A BRIEF ACCOUNT OF THE HISTORY OF EXACT FORMULAE IN ARITHMETIC GEOMETRY

JOHN COATES

1. Introduction

The exact formulae of arithmetic geometry, most of which are still largely conjectural, are unquestionably some of the most mysterious, beautiful, and important questions in mathematics today. In these three lectures, I want to discuss the history of these exact formulae and the work which has been done on them in the past. Specifically, I intend to discuss the history of some key discoveries related to the following questions:- (i) The procedure of infinite descent, (ii) Dirichlet's class number formula, (iii) Kummer's work on cyclotomic fields and Iwasawa's profound extension of it, (iv) the conjecture of Birch and Tate, and (v) the conjecture of Birch and Swinnerton-Dyer. I will concentrate on the first key examples, rather than discussing the more general theorems which followed.

2. Lecture 1

In this first lecture, I will briefly discuss several major discoveries made before 1900, which were enormously influential in all subsequent work on the exact formulae of arithmetic geometry made in the 20th century.

The first is Fermat's discovery in the 1630's of what became known as his method of infinite descent, which enabled him to prove that there is no right-angled triangle of area 1, all of whose sides have lengths in $\mathbb{Q}$, the field of rational numbers. His argument can be explained to high school students, and I will sketch it now, leaving you to fill in the details. Denote a right-angled triangle with side lengths a, b, c (c will be the length of the hypothenuse) by $[a, b, c]$. We say the triangle $[a, b, c]$ is *primitive* if a, b, c are all integers with greatest common divisor 1. If $[a, b, c]$ is primitive, it is not difficult to see that one of a and b, say b, must be even, and that there then exist relatively prime positive integers m, n with

$$a = n^2 - m^2, b = 2nm, c = n^2 + m^2.$$

Suppose now that there exists a primitive $\Delta_1 = [a_1, b_1, c_1]$ whose area is the square of an integer. Hence, by the previous remark, there exist relatively prime positive integers m_1, n_1 such that

$$a_1 = {n_1}^2 - {m_1}^2,\ b_1 = 2n_1m_1,\ c_1 = {n_1}^2 + {m_1}^2.$$

Now the area of Δ_1 is equal to $n_1m_1(n_1+m_1)(n_1-m_1)$. We next claim that all of these last four factors are relatively prime in pairs. Because n_1 and m_1 are relatively prime, this only needs checking for the pair $n_1 + m_1$ and $n_1 - m_1$, and for this pair any common divisor must divide both $2n_1$ and $2m_1$, which is impossible because a_1, b_1, c_1 do not have 2 as a common divisor. Hence there exist positive integers t, w, u, v such that $n_1 = t^2, m_1 = w^2, n_1 + m_1 = u^2, n_1 - m_1 = v^2$. Since a_1 is odd, we see that both u and v are odd. Also we have

$$2w^2 = (u + v)(u - v). \tag{2.1}$$

Now u and v are relatively prime, and so the greatest common divisor of $u + v$ and $u - v$ must be 2. Thus the equation (2.1) implies that one of $u + v$ and $u - v$ must be of the form $2r^2$ and the other must be of the form $4s^2$ for some positive integers r and s. It follows that $u = r^2 + 2s^2$ and $\pm v = r^2 - 2s^2$, whence

$$t^2 = (u^2 + v^2)/2 = r^4 + 4s^4. \tag{2.2}$$

But then we have a right-angled triangle $\Delta_2 = [r^2, 2s^2, t]$, with sides of integral length, whose area is the square of the integer rs, and whose hypothenuse has length t, which is strictly less than the length $c_1 = t^4 + w^4$ of the hypothenuse of our initial triangle Δ_1. We can now repeat the argument with Δ_2 (divide out the greatest common divisor of the lengths of its sides to get a primitive triangle) instead of Δ_1, and so on. The hypothenuses of the triangles obtained by this descent procedure form a strictly decreasing sequence of positive integers, and so we obtain a contradiction from this "infinite descent".

Fermat himself noted that his proof also showed that, in the special case $n = 4$, the equation

$$x^n + y^n = z^n \tag{2.3}$$

has no solution in integers x, y, z with $xyz \neq 0$. Indeed, if there was such a non-trivial solution of (2.3), when $n = 4$, then defining $a = z^4 - x^4$, $b = 2x^2z^2$, $c = x^4 + z^4$, we would obtain a right-angled triangle $[a, b, c]$ with area $x^2y^4z^2$, contradicting Fermat's proof above. He claimed that the equation (2.3) has no solution in integers x, y, z with $xyz \neq 0$ for all integers $n \geq 3$, and attempts to prove this conjecture was one of the main driving forces behind the great developments in number theory made in Germany in the 19th century.

Fermat had presumably come across the problem of proving that there is no right-angled triangle of area 1 whose sides all have rational lengths, because Arabic, and probably Asian mathematicians, were interested in the general question of which positive integers N are the area of a right-angled triangle all of whose sides have *rational* length. The smallest integer N with this property which the ancients had found was $N = 5$ (for example, the triangle $[9/6, 40/6, 41/6]$ has area 5). If N is the area of a rational right-angled triangle, we will follow the ancients and say N is a *congruent number*. As must have been known to the ancients, N will be a congruent number if and only if the elliptic curve

$$C^{(N)} : y^2 = x^3 - N^2x \tag{2.4}$$

has a point (x, y) with $x, y \in \mathbb{Q}$ and $y \neq 0$. While Fermat's argument of infinite descent proves that 1 is not a congruent number, no algorithm has ever been proven that will infallibly show in a finite number of steps whether or not an arbitrary positive integer N is a congruent number. We shall discuss more fully the fundamental difficulty which arises in Lecture 2. There was a second mystery about congruent numbers which appeared in the ancient numerical tables. These tables showed that $5, 6, 7, 13, 14, 15, 21, 22, 23, \ldots$ are all congruent numbers, and it appeared that all positive integers of the form $8n+5, 8n+6, 8n+7$ ($n = 0, 1, \ldots$) are always congruent numbers. There are some positive integers which are congruent numbers which are not of this form, the smallest of which is 34, but these exceptions are few and far between. From the point of view of elementary arithmetic this seems utterly mysterious, and the assertion remains unproven today. However, as we shall explain later, this mystery is explained by the conjecture of Birch and Swinnerton-Dyer relating the statement to the value at $s = 1$ of the complex L-series of the elliptic curve (2.4)

We now turn to the work of Dirichlet. After some early work on (2.3) (in particular, he was the first to prove the Fermat conjecture for $n = 5$), Dirichlet in 1837 made the first great discovery linking purely arithmetic questions to special values of L-functions. Let p be any prime > 3 such that $p \equiv 3 \bmod 4$, define $K = \mathbb{Q}(\sqrt{-p})$, and let h denote the class number of K. Let R (resp. N) denote the number of quadratic residues (resp. quadratic non-residues) modulo p lying in the set $\{1, 2, \ldots, (p-1)/2\}$.

Theorem 2.1. *(Dirichlet) For all primes $p \equiv 7 \bmod 8$, we have $h = R - N$, and for all primes $p > 3$ with $p \equiv 3 \bmod 8$, we have $h = (R - N)/3$.*

An essential part of Dirichlet's proof was based on his introduction of what he called *L-series*. Let $\chi(n) = (n/p)$ be the quadratic character modulo p, and define

$$L(\chi, s) = \prod_{q \neq p} (1 - \chi(q)/q^s)^{-1},$$

where the product is taken over all primes $q \neq p$, and s is a complex variable with the real part > 1. Dirichlet proved that $L(\chi, s)$ has an entire analytic continuation, and proved that $L(\chi, 0) = R - N$ or $(R - N)/3$, according as the prime $p > 3$ satisfies $p \equiv 7$ or $3 \bmod 8$. Thus his class number formula amounts to the assertion that

$$h = L(\chi, 0).$$

Note that Dirichlet's theorem implies, in particular, that $R > N$ for all primes $p > 3$ with $p \equiv 3 \bmod 4$. We remark that no purely arithmetic proof, which does not involve L-values, of this seemingly simple assertion has ever been found.

The second great discovery of a mysterious link between special values of L-series and arithmetic was made by Kummer, who had been a school teacher until Dirichlet and Jacobi helped him to be appointed to a professorship at the University of Breslau. Kummer had also been working on the Fermat equation (2.3) for $n = p$ any odd prime, using the newly discovered factorization ideas in the ring of algebraic integers of the field $F = \mathbb{Q}(\rho)$, where ρ denotes a primitive p-th root of unity. Kummer's work led him to break up the set of all odd prime numbers into two classes. He defined an odd prime p to be *regular*, or

irregular, according as p does not, or does, divide the class number of the field F. For example, $p = 37$ is the first irregular prime. For regular primes p, Kummer was able to settle the Fermat conjecture.

Theorem 2.2. *(Kummer) If p is an odd regular prime, the equation $x^p + y^p = z^p$ has no solution in integers x, y, z with $xyz \neq 0$.*

Kummer's ideas were subsequently taken up by many other mathematicians in the latter part of the 19th century, but with no success for proving the Fermat conjecture for all irregular primes p.

Even today using high speed computers, it becomes prohibitively difficult numerically to compute the class number of number fields of even moderately large degree over $\mathbb{Q}$, let alone in the time of Kummer when only hand computations were possible. Thus, in applying the above theorem to numerical examples, Kummer was confronted with the problem of deciding which odd prime numbers p are regular. Presumably, it was this fact which led him to perhaps his most striking discovery. We recall that the Riemann zeta function is defined, for the real part of the complex variable $s > 1$, by the Euler product

$$\zeta(s) = \prod_q (1 - 1/q^s)^{-1}$$

where the product is taken over all primes q. It was well known that $\zeta(s)$ has an analytic continuation over the whole complex plane, apart from a simple pole at $s = 1$. Moreover, it had been proven by Euler in the 18th century that

$$\zeta(-n) = -B_{n+1}/(n+1) \quad (n = 1, 3, 5, ...),$$

where the Bernoulli numbers B_n are the rational numbers defined by the expansion

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n t^n / n! \ .$$

Theorem 2.3. *(Kummer) The odd prime p is regular if and only if p does not divide the numerator of any of the rational numbers $\zeta(-1), \zeta(-3), \ldots, \zeta(4-p)$.*

Using this theorem, it is easy numerically to make a list of regular primes up to a rather large limit (in fact, today all regular primes up to 12×10^6 are known). For example, since

$$\zeta(-31) = \frac{37 \times 208360028141}{16320},$$

it follows that 37 is indeed an irregular prime. Today, it still has not been proven that there are infinitely many regular primes, although the numerical data strongly suggests that about 60.61% of all primes are regular.

3. Lecture 2

We now begin discussing 20th century work which grew out of these great earlier discoveries of Fermat, Dirichlet and Kummer.

We recall that an elliptic curve E over $\mathbb{Q}$ is a curve of genus 1 defined over $\mathbb{Q}$ with a given rational point on it. The set $E(\mathbb{Q})$ of rational points on E then has a natural abelian group law on it, with the given rational point being the zero element. By the Riemann–Roch theorem, we can always find an equation for E of the form

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6, \tag{3.1}$$

where the coefficients $a_i \in \mathbb{Z}$.

Theorem 3.1. *(Mordell) $E(\mathbb{Q})$ is a finitely generated abelian group.*

Mordell's proof is deceptively simple, and we sketch a not essentially different later version of it. If $u = m/n$ is a rational number with $m, n \in \mathbb{Z}$ and relatively prime, we define $H(u) = log\, max(|m|, |n|)$. Then Neron and Tate showed that there exists a unique real-valued function H on $E(\mathbb{Q})$ such that $H(2P) = 4H(P)$ and $H(P) - H(x(P))$ is bounded for P running over $E(\mathbb{Q})$. One then shows that, if m is any integer ≥ 2 and the P_i form a set of coset representatives of $mE(\mathbb{Q})$ in $E(\mathbb{Q})$, and if we can find a real number b such that $H(P_i) < b$ for all i, then $E(\mathbb{Q})$ is generated by all P such that $H(P) < b$. Thus Mordell's theorem will follow if we can prove that $E(\mathbb{Q})/mE(\mathbb{Q})$ is finite. As we now briefly explain, it is in the proof of this last assertion where probably the most mysterious group in the whole of mathematics inextricably arises. Let p be any prime, and k a positive integer. Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$,

and write $G_{\mathbb{Q}}$ for the Galois group of $\overline{\mathbb{Q}}$ over $\mathbb{Q}$. Then multiplication by p^k gives the exact sequence of $G_{\mathbb{Q}}$-modules

$$0 \to E_{p^k} \to E(\overline{\mathbb{Q}}) \to E(\overline{\mathbb{Q}}) \to 0,$$

whence, taking $G_{\mathbb{Q}}$-cohomology, we obtain the exact sequence

$$0 \to E(\mathbb{Q})/p^k E(\mathbb{Q}) \to H^1(G_{\mathbb{Q}}, E_{p^k}) \to (H^1(G_{\mathbb{Q}}, E(\overline{\mathbb{Q}})))_{p^k} \to 0. \tag{3.2}$$

We now define a subgroup $Sel(E_{p^k}/\mathbb{Q})$ of $H^1(G_{\mathbb{Q}}, E_{p^k})$ by the following local considerations. For each finite or infinite prime v of $\mathbb{Q}$, let $\mathbb{Q}_v$ denote the completion of $\mathbb{Q}$ at v, and write G_v for its absolute Galois group. We have a local exact sequence entirely analogous to (3.2) with $\mathbb{Q}$ replaced by $\mathbb{Q}_v$, and $G_{\mathbb{Q}}$ replaced by G_v, and, since we can view G_v as a subgroup of $G_{\mathbb{Q}}$, there is a restriction map $r_{v,k} : H^1(G_{\mathbb{Q}}, E_{p^k}) \to H^1(G_v, E_{p^k})$. We define $Sel(E_{p^k}/\mathbb{Q})$ to be the set of all $z \in H^1(G_{\mathbb{Q}}, E_{p^k})$ such that $r_{v,k}(z) \in E(\mathbb{Q}_v)/p^k E(\mathbb{Q}_v)$ for all places v of $\mathbb{Q}$. Here "*Sel*" is used in honour of the Norwegian mathematician Ernst Selmer, whose ideas systematised earlier work due to Fermat, Mordell, Weil and others. It is then easily seen that we have an exact sequence

$$0 \to E(\mathbb{Q})/p^k E(\mathbb{Q}) \to Sel(E_{p^k}/\mathbb{Q}) \to (Ш(E/\mathbb{Q}))_{p^k} \to 0, \tag{3.3}$$

where

$$Ш(E/\mathbb{Q}) = Ker(H^1(G_{\mathbb{Q}}, E(\overline{\mathbb{Q}})) \to \prod_v H^1(G_v, E(\overline{\mathbb{Q}_v}))$$

is the Tate–Shafarevich group of $E/\mathbb{Q}$. Classical arguments in algebraic number theory then show:-

Theorem 3.2. *For all primes p and all integers $k \geq 1$, $Sel(E_{p^k}/\mathbb{Q})$ is a finite group which can be effectively determined by a theoretical algorithm.*

It should be stressed that the theoretical algorithm referred to in this theorem is of little practical use in numerical examples when $p^k > 5$. Nevertheless, in view of (3.3), this result not only proves theoretically that $E(\mathbb{Q})/p^k E(\mathbb{Q})$ is finite, but in addition shows that $(Ш(E/\mathbb{Q}))_{p^k}$ is finite, for all primes p and all integers $k \geq 1$. However, the mystery is that the proof does not tell us the orders of the two individual parts of $Sel(E_{p^k}/\mathbb{Q})$. Classically, number-theorists tried to overcome this difficulty by carrying out what they called *higher descents*. For each integer $k \geq 1$, multiplication by p^{k-1} defines a $G_{\mathbb{Q}}$-homomorphism from E_{p^k} onto E_p, and this in turn gives rise to a homomorphism $d_{p,k} : Sel(E_{p^k}/\mathbb{Q}) \to Sel(E_p/\mathbb{Q})$. Writing $Sel_{p,k}(E/\mathbb{Q}) =$ image of $d_{p,k}$, we see easily that we then have an exact sequence

$$0 \to E(\mathbb{Q})/pE(\mathbb{Q}) \to Sel_{p,k}(E/\mathbb{Q}) \to p^{k-1}(Ш(E/\mathbb{Q}))_{p^k} \to 0,$$

where again the decreasing sequence of subgroups $Sel_{p,k}(E/\mathbb{Q})$ $(k = 1, 2, \ldots)$ are all theoretically effectively computable. Hence we will be able to find an integer $k \geq 1$ such that $Sel_{p,k}(E/\mathbb{Q}) = E(\mathbb{Q})/pE(\mathbb{Q})$ if and only if the p-primary subgroup $Ш(E/\mathbb{Q})(p)$ of $Ш(E/\mathbb{Q})$ is finite. Here we arrive at one of the remaining great mysteries of number theory. How can we prove that, for every elliptic curve $E/\mathbb{Q}$ there exists at least one prime p such that $Ш(E/\mathbb{Q})(p)$ is finite? Even a proof of this weak statement would have profound implications for some of the most important open classical questions in arithmetic geometry. However, it has long been conjectured that, we have, much more strongly:-

Conjecture 3.3. *For all elliptic curves E, $Ш(E/\mathbb{Q})$ is finite.*

The only general result proven to date about the Tate–Shafarevich group is the following theorem, whose discovery had important consequences for the formulation of the analytic conjecture which we shall discuss next.

Theorem 3.4. *(Cassels) If $Ш(E/\mathbb{Q})$ is finite, then its order must be a perfect square.*

In retrospect, it now seems surprising that about 120 years were to pass from Dirichlet's proof of his wonderful exact formula for the order of the ideal class group of an imaginary field until number-theorists discovered a precise conjectural analogue of it for elliptic curves defined over $\mathbb{Q}$. This was the great discovery of Birch and Swinnerton-Dyer in the early 1960's, which grew out of a series of brilliant numerical experiments on the early EDSAC computers in Cambridge. We now fix any equation for $E/\mathbb{Q}$ of the form (3.1), which is minimal in the sense that the discriminant Δ of this equation is as small as possible in absolute value amongst all such equations. For each prime number p, define N_p to be the integer such that $N_p - 1$ is the number of solutions of the equation (3.1) when viewed as a congruence modulo p. The complex L-series of $E/\mathbb{Q}$ is then defined by the Euler product

$$L(E, s) = \prod_{p|\Delta} (1 - t_p p^{-s})^{-1} \prod_{(p,\Delta)=1} (1 - t_p p^{-s} + p^{1-2s})^{-1}, \tag{3.4}$$

where $t_p = p + 1 - N_p$. This Euler product converges only in the half plane $R(s) > 3/2$, and when Birch and Swinnerton-Dyer [1] carried out their work the entire analytic continuation and functional equation for $L(E, s)$ was only known (thanks to work of Eisenstein, Kronecker, and Deuring) for elliptic curves with complex multiplication, that is for curves E whose ring of $\mathbb{C}$-endomorphisms $End_{\mathbb{C}}(E)$ is an order in an imaginary quadratic field K (this family includes, in particular, the curves $C^{(N)}$ defined by (2.4)). Today, we know the analytic continuation and the following functional equation for $L(E, s)$ for all elliptic curves E over $\mathbb{Q}$, thanks to the deep work of Andrew Wiles, aided by Richard Taylor [21] and many others (see the book [7]). The conductor N_E of E is defined by $N_E = \prod_{p|\Delta} p^{f_p}$, where Δ is the discriminant of the minimal equation (3.1) for E, and $f_p \geq 1$ is an integer arising naturally from the geometry of the minimal equation (3.1) over $\mathbb{Z}_p$. Put

$$\Lambda(E, s) = N_E^{-s/2}(2\pi)^{-s}\Gamma(s)L(E, s).$$

Theorem 3.5. *$\Lambda(E, s)$ is entire, and satisfies the functional equation*

$$\Lambda(E, s) = w_E\Lambda(2 - s), \tag{3.5}$$

where $w_E = \pm 1$ can be computed from the geometry of E at the primes $p|\Delta$.

Moreover, in view of earlier work of Ribet [18], Wiles' theorem also proves at last the Fermat conjecture that, for all integers $n \geq 3$, the equation (2.4) has no solution in integers x, y, z with $xyz \neq 0$. Now let g_E denote the rank of the finitely generated abelian group $E(\mathbb{Q})$.

Conjecture 3.6. *(Birch and Swinnerton-Dyer) $L(E, s)$ has a zero at $s = 1$ of order exactly g_E.*

In particular, this conjecture predicts that if $L(E, 1) \neq 0$, then $E(\mathbb{Q})$ must be finite, and this is known today thanks to ideas which grew out of Iwasawa's beautiful extension of the work of Kummer on cyclotomic fields in the case of elliptic curves with complex multiplication [3], and thanks to the work of Kolyvagin [16] and Gross-Zagier [10] in the case of elliptic curves without complex multiplication. It is still unknown how to prove that $E(\mathbb{Q})$ is finite implies that $L(E, 1) \neq 0$. Note also that the functional equation (3.5) shows that $L(E, s)$ will have a zero at $s = 1$ of *odd* order precisely when $w_E = -1$. Hence Conjecture 3.6 predicts that we must have $E(\mathbb{Q})$ is infinite whenever $w_E = -1$. When E is the congruent number curve $C^{(N)}$, it has long been known that, assuming N is square free, we have $w_{C^{(N)}} = 1$ when $N \equiv 1, 2, 3 \bmod 8$, and $w_{C^{(N)}} = -1$ when $N \equiv 5, 6, 7 \bmod 8$. Thus we see that the Birch-Swinnerton-Dyer conjecture explains conjecturally the ancient mystery of why all square free positive integers $\equiv 5, 6.7 \bmod 8$ seem to be congruent. Unfortunately, at present, it is only known how to prove this very special consequence of the Birch-Swinnerton-Dyer conjecture if we impose the additional hypothesis that, for some prime p, the p-primary subgroup of the Tate-Shafarevich group $Ш(C^{(N)}/\mathbb{Q})$ is finite.

But what of exact formulae in this settting? Making use of deep earlier work of Gross-Zagier, Kolyvagin [16] found a remarkable proof that $Ш(E/\mathbb{Q})$ is finite for those elliptic curves $E/\mathbb{Q}$ with $L(E, 1) \neq 0$, and Rubin [19] subsequently found a different proof of the same result for elliptic curves with complex multiplication using an ingenious earlier idea of Thaine [22]. This leads us to the first key case of the exact formula for the order of $Ш(E/\mathbb{Q})$ conjectured by Birch and Swinnerton-Dyer. Let $\omega_E = dx/(2y+a_1x+a_3)$ be the canonical holomorphic differential attached to the equation (3.1), and write Ω_E for the least positive real period of ω_E. It is known that $L(E, 1)/\Omega_E$ is a rational number.

Conjecture 3.7. *(Birch and Swinnerton-Dyer) If $L(E, 1) \neq 0$, we have*

$$\#(Ш(E/\mathbb{Q})) = \frac{L(E, 1)}{\Omega_E} \times \frac{\#(E(\mathbb{Q}))^2}{t_E}, \text{ where } t_E = c_\infty(E)\prod_{p|\Delta} c_p(E). \tag{3.6}$$

Here the Tamagawa factors $c_v(E)$ are defined as follows. If $v = \infty$, then $c_\infty(E)$ is the number of connected components of $E(\mathbb{R})$, and for $v = p$ a finite prime dividing the discriminant Δ of our minimal equation (3.1), we have $c_p = [E(\mathbb{Q}_p) : E_0(\mathbb{Q}_p)]$, where $E_0(\mathbb{Q}_p)$ is the subgroup consisting of those points whose reduction modulo p is a non-singular point on the curve over $\mathbb{F}_p$ obtained by reducing (3.1) modulo p. Numerically, the right side of (3.6) has been computed for a vast number of elliptic curves $E/\mathbb{Q}$, and happily it has always turned out to be the square of an integer, in accord with the theorem of Cassels. More generally, Birch and Swinnerton-Dyer also conjecture an exact formula for the order of $Ш(E/\mathbb{Q})$ in terms of the leading term of the Taylor expansion of $L(E, s)$ about $s = 1$ for all elliptic curves $E/\mathbb{Q}$. However this formula involves a regulator term coming from the canonical height H on E when $g_e \geq 1$. We also stress that so far $Ш(E/\mathbb{Q})$ has never been proven to be finite for a single elliptic curve $E/\mathbb{Q}$

with $g_E \geq 2$, although the work of Kolyvagin and Gross-Zagier also proves that $Ш(E/\mathbb{Q})$ is finite for those E such that $L(E,s)$ has a simple zero at $s=1$.

We end this lecture with a brief discussion of yet another exact formula in arithmetic geometry, conjectured by Birch and Tate. The interest of this formula is that it suggests a possible bridge between the ideas of Dirichlet, Kummer and Iwasawa on the one hand, and the conjecture of Birch and Swinnerton-Dyer on the other hand. Let J be any field, and write $J^\times$ for the multiplicative group of J. The Milnor K_2 of J is defined by

$$K_2 J = J^\times \otimes_{\mathbb{Z}} J^\times / W,$$

where W is the subgroup of the tensor product generated by all elements $a \otimes b$ with $a+b=1$. If v is any discrete valuation of J with residue field j_v, the formula $\lambda_v(a,b) =$ residue class of $(-1)^{v(a)v(b)} a^{v(b)}/b^{v(a)}$ defines a homomorphism $\lambda_v : K_2 J_v \to j_v^\times$ called *the tame symbol* at v. Suppose now that J is a finite extension of $\mathbb{Q}$, and let $\phi_J : K_2 J \to \prod_v j_v^\times$ be the map given by the tame symbols at all finite places v of J. The tame kernel $R_2 J$ is then defined to be the kernel of the map ϕ_J. By a deep theorem of Garland [9] using methods from differential geometry, $R_2 J$ is a finite group. We recall that the number field J is said to be *totally real* if each embedding of J in $\mathbb{C}$ maps J into $\mathbb{R}$. If J is any number field, we recall that its complex ζ-function $\zeta(J,s)$ is defined for $\mathbb{R}(s) > 1$ by the Euler product

$$\zeta(J,s) = \prod_v (1 - (Nv)^{-s})^{-1},$$

where v runs over all finite places of F, and Nv denotes the cardinality of the residue filed j_v of v. Then $\zeta(F,s)$ has a meromorphic continuation to the whole complex plane, with a simple pole at $s=1$. Moreover, $\zeta(J,s)$ satisfies a simple functional equation relating its values at s and $1-s$. In particular, this functional equation shows that, when J is a totally real number field, we have $\zeta(J,-n) \neq 0$ for all odd positive integers n, and work of Klingen and Siegel shows that these values are rational numbers. Define $w_2(J)$ to be the largest positive integer m such that the extension obtained by adjoining the m-th roots of unity to J has Galois group over J which is annihilated by 2.

Conjecture 3.8. *(Birch and Tate). For all totally real number fields J, the tame kernel $R_2 J$ has order equal to the absolute value of $w_2(J)\zeta(J,-1)$.*

As we shall explain in the third lecture, the ideas of Iwasawa theory, growing out of Iwasawa's beautiful generalization of Kummer's ideas, do enable us to prove weak forms of both Conjectures 3.7 and 3.8. More precisely, one can prove the p-part of Conjecture 3.8 for all odd primes p, and the p-part Conjecture 3.7 for all sufficiently large primes p. However, it should also be stressed that the p-part of Conjecture 3.7 for these exceptional small primes p is often the most interesting in the study of classical problems in number theory.

4. Lecture 3

Let p be any prime, let F_∞ be the field obtained by adjoining all p-power roots of unity to $\mathbb{Q}$, and let $\mathfrak{F}_\infty$ be the maximal real subfield of F_∞. We owe to Iwasawa (see [12], [13]) the great discovery that the p-adic analogue of the Riemann zeta function (the existence of this p-adic analogue had been proven earlier by Leopoldt and Kubota) has a remarkable Galois module theoretic connexion with the Galois group

$$X_\infty = \mathrm{Gal}(M_\infty/\mathfrak{F}_\infty), \tag{4.1}$$

where M_∞ denotes the maximal abelian p-extension of $\mathfrak{F}_\infty$ which is unramified outside the primes above p and ∞. In particular, this new insight into the arithmetic of cyclotomic fields yielded, for the first time, an arithmetic interpretation of the special values $\zeta(-n)$ for $n = 1, 3, ...$, which, thanks to work of Tate [20], boiled down to the Birch–Tate conjecture for $\mathbb{Q}$ when $n=1$. It also gave a simple conceptual explanation for Kummer's criterion (Theorem 2.2) for an odd prime p to be regular. Moreover, it was very soon clear that Iwasawa's remarkable discovery ought to hold in much greater generality, and that such generalizations could then lead to a path for attacking the conjecture of Birch and Tate for all totally real number fields J, and also some of the key cases of the conjecture of Birch and Swinnerton-Dyer for elliptic curves defined over $\mathbb{Q}$. A large body of subsequent work has now born out fully these early hopes.

We first briefly explain Iwasawa's radical new construction of the p-adic analogue of the Riemann zeta function given in [13]. While [13] is based on relatively elementary arguments with the classical Stickelberger ideal, Iwasawa was in fact led to this construction by his earlier deep study in [12] of an

object closely related to the Galois group $X_\infty = \mathrm{Gal}(M_\infty/\mathfrak{F}_\infty)$, which will be discussed in the second part of this lecture. Let

$$G = \mathrm{Gal}(\mathfrak{F}_\infty/\mathbb{Q}). \tag{4.2}$$

By the irreducibility of the p-power cyclotomic equation, the action of $\mathrm{Gal}(F_\infty/\mathbb{Q})$ on the group of all p-power roots of unity gives an isomorphism $\chi : \mathrm{Gal}(F_\infty/\mathbb{Q}) \to \mathbb{Z}_p^\times$, which we call the cyclotomic character. In particular, for every even integer k, χ^k will be a character of G. Iwasawa realized that a basic role is played by the topological $\mathbb{Z}_p$-algebra

$$\Lambda(G) = \varprojlim_U \mathbb{Z}_p[G/U], \tag{4.3}$$

where U runs over the set of open subgroups of G, and $\mathbb{Z}_p[G/U]$ is the $\mathbb{Z}_p$-group ring of the finite group G/U. In fact, it is easily seen we can view $\Lambda(G)$ as the algebra of $\mathbb{Z}_p$-valued measures on G. If f is any continuous valued function on G with values in $\mathbb{Q}_p$, we write $\int_G f d\mu$ for the integral of f against an element μ of $\Lambda(G)$. To take account of the fact that the p-adic analogue ζ_p of the complex Riemann zeta function $\zeta(s)$ also has a pole at $s = 1$, one defines a *pseudo-measure* on G to be any element μ of the ring of fractions at $\Lambda(G)$ such that $(\sigma - 1)\mu$ belongs to $\Lambda(G)$ for all $\sigma \in G$. We recall that, for each even integer $k > 0$, Euler proved that $\zeta(1-k)$ is a rational number, and so it can be viewed as lying in the field $\mathbb{Q}_p$ of p-adic numbers.

Theorem 4.1. *(Iwasawa). There exists a unique pseudo-measure ζ_p on G such that*

$$\int_G \chi^k d\zeta_p = (1 - p^{k-1})\zeta(1-k), \tag{4.4}$$

for all even integers $k \geq 2$.

Of course, we know from the functional equation of the complex zeta function that $\zeta(1-k) \neq 0$ for all even integers $k \geq 2$. When p is an irregular prime, the p-adic zeta function ζ_p will, in fact, always have p-adic zeroes, but, unlike the Riemann Hypothesis for the complex zeta function $\zeta(s)$, no reasonable conjecture has yet been formulated about the possible location of p-adic zeroes of ζ_p. It is not even known at present that the integral on the left hand side of (4.4) is non-zero for all even integers $k < 0$.

To keep notation simple, let us assume from now on that p is an odd prime. As before, let ρ be a primitive p-th root of unity. While most of the earlier work on cyclotomic fields, starting from Kummer, had centred around the study of the arithmetic of the field $F = \mathbb{Q}(\rho)$ and its maximal real subfield $\mathfrak{F} = \mathbb{Q}(\rho + \rho^{-1})$, Iwasawa was the first to realize that one obtained a radically new view if one considered instead the arithmetic of the infinite towers of fields $F_\infty/\mathbb{Q}$ and $\mathfrak{F}_\infty/\mathbb{Q}$. For each integer $n \geq 0$, let ρ_n denote a primitive p^{n+1}-th root of unity, and define

$$F_n = \mathbb{Q}(\rho_n)\,,\ \mathfrak{F}_n = \mathbb{Q}(\rho_n + \rho_n^{-1}). \tag{4.5}$$

Let e be a primitive root mod p such that $e^{p-1} \not\equiv 1 \bmod p^2$, so that e is a primitive root mod p^{n+1} for all integers $n \geq 0$. It is easily seen that

$$c_n = \frac{\rho_n^{-e/2} - \rho_n^{e/2}}{\rho_n^{-1/2} - \rho_n^{1/2}} \tag{4.6}$$

lies in $\mathfrak{F}_n$, and is a unit of the ring of algebraic integers lying in this field. We define the group D_n of *cyclotomic units* of $\mathfrak{F}_n$ to be the subgroup of $\mathfrak{F}_n^\times$ generated by -1, and all conjugates of the elements (4.6) under the action of the Galois group of $\mathfrak{F}_n/\mathbb{Q}$. Now the prime p is totally ramified in the field $\mathfrak{F}_n$, and we write $\mathfrak{p}_n$ for the unique prime of $\mathfrak{F}_n$ lying above it, and Φ_n for the completion of $\mathfrak{F}_n$ at $\mathfrak{p}_n$. Let U_n denote the group of units $\equiv 1 \bmod \mathfrak{p}_n$ of the ring of integers of the completion of $\mathfrak{F}_n$ at $\mathfrak{p}_n$. We then define C_n to be the subgroup of all elements of D_n which are $\equiv 1 \bmod \mathfrak{p}_n$, and let $\overline{C}_n$ be the closure of the image of C_n in U_n in the $\mathfrak{p}_n$-adic topology (equivalently, $\overline{C}_n$ is the $\mathbb{Z}_p$-submodule generated by the image of C_n in U_n). Now for all $m \geq n$, $\mathfrak{F}_m$ (resp. Φ_m) is a Galois extension of $\mathfrak{F}_n$ (resp. Φ_n) of degree p^{m-n}, and we write $N_{m,n}$ for the corresponding global (resp. local) norm map for this extension. Now Φ_m/Φ_n is a totally ramified cyclic extension, whence $N_{m,n}(U_m) = U_n$. Moreover, $N_{m,n}(\overline{C}_m) = \overline{C}_n$ because $N_{m,n}(c_m) = c_n$. Taking projective limits with respect to the norm maps, we then define

$$U_\infty = \varprojlim_n U_n\,,\ \overline{C}_\infty = \varprojlim_n \overline{C}_n. \tag{4.7}$$

Now, as $\overline{C}_n \subset U_n$ are both $\mathbb{Z}_p$-modules which are stable under the action of $\mathrm{Gal}(\Phi_n/\mathbb{Q}_p) = \mathrm{Gal}(\mathfrak{F}_n/\mathbb{Q})$ for all $n \geq 0$, it follows that $\overline{C}_\infty$ is a $\Lambda(G)$-submodule of the $\Lambda(G)$-module U_∞, and so $U_\infty/\overline{C}_\infty$ has a natural

structure as a $\Lambda(G)$-module. Now there is a natural augmentation map $\Lambda(G) \to \mathbb{Z}_p$, and we define $I(G)$ to be the kernel of this augmentation map. Since ζ_p is a pseudo-measure, we have $\zeta_p I(G) \subset \Lambda(G)$.

Theorem 4.2. *(Iwasawa). For all primes p, $U_\infty/\bar{C}_\infty$ is canonically isomorphic as a $\Lambda(G)$-module to $\Lambda(G)/\zeta_p I(G)$.*

Iwasawa's original proof of this theorem was highly ingenious, and depended on mysterious explicit choices of elements in the field $\mathfrak{F}_\infty$. Faced with the problem of generalizing this theorem to elliptic curves with complex multiplication, Wiles and I discovered a much simpler proof which starts from the simple observation that

$$g_e(T) = \frac{(1+T)^{-e/2} - (1+T)^{e/2}}{(1+T)^{-1/2} - (1+T)^{1/2}}$$

is a formal power series in $\mathbb{Z}_p[[T]]$ with the property that $g_e(\rho_n - 1) = c_n$ for all integers $n \geq 0$. We realized that, more generally, for any $u_\infty = (u_n)$ in U_∞, there exists a power series $g_{u_\infty}(T)$ in $\mathbb{Z}_p[[T]]$ such that $g_{u_\infty}(\rho_n - 1) = u_n$ for all $n \geq 0$. It was then possible to give a simpler proof of Theorem 4.2 using these power series, which also could be generalized to elliptic curves with complex multiplication, [4]. Coleman was a Master's student in Cambridge at the time, and he shortly afterwards found a much more beautiful proof of the existence of these interpolating power series $g_{u_\infty}(T)$ which worked more generally for the local division towers arising from arbitrary Lubin–Tate formal groups [6]

It is hard to exaggerate the large and varied body of research in arithmetic geometry which has grown, and is continuing to grow, out of Iwasawa's discovery of Theorem 4.2. We only have time to mention a few of the early discoveries here. Firstly, within the theory of cyclotomic fields, what one really needs, to attack for example the Birch–Tate conjecture, is an analogue of Theorem 4.2 for the Galois group X_∞ rather than $U_\infty/\bar{C}_\infty$. For each $n \geq 0$, let $\mathfrak{E}_n$ be the group of global units of $\mathfrak{F}_n$, which are $\equiv 1 \bmod \mathfrak{p}_n$, and let $\overline{\mathfrak{E}}_n$ be the closure in the $\mathfrak{p}_n$-adic topology of the image of $\mathfrak{E}_n$ in U_n. We then define $\overline{\mathfrak{E}}_\infty$ to be the projective limit of the $\overline{\mathfrak{E}}_n$ with respect to the norm maps. Also, let $\mathfrak{A}_n$ denote the p-primary subgroup of the ideal class group of $\mathfrak{F}_n$, and define $\mathfrak{A}_\infty = \varprojlim_n \mathfrak{A}_n$ where the projective limit is again taken with respect to the norm maps. Recall that X_∞ is the Galois group defined by (4.1). Thanks to Artin's global reciprocity law, we then have the exact sequence

$$0 \to U_\infty/\overline{\mathfrak{E}}_\infty \to X_\infty \to \mathfrak{A}_\infty \to 0. \tag{4.8}$$

Now $G = \mathrm{Gal}(\mathfrak{F}_\infty/\mathfrak{F})$ has a natural action on each of the groups in the exact sequence (4.8) (the natural action of G on X_∞ comes from the fact that M_∞ is obviously Galois over $\mathbb{Q}$, and thus G will act by lifting inner automorphisms), and (4.8) is then an exact sequence of compact G-modules. It is easily seen that this G-action on each of the modules in (4.8) extends by linearity and continuity to an action of the whole Iwasawa algebra $\Lambda(G)$. Now, when $\mathfrak{A}_0 = 0$ or equivalently the class number of $\mathfrak{F}$ is prime to p (in fact, this happens for all primes $p < 12 \times 10^6$, and no counter example has yet been found), Iwasawa showed by a simple argument that we have $\mathfrak{A}_\infty = 0$. Moreover, it also follows from the analytic class number forumula for the fields $\mathfrak{F}_n$ for all $n \geq 0$ that $\overline{\mathfrak{E}}_\infty = \overline{C}_\infty$, and thus $X_\infty = U_\infty/\bar{C}_\infty$. It was therefore natural to ask whether or not there is an analogue of Theorem 4.2 without the hypothesis that the class number of $\mathfrak{F}$ is prime to p, and this question became rapidly known as the "*main conjecture*" on cyclotomic fields. More precisely, Iwasawa [14] had already shown that X_∞ is a finitely generated torsion $\Lambda(G)$-module, and the general algebraic structure theory for such modules then shows that there is an exact sequence of $\Lambda(G)$-modules.

$$0 \to \sum_{i=1}^{r} \Lambda(G)/f_i\Lambda(G) \to X_\infty \to W \to 0,$$

where W is a finite $\Lambda(G)$-module, and $f_1, \ldots, f_r$ are non-zero divisors in $\Lambda(G)$. We then define $ch_G(X_\infty) = f_1 \ldots f_r \Lambda(G)$. We owe to Mazur and Wiles [17] the first proof of this "*main conjecture*".

Theorem 4.3. *(Mazur–Wiles). For all odd primes p, we have $ch_G(X_\infty) = \zeta_p I(G)$.*

The simplest consequence of this result is Kummer's criterion for odd primes p. Indeed, on the one hand it was shown by Iwasawa that if X_∞ finite then necessarily $X_\infty = 0$, while on the other hand p not dividing the numerator of any of $\zeta(-1), \zeta(-3), \ldots, \zeta(4-p)$ is easily seen to be equivalent to the assertion that $\zeta_p I(G) = \Lambda(G)$.

To attack the Birch–Tate conjecture for any totally real number field J, we need the analogue of Theorem 4.3 for J, in which we now take the field $\mathfrak{J}_\infty$, which is defined to be the maximal real subfield

of the field obtained by adjoining all p-power roots of unity to J. In this general case, the construction of a p-adic zeta function as a pseudo-measure on G_J, where $G_J = \mathrm{Gal}(\mathfrak{J}_\infty/J)$, and satisfying an analogue Theorem 4.1 was proven independently by Cassou-Nogues [2] and Deligne–Ribet [8]. Let $M_{\infty,J}$ be the maximal abelian p-extension of $\mathfrak{J}_\infty$ which is unramified outside the primes lying above p and ∞, and define

$$X_{\infty,J} = \mathrm{Gal}(M_{\infty,J}/\mathfrak{J}_\infty).$$

Then $X_{\infty,J}$ has its natural structure as a module over the Iwasawa algebra $\Lambda(G_J)$ of G_J. In deep work, Wiles [23] then established for $X_{\infty,J}$ the precise analogue of Theorem 4.3. A consequence of this analogue of Theorem 4.3 is the proof of the p-part of the Birch-Tate conjecture for J and all odd primes p. Indeed it follows easily [5] from work of Iwasawa [14] and Tate [20] that, for all odd primes p, the dual of the p-primary subgroup of $R_2 J$ is isomorphic to the group of G_J-coinvariants of the module $X_{\infty,J}(-2)$; here $X_{\infty,J}(-2)$ means that the natural action of G_J on $X_{\infty,J}$ has been twisted to χ_J^{-2}, where χ_J is the p-cyclotomic character for the field J. Granted this, it is then just an exercise to deduce the p-part of the Birch-Tate conjecture from the analogue of Theorem 4.3 for J. However, to the best of my knowledge, technical difficulties still remain in giving a fully detailed proof of the 2-primary part of the Birch–Tate conjecture for all totally real J. Turning to elliptic curves $E/\mathbb{Q}$, it is now known how to prove a precise analogue of Theorem 4.3 for all sufficiently large primes p of potentially ordinary reduction. However, new aspects of the Iwasawa theory, which we do not have time to discuss here, emerge for those primes p where E has potentially supersingular reduction. A consequence of this large body of work for elliptic curves is that, when $L(E,1) \neq 0$, it is now known how to prove the p-part of Conjecture 3.7 for all sufficiently large primes p, thanks to work of Rubin [19] when E has complex multiplication, and to work of Kato [15] and others when E has no complex multiplication. Hopefully, number-theorists will be able to handle the many interesting issues about the Iwasawa theory of these small primes for E in the not too distant future, and give at last a full proof of Conjecture 3.7 when $L(E,1) \neq 0$.

References

[1] B. Birch, P. Swinnerton-Dyer, *Notes on elliptic curves (II)*, Crelle 218, 79-108, 1965.
[2] P. Cassou-Nogues, *Valeurs aux entiers negatifs des fonctions zeta et fonctions zeta p-adiques*, Invent. Math. 51, 29-59, 1979.
[3] J. Coates, A. Wiles *On the conjecture of Birch and Swinnerton-Dyer*. Invent. Math. 39, 223-251, 1977.
[4] J. Coates, A. Wiles, *On p-adic L-functions and elliptic units*, J. Australian Math. Soc. 26, 1–25, 1978.
[5] J. Coates, *On K_2 and some classical conjectures in algebraic number theory*, Ann. Math. 95, 99-116, 1972.
[6] R. Coleman, *Division values in local fields*, Invent. Math. 53, 91–116, 1979.
[7] G. Cornell, J. Silverman, G. Stevens, *Modular Forms and Fermat's Last Theorem*, Springer, New York, 1997.
[8] P. Deligne, K. Ribet, *Values of abelian L-functions at negative integers over totally real fields*, Invent. Math. 59, 227-286, 1980.
[9] H. Garland, *A finiteness theorem for K_2 of a number field*, Ann. Math. 94, 534-548, 1971.
[10] B. Gross, D. Zagier, *Heegner points and derivatives of L-series*, Invent. Math. 84, 225-320, 1986.
[11] B. Gross, *Kolyvagin's work on modular elliptic curves* in *L-functions and arithmetic (Durham 1989*, London Math. Soc. Lecture Notes 153, 235-256, 1991.
[12] K. Iwasawa, *On some modules in the theory of cyclotomic fields*, J. Math. Soc. Japan, 16, 42–82, 1964.
[13] K. Iwasawa, *On p-adic L-functions*, Ann. Math., 89, 198–205, 1969.
[14] K. Iwasawa, *On $\mathbb{Z}_l$-extensions of algebraic number fields*, Ann. Math. 98, 246–326, 1973.
[15] K. Kato, *Euler systems, Iwasawa theory, and Selmer groups*, Kodai Math. J. 22, 313–372, 1999.
[16] V. Kolyvagin, *Finiteness of $E(\mathbb{Q})$ and $Ш(E/\mathbb{Q})$ for a class of Weil curves*, Izv. Akad. Nauk. 52, 1988.
[17] B. Mazur, A. Wiles, *Class fields of abelian extensions of $\mathbb{Q}$*, Invent. Math. 76 , 179–330, 1984.
[18] K. Ribet, *On modular representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms*, Invent. Math. 100, 431-476, 1990.
[19] K. Rubin, *The main conjectures of Iwasawa theory for imaginary quadratic fields*, Invent. Math. 103, 25–68, 1991.
[20] J. Tate, *Letter from Tate to Iwasawa on the relation between K_2 and Galois cohomology*, *Algebraic K-theory II*, Springer Lecture Notes, 342, 524–527, Berlin, 1973.
[21] R. Taylor, A. Wiles, *Ring theoretic properties of certain Hecke algebras*, Ann. Math. 141, 553-572, 1995.
[22] F. Thaine, *On the ideal class groups of real abelian number fields*, Ann. Math. 128, 1-18, 1988.
[23] A. Wiles, *The Iwasawa conjecture for totally real fields*, Ann. Math. 131, 493–540, 1990.

John Coates,
Emmanuel College, Cambridge,
England.
jhc13@dpmms.cam.ac.uk