P_3+P_4)}, \ldots,$$

so that

$$\begin{aligned} &\frac{P_1P_2}{O(P_1+P_2)} \cdot \frac{(P_1+P_2)P_3}{O(P_1+P_2+P_3)} \cdot \frac{(P_1+P_2+P_3)P_4}{O(P_1+P_2+P_3+P_4)} \cdots \\ =&\frac{P_1 \cdots P_m}{O^{m-1}(P_1+\cdots+P_m)} \end{aligned} \tag{1}$$

is the divisor of a quantity in the curve. This leads to the following general result about divisors and addition:

Theorem *Given points $P_1, \ldots, P_m$ and $Q_1, \ldots, Q_m$ (with repetitions allowed) that lie on the elliptic curve, the divisor*

$$\frac{P_1 \cdots P_m}{Q_1 \cdots Q_m}$$

is the divisor of a quantity in the curve field if and only if

$$P_1 + \cdots + P_m = Q_1 + \cdots + Q_m.$$

Proof If $\frac{P_1 \cdots P_m}{Q_1 \cdots Q_m}$ is the divisor of a quantity in the curve field, then so is

$$\frac{P_1\cdots P_m}{Q_1\cdots Q_m}\cdot\frac{Q_1\cdots Q_m}{O^{m-1}(Q_1+\cdots+Q_m)}=\frac{P_1\cdots P_m}{O^{m-1}(Q_1+\cdots+Q_m)}$$

by (1) applied to $Q_1,\ldots,Q_m$. Using (1) with $P_1,\ldots,P_m$ and Lemma 3, it follows that $P_1+\cdots+P_m=Q_1+\cdots+Q_m$.

Conversely, assume that $P_1+\cdots+P_m=Q_1+\cdots+Q_m$. Then (1) implies that

$$\frac{P_1\cdots P_m}{O^{m-1}(P_1+\cdots+P_m)}\text{ and }\frac{Q_1\cdots Q_m}{O^{m-1}(Q_1+\cdots+Q_m)}=\frac{Q_1\cdots Q_m}{O^{m-1}(P_1+\cdots+P_m)}$$

are divisors of quantities in the curve field. Hence the same is true for their quotient, which is $\frac{P_1\cdots P_m}{Q_1\cdots Q_m}$. □

The theorem captures the relation between divisors and addition on an elliptic curve but seems far removed from the theorems proved in earlier essays of Chapter 9. The strong link can be seen as follows. Since $y^2=f(x)$ has degree $n=2$ in y, the equations

$$y^2=f(x)\ \text{ and }\ \Theta_\nu(x,y)=0$$

have 2ν solutions $P_i=(x_i,y_i)$ that were constructed and studied in Essay 9.6. These are the zeros of $\Theta_\nu(x,y)$ when regarded as a quantity in the curve field, and they all have multiplicity 1. Also, as shown in the proof of Lemma 1 of Essay 9.6, the poles of $\Theta_\nu(x,y)$ occur where $x=\infty$. Since O is the only point on $y^2=f(x)$ with $x=\infty$, the divisor

$$\frac{P_1\cdots P_{2\nu}}{O^{2\nu}}.$$

is the divisor of $\Theta_\nu(x,y)$ as a quantity in the curve field. By the theorem,

$$(2)\qquad (x_1,y_1)+\cdots+(x_{2\nu},y_{2\nu})=P_1+\cdots+P_{2\nu}=\underbrace{O+\cdots+O}_{2\nu\text{ times}}=O$$

since O is the additive identity.

Example 10 The elliptic curve $y^2=x^3+1$ has been used in several examples to illustrate theorems proved in this chapter. Some of the key ideas used in the proofs have a direct connection to the addition law in the case of an elliptic curve.

When $\nu=2$, $\Theta_2(x,y)=x^2+ax+b+cy$, and the solutions (x_1,y_1), (x_2,y_2), (x_3,y_3), and (x_4,y_4) of the equations

$$y^2=x^3+1\ \text{ and }\ x^2+ax+b+cy=0$$

were introduced in Example 3 of Essay 9.6. The parameter change in Essay 9.9 implies that a,b,c can be expressed in terms of (x_1,y_1), (x_2,y_2), (x_3,y_3). Explicit formulas were computed in Example 6 of Essay 9.8, where it was noted that (x_4,y_4) can be expressed in terms of (x_1,y_1), (x_2,y_2), (x_3,y_3). By (2), the addition law gives the explicit formula

$$(3) \qquad (x_4, y_4) = -(x_1, y_1) - (x_2, y_2) - (x_3, y_3),$$

confirming the claim made at the end of Example 6.

Preliminary Lemma 2 of Essay 9.9 used Abel's idea of replacing some of the (x_i, y_i) with constant solutions in order to get results that apply to an arbitrary number of solutions. In Example 7 of Essay 9.9, this was implemented for $y^2 = x^3 + 1$ by letting (x_1, y_1) be the point O at infinity, with the result that (3) becomes

$$(x_1', y_1') = -O - (x_2, y_2) - (x_3, y_3) = -(x_2, y_2) - (x_3, y_3).$$

(Since (x_1, y_1), (x_2, y_2), and (x_3, y_3) are independent, replacing (x_1, y_1) with O has no effect on (x_2, y_2) and (x_3, y_3) but changes (x_4, y_4) to (x_1', y_1').) Note also that in Example 7, the solutions (x_1', y_1'), (x_2, y_2), and (x_3, y_3) lie on a line, so that the above equation is an immediate consequence of the corollary proved in this essay.

Finally, the proof of Abel's Theorem 3 in Essay 9.10 used another idea due to Abel in order to change the minus signs in Preliminary Lemma 2 into the plus signs in Abel's Theorem 3. In Example 8 of Essay 9.10, this was implemented for $y^2 = x^3 + 1$ by replacing $S = (x_1', y_1')$ with $(x_1', -y_1')$, which is the second point where the vertical line through S intersects the curve $y^2 = x^3 + 1$. It follows that $(x_1', -y_1')$ is the sum of the points P and Q as defined in this essay. Example 8 relabels P, Q, and $(x_1', -y_1')$ as (x_1, y_1), (x_2, y_2), and (x_3, y_3) respectively, so that the formula given in Example 8 is *precisely* the addition law $(x_3, y_3) = (x_1, y_1) + (x_2, y_2)$.

Example 10 shows that the ideas of Abel's Paris Memoir [2], when adapted to elliptic curves, lead naturally to the addition law. Yet these ideas apply to all algebraic curves, which is nothing short of amazing. There is much in [2] that is unclear and in need of clarification. Nonetheless, Abel clearly had a profound understanding of the phenomena involved but did not have time to work out his thoughts in detail. The Synopsis to Part II quoted the poignant last sentence of his last paper [1]:

> I intend on another occasion to develop the numerous applications of this theorem, which will cast light on the nature of the transcendental functions in question.

The essays in this chapter have endeavored to "cast light" on Abel's theorem.

Endnotes to Chapter 9

Endnote 9.1 (Partial Fractions) In the setting of Essay 9.4, consider a differential $R(x)\,dx$ for a rational function $R(x)$ with coefficients in a c-field $\mathcal{K}$. There are constructive methods (see [13, Chapter 2]) for finding a rational function $S(x)$ and polynomials $A(x)$, $B(x)$, all with coefficients in $\mathcal{K}$, with the following properties:

1. $A(x)$ and $B(x)$ are relatively prime with $\deg A(x) < \deg B(x)$.
2. $B(x)$ has no multiple roots in a splitting field, i.e., $B(x)$ is relatively prime to its derivative $\frac{dB}{dx}$.
3. $R(x) = \dfrac{d}{dx}S(x) + \dfrac{A(x)}{B(x)}$.

Furthermore, if $\alpha_1, \ldots, \alpha_m$ are the roots of $B(x)$ in a splitting field, then

$$\frac{A(x)}{B(x)} = \sum_{i=1}^{m} \gamma_i \frac{1}{x - \alpha_i}, \quad \gamma_i = \frac{A(\alpha_i)}{\frac{dB}{dx}(\alpha_i)}$$

by a classic formula of Hermite. Thus,

$$R(x) = \frac{d}{dx}S(x) + \sum_{i=1}^{m} \gamma_i \frac{1}{x - \alpha_i},$$

which implies

$$\int R(x)\,dx = S(x) + \sum_{i=1}^{m} \gamma_i \log(x - \alpha_i) + C.$$

To connect this with the formulas in Endnote 9.2, note that the above formula for $R(x)$ can be written

$$R(x) = \frac{d}{dx}S(x) + \sum_{i=1}^{m} \gamma_i \frac{\frac{dg_i}{dx}(x)}{g_i(x)}$$

where $g_i(x) = x - \alpha_i$. There are more compact versions of this decomposition that allow the g_i to have higher degree (see [13, Chapter 2]).

Endnote 9.2 (Closed Rational Differentials) Let $\sum_{j=1}^{N} R_j(a_1, \ldots, a_N)\,da_j$ be a closed rational differential with coefficients in a c-field $\mathcal{K}_0$. Thus

$$\frac{\partial R_j}{\partial a_k} = \frac{\partial R_k}{\partial a_j}$$

for all $j < k$. In this situation, Theorem 8 of [17] implies that there is a finite extension $\mathcal{K}_0'$ of $\mathcal{K}_0$, rational functions $g_0, g_1, \ldots, g_L$ of $a_1, \ldots, a_N$, and constants $\gamma_1, \ldots, \gamma_L$ such that:

1. g_0 has coefficients in $\mathcal{K}_0$, $g_1, \ldots, g_L$ have coefficients in $\mathcal{K}_0'$, and $\gamma_1, \ldots, \gamma_L$ also lie in $\mathcal{K}_0'$.

2. $R_j = \dfrac{\partial g_0}{\partial a_j} + \displaystyle\sum_{k=1}^{L} \gamma_k \dfrac{\frac{\partial g_k}{\partial a_j}}{g_k}$ for every $j = 1, \ldots, N$.

If one formally defines the "rational and logarithmic" function

$$v = g_0(a_1, \ldots, a_N) + \sum_{k=1}^{L} \gamma_k \log\bigl(g_k(a_1, \ldots, a_N)\bigr),$$

then the above formula for R_j implies that $dv = \sum_{j=1}^{N} R_j(a_1, \ldots, a_N)\, da_j$.

The proof in [17] is written in the language of algebraically closed fields, but the argument can be adapted to give a constructive proof that applies to a c-field $\mathcal{K}_0$.