Lecture 1: Periods

Following the paper *Periods* [KZ] by M. Konstsevich and D. Zagier we define a period as "a complex number whose real and imaginary parts are values of absolutely convergent integrals of rational functions with rational coefficients, over domains in $\mathbb{R}^n$ given by polynomial inequalitites with rational coefficients." The name period stems from the fact that the periods associated with trigonometric functions and elliptic curves over $\mathbb{C}$ are periods in the sense just defined.

It can be shown that every algebraic number is a period. For example,

$$\sqrt{2} = \int_{2x^2 \leq 1} dx.$$

On a slightly less obvious note, if α is a complex root of $z^3 + z + 1$ then

$$\Re(\alpha) = \int_{\Omega} dx,$$

where Ω is the interval defined by $0 \leq 8x^3 + 2x \leq 1$. The periods are clearly a countable subset of the complex numbers, and hence a proper subset. However, it can be tricky to demonstrate that a given number is not a period. For example, it is unknown whether or not e or $1/\pi$ are periods. The set of periods do include many other noteworthy mathematical constants, such as

$$\pi = \int_{x^2+y^2 \leq 1} dx\, dy \quad = \int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} = \int_{-\infty}^{\infty} \frac{dx}{1+x^2},$$

$$\log 2 = \int_1^2 \frac{dx}{x}, \quad \text{and} \quad \zeta(3) = \int_{\Omega} \frac{dx\, dy\, dz}{(1-x)yz},$$

where $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ for $\Re(s) > 1$ is the standard Riemann zeta-function, and Ω is the region $0 \leq x \leq y \leq z \leq 1$. The latter integral may be verified by expanding $1/(1-x)$ in a geometric series, then integrating term by term.

The middle expression for π illustrates the general principle that employing algebraic functions (or in fact algebraic coefficients) within our integral will still produce a period. Thus the Beta function evaluated at rational values yields a period. Take, for instance,

$$\beta(\tfrac{1}{3}, \tfrac{4}{5}) = \int_0^1 t^{1/3}(1-t)^{-4/5} dt = \int_{\Omega} dx\, dy\, dt,$$

for $\Omega = \{(x, y, t) | 0 \leq t \leq 1,\ 0 \leq x^3 \leq t,\ 0 \leq y^5(1-t)^4 \leq 1\}$. Less obvious examples of periods are given by $\Gamma(p/q)^q$ for $p, q \in \mathbb{N}$, where $\Gamma(s)$ is the Gamma function defined by

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^s \frac{dt}{t}, \quad \Re(s) > 1.$$

It is well known that $\Gamma(1/2)^2 = \pi$, while one can verify that

$$\Gamma(\tfrac{1}{3})^3 = 2^{4/3} 3^{1/2} \pi \int_0^1 \frac{dx}{\sqrt{1-x^3}}$$

using the Beta function and the Gauss multiplication formula.

Finally, we consider the zeta-functions of multiple arguments defined by

$$\zeta(s_1, \ldots, s_k) = \sum_{0 < n_1 < \cdots < n_k} n_1^{-s_1} \cdots n_k^{-s_k},$$

which will converge for $\Re(s_k) > 1$ and $\Re(s_i) \geq 1$ when $i < k$. (See [Za] for an introduction to these multizeta values.) If the s_i are integers satisfying $s_k \geq 2$ and $s_i \geq 1$ for $i < k$ then we obtain further examples of periods. Thus

$$\zeta(1,2) = \int_0^1 \int_0^1 \int_y^1 \frac{y\, dx\, dy\, dz}{(1-xy)(1-y)z},$$

which may be verified in the same manner as above. This quantity leads to our first example of an identity involving periods, for it turns out that $\zeta(1,2) = \zeta(3)$. The proof of this fact illustrates the dearly held belief that one should be able to prove any identity between periods using only standard algebraic manipulations, elementary properties of integrals, the change of variables theorem, and Stokes's formula. (Unfortunately, no bound is placed on the degree of complexity these proofs might achieve!) In our case, making the substitution $(u,v,w) = (xy, y, z)$ in the above expression and performing the integrations with respect to u and w leads to an integral which is equivalent to the integral for $\zeta(3)$.

A more sophisticated such argument, due to Calabi, can be used to demonstrate that $\zeta(2) = \pi^2/6$. We will evaluate the integral

$$I = \int_0^1 \int_0^1 \frac{dx\, dy}{(1-xy)\sqrt{xy}}$$

in two ways. By expanding $1/(1-xy)$ in a geometric series and integrating term by term we discover that $I = \sum_0^\infty (n+\frac{1}{2})^{-2} = (4-1)\zeta(2)$. On the other hand, making the change of variables

$$x = v^2 \frac{1+u^2}{1+v^2}, \quad y = u^2 \frac{1+v^2}{1+u^2}$$

produces the Jacobian

$$\left|\frac{d(x,y)}{d(u,v)}\right| = \frac{4uv(1-u^2)(1-v^2)}{(1+u^2)(1+v^2)} = \frac{4(1-xy)\sqrt{xy}}{(1+u^2)(1+v^2)}$$

and the new domain Ω of integration $u \geq 0$, $v \geq 0$, and $uv < 1$. Therefore

$$\begin{aligned} I &= 4 \int_\Omega \frac{du}{1+u^2} \frac{dv}{1+v^2} \\ &= 2 \left(\int_0^\infty \frac{du}{1+u^2} \right) \left(\int_0^\infty \frac{dv}{1+v^2} \right) \\ &= \frac{\pi^2}{2}. \end{aligned}$$

We present an assortment of other identities involving periods to demonstrate the wide variety of forms they can take.

$$\sqrt{5} + \sqrt{22 + 2\sqrt{5}} = \sqrt{11 + 2\sqrt{29}} + \sqrt{16 - 2\sqrt{29} + 2\sqrt{55 - 10\sqrt{29}}}$$

$$\zeta(1,3) = \frac{2\pi^4}{6!}, \quad \zeta(1,3,1,3) = \frac{2\pi^8}{10!}$$

$$\left(\int_0^1 \frac{dx}{\sqrt{1-x^4}} \right) \left(\int_0^1 \frac{x^2 dx}{\sqrt{1-x^4}} \right) = \frac{\pi}{4}$$

$$\int_0^\pi \int_0^\pi \int_0^\pi \frac{du dv dw}{1 - \cos(u)\cos(v)\cos(w)} = \frac{\Gamma(1/4)^4}{4}$$

$$\int_0^\pi \int_0^\pi \int_0^\pi \frac{dudvdw}{3-\cos(u)cos(v)-\cos(u)\cos(w)-\cos(v)\cos(w)} = \frac{3\Gamma(1/3)^6}{2^{14/3}\pi}$$

$$\int_0^{\pi/2} \arccos\left(\frac{\cos x}{1+2\cos x}\right) dx = \frac{5\pi^2}{24}$$

$$\int_0^{\pi/3} \arccos\left(\frac{1}{1+2\cos x}\right) dx = \frac{\pi^2}{8}$$

$$\int_0^{\pi/3} \arccos\left(\frac{1-\cos x}{2\cos x}\right) dx = \frac{11\pi^2}{72}$$

$$\int_0^1 \frac{\log\left(1+x^{4+\sqrt{15}}\right)}{1+x} dx =$$

$$= \frac{\pi^2}{12}\left(\sqrt{15}-2\right) + \log 2 \log\left(\sqrt{3}+\sqrt{5}\right) + \log\left(\frac{1+\sqrt{5}}{2}\right)\log\left(2+\sqrt{3}\right)$$

The formulas for $\zeta(1,3)$, $\zeta(1,3,1,3)$ as well as the general case $\zeta(1,3,\ldots,1,3)$ were found numerically by Zagier and the general case was proved several years later by Broadhurst; we will present a streamlined proof below from [KZ]. The integrals involving $\cos(u)$ are two of the *Three Triple Integrals* of Watson [Wa], which arose from a problem in ferromagnetism. The inverse trig integrals are due to Coxeter [Cx] and arose from his consideration of "four-dimensional figures." The final example is one among other identities of this type due to Herglotz [He] and are a consequence of his formulation of Kronecker's limit formula for real quadratic fields. (It does not seem clear, apriori, why the left hand side should be a period.) In a later lecture we will define the Mahler measure of a polynomial and discuss some remarkable identities involving these quantities.

A further appearance of a period with historical significance involves the arithmetic-geometric mean $M(a,b)$ of two positive real numbers. Starting with $a_1 = a$ and $b_1 = b$, recursively define $a_n = \sqrt{a_{n-1}b_{n-1}}$ and $b_n = (a_{n-1}+b_{n-1})/2$. Then $\lim a_n$ and $\lim b_n$ exist and are equal; we let $M(a,b)$ denote their common value. Gauss's numerical discovery of the identity

$$M(1,\sqrt{2}) = \frac{\pi}{\omega}, \quad \text{where} \quad \omega = \int_0^1 \frac{dx}{\sqrt{1-x^4}},$$

motivated his extensive study of the arithmetic-geometric mean (AGM) and prompted the following entry in his diary on May 30, 1799: "We have established that the AGM of 1 and $\sqrt{2}$ is π/ω to the eleventh decimal place; the demonstration of this fact will surely open an entirely new field of analysis." For a fascinating account of this discovery, see Cox's beautiful paper [Co].

The question of how to identify periods, or even rationals for that matter, from their decimal expansions suggests the following game. Suppose that I choose a rational number between 0 and 1 of some predetermined complexity, say with numerator and denominator of three digits or less. You can buy successive digits of the decimal expansion of my number for \$1. What is your best strategy if your goal is to correctly guess the fraction with the least expenditure of money? (To be more precise, if you played this game many times and computed the average amount of money spent per correct guess, how low could you drive this average?) What is the most efficient algorithm for making your guesses? (Hint: do *not* wait for the periodic decimal expansion to become evident!) What if we played this game with algebraic numbers, or periods in general? To whet your appetite, see if you can identify the following periods from their decimal approximations, given the information about the number to the right.

$0.7419354838709677\cdots$	rational
$0.99848465715368102\cdots$	rational
$0.68232780382801932736948 37\cdots$	algebraic number of degree 3
$1.0757660660868371580595995\cdots$	algebraic number of degree 10
$2.3080423740661582448037805\cdots$	rational linear combination of π and $\zeta(3)$
$4.8957278968894452908646 8\cdots$	rational multiple of a power of π

In general there is a very efficient algorithm (the LLL basis reduction algorithm) which will find small integer dependencies among any number of constants [Ch].

Nonetheless, decimal approximations can sometimes be misleading. One of the most infamous non-equalities is given by $\pi\sqrt{163}/3 \approx \log(640320)$. These two real numbers share sixteen initial digits before disagreeing! Several other mathematical near misses are worth mentioning. In [Eu] Euler gives such an example, with the title *Exemplum Memorabile Inductionis Fallacis.* He defines the sequence of *trigonal numbers* as

$$a_n = \left[\left(x+1+x^{-1}\right)^n\right]_0, \qquad n \geq 0$$

where $[\,\cdot\,]_0$ stands for constant term. This sequence begins 1, 1, 3, 7, 19, 51, and so on. Euler noticed that

$$-a_{n+1} + 3a_n = F_{n-1}^2 + F_{n-1}, \qquad n = 0, \ldots, 8$$

but fails for $n = 9$, where $F_{-1} = 1$, $F_0 = 0$, $F_1 = 1$, and $F_{n+1} = F_n + F_{n-1}$ is the Fibonacci sequence. Why is it that an identity like this cannot possibly hold for all n?

Even more fantastic still, Boyd [Bo] considers the sequence defined by

$$a_{n+2} = \left[\frac{a_{n+1}^2}{a_n} + 1\right], \qquad a_0 = 8,\ a_1 = 55$$

where $[x]$ is the integer part of x. This sequence matches the coefficients in the Taylor expansion about the origin of

$$\frac{8+7x-7x^2-7x^3}{1-6x-7x^2+5x^3+6x^4} = 8 + 55x + 379x^2 + 2612x^3 + \cdots$$

for $n = 0$, 1, ... , 11055 but fails for $n = 11056$. (An equivalent formulation is that a_n satisfies a linear recurrence relation with constant coefficients for some (long) range $n = 0, 1, 2, \ldots$ but not for all n.)

There exist similar sequences that *do* satisfy linear recurrences. For instance [Ca] shows that if a_n is defined by

$$a_n = \left[\frac{a_{n+1}^2}{a_n} + \frac{1}{2}\right], \qquad a_0 = 3,\ a_1 = 10$$

then

$$\sum_{n=0}^{\infty} a_n x^n = \frac{1}{1-x(3+x)}.$$

These examples arose from the work of Pisot on his special numbers.

Finally there is the curious phenomenon known as Langston's ant. The web site `http://www.math.sunysb.edu/~scott/ants/` describes it as follows

> Briefly, an "ant" moves around on an infinite checkerboard, each square of which we refer to as a "cell". Each cell in the plane is labeled as either an L-cell or an R-cell. (Usually, one fills the plane

> with L-cells to start.) The ant starts out on the boundary between two cells, and as it passes through each cell, it makes a 90 degree turn, turning to the left in L-cells and to the right in R-cells, and it changes the state of the cell it just left, switching L-cells to R-cells, and vice versa. Following this simple set of rules gives rise to some rather complicated behavior; the pattern of the ant's track alternates between apparent chaos and symmetry, but eventually it starts to build a "highway" moving off in a single direction.

We return to our discussion of identities among periods, which often arise in a geometric setting. The arc length of a curve in $\mathbb{R}^2$ is typically given by an intractable integral, but occasionally the length can be computed exactly in terms of standard mathematical constants. The circle, astroid, and lemniscate with equations $x^2 + y^2 = 1$, $x^{2/2} + y^{2/3} = 1$, and $(x^2 + y^2)^2 = 2(x^2 - y^2)$ have arc length 2π, 6, and 2ω, respectively, where ω is the constant defined above in conjunction with Gauss and the AGM. (It can be shown that $\omega = \Gamma(1/4)^2 2^{-3/2}\pi^{-1/2}$.) A more surprising result, known as Vivianni's problem, states that the surface area and volume of the sphere $x^2 + y^2 + z^2 \leq 1$ after the two cylinders $(x \pm 1/2)^2 + y^2 \leq 1/4$ have been removed are each rational.

We now present the promised streamlined proof of Zagier's conjecture that

$$\zeta(\underbrace{1, 3, \ldots, 1, 3}_{2m}) = \frac{2\pi^{4m}}{(4m+2)!}$$

based on his observation of this equality for $m = 1$ and 2. A demonstration of the general formula took several years to find; the following proof is based on the one found by Broadhurst but is considerably more streamlined than the original argument. Define the hypergeometric series for any three complex parameters a, b, and c and a complex variable x as

$$F(a, b; c; x) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n \; n!} x^n,$$

which is absolutely convergent for $|x| < 1$. Here $(a)_n = a(a+1)\cdots(a+n-1)$ is a "rising factorial." Now form the generating function

$$1 + \sum_{m=1}^{\infty} \left(\sum_{0<a_1<b_1<\cdots<a_m<b_m} \frac{x^{b_m}}{a_1 b_1^3 \cdots a_m b_m^3} \right) (-4t^4)^m.$$

Incredibly, this sum factors as $F(t, -t; 1; x)F(it, -it; 1; x)$, a fact which can be established by proving that as a power series in x, each expression has constant term 1 and is annhilated by the operator $\left((1-x)\frac{d}{dx}\right)^2 \left(x\frac{d}{dx}\right)^2 + 4t^4$. Setting $x = 1$ in the resulting identity yields

$$\begin{aligned} 1 + \sum_{m=1}^{\infty} \zeta(\underbrace{1, 3, \ldots, 1, 3}_{2m})(-4t^4)^m &= F(t, -t; 1; 1)F(it, -it; 1; 1) \\ &= \frac{\sin \pi t}{\pi t} \frac{\sinh \pi t}{\pi t} \\ &= \sum_{0}^{\infty} \frac{2\pi^{4m}}{(4m+2)!}(-4t^4)^m, \end{aligned}$$

which completes the proof. Each step outlined above conceals some clever and often intricate manipulations. As a worthwhile exercise the reader is encouraged to fill in as many details as possible.

As a final example of relationships among periods we introduce the dilogarithm and present the five term identity which it satisfies. Following Abel we define

$$\mathrm{Li}_2(x) = -\int \log(1-x)\frac{dx}{x} = \sum_{n=1}^{\infty} \frac{x^n}{n^2}, \qquad |x| < 1,$$

where we choose the anti-derivative satisfying the condition $\mathrm{Li}_2(0) = 0$. We treat the indefinite integrals in a formal manner, without attending to issues of convergence. So replace x by $\frac{a}{1-a}\frac{y}{1-y}$, treating a as a parameter and y as a variable. This substitution yields

$$\mathrm{Li}_2\left(\frac{a}{1-a}\frac{y}{1-y}\right) = -\int \log\left(\frac{1-a-y}{(1-a)(1-y)}\right)\cdot\left(\frac{1}{y}+\frac{1}{1-y}\right)\,dy.$$

The integrand may be expanded to produce the four terms

$$-\int \log\left(1-\frac{y}{1-a}\right)\frac{dy}{y} + \int \log(1-y)\frac{dy}{y} - \int \log\left(1-\frac{a}{1-y}\right)\frac{dy}{1-y} + \int \log(1-a)\frac{dy}{1-y}.$$

The first two terms are easily recognized as values of the dilogarithm, the third becomes recognizable upon substituting $z = \frac{a}{1-y}$, and the final term can be directly integrated. Therefore

$$\mathrm{Li}_2\left(\frac{a}{1-a}\frac{y}{1-y}\right) = \mathrm{Li}_2\left(\frac{y}{1-a}\right) - \mathrm{Li}_2(y) - \mathrm{Li}_2\left(\frac{a}{1-y}\right) - \log(1-a)\log(1-y) + C.$$

Either by symmetry or by setting $y = 0$ we deduce the constant of integration must be $C = -\mathrm{Li}_2(a)$. The resulting identity gives the analog of $\log(xy) = \log x + \log y$ for the dilogarithm of a product. This five term relation for Li_2 has been discovered and rediscovered many times. Note that once we guess the correct identity it is a simple matter to verify it by differentiation. Later we shall see how this formula arises by considering sums of tetrahedra formed on five vertices.

References

[Bo] D. Boyd, *Linear recurrence relations for some generalized Pisot sequences,* Advances in number theory (Kingston, ON, 1991), Oxford University Press, New York, 1993, pp. 333–340

[Ca] G. Cantor, *On families of Pisot E-sequences,* Ann. Sci. École Norm. Sup. **9** (1976), 283–308

[Ch] H. Cohen, *Computational aspects of Number Theory* Mathematics unlimited—2001 and beyond, Contemp. Math., Springer, Berlin, 2001, pp. 301–329

[Co] D. Cox, *The Arithmetic-Geometric Mean of Gauss,* L'enseignement Mathématique **30** 1984, 275–330

[Cx] H.S.M. Coxeter, *The beauty of geometry. Twelve essays,* Dover Publications, Inc., Mineola, NY, 1999

[Eu] L. Euler, *Observationes analyticae,* Opera Omnia, vol. 15, 1765, pp. 50–69

[He] G. Herglotz, *Über die Kroneckersche Grenzformel für reelle quadratische Körper,* Gessammelte Schriften, Vandenhoeck and Ruprecht, Götingen, 1979, pp. 466–484

[KZ] M. Kontsevich and D. Zagier, *Periods,* Mathematics unlimited—2001 and beyond, Contemp. Math., Springer, Berlin, 2001, pp. 771-808

[Wa] G.N. Watson, *Three triple integrals,* Quart. J. Math., Oxford Ser. **10** (1939), 266–276

[Za] D. Zagier, *Values of zeta functions and their applications,* First European Congress of Mathematics II (Paris, 1992), Progr. Math., vol. 120, Birkhäuser, Basel, 1994, pp. 497–512

Lectures 2–3: Zeta and L-Functions

Let X be an algebraic variety over a finite field, say defined by the equations $f_j(x_1, \ldots, x_n) = 0$ for $1 \leq j \leq m$ with coefficients in $\mathbb{F}_q$. Motivated by the principle that geometry over a finite field will involve counting, as opposed to pictures, we proceed to tally the number of solutions to the equations defining X. This approach makes sense within the finite extensions of $\mathbb{F}_q$, namely the fields $\mathbb{F}_{q^r}$ for any natural number r. (We will assume throughout that all the action takes place in some fixed algebraic closure $\overline{\mathbb{F}_q}$ of $\mathbb{F}_q$.) Therefore we define N_r to be the number of ordered n-tuples $(x_1, \ldots, x_n)$ with $x_i \in \mathbb{F}_{q^r}$ satisfying the m equations defining X. The most useful form in which to package this data turns out to be the formal power series

$$Z(X,T) = \exp\left(\sum_{r=1}^{\infty} \frac{N_r}{r} T^r\right),$$

known as the zeta function of $X/\mathbb{F}_q$. For example, if X is defined by $f(x) = x - \alpha$ for $\alpha \in \mathbb{F}_q$, then clearly $N_r = 1$ for all r and we find that

$$Z(X,T) = \exp\left(-\log(1-T)\right) = \frac{1}{1-T}.$$

The general dimension zero case can be analyzed without much more difficulty. Hence suppose that X is defined by a polynomial $f(x)$ of degree d in one variable. We will assume that $f(x)$ has no repeated roots in $\overline{\mathbb{F}_q}$, since none of the N_r are affected by the presence of multiple roots. Let $A = \mathbb{F}_q[x]/(f)$ be the quotient ring, which is an étale algebra. Then A has no nilpotent elements and it can be shown that

$$A = \prod_{i=1}^{t} k_i$$

is a product of fields of finite degree over $\mathbb{F}_q$, each k_i corresponding to an irreducible factor $f_i(x)$ of $f(x)$. As expected, $k_i \cong \mathbb{F}_q[x]/(f_i)$, so k_i has degree d_i over $\mathbb{F}_q$, where d_i is the degree of $f_i(x)$.

Now the number of solutions to $f(x) = 0$ in a given extension of $\mathbb{F}_q$ is simply the sum of the number of solutions to each $f_i(x)$. So let X_i be the variety defined by the equation $f_i(x) = 0$, and let N_{ri} be the number of solutions to $f_i(x) = 0$ over the field $\mathbb{F}_{q^r}$. Then we have

$$Z(X,T) = \exp\left(\sum_{r=1}^{\infty} \frac{N_{r1} + \cdots + N_{rt}}{r} T^r\right) = \prod_{i=1}^{t} Z(X_i, T).$$

Since $f_i(x)$ is irreducible it has exactly d_i roots, each contained in a field of degree d_i over $\mathbb{F}_q$, namely the field with q^{d_i} elements. Thus $f_i(x)$ splits completely in any field $\mathbb{F}_{q^e}$ containing the field with q^{d_i} elements, that is, whenever $d_i | e$. Otherwise $f_i(x)$ has no solutions in $\mathbb{F}_{q^e}$. It follows that

$$Z(X_i,T) = \exp\left(\sum_{r=1}^{\infty} \frac{d_i}{rd_i} T^{rd_i}\right) = \frac{1}{1-T^{d_i}}.$$

We conclude that

$$Z(X,T) = \prod_{i=1}^{t} \frac{1}{1-T^{d_i}}.$$

This example illustrates several features (known as the Weil conjectures, proven by Deligne) which the zeta functions $Z(X,T)$ all share. First, $Z(X,T)$

is a rational function of T with coefficients in $\mathbb{Q}$. Next, there is a functional equation, which in the dimension zero case reduces to

$$Z(X,1/T)=(-1)^t T^d Z(X,T), \qquad d=\sum_{i=1}^{t} d_i.$$

Finally, there is a statement regarding the absolute values of the roots (over $\mathbb{C}$) of the numerator and denominator of $Z(X,T)$; in our case it simply says that they are roots of unity. We also mention that if $Z(X,T)$ has the form

$$Z(X,T)=\frac{\prod(1-\beta_j T)}{\prod(1-\alpha_k T)}$$

then it follows that $N_r=\sum\alpha_k^r-\sum\beta_j^r$. In particular, if there are a total of l factors in the expression for $Z(X,T)$ then given $N_1, \ldots, N_l$ we can determine the general formula for N_r, and hence the zeta function.

As an application of the above ideas, suppose $f(x)\in\mathbb{Z}[x]$ has no repeated roots, so that $A=\mathbb{Q}[x]/(f)$ is étale, say with degree d over $\mathbb{Q}$. Then for all but a finite number of primes the reduction modulo p to $A_p=\mathbb{F}_p[x]/(\overline{f})$ will also be étale with degree d over $\mathbb{F}_p$. Let S be the finite set of primes for which this does not occur. For $p\notin S$, let X_p denote the variety corresponding to A_p, and declare

$$\zeta_S(X,s)=\prod_{p\notin S} Z(X_p,p^{-s}),$$

where we have substituted p^{-s} for T. This zeta function effectively collects all the local information at each prime p. Hence when $f(x)=x-a$ for $a\in\mathbb{Z}$ we find $S=\emptyset$ and $\zeta(X,s)=\prod(1-p^{-s})^{-1}=\zeta(s)$, the Riemann zeta-function.

It is instructive to detail this procedure for a less trivial polynomial, such as $f(x)=x^2+x-1$, for which $S=\{5\}$. Rewriting the equation $f(x)\equiv 0 \bmod p$ as $(2x-1)^2\equiv 5 \bmod p$ for p odd, $p\neq 5$, we see that $f(x)$ splits in $\mathbb{F}_p$ precisely when 5 is a quadratic residue mod p, otherwise it splits in the quadratic extension of $\mathbb{F}_p$. (The same holds for $p=2$, as the reader may verify.) According to the above discussion, $Z(X_p,T)=(1-T)^{-2}$ in the former case, while $Z(X_p,T)=(1-T^2)^{-1}$ in the latter case. Hence

$$\zeta_S(X,s)=\prod_{\left(\frac{5}{p}\right)=1}(1-p^{-s})^{-2}\prod_{\left(\frac{5}{p}\right)=-1}(1-p^{-2s})^{-1}.$$

Notice that each term contains a factor of $(1-p^{-s})^{-1}$. We may complete this zeta-function by including a factor of $(1-p^{-s})^{-1}$ for $p=5$, which is reasonable, since $f(x)$ has exactly one solution in each finite extension of $\mathbb{F}_5$. We denote the result by $\zeta(X,s)$, namely

$$\zeta(X,s)=\zeta(s)L(X,s), \qquad L(X,s)=\prod_p\left(1-\left(\frac{5}{p}\right)p^{-s}\right)^{-1},$$

where $\left(\frac{5}{5}\right)=0$. The factor $\left(\frac{5}{p}\right)$ is an example of a Dirichlet character, that is, a multiplicative homomorphism $\chi:(\mathbb{Z}/N\mathbb{Z})^*\longrightarrow\mathbb{C}^*$ where $\chi(a)=0$ if $(a,n)>1$. It is primitive if the value of N is as small as possible; this value is called the conductor of χ. The associated L-function is given by

$$L(\chi,s)=\prod_p(1-\chi(p)p^{-s})^{-1}.$$

In order to study the analytic aspects of these L-functions we first review some elementary properties of the gamma function. In the region $\Re(s) > 1$ it is defined as $\Gamma(s) = \int_0^\infty e^{-t} t^s \frac{dt}{t}$. It extends to a meromorphic function on all of $\mathbb{C}$ having simple poles at $\{0, -1, -2, \dots\}$ with residue $(-1)^n/n!$ at $s = -n$. In addition, the gamma function satisfies the functional equation $\Gamma(s+1) = s\Gamma(s)$. Since $\Gamma(1) = 1$ follows easily from the definition, we find that $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$. We also mention the duplication formula

$$\Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{s+1}{2}\right) = 2^{1-s}\pi^{1/2}\Gamma(s).$$

Furthermore, the change of variables $t \mapsto at$ yields $a^{-s}\Gamma(s) = \int_0^\infty e^{-at} t^s \frac{dt}{t}$. Finally, one of the crucial properties of the Gamma function is that $\Gamma(s) \neq 0$ for all complex numbers.

The function $L(\chi, s)$ also has an analytic continuation and satisfies a functional equation. As a model for our method of proof we first handle the case of the trivial character, namely $L(1, s) = \zeta(s)$.

Theorem: *The function $\zeta^\star(s) = \pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$ defined for $\Re(s) > 1$ extends to a meromorphic function on all of $\mathbb{C}$ satisfying $\zeta^\star(1-s) = \zeta^\star(s)$.*

Thus $\zeta^\star(s)$ provides an analytic continuation of $\zeta(s)$, and will provide infomation about the zeros, poles, and residues of $\zeta(s)$ as well.

Proof: Using the series for $\zeta(s)$ we write $\zeta^\star(s) = \sum_{n=1}^\infty (\pi n^2)^{-s/2}\Gamma\left(\frac{s}{2}\right)$. Then applying the formula for $a^{-s}\Gamma(s)$ given above we obtain

$$\zeta^\star(s) = \int_0^\infty \sum_{n=1}^\infty e^{-\pi n^2 t} t^{s/2} \frac{dt}{t} = \int_0^\infty \frac{1}{2}(\theta(t) - 1) t^{s/2} \frac{dt}{t},$$

where we have defined

$$\theta(t) = \sum_{n \in \mathbb{Z}} e^{-\pi n^2 t}, \qquad t > 0$$

based on the premise that sums over a complete lattice (such as $\mathbb{Z}$) are easier to handle. As we shall see, $\theta(t)$ satisfies the functional equation $\theta(\frac{1}{t}) = t^{1/2}\theta(t)$. This relationship comes into play when we split the integral into the two pieces $\int_0^\infty = \int_0^1 + \int_1^\infty$ and make the change of variables $t \mapsto \frac{1}{t}$ in the first one, yielding

$$\begin{aligned}
\zeta^\star(s) &= \frac{1}{2}\int_1^\infty (t^{1/2}\theta(t) - 1) t^{-s/2}\frac{dt}{t} + \frac{1}{2}\int_1^\infty (\theta(t) - 1) t^{s/2}\frac{dt}{t} \\
&= \frac{1}{2}\int_1^\infty (\theta(t) - 1)\left(t^{s/2} + t^{(1-s)/2}\right)\frac{dt}{t} - \frac{1}{2}\int_1^\infty \left(t^{-s/2} - t^{(1-s)/2}\right)\frac{dt}{t} \\
&= \frac{1}{2}\int_1^\infty (\theta(t) - 1)\left(t^{s/2} + t^{(1-s)/2}\right)\frac{dt}{t} - \left(\frac{1}{s} + \frac{1}{1-s}\right).
\end{aligned}$$

Up to this point we have been working under the assumption that $\Re(s) > 1$. However, the final expression for $\zeta^\star(s)$ is defined for all $s \in \mathbb{C}$ except $s = 0$, 1 because the function $(\theta(t) - 1)$ decreases very rapidly as t goes to infinity. This proves that $\zeta^\star(s)$ is meromorphic on the entire complex plane with simple poles at $s = 0$ and 1 with residues -1 and 1, respectively. Moreover, it is clear by inspection that $\zeta^\star(s) = \zeta^\star(1-s)$, thus proving the functional equation.

Therefore $\zeta(s) = \pi^{s/2}\zeta^\star(s)\Gamma\left(\frac{s}{2}\right)^{-1}$ gives an expression for the Riemann zeta-function which agrees with the series representation when $\Re(s) > 1$ but is valid on all of $\mathbb{C}$. Using the fact that neither $\Gamma(s)$ nor $\zeta(s)$ have zeros for $\Re(s) > 1$, we conclude the same is true of $\zeta^\star(s)$. The functional then implies that $\zeta^\star(s)$ has no zeros for $\Re(s) < 0$ either. Hence the zeros of $\zeta(s)$ correspond to poles of $\Gamma\left(\frac{s}{2}\right)$ in this domain, so $\zeta(s)$ has a simple zero at $s = -2, -4, -6, \ldots$. Clearly $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. Based on our knowledge of $\zeta^\star(s)$, the only other candidate for a pole is $s = 0$. However, we find that

$$\zeta(0) = \lim_{s\to 0} \frac{\pi^{s/2}\zeta^\star(s)}{\Gamma\left(\frac{s}{2}\right)} = \lim_{s\to 0} \frac{s\zeta^\star(s)}{2\cdot\frac{s}{2}\Gamma\left(\frac{s}{2}\right)} = -\frac{1}{2},$$

since $\frac{s}{2}\Gamma\left(\frac{s}{2}\right) = \Gamma\left(\frac{s}{2}+1\right)$ and hence goes to $\Gamma(1) = 1$ as $s \to 0$. It turns out that the behavior of $\zeta(s)$ near $s = 0$ provides information about the unit group of the integers! The 2 appearing in the denominator corresponds to the fact that $\mathbb{Z}$ contains two roots of unity, and the order of vanishing of $\zeta(s)$ at $s = 0$ (which is zero, in this case) predicts the rank of the unit group.

We now provide the proof that $\theta(\frac{1}{t}) = t^{1/2}\theta(t)$, which requires the following clever tool.

Theorem: (Poisson Summation Formula) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a rapidly decreasing C^∞ function with Fourier transform $\hat{f}(y) = \int f(x)e^{-2\pi i x\cdot y}\, dx$. Then*

$$\sum_{m\in\mathbb{Z}^n} f(m) = \sum_{m\in\mathbb{Z}^n} \hat{f}(m).$$

Proof: We provide an outline of the argument. Let

$$h(x) = \sum_{m\in\mathbb{Z}^n} f(x+m) = \sum_{m\in\mathbb{Z}^n} c_m e^{2\pi i m\cdot x},$$

where the coefficients c_m are given by

$$\begin{aligned} c_m &= \int_{[0,1]^n} h(x)e^{2\pi i m\cdot x}\, dx \\ &= \sum_{l\in\mathbb{Z}^n} \int_{[0,1]^n} f(x+l)e^{2\pi i m\cdot x}\, dx \\ &= \int_{\mathbb{R}^n} f(x)e^{2\pi i m\cdot x}\, dx \\ &= \hat{f}(m). \end{aligned}$$

Now plug in $x = 0$ to obtain the Poisson summation formula.

Finally, to prove the functional equation for $\theta(t)$ we implement the function $g(x) = e^{-\pi t x^2}$, whose Fourier transform is $\hat{g}(y) = t^{-1/2}e^{\pi y^2/t}$. (This follows from the fact that the function $e^{-\pi x^2}$ is its own Fourier transform, and then scaling by $x \mapsto \sqrt{t}x$.) Applying Poisson summation to $g(x)$ immediately gives $\theta(\frac{1}{t}) = t^{1/2}\theta(t)$, as desired.

The technique outlined above demonstrates a standard approach to proving analytic continuation and finding a functional equation. For instance, it can be used when $F = \mathbb{Q}[x]/(f)$ is a number field. The prime ideals of the ring of integers $\mathcal{O}_F$ give rise to a zeta-function

$$\zeta_F(s) = \prod_{\mathfrak{p}} \left(1 - (N\mathfrak{p})^{-s}\right)^{-1}.$$

Hecke observed that in order to use the above strategy for analytic continuation in this setting one must heed the infinite primes of F as well. Therefore we write

$$V = F \otimes \mathbb{R} \cong \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$$

corresponding to the number of real and pairs of complex roots. (I.e. the degree of $f(x)$ is $r_1 + 2r_2$.) Complex conjugation acts on V providing a decomposition $V = V^+ \oplus V^-$ into subspaces with eigenvalues 1 and -1 having dimensions $n_+ = r_1 + r_2$ and $n_- = r_2$. We then define $\Gamma_+(s) = \pi^{-s/2}\Gamma\left(\frac{s}{2}\right)$ and $\Gamma_-(s) = \pi^{-(s+1)/2}\Gamma\left(\frac{s+1}{2}\right)$. Now set

$$\zeta_F^\star(s) = |\Delta_F|^{s/2}\Gamma_+(s)^{n_+}\Gamma_-(s)^{n_-}\zeta_F(s) = \gamma(s)\zeta_F(s).$$

Here Δ_F stands for the discriminant of F. Then just as before it can be shown that $\zeta_F^\star(s)$ provides an analytic continuation of $\zeta_F(s)$ to the entire complex plane with simple poles at $s = 0$ and 1, which satisfies the functional equation $\zeta_F^\star(s) = \zeta_F^\star(1-s)$. As above we know that $\zeta_F^\star(s) \neq 0$ for $\Re(s) > 1$ and hence for $\Re(s) < 0$ as well by the functional equation. Furthermore, we know the orders of the zeros and poles of $\gamma(s)$ since we understand $\Gamma(s)$. Therefore we may deduce the behavior of $\zeta_F(s)$ at the integers. The table below summarizes our results; the entries record the order of the zero of the corresponding function. Thus '2' means a zero of order two, '-1' denotes a simple pole, and '0' indicates a non-zero value.

s	-3	-2	-1	0	1	2	3	4
$\zeta_F^\star(s)$	0	0	0	-1	-1	0	0	0
$\gamma(s)$	$-n_-$	$-n_+$	$-n_-$	$-n_+$	0	0	0	0
$\zeta_F(s)$	n_-	n_+	n_-	$n_+ - 1$	-1	0	0	0

After Deligne, we call $m \in \mathbb{Z}$ a critical value for F if neither $\gamma(m)$ nor $\gamma(1-m)$ has a pole at $s = m$. For example, if F is a totally real field, so that $r_2 = 0$, then the critical values of F are the positive even and negative odd integers. In a sense the critical values are the "easy integers," in that we are able to compute $\zeta_F(s)$ at the critical values in relatively simple terms. For example, in the trivial case $L(1, s) = \zeta(s)$ the critical values include the positive even integers, and $\zeta(m) \in \mathbb{Q}\pi^m$ for $m > 0$ even.

We now return to the situation in which X is an ètale variety of degree two, dimension zero defined by a separable quadratic equation $f(x) = 0$ with integral coefficients, as was the case in our example above when $f(x) = x^2+x-1$. In this setting the law of quadratic reciprocity can be formulated as follows: there is a one-to-one correspondence between these varieties X and quadratic Dirichlet characters χ via the relationship $\zeta(X, s) = \zeta(s)L(\chi, s)$. Let $F = \mathbb{Q}[x]/(f)$ be the coordinate ring of X, which will be a field unless f is reducible, which is the first case below. Otherwise we can classify the field F by the parity of χ as shown in the table.

$F \cong \mathbb{Q} \oplus \mathbb{Q}$	$\chi \equiv 1$ (χ trivial)
F is real quadratic ($n_+ = 2$, $n_- = 0$)	$\chi(-1) = 1$ (χ even)
F is imaginary quadratic ($n_+ = 1$, $n_- = 1$)	$\chi(-1) = -1$ (χ odd)

It should come as no surprise that the associated L-functions have analytic continuations given by

$$L^\star(\chi, s) = \gamma(s)L(\chi, s) = \begin{cases} |\Delta_F|^{s/2}\Gamma_+(s)L(\chi, s) & \chi \text{ even} \\ |\Delta_F|^{s/2}\Gamma_-(s)L(\chi, s) & \chi \text{ odd} \end{cases},$$

satisfying the functional equation $L^\star(\chi, s) = L^\star(\overline{\chi}, 1-s)$. Using $\gamma(s)$ we define the critical values $m \in \mathbb{Z}$ exactly as before. We then obtain the neat result that

$$(-1)^m\chi(-1)\,\mathrm{sgn}(m) = 1 \quad \Longleftrightarrow \quad m \text{ is critical for } L(\chi, s).$$

Let us now examine the behavior of $\zeta_F(s)$ near $s = 0,\ 1$ for a number field F. Since these are non-critical values, we expect the analysis to be more complicated. In fact, Dirichlet's class number formula will appear. Recall that if $F = \mathbb{Q}[x]/(f)$ with ring of integers $\mathcal{O}_F$ then

$$\zeta_F(s) = \prod_{\mathfrak{p}\subset\mathcal{O}_F} \left(1 - (N\mathfrak{p})^{-s}\right)^{-1} = \sum_{\mathfrak{a}\subset\mathcal{O}_F} (N\mathfrak{a})^{-s}$$

for $\Re(s) > 1$. We have seen that there is a simple pole at $s = 1$; we now wish to determine its residue. In general this quantity will involve the class group $\mathrm{Cl}(F)$ and the unit group $\mathcal{O}_F^*$. However, to provide an illustration of the overall sequence of ideas without raising too many technical issues, we will concentrate on the case in which F has class number one, i.e. $\mathcal{O}_F$ is a principal ideal domain.

To fix our ideas, suppose that $F = \mathbb{Q}(\sqrt{3})$. There are two real embeddings $\sigma_1, \sigma_2 : F \hookrightarrow \mathbb{R}$ which provide an embedding of $\mathcal{O}_F \hookrightarrow F \otimes \mathbb{R} \cong \mathbb{R}^2$ as a full lattice, namely $\alpha \mapsto (\alpha^{\sigma_1}, \alpha^{\sigma_2})$ for $\alpha \in \mathcal{O}_F$. In our case $\mathcal{O}_F \cong \mathbb{Z} + \sqrt{3}\,\mathbb{Z}$ so the image lattice has $(1, 1)$ and $(\sqrt{3}, -\sqrt{3})$ as a basis. Recall that a fundamental unit for $\mathcal{O}_F$ is $\epsilon = 2 + \sqrt{3}$, so the unit group is $\mathcal{O}_F^* = \pm\epsilon^n$ for $n \in \mathbb{Z}$. In our case the units are exactly those elements with norm one, hence precisely the points of the lattice on the hyperbola $xy = 1$. There is a cone $C \subset \mathbb{R}^2$ with the property that every ideal $\mathfrak{a} \subset \mathcal{O}_F$ is represented by exactly one lattice point within C, in that $\mathfrak{a} = (\alpha)$ for exactly one $\alpha \in \mathcal{O}_F$ whose image lies in C. (The reader may verify that the region bounded by two rays emanating from the origin, one passing through $(1, 1)$ and the other through $(\epsilon, \epsilon^{-1})$ together with the region bounded by rays through $(\sqrt{3}, -\sqrt{3})$ and $(\epsilon\sqrt{3}, -\epsilon^{-1}\sqrt{3})$ is a cone with this property. The set C should include the first and third rays but not the second or fourth.)

The same occurs for arbitrary number fields: we have $\mathcal{O}_F \hookrightarrow F \otimes \mathbb{R}$ as a full lattice L, a cone $C \subset \mathbb{R}^n$ containing exactly one representative of each ideal $\mathfrak{a} \subset \mathcal{O}_F$, and a norm $N : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0}$ satisfying $N(L) \subset \mathbb{Z}_{\geq 0}$ and $N(tx) = t^n N(x)$ for $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$. (See Shavarevich and Borevich for a more detailed exposition of these ideas.) Therefore the zeta-function may be written as

$$\zeta_F(s) = \sum_{\alpha\in L\cap C} \frac{1}{N(\alpha)^s}.$$

If we set $a_n = \{\alpha \in L \cap C | N(\alpha) = n\}$ then the previous formula becomes the Dirichlet series $\zeta_F(s) = \sum_{n=1}^{\infty} a_n n^{-s}$.

To proceed further we will need to approximate the quantity $A(x) = \sum_{n\leq x} a_n$, the number of lattice points in the cone C with norm at most x. Let C_x denote this region in $\mathbb{R}^n$. But counting lattice points in C_x approximates its volume; to be precise, $\mathrm{vol}(C_x) \sim A(x)|L|$, where $|L| = \mathrm{vol}(\mathbb{R}^n/L)$ is the volume of the lattice. (This volume can be computed by taking the determinant of a matrix of basis vectors of L.) On the other hand, because of the manner in which the norm map N scales, $\mathrm{vol}(C_x) = xV_1$, where $V_1 = \mathrm{vol}(C_1)$. (For example, in the instance $F = \mathbb{Q}(\sqrt{3})$ described above V_1 is the area of the region inside C bounded by the hyperbolas $xy = \pm 1$.) We conclude that $A(x) \sim \kappa x$ for $\kappa = V_1/|L|$.

One of basic tools for understanding Dirichlet series is Abel summation. In general, suppose that c_n for $n \geq 1$ is a sequence of complex numbers and define a function $C(t) = \sum_{n\leq t} c_n$ for $t \geq 1$. Then given a differentiable function $f(x)$ defined for $x \geq 1$ we find that

$$\begin{aligned} \sum_{1\leq n\leq x} c_n f(n) &= \sum_{1\leq n\leq x-1} C(n)(f(n) - f(n-1)) + C(x)f([x]) \\ &= C(x)f([x]) - \int_1^x f'(t)C(t)\,dt. \end{aligned}$$

We apply this formula by taking $f(t) = t^{-s}$ for a fixed s with $\Re(s) > 1$ and using the sequence a_n from above with partial sums $A(x) \sim \kappa x$. These substitutions yield

$$\sum_{1 \leq n \leq x} a_n n^{-s} = A(x)[x]^{-s} + \int_1^x st^{-s-1} A(t)\, dt.$$

Taking the limit as $x \to \infty$ produces

$$\sum_{n=1}^{\infty} a_n n^{-s} = \lim_{x \to \infty} \frac{A(x)}{x^s} + \int_1^{\infty} st^{-s} A(t) \frac{dt}{t}.$$

Since $\Re(s) > 1$ the first term in the sum goes to zero. If we were to have $A(x) = \kappa x$ then the second term would equal $\frac{\kappa s}{s-1}$ by direct integration. Since we only have $A(x) \sim \kappa x$ the integral will differ from this expression, but only by an amount which remains bounded as $s \to 1$ from above. (This last statement requires a little care to prove; see Frölich and Taylor, section VIII.2.) In summary, we have shown that

Theorem: *The residue of $\zeta_F(s)$ at $s = 1$ is given by*

$$\lim_{s \to 1^+} (s-1)\zeta_F(s) = \kappa = \frac{V_1}{|L|}.$$

The numbers V_1 and $|L|$ can be computed in terms of simple constants and intrinsic quantities associated with our field F. In general there exists a homomorphism $\log : \mathcal{O}_F^* \longrightarrow V_0$ from the unit group to a vector space V_0 of dimension $n_+ - 1$ whose image is a full lattice and whose kernel is $\mu(F)$, the roots of unity in F. In the case $F = \mathbb{Q}(\sqrt{3})$ the map looks like $\alpha \mapsto (\log|\alpha^{\sigma_1}|, \log|\alpha^{\sigma_2}|)$ and V_0 is the one-dimensional subspace in which the sum of the coordinates is zero. (Note the image of the log map lies in V_0.) The volume of this lattice divided by $\sqrt{r_1 + r_2}$ is called the regulator R. One can show via somewhat laborious computation that $V_1 = 2^{r_1}\pi^{r_2} R/w$, where w counts the number of roots of unity in F. Another calculation reveals that $|L| = 2^{r_2}|\Delta_F|^{-1/2}$. Finally, the class number h will make an appearance since in general the sum defining $\zeta_F(s)$ extends over all ideals $\mathfrak{a} \subset \mathcal{O}_F$, not just principal ideals. In summary, the Dirichlet class number formula gives the residue of $\zeta_F(s)$ at $s = 1$:

$$\lim_{s \to 1} (s-1)\zeta_F(s) = \frac{2^{r_1}(2\pi)^{r_2} hR}{w\sqrt{|\Delta_F|}}.$$

It is instructive to use this result to deduce the behavior of $\zeta_F(s)$ near $s = 0$ using the functional equation and various facts established above. One should find that $\zeta_F(s)$ has a zero of order $n_+ - 1$ with leading coefficient $-\frac{hR}{w}$. Testing this in the case $F = \mathbb{Q}$ we have $n_+ = 1$, $h = R = 1$, $w = 2$, and $\zeta(0) = -\frac{1}{2}$.

In a later lecture we will exhibit an analogous formula due to G. Humbert which involves the value of $\zeta_F(2)$ when F is an imaginary quadratic field. More precisely, the group $\Gamma = PGL_2(\mathcal{O}_F)$ acts on hyperbolic 3-space $\mathbb{H}^3$, and the volume of the quotient is

$$\mathrm{vol}(\Gamma \backslash \mathbb{H}^3) = \frac{3|\Delta_F|^{3/2}\zeta_F(2)}{4\pi}.$$

The formula is rather remarkable in that it connects two seemingly unrelated quantities.

Lectures 4–5: K-Theory, Part I

Our motivation for studying K-theory comes from a result of Borel involving a sequence of groups $K_n(A)$, $n \geq 0$ associated with a ring A. For number fields F we have defined $\zeta_F(s)$, a meromorphic function on the entire complex plane. Let ρ_m be the order of the zero of $\zeta_F(s)$ at $s = 1 - m$ for $m > 1$. Borel has proved that the rank of $K_{2m-1}(F)$ is ρ_m and that there is a homomorphism $\mathrm{reg} : K_{2m-1}(F) \longrightarrow \mathbb{R}^{\rho_m}$ whose image is a full lattice L with covolume

$$|L| = C\pi^N |\Delta_F|^{1/2} \zeta_F(m),$$

where C is a product of constants associated with the field F, N is an integer, and Δ_F is the discriminant of the field. (Using the functional equation, the above formula can also be stated in terms of the leading coefficient of $\zeta_F(s)$ expanded about $s = 1 - m$.) In other words, there is a formula analogous to the Dirichlet class number formula for $m > 1$.

The subject of K-theory has grown into an extensive, complicated branch of mathematics. It had its roots in topological considerations, which motivated the definitions of $K_0(A)$ and $K_1(A)$ given below. A fair amount of debate ensued over how best to define functors K_n for $n \geq 2$ from the category of associative (usually commutative, in practice) rings with 1 to the category of abelian groups. (Naturally, one would hope for a definition giving rise to nice unifying properties, such as the existence of long exact sequences.) In what follows, we will develop the foundation of the theory to the extent that we need it for future applications.

We define $K_0(A)$ to be the abelian group with one generator $[P]$ for each finitely generated projective module over A, subject to the relations

$$[P] + [Q] = [P \oplus Q].$$

Recall that an A-module P is projective if there exists an A-module Q such that $P \oplus Q \cong A^r$ for some $r \in \mathbb{N}$. Furthermore, we say that P and Q are stably isomorphic if $P \oplus A^r \cong Q \oplus A^r$ for some $r \in \mathbb{N}$.

Lemma: *$[P] = [Q]$ in $K_0(A)$ if and only if P and Q are stably isomorphic.*

The group $K_0(A)$ is actually a commutative ring, with the product operation given by the tensor product over A, so that $[P] \cdot [Q] = [P \otimes Q]$.

In the case that $A = \mathcal{O}_F$ is the ring of integers of a number field F we can state more precisely that

$$K_0(A) \cong \mathbb{Z} \oplus \mathrm{Cl}(F).$$

The correspondence is set up by the fact that any finitely generated $\mathcal{O}_F$-module is of the form $\mathfrak{a}_1 \oplus \cdots \oplus \mathfrak{a}_r$ for ideals $\mathfrak{a}_i \subset \mathcal{O}_F$. We map this element to $(r, \{\mathfrak{a}_1\mathfrak{a}_2 \cdots \mathfrak{a}_r\})$, involving the class of the product of the ideals. This correspondence is well-defined and provides the isomorphism. With this description the ring structure is given by the addition and multiplication rules

$$(r, \{\mathfrak{a}\}) + (s, \{\mathfrak{b}\}) = (r + s, \{\mathfrak{a}\}\{\mathfrak{b}\}),$$

$$(r, \{\mathfrak{a}\}) \cdot (s, \{\mathfrak{b}\}) = (rs, \{\mathfrak{a}\}^s\{\mathfrak{b}\}^r),$$

for r, $s \in \mathbb{Z}$. In particular, notice that multiplying any two rank zero elements yields 0, since taking $r = s = 0$ above gives a product of $(0, 1)$, which is 0 in $K_0(A)$. The product formula can be derived by choosing representatives $[\mathfrak{a}] + [A^{r-1}]$ and $[\mathfrak{b}] + [A^{s-1}]$, where $[A^{-n}]$ is interpreted as $-[A^n]$, then taking the tensor product. The relationship $\mathfrak{a} \oplus \mathfrak{b} \cong A \oplus \mathfrak{a}\mathfrak{b}$ is useful here.

We next describe the abelian group $K_1(A)$. Denote the direct limit (union) of the multiplicative groups $GL_n(A)$ as $GL(A)$, where $GL_n(A)$ embeds in $GL_{n+1}(A)$ via the homomorphism

$$M \mapsto \begin{bmatrix} M & 0 \\ 0 & 1 \end{bmatrix}.$$

On an intuitive level elements of $GL(A)$ are "infinite matrices" of the form $I_\infty + M_0$, where I_∞ is the "infinite identity" and M_0 has only a finite number of non-zero entries.

Given distinct positive integers i, j and $\lambda \in A$ we next define the elementary matrix $e_{ij}(\lambda) \in GL_n(A)$ to be the matrix which differs from the identity by a single off-diagonal entry λ in the ij^{th} position. The elementary matrices satisfy several simple relationships. For example, $e_{ij}(\lambda)e_{ij}(\mu) = e_{ij}(\lambda + \mu)$. Therefore $e_{ij}(\lambda)^{-1} = e_{ij}(-\lambda)$. We also have the commutator relationships:

$$[e_{ij}(\lambda), e_{kl}(\mu)] = \begin{cases} 1 & i \neq l, j \neq k \\ e_{il}(\lambda\mu) & j = k, i \neq l \\ e_{kj}(-\mu\lambda) & j \neq k, i = l \end{cases}.$$

Notice that the case $i = l$, $j = k$ is absent, as the commutator in this case is not equal to a single elementary matrix.

Let $E_n(A)$ be the subgroup of $GL_n(A)$ generated by all elementary matrices $e_{ij}(\lambda)$ with $i, j \leq n$, and in the same manner as before define $E(A)$ to be the direct limit of these groups. Since every elementary matrix has determinant equal to 1, we have $E(A) \subset SL(A) \subset GL(A)$. We now set $K_1(A) = GL(A)/E(A)$. The following result due to Whitehead will show that $K_1(A)$ is in fact the abelianization of $GL(A)$.

Lemma: *$E(A)$ is exactly the commutator subgroup of $GL(A)$.*

Proof: We first show that every elementary matrix is contained in the commutator subgroup. By the above relationships this fact is immediate, since $e_{ij}(\lambda) = [e_{ik}(\lambda), e_{kj}(1)]$ for any k distinct from i and j. We now show that the commutator of any two elements M and N of $GL(A)$ is contained in $E(A)$. We will treat M and N as matrices in $GL_n(A)$ for n sufficiently large. Note that the identity

$$\begin{bmatrix} M & 0 \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} N & 0 \\ 0 & N^{-1} \end{bmatrix} \begin{bmatrix} (NM)^{-1} & 0 \\ 0 & NM \end{bmatrix} = \begin{bmatrix} [M,N] & 0 \\ 0 & I \end{bmatrix}$$

shows that $[M, N]$ can be written as a product of matrices of a certain form, so it suffices to show that such matrices are in $E(A)$. But one can verify that

$$\begin{bmatrix} M & 0 \\ 0 & M^{-1} \end{bmatrix} = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ L & I \end{bmatrix} \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix} \begin{bmatrix} I & M^{-1}L \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -ML & I \end{bmatrix},$$

where $M = I + L$. Since an upper (or lower) triangular matrix with ones on the diagonal is clearly a product of elementary matrices, we have shown that $[M, N] \in E(A)$, completing the proof. Observe how the passage to the direct limit allowed us to write $[M, N]$ in the desired form; in general it is not true that $E_n(A)$ is the commutator subgroup of $GL_n(A)$.

If A is a commutative ring then we have a split exact sequence

$$1 \longrightarrow SL(A) \longrightarrow GL(A) \xrightarrow{\det} A^* \longrightarrow 1,$$

due to the natural embedding $A^* \cong GL_1(A) \hookrightarrow GL(A)$. Hence

$$K_1(A) \cong A^* \oplus SL(A)/E(A).$$

We know from linear algebra that $SL(A)$ is generated by elementary matrices when A is a field, so that $SL(A) = E(A)$ yielding $K_1(A) \cong A^*$. This also occurs when A is a local ring or a Dedekind domain, for example if $A = \mathcal{O}_F$ is the ring of integers of a number field. In these situations it is also true that $SL_n(A) = E_n(A)$ for $n \geq 3$. However, Swan points out that $A = \mathbb{Z}[\sqrt{-5}]$ is a counterexample for $n = 2$.

As noted above, there are several relationships among the elementary matrices which are valid regardless of the ring A involved. There may be others, though, which are dependent on the structure of A. Loosely speaking, the abelian group $K_2(A)$ represents these non-trivial relations. To formalize this construction, we introduce the Steinberg group $\mathrm{St}_n(A)$ for $n \geq 3$, which has one generator $x_{ij}(\lambda)$ for each pair $1 \leq i, j \leq n$ with $i \neq j$ and each $\lambda \in A$, subject to the same relations satisfied by the $e_{ij}(\lambda)$ above. For example, one has $x_{ij}(\lambda)x_{ij}(\mu) = x_{ij}(\lambda + \mu)$ and $[x_{ij}(\lambda), x_{jk}(\mu)] = x_{ik}(\lambda\mu)$ for distinct indices i, j, and k. There is a natural embedding $\mathrm{St}_n(A) \hookrightarrow \mathrm{St}_{n+1}(A)$ by which we can form the direct limit (union), denoted by $\mathrm{St}(A)$. Clearly the homomorphism from $\mathrm{St}_n(A)$ to $E_n(A)$ taking $x_{ij}(\lambda)$ to $e_{ij}(\lambda)$ is surjective, so passing to the limit we have a homomorphism $\phi : \mathrm{St}(A) \longrightarrow GL(A)$ whose image is exactly $E(A)$. We define $K_2(A)$ to be the kernel of ϕ. These constructions are summarized in the exact sequence

$$1 \longrightarrow K_2(A) \longrightarrow \mathrm{St}(A) \longrightarrow GL(A) \longrightarrow K_1(A) \longrightarrow 1.$$

The fact that $K_2(A)$ is in fact abelian is due to a theorem of Steinberg.

Theorem: *The center of* $\mathrm{St}(A)$ *is precisely* $K_2(A)$.

Proof: We first claim that the only element in the center of $E(A)$ is the identity. For if $y \in Z(E(A))$, then $y \cdot e_{ij}(1) = e_{ij}(1) \cdot y$ implies that $y_{ij} = 0$ and $y_{ii} = y_{jj}$. Hence y must be a scalar multiple of the identity in $E(A)$. But the only element with this property in $E(A)$ is the identity itself, since elements can only differ from the identity in a finite number of positions. It follows that if $y \in Z(\mathrm{St}(A))$ then $\phi(y) \in Z(E(A))$, and hence $\phi(y) = 1$, so $y \in K_2(A)$ by definition. This shows that $Z(\mathrm{St}(A)) \subset K_2(A)$.

Demonstrating the reverse inclusion will require a little more work. Hence suppose that $y \in K_2(A)$, meaning $\phi(y) = 1$. Choose n large enough so that y is a product of elements $x_{kl}(\lambda)$ with $k.l < n$. Let P_n be the subgroup of $\mathrm{St}_n(A)$ generated by the elements $x_{in}(\mu)$ with $i < n$. Then P_n is abelian by construction and essentially free, in that an element of P_n can be written uniquely in the form $x_{1n}(\mu_1) \cdots x_{n-1,n}(\mu_{n-1})$. Consequently, ϕ maps P_n isomorphically onto the subgroup of matrices in $E_n(A)$ of the form

$$\begin{bmatrix} 1 & & & & \mu_1 \\ & 1 & & & \mu_2 \\ & & \ddots & & \\ & & & 1 & \mu_{n-1} \\ & & & & 1 \end{bmatrix}.$$

We claim that if $p = x_{in}(\mu)$ for $i < n$ then $[y, p] \in P_n$. This follows from the observation that when $k, l < n$ we have $[x_{kl}(\lambda), x_{in}(\mu)] = 1$ or $x_{kn}(\lambda\mu)$, an element of P_n either way. Since y is a product of the $x_{kl}(\lambda)$, it follows that $[y, p] \in P_n$ as desired. But $\phi([y, p]) = 1$ because $\phi(y) = 1$, and since ϕ maps P_n isomorphically onto the matrix subgroup described above, we conclude that $\mu_1 = \mu_2 = \cdots = \mu_{n-1} = 1$, which means that $[y, p] = 1$. By a symmetrical argument we deduce that $[y, p] = 1$ for $p = x_{nj}(\mu)$ with $j < n$ as well. But elements of the form $x_{in}(\mu)$ or $x_{nj}(\mu)$ generate $\mathrm{St}_n(A)$, hence y is in the center of $\mathrm{St}_n(A)$. Passing to the limit as $n \to \infty$, we conclude that $\phi(y) = 1$ implies that $y \in Z(\mathrm{St}(A))$, which completes the proof.

Before continuing, we remark that if $E(A) \cong SL(A)$ and $K_2(A) = 1$ (which occurs when A is a finite field, as we shall see) then $\mathrm{St}(A) \cong SL(A)$. In fact, $\mathrm{St}_n(A) \cong SL_n(A)$ for $n \geq 5$ in this case. Thus we have a presentation for the groups $SL_n(A)$, which may have been Steinberg's original motivation for investigating the groups $\mathrm{St}_n(A)$. We also mention that the same does not apply when $A = \mathbb{Z}$: we have $E(\mathbb{Z}) = SL(\mathbb{Z})$, but it turns out that $K_2(\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$. The latter fact is a bit surprising, perhaps, since one expects functors to be "obvious" on $\mathbb{Z}$.

The exact sequence $1 \to K_2(A) \to \mathrm{St}(A) \to E(A) \to 1$ is an example of a central extension, which we now define. If G is an arbitrary group, and $\phi : X \to G$ a surjective group homomorphism with $\ker(\phi) \subset Z(X)$, then we call the pair (X, ϕ) a central extension of G. This concept arose naturally in several contexts. For example, in 1904 Schur showed that given a vector space V over K and a projective representation $\psi : G \to PGL(V) = GL(V)/K^*$, there exists a central extension $(\tilde{G}, \pi)$ of G which permits a lifting of ψ to $\tilde{\psi} : \tilde{G} \to GL(V)$. Briefly, the construction of $\tilde{G}$ goes as follows. For each $u \in G$ let $\tilde{f}(u) \in GL(V)$ be a representative for $\psi(u) \in PGL(V)$. (We declare $\tilde{f}(1) = 1$.) Because ψ is a homomorhpish we know that $\tilde{f}(u)\tilde{f}(v) = \tilde{f}(uv)\delta(u, v)$ for $\delta(u, v) \in K^*$. Then $\delta : G \times G \to K^*$ is a 2-cocycle (from group cohomology) and this gives rise to the central extension $\tilde{G} = \{(\alpha, u) | \alpha \in K^*, u \in G\}$, where $\pi : \tilde{G} \to G$ is projection on the second coordinate. The group law is given by $(\alpha, u) \cdot (\beta, v) = (\alpha\beta\delta(u, v), uv)$, and the lifting is given by $\tilde{\psi} : (\alpha, u) \mapsto \tilde{f}(u)\alpha$.

The concept of central extensions also appears in topology. Suppose that G is a connected topological group (such as $SL_n(\mathbb{R})$), and let $\pi : X \to G$ be a covering of G by a connected topological space X. We claim that (X, π) is a central extension of G. Suppose that $n \in \ker(\pi)$ and consider the homeomorphism $\psi : X \to X$ defined by $\psi(x) = nxn^{-1}$. Since $\pi \circ \psi = \pi$ and $\psi(1) = 1$ we conclude by properties of covering maps that ψ is the identity on the connected component of X containing 1, i.e. $nxn^{-1} = x$ on X, so $n \in Z(X)$. Furthermore, if G is a connected Lie group, then we know there exists a universal cover $\phi : \tilde{G} \to G$ which has the universal property that given any cover $\pi : X \to G$, there exists a unique morphism $h : G \to X$ for which $\phi = \pi \circ h$.

We will look for a similar such algebraic object (U, ν) which is a central extension of G and satisfies a corresponding universal condition, namely given any central extension (X, ϕ) of G there is a unique homomorphism $h : U \to X$ over G, i.e. such that $\phi \circ h = \nu$. As usual, this universal property allows us to conclude that (U, ν) is unique up to unique isomorphism, if it exists. To describe universal central extensions we will need the concepts of perfect groups and split extensions. We say that a group is perfect if it equals its commutator subgroup, so that $G = [G, G]$. This condition will be analogous to the requirement of connectedness imposed above. Next, we say that a central extension $\phi : X \to G$ splits if it admits a section $s : G \to X$ with $\phi \circ s = \mathrm{id}_G$. In this case the extension is essentially trivial, since the sequence $1 \to \ker(\phi) \to X \to G \to 1$ splits, yielding $X \cong G \times \ker(\phi)$. With this terminology we may now assert that

Theorem: *A central extension (U, ν) of G is universal if and only if U is perfect and every central extension of U splits.*

Proof: Let (X, ϕ) be any central extension of G, and first assume that every central extension of U splits. This will imply the existence of a homomorphism $h : U \to X$ over G. Look at the pull-back $U \times_G X = \{(u, x) | \nu(u) = \phi(x)\}$, a subgroup of $U \times X$. Let π_1 and π_2 be the projections onto the first and second coordinates. Clearly $\ker(\pi_1)$ is the set of elements $(1, x)$ with $\phi(x) = 1$. Hence

$x \in Z(X)$ since (X, ϕ) is a central extension. It follows that $(1, x) \in Z(U \times_G X)$, thus $(U \times_G X, \pi_1)$ is a central extension of U. By hypothesis there is a section $s : U \to U \times_G X$, and setting $h = \pi_2 \circ s$ provides the desired homomorphism. This construction is summarized in the commutative diagram below.

$$\begin{tikzcd}
U \times_G X \arrow[r, phantom, "\pi_1", shift left=2] \arrow[r, phantom, "s"', shift right=2] \arrow[d, phantom, "\pi_2"'] & U \arrow[d, phantom, "\nu"] \arrow[dl, phantom, "h"] \\
X \arrow[r, phantom, "\phi"] & G
\end{tikzcd}$$

We next show that h is unique if U is perfect.

Lemma 1: *Let (X, ϕ) and (Y, ψ) be two central extensions of G and suppose $h : Y \to X$ is a homomorphism over G, so that $\phi \circ h = \psi$. If Y is perfect, then h is unique.*

Proof: Let h_1, h_2 be two such maps over G. Then for all $y \in Y$, $h_1(y) = h_2(y)x$ for some $x \in \ker(\phi) \subset Z(X)$. The same is true for any $y' \in Y$, so we find by an easy argument that $h_1([y, y']) = h_2([y, y'])$. But Y is perfect, so generated by elements of the form $[y, y']$. This shows that $h_1 = h_2$.

An immediate application of the lemma shows that our map h above is unique. We remark that given a central extension (Y, ψ) in which Y is not perfect, it is not hard to produce an example of another central extension (X, ϕ) and distinct homomorphisms h_1, h_2 over G. This shows that if U is universal, then it must be perfect.

Lemma 2 *If (X, ϕ) is a central extension of a perfect group G, then $X' = [X, X]$ is also perfect and ϕ restricted to X' is still surjective onto G.*

Proof: Clearly $\phi(X') = \phi([X, X]) = [\phi(X), \phi(X)] = [G, G] = G$, which demonstrates the surjectivity. Now for any $x \in X$ we have $\phi(x) = \phi(x')$ for some $x' \in X'$. Hence $x = x'n$ for an $n \in \ker(\phi) \subset Z(X)$. So for arbitrary $x_1, x_2 \in X$ it follows that $[x_1, x_2] = [x'_1, x'_2]$. But this is exactly what we need to show that X' is perfect.

To complete the proof of the theorem it remains to show that if (U, ν) is a universal extension of G, then every central extension of U splits. (We already know that U is perfect.) So suppose that (X, ϕ) is a central extension of U. We claim that $(X, \nu \circ \phi)$ is a central extension of G. The composite map is clearly surjective. Now let $\nu(\phi(x_0)) = 1$, which implies that $\phi(x_0) \in Z(U)$. Consider the homomorphism $h_0 : X \to X$ defined by $h_0(x) = x_0 x x_0^{-1}$. Note that $\phi(x) = \phi(h_0(x))$ since $\phi(x_0)$ is in the center of U. We now restrict to $h_0 : X' \to X'$, because X' is perfect by lemma two. Then h_0 and the identity map both commute with $\phi : X' \to U$, so by lemma one they must be the same homomorphism, which shows that x_0 commutes with all elements x' of X'. But by lemma two $\phi : X' \to U$ is surjective, so given any $x \in X$ we have $x = x'n$ for some $n \in \ker(\phi) \subset Z(X)$. Since x_0 commutes with both x' and n, it commutes with x, so $x_0 \in Z(X)$, proving that $(X, \nu \circ \phi)$ is a central extension of G.

Finally, we employ the universal property of U to obtain the map $h : U \to X$, which will be the desired section. We know that $\nu \circ \phi \circ h = \nu$, hence $\phi \circ \nu$ and $\mathrm{id} : U \to U$ are both homomorphisms over G, and hence are the same since U

is perfect, by lemma one. This proves that h is a section, completing the proof.

$$\begin{tikzcd}
X \arrow[dr, "\phi"'] & & U \arrow[ll, "h"'] \arrow[dl, "\text{id}"'] \arrow[dd, "\nu"] \\
 & U \arrow[dr, "\nu"'] & \\
 & & G
\end{tikzcd}$$

Theorem: *G has a universal central extension if and only if G is perfect.*

Proof: To begin with, suppose that G has a universal central extension (U, ν). Then because $\nu(U) = G$ and U is perfect it follows at once that G is perfect.

Conversely, suppose that G is perfect. There is a standard resolution of G given by $1 \to R \to F \to G \to 1$, where F is a free group with one generator for each element of G, and the kernel R represents the relations among the generators of F dictated by the group structure of G. Consider the subgroup $[R, F]$ of F generated by all elements of the form $rfr^{-1}f^{-1}$. Since R is normal in F then so is $[R, F]$, and in fact $[R, F] \subset R$. Therefore we have a surjective map

$$\phi : F/[R, F] \longrightarrow F/R \cong G.$$

By construction $\ker(\phi)$ is in the center of $F/[R, F]$. It follows from lemma two that $(F/[R, F])' \cong [F, F]/[R, F]$ is perfect. The restriction of ϕ to this subgroup is still surjective, so it is a central extension.

We claim that $([F, F]/[R, F], \phi)$ is in fact the universal central extension of G. Suppose that (X, ψ) is any other central extension of G. Because F is a free group on the elements of G and ψ is surjective, there is certainly a homomorphism $h : G \to X$ which satisfies $\psi \circ h = \pi$, where $\pi : F \to G$ is the projection map. Now given $r \in R$ we have $\pi(r) = 1$, so $\psi(h(r)) = 1$ implying that $h(r) \in \ker(\phi) \subset Z(X)$. Thus $h(rfr^{-1}f^{-1}) = 1$, so that $h([R, F]) = 1$ and h factors to a homomorphism from $F/[R, F]$ to X. Finally, restricting h to the subgroup $[F, F]/[R, F]$ gives a homomorphism to X over G which commutes with ϕ. By lemma one h must be unique, which establishes the universal property.

The kernel of $\phi : [F, F]/[R, F] \to G$ is the subgroup $(R \cap [F, F])/[R, F]$, often called the Schur multiplier of G, after the man who first introduced the concept of a central extension. It is isomorphic to $H_2(G)$, as we shall see later when we present a more general method for constructing the groups $K_n(A)$ using tools from algebraic topology.

We are now prepared to explain why $\text{St}(A)$ is the universal central extension of $E(A)$. Clearly $\text{St}(A)$ is perfect, since every generator $x_{ij}(\lambda) = [x_{ik}(\lambda), x_{kj}(1)]$ for any k distinct from i and j. Now suppose that (X, ϕ) is a central extension of $\text{St}(A)$; we must show there exists a section $s : \text{St}(A) \to X$. Given $i \neq j$ and $\lambda \in A$ choose h distinct from i and j and $y, y' \in X$ such that $\phi(y) = x_{ih}(1)$ and $\phi(y') = x_{hj}(\lambda)$. Then setting $s_{ij}(\lambda) = [y, y']$ is well-defined and by construction $\phi(s_{ij}(\lambda)) = x_{ij}(\lambda)$. We now obtain the desired section by mapping $x_{ij}(\lambda)$ to $s_{ij}(\lambda)$. (Note that this proof outline omits a page or two of algebra!) In general $\text{St}_n(A)$ is also the universal central extension of $E_n(A)$ for $n \geq 5$, but there are a few exceptions for smaller n. For example, we shall soon see that $K_2(A)$ is trivial when A is a finite field. Therefore $\text{St}_2(\mathbb{F}_4) = E_2(\mathbb{F}_4) = SL_2(\mathbb{F}_4)$. But $SL_2(\mathbb{F}_4) \cong PSL_2(\mathbb{F}_5)$, which has $SL_2(\mathbb{F}_5)$ as a non-trivial cover. In other words, no section exists for this central extension, so $\text{St}_2(\mathbb{F}_4)$ cannot be the universal central extension of $E_2(\mathbb{F}_4)$. Similarly, the fact that $SL_3(\mathbb{F}_2) \cong PSL_2(\mathbb{F}_7)$ shows that $\text{St}_3(\mathbb{F}_2)$ cannot be the universal central extension of $E_3(\mathbb{F}_2)$.

Lectures 6–8: K-Theory, Part II

Our primary goal in what follows will be to construct a skew-symmetric, bimultiplicative symbol $\{\cdot,\cdot\} : A^* \times A^* \to K_2(A)$ when A is commutative and prove that it satisfies the Steinberg relation $\{u, 1-u\} = 1$ when $u, 1-u \in A^*$. We begin by developing a few relationships among the generators of the Steinberg group $\mathrm{St}(A)$. Let

$$w_{ij}(u) = x_{ij}(u)x_{ji}(-u^{-1})x_{ij}(u), \qquad h_{ij}(u) = w_{ij}(u)w_{ij}(-1).$$

As usual $\phi : \mathrm{St}(A) \to E(A)$ is the homomorphism taking $x_{ij}(u)$ to $e_{ij}(u)$. One can easily verify that

$$\phi(w_{ij}(u)) = W_{ij}(u) = \begin{bmatrix} \ddots & & & & \\ & 0 & & u & \\ & & \ddots & & \\ & -u^{-1} & & 0 & \\ & & & & \ddots \end{bmatrix}$$

differs from the identity matrix only in the four positions shown. (We are assuming that $i < j$ for these illustrations.) Similarly

$$\phi(h_{ij}(u)) = H_{ij}(u) = \begin{bmatrix} \ddots & & & & \\ & u & & & \\ & & \ddots & & \\ & & & u^{-1} & \\ & & & & \ddots \end{bmatrix}.$$

By definition $W_{ij}(u)$ and $H_{ij}(u)$ are a product of elementary matrices, hence lie in $E(A)$.

In general suppose that (X, ϕ) is a central extension of G and denote $\ker(\phi)$ by N. There is a method for constructing elements of N given any two commuting elements g and h of G. Choose $x, y \in X$ with $\phi(x) = g$ and $\phi(y) = h$. It is easy to check that the commutator $[x, y]$ does not depend on the choice of x and y; we denote the result by $g \star h$. Since g and h commute, $\phi([x, y]) = 1$, hence $g \star h \in N$. We claim that $\star$ is skew-symmetric, bimultiplicative, and invariant under inner automorphisms. These facts will follow from the commutator identities

i. $[x, y]^{-1} = [y, x]$,

ii. $[x_1x_2, y] = [x_1, [x_2, y]][x_2, y][x_1, y]$,

iii. $[wxw^{-1}, wyw^{-1}] = w[x, y]w^{-1}$.

For example, to prove bimultiplicativity suppose that g_1 and g_2 each commute with h, and let x_1, x_2, and y be lifts of these elements to X. Then

$$\begin{aligned} (g_1g_2) \star h &= [x_1x_2, y] \\ &= [x_1, [x_2, y]][x_2, y][x_1, y] \\ &= [x_1, y][x_2, y] = (g_1 \star h)(g_2 \star h). \end{aligned}$$

The factor of $[x_1, [x_2, y]]$ vanishes because the inner commutator is in N (since g_2 and h commute), so it lies in the center of X. The other statements may be proved in a similar fashion.

Now suppose that A is a commutative ring and consider the universal central extension

$$1 \longrightarrow K_2(A) \longrightarrow \mathrm{St}(A) \xrightarrow{\phi} E(A) \longrightarrow 1.$$

Clearly $H_{12}(u)$ and $H_{13}(v)$ are commuting elements of $E(A)$ for any $u, v \in A^*$, so we obtain the element $H_{12}(u) \star H_{13}(v) = [h_{12}(u), h_{13}(v)]$ in $K_2(A)$, which we denote by $\{u, v\}$. All the properties of $\star$ transfer immediately, so we have a skew-symmetric, bimultiplicative symbol $\{\cdot, \cdot\} : A^* \times A^* \to K_2(A)$. We shall soon see that $\{u, v\} = [h_{ij}(u), h_{ik}(v)]$ for any distinct i, j, and k, so our definition was fully general, in a sense. We will also see that $[h_{ij}(u), h_{kl}(v)] = 1$, while $[h_{ij}(u), h_{ij}(v)] = \{u, v\}^2$, which explains why we avoided those cases.

The proof of these facts hinges on a key observation. First recall that the permutation matrix P_σ associated with a permutation σ is the matrix with a 1 in row $\sigma(i)$, column i for $1 \leq i \leq n$ and 0's elsewhere. We embed P_σ in $GL(A)$ in the usual manner. Multiplication by P_σ on the left permutes the first n rows according to σ; multiplication by P_σ^{-1} on the right has the same effect on the first n columns. We adopt the shorthand $\sigma(ij)$ to represent the pair $\sigma(i), \sigma(j)$. Finally, let $\mathrm{diag}(v_1, \ldots, v_n)$ denote the matrix in $GL(A)$ with entries $v_1, \ldots, v_n$, 1, 1, ... along the diagonal.

Theorem: *Suppose $W = P_\sigma \mathrm{diag}(v_1, \ldots, v_n)$ is an element of $E(A)$, and choose $w \in \mathrm{St}(A)$ with $\phi(w) = W$. Then conjugation by w is given by*

$$w x_{ij}(\lambda) w^{-1} = x_{\sigma(ij)}(v_i \lambda v_j^{-1}).$$

Proof: First note that the requirement that $P_\sigma \mathrm{diag}(v_1, \ldots, v_n) \in E(A)$ forces $v_1 v_2 \cdots v_n = \mathrm{sign}(\sigma) = \pm 1$ since $E(A) \subset SL(A)$, so in particular the $v_i \in A^*$ are invertible. We also point out that properties of central extensions ensure that conjugation by w is independent of the choice of w. We claim that the map

$$\psi : x_{ij}(\lambda) \mapsto x_{\sigma(ij)}(v_i \lambda v_j^{-1})$$

gives rise to an automorphism of $\mathrm{St}(A)$. The map ψ is clearly a bijection on the generators of $\mathrm{St}(A)$, so we need only check that ψ respects the relations among them. This is routine to verify; for instance,

$$\psi([x_{ij}(\lambda), x_{jk}(\mu)]) = \psi(x_{ik}(\lambda\mu)) = x_{\sigma(ik)}(v_i \lambda \mu v_k^{-1}).$$

On the other hand,

$$\begin{aligned} [\psi(x_{ij}(\lambda)), \psi(x_{jk}(\mu))] &= [x_{\sigma(ij)}(v_i \lambda v_j^{-1}), x_{\sigma(jk)}(v_j \mu v_k^{-1})] \\ &= x_{\sigma(ik)}(v_i \lambda v_j^{-1} v_j \mu v_k^{-1}) \\ &= x_{\sigma(ik)}(v_i \lambda \mu v_k^{-1}), \end{aligned}$$

the same result.

We now observe that there are two possible automorphisms of $\mathrm{St}(A)$ which will make the diagram below commute, where the bottom homomorphism is conjugation by W.

$$\begin{tikzcd} \mathrm{St}(A) \arrow[r, dashed, "?"] \arrow[d, "\phi"'] & \mathrm{St}(A) \arrow[d, "\phi"'] \\ E(A) \arrow[r, dashed, "W \cdot W^{-1}"] & E(A) \end{tikzcd}$$

Clearly conjugation by w across the top makes the diagram commute since $\phi(w) = W$. To see that $\psi : \mathrm{St}(A) \to \mathrm{St}(A)$ also works we need only verify commutativity on the generators $x_{ij}(\lambda)$. On the one hand,

$$\phi(\psi(x_{ij}(\lambda)) = \phi(x_{\sigma(ij)}(v_i \lambda v_j^{-1})) = e_{\sigma(ij)}(v_i \lambda v_j^{-1}).$$

On the other hand, the reader may check that $We_{ij}(\lambda)W^{-1}$ produces precisely the same result by multiplying the corresponding matrices. But now we have two central extensions of $E(A)$, namely $(\mathrm{St}(A), W \cdot W^{-1} \circ \phi)$ and $(\mathrm{St}(A), \phi)$, and two homomorphisms between them over $E(A)$. Since $\mathrm{St}(A)$ is perfect, we conclude by lemma one that the homomorphisms are the same, which is the statement of the theorem. Since $St_n(A)$ is the universal central extension of $E_n(A)$ for $n \geq 5$, an entirely analogous argument shows that the theorem holds in the finite matrix situation as well when $n \geq 5$.

Thus to determine the effect of conjugation by $w_{ij}(u)$ in the Steinberg group we need only write $\phi(w_{ij}(u)) = W_{ij}(u)$ in the form $P_\sigma D$. This is accomplished by $\sigma = (i\ j)$ and $D = \mathrm{diag}(\dots, -u^{-1}, \dots, u, \dots)$, where $v_i = -u^{-1}$, $v_j = u$, and all other $v_k = 1$. Denoting this conjugation by ψ we discover that

$$\begin{aligned} w_{ij}(u) = \psi(w_{ij}(u)) &= \psi(x_{ij}(u))\psi(x_{ji}(-u^{-1}))\psi(x_{ij}(u)) \\ &= x_{ji}(-u^{-1})x_{ij}(u)x_{ji}(-u^{-1}) \\ &= w_{ji}(-u^{-1}), \end{aligned}$$

where we used $\psi(x_{ij}(u)) = x_{\sigma(ij)}(-u^{-1}uu^{-1}) = x_{ji}(-u^{-1})$ and similar computations in the middle step. This provides our first non-trivial relation among the $x_{ij}(u)$. To be precise, letting $\alpha = x_{ij}(u)$ and $\beta = x_{ji}(-u^{-1})$ we have shown that $\alpha\beta\alpha = \beta\alpha\beta$.

The equality just mentioned is an example of a braid relation, which we pause to describe. Physically, a braid on three vertices consists of three pieces of string fastened to a set of three pegs on either end, with any number of crossings inbetween, as illustrated below. Two braids are equivalent if one can be continuously shifted to match the second, such as the two shown.

A braid

Equivalent braids

The set of all equivalence classes of braids form a group, with the group operation given by attaching the end of one braid to the start of the next. We illustrate this procedure on three simple braids labeled α, β, and ϵ.

$\alpha =$

$\beta =$

$\epsilon =$

Here we have drawn the braids $\alpha \circ \beta$, $\beta \circ \alpha$, and $\epsilon \circ \alpha$.

$\alpha \circ \beta =$ $\beta \circ \alpha =$ $\epsilon \circ \alpha =$

From these examples it should be clear that this braid group is not commutative and has ϵ as the identity. The reader is invited to construct the braid representing α^{-1} and also verify that $\alpha\beta\alpha = \beta\alpha\beta$, hence the name "braid relation." In fact, α and β generate this braid group and satisfy $\alpha\beta\alpha = \beta\alpha\beta$ as the sole relation. In general, the braid group on n vertices is generated by the $n-1$ left-over-right crossings on adjacent vertices, subject to the relations that crossings with no vertices in common commute, while crossings with a vertex in common, such as α and β, satisfy the braid relation.

As a second application of the theorem, we show that our definition of $\{u, v\} = [h_{12}(u), h_{13}(v)]$ is independent of the indices chosen. Given distinct positive integers i, j, and k select an even permutation $\sigma \in S_n$ with $\sigma(i) = 1$, $\sigma(j) = 2$, and $\sigma(k) = 3$. (Such a permutation always exists for n large enough.) Let $D = \operatorname{diag}(1, \ldots, 1)$ be the identity matrix, so $v_m = 1$ for all m. Then $P_\sigma D \in E(A)$, and choosing $w \in \operatorname{St}(A)$ with $\phi(w) = P_\sigma D$ we conclude by our theorem that conjugation by w maps $x_{lm}(u) \mapsto x_{\sigma(lm)}(u)$. Using property *iii.* of the commutator identities listed above we deduce that

$$\begin{aligned}
[h_{ij}(u), h_{ik}(v)] &= w[h_{ij}(u), h_{ik}(v)]w^{-1} \\
&= [wh_{ij}(u)w^{-1}, wh_{ik}(v)w^{-1}] \\
&= [h_{\sigma(ij)}(u), h_{\sigma(ik)}(v)] \\
&= [h_{12}(u), h_{13}(v)] = \{u, v\}.
\end{aligned}$$

Here we used the fact that $h_{ij}(u)$ is a product of six generators, so $wh_{ij}(u)w^{-1}$ is a product of corresponding generators with the indices i, j, and k replaced by 1, 2, and 3, yielding $h_{12}(u)$, and similarly for $h_{13}(v)$.

In general, let $U = \operatorname{diag}(u_1, \ldots, u_n)$ and $V = \operatorname{diag}(v_1, \ldots, v_n)$ with the stipulation that $\prod u_i = \prod v_i = 1$. Setting $m = n+1$ and $m' = n+2$ for ease of notation we see that $U, V \in E(A)$ by writing

$$U = H_{1m}(u_1)H_{2m}(u_2)\cdots H_{nm}(u_n), \quad V = H_{1m'}(v_1)H_{2m'}(v_2)\cdots H_{nm'}(v_n).$$

Observe that $H_{ij}(u) \star H_{kl}(v) = [h_{ij}(u), h_{kl}(v)] = 1$ when the indices are all distinct, because every generator in the product for $h_{ij}(u)$ commutes with every one in $h_{kl}(v)$. Combining this fact with the bimultiplicativity of $\star$ we find that

$$\begin{aligned}
U \star V &= \left(\prod_{i=1}^{n} H_{im}(u_i)\right) \star \left(\prod_{j=1}^{n} H_{jm'}(v_j)\right) \\
&= \prod_{i,j=1}^{n} H_{im}(u_i) \star H_{jm'}(v_j) \\
&= \prod_{k=1}^{n} H_{km}(u_k) \star H_{km'}(v_k) \\
&= \prod_{k=1}^{n} \{u_k, v_k\}.
\end{aligned}$$

Therefore the subgroup of $K_2(A)$ containing $U \star V$ for all pairs of diagonal matrices is generated by the elements $\{u, v\}$. Finally, we mention that in many cases we lose nothing by restricting ourselves to elements of the form $\{u, v\}$ in $K_2(A)$, since in these cases any pair of commuting matrices U and V can either be reduced to the case of diagonal matrices, or will automatically give $U \star V = 1$ in $K_2(A)$ by the nature of its construction. For instance, Matsumoto has shown that $K_2(F)$ is generated by the elements $\{u, v\}$ when F is a field or skew-field.

Let us now prove a fundamental property of the symbol $\{\cdot, \cdot\}$.

Theorem: (Steinberg) *Let A be a commutative ring, and let $u, v \in A^*$ be invertible elements with $u + v = 1$. Then $\{u, v\} = 1$.*

Proof: By definition, $\{u, v\} = [h_{12}(u), h_{13}(v)]$, so showing that $\{u, v\} = 1$ is equivalent to proving that conjugation by $h_{12}(u)$ fixes $h_{13}(v)$. Let ψ denote conjugation by $h_{12}(u)$. Since $\phi(h_{12}(u)) = P_\sigma D$ for the identity permutation σ and $D = \operatorname{diag}(u, u^{-1})$, we conclude that the action of ψ on generators is $x_{ij}(v) \mapsto x_{ij}(v_i v v_j^{-1})$ where $v_1 = u$, $v_2 = u^{-1}$, and $v_i = 1$ for $i \geq 3$. Therefore $\psi(w_{13}(\lambda)) = w_{13}(u\lambda)$ and we have

$$\psi(h_{13}(v)) = \psi(w_{13}(v))\psi(w_{13}(-1) = w_{13}(uv)w_{13}(-u).$$

So checking that ψ fixes $h_{13}(v)$ boils down to verifying that $w_{13}(v)w_{13}(-1) = w_{13}(uv)w_{13}(-u)$, or that

$$w_{13}(v)w_{31}(1)w_{13}(u) = w_{13}(uv),$$

since $w_{13}(u)^{-1} = w_{13}(-u)$ from the definitions and $w_{ij}(\lambda) = w_{ji}(-\lambda^{-1})$. Before plunging into the algebraic morass we first recall that conjugation by $w_{13}(\lambda)$ corresponds to the transposition $\sigma = (1\ 3)$ and the diagonal entries $v_1 = -\lambda^{-1}$, $v_3 = \lambda$. Thus

$$w_{13}(v)x_{31}(1) = x_{13}(-v^2)w_{13}(v), \quad \text{and} \quad w_{13}(u)x_{13}(-u^2) = x_{31}(1)w_{13}(u).$$

We are now able to compute

$$\begin{aligned}
&w_{13}(v)w_{31}(1)w_{13}(u) \\
&\quad = w_{13}(v)x_{31}(1)x_{13}(-1)x_{31}(1)w_{13}(u) \\
&\quad = x_{13}(-v^2)w_{13}(v)x_{13}(-1)w_{13}(u)x_{13}(-u^2) \\
&\quad = x_{13}(-v^2)x_{13}(v)x_{31}(-v^{-1})x_{13}(v)x_{13}(-1)x_{13}(u)x_{31}(-u^{-1})x_{13}(u)x_{13}(-u^2) \\
&\quad = x_{13}(v - v^2)x_{31}(-v^{-1})x_{13}(u + v - 1)x_{31}(-u^{-1})x_{13}(u - u^2) \\
&\quad = x_{13}(uv)x_{31}(-u^{-1} - v^{-1})x_{13}(uv) \\
&\quad = x_{13}(uv)x_{31}(-(uv)^{-1})x_{13}(uv) \\
&\quad = w_{13}(uv).
\end{aligned}$$

Here we used the fact that $x_{13}(u + v - 1) = x_{13}(0) = 1$ and $-u^{-1} - v^{-1} = -(uv)^{-1}$, which follows from $u + v = 1$.

Matsumoto proved in his thesis that if F is a field, then $K_2(F)$ is generated by elements of the form $\{u, v\}$, with $u, v \in F^*$. Furthermore, he demonstrated at the same time that the generators $\{u, v\}$ are subject only to the relations (and their consequences) that we have seen so far, namely

i. $\{u, v\}^{-1} = \{v, u\}$,

ii. $\{u_1 u_2, v\} = \{u_1, v\}\{u_2, v\}$,

iii. $\{u, 1 - u\} = 1$.

Let us derive a few of the formal consequences of these relations. The first two conditions easily imply that $\{\cdot,\cdot\}$ is multiplicative in the second variable. Equally clearly we find that $\{u,1\} = \{1,v\} = 1$ and $\{u^{-1},v\} = \{u,v^{-1}\} = \{u,v\}^{-1}$. The Steinberg relation also implies that $\{u,-u\} = 1$. For if $u \in F$, $u \neq 0,1$ then the identity $-u = (1-u)/(1-u^{-1})$ yields

$$\{u,-u\} = \{u,1-u\}\{u,1-u^{-1}\}^{-1} = \{u,1-u\}\{u^{-1},1-u^{-1}\} = 1.$$

The same argument shows that if A is a commutative ring and $u, 1-u \in A^*$ then $\{u,-u\} = 1$.

Now suppose that F is a field, G is an abelian group, and $c : F^* \times F^* \to G$ is any map satisfying the three relations above, and hence the consequences just derived as well. We call c a Steinberg symbol. Matsumoto's theorem then states that $K_2(F)$ is universal in the sense that there is a unique homomorphism $h : K_2(F) \to G$ for which $c = h \circ \{\cdot,\cdot\}$. Equivalently, the set of symbols $c : F^* \times F^* \to G$ is in one-to-one correspondence with the set $\mathrm{Hom}(K_2(F), G)$.

Up to this point we have defined $K_2(A)$ and developed a number of interesting properties concerning this abelian group, but have yet to actually compute $K_2(A)$ for any ring of interest, or even show that it is non-trivial. We shall now rectify this state of affairs. By definition, if a product of elementary matrices $e_{ij}(\lambda)$ is the identity in $E(A)$, then the corresponding product of generators $x_{ij}(\lambda)$ in $\mathrm{St}(A)$ will lie in $K_2(A)$. Because the relations among the generators were designed to mimic those among the elementary matrices, the product will usually, but not always, be the identity in $K_2(A)$. For instance, consider

$$T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = W_{12}(1) = \phi(w_{12}(1)) = \phi(x_{12}(1)x_{21}(-1)x_{12}(1)).$$

Since $T^4 = I$ is the identity matrix, then $(x_{12}(1)x_{21}(-1)x_{12}(1))^4 \in K_2(A)$ is such a product. As an exercise, show that this product equals $\{-1,-1\}$. In the case $A = \mathbb{Z}$ we obtain a non-trivial element, which in fact generates $K_2(\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$.

With a little more effort we can demonstrate that $K_2(\mathbb{R})$ also has $\{-1,-1\}$ as a non-trivial element. Let $\mathbb{H}$ be the Hamiltonian quaternion algebra with basis elements $\mathbf{1}$, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ over $\mathbb{R}$ satisfying the usual multiplication rules. If $\alpha = \alpha_0\mathbf{1}+\alpha_1\mathbf{i}+\alpha_2\mathbf{j}+\alpha_3\mathbf{k}$ is a quaternion then its norm is $N(\alpha) = \alpha_0^2+\alpha_1^2+\alpha_2^2+\alpha_3^2$. Let $\mathbb{H}_1$ be the subgroup of the quaternions with norm one. Then there is an exact sequence

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathbb{H}_1 \longrightarrow SO_3(\mathbb{R}) \longrightarrow 1$$

in which $\alpha \in \mathbb{H}_1$ is taken to the transformation $x \mapsto \alpha x \alpha^{-1}$. Here we treat $x = (x_1,x_2,x_3) \in \mathbb{R}^3$ as a pure quaternion $x = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ to define the image in the special orthogonal group. The elements $\mathbf{i}$ and $\mathbf{j}$ are mapped to the transformations represented by the matrices

$$\mathbf{i} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \mathbf{j} \mapsto \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Using the natural embedding $SO_3(\mathbb{R}) \hookrightarrow SL_3(\mathbb{R}) \subset SL(\mathbb{R})$ and employing topological arguments, one can show that $[\mathbf{i},\mathbf{j}] = -\mathbf{1}$ implies that $\{-1,-1\} = -1$ in $K_2(\mathbb{R})$. The details are left to the reader. Since $\mathbb{Z} \hookrightarrow \mathbb{R}$ there is a natural map $K_2(\mathbb{Z}) \to K_2(\mathbb{R})$ sending $\{-1,-1\}_{\mathbb{Z}} \mapsto \{-1,-1\}_{\mathbb{R}}$, which shows once again that $\{-1,-1\}$ is non-trivial in $K_2(\mathbb{Z})$. It turns out that $K_2(\mathbb{R}) \cong \mathbb{Z}/2\mathbb{Z} \times D$ for a divisible group D with the same cardinality as $\mathbb{R}$.

Let us now prove that

Theorem: $K_2(\mathbb{F}_q) = 0$ *for a finite field.*

This means that $\mathrm{St}(\mathbb{F}_q) \cong SL(\mathbb{F}_q)$, giving a description of $SL(\mathbb{F}_q)$ in terms of generators and relations. (We also have $\mathrm{St}_n(\mathbb{F}_q) \cong SL_n(\mathbb{F}_q)$ for $n \geq 5$.)

Proof: By Matsumoto's theorem it suffices to show that $\{u, v\} = 1$ for all $u, v \in \mathbb{F}_q^*$. We know that $\mathbb{F}_q^*$ is cyclic of order $q - 1$; let t be a generator of this group. For arbitrary $u, v \in \mathbb{F}_q$ write $u = t^i$, $v = t^j$, so that $\{u, v\} = \{t, t\}^{ij}$. Therefore it suffices to show that $\{t, t\} = 1$. By skew-symmetry, $\{t, t\} = \{t, t\}^{-1}$, so $\{t, t\} = \pm 1$. We first suppose that $q = 2^r$, so that $q - 1$ is odd. In this case

$$\{t, t\}^{q-1} = \{t, t^{q-1}\} = \{t, 1\} = 1,$$

which rules out $\{t, t\} = -1$, so $\{t, t\} = 1$ as desired. Otherwise we have $q = p^r$ for an odd prime p, so $q - 1$ is even. We claim there are elements $u = t^i$ and $v = t^j$ in $\mathbb{F}_q^*$ such that $u + v = 1$ and i, j are both odd. (I.e. u and v are non-squares in $\mathbb{F}_q$.) Assuming this, we would have

$$1 = \{u, v\} = \{t, t\}^{ij},$$

which forces $\{t, t\} = 1$ as before.

To establish the claim, consider the q ordered pairs $(u, 1 - u)$ as u varies over all elements of $\mathbb{F}_q$. Each element of $\mathbb{F}_q$ appears exactly twice in this list, once in the first place and once in the second. Including 0, there are exactly $\frac{1}{2}(q + 1)$ squares in $\mathbb{F}_q$, hence $q + 1$ appearances of squares in our list. Four of them appear in the pairs $(0, 1)$ and $(1, 0)$; the remaining $q - 3$ are distributed among the other $q - 2$ pairs in some fashion, which proves the existence of at least one pair $(u, 1 - u)$ in which neither entry is a square in $\mathbb{F}_q$.

One of the simplest instances of a Steinberg symbol, called a tame symbol, arises when we impose a discrete valuation ν on a field F. Let $A \subset F$ be the valuation ring and $\mathfrak{m} \subset A$ its unique maximal ideal with residue field $k = A/\mathfrak{m}$.

Proposition: *The tame symbol* $(\cdot, \cdot)_\nu : F^* \times F^* \to k^*$ *defined by*

$$(x, y)_\nu = (-1)^{\nu(x)\nu(y)} \frac{x^{\nu(y)}}{y^{\nu(x)}} \bmod \mathfrak{m}$$

is a Steinberg symbol.

For example, if X is a Riemann surface with $w \in X$ then we can take $F = \mathbb{C}(X)$ to be the field of complex-valued meromorphic functions on X and the valuation, which we also call w, to measure the order of vanishing of a function $x \in F$ at w. Thus $w(x) = -1$ means that x has a simple pole at w. So if t is a local parameter for two functions x and y in a neighborhood of w and $x = a_n t^n + \cdots$, $y = b_m t^m + \cdots$ for $m, n \in \mathbb{Z}$ then

$$(x, y)_w = \pm \frac{a_n^m t^{mn} + \cdots}{b_m^n t^{mn} + \cdots} \bmod \mathfrak{m} = \pm \frac{a_n^m}{b_m^n} \in \mathbb{C}^*,$$

since $\mathfrak{m}$ is the ideal generated by t, that is, functions with a zero at w.

Proof: Verification of skew-symmetry and bimultiplicativity follow at once from properties of valuations. To show that $(x, y)_\nu = 1$ when $x + y = 1$ we consider several cases. Since $\nu(x + y) = 0$, so we can't have both $\nu(x) > 0$ and $\nu(y) > 0$. First suppose that $\nu(x) = 0$, $\nu(y) > 0$, meaning $y \in \mathfrak{m}$. Then we compute

$$(x, y)_\nu = (-1)^0 \frac{(1 - y)^{\nu(y)}}{y^0} \bmod \mathfrak{m} = 1.$$

Similarly, we discover that $(y, x)_\nu = 1$ when $\nu(x) > 0$, $\nu(y) = 0$, which shows $(x, y)_\nu = 1$ by skew-symmetry. The statement is clearly true for $\nu(x) = 0$, $\nu(y) = 0$, so we next consider $\nu(x) < 0$, say $\nu(x) = -n$ for $n \in \mathbb{N}$. Then $\frac{1}{x} \in \mathfrak{m}$, and properties of ν require $\nu(1 - x) = \nu(y) = -n$ as well, hence

$$(x, y)_\nu = (-1)^{n^2} \frac{x^{-n}}{(1-x)^{-n}} \bmod \mathfrak{m} = \left(1 - \frac{1}{x}\right)^n \bmod \mathfrak{m} = 1.$$

This exhausts all possible cases, so the assertion is proved.

A second example of a Steinberg symbol is given by the Hilbert symbol. Let $F = \mathbb{Q}_p$ (resp. $F = \mathbb{R}$), and for $a, b \in F^*$ define $(a, b)_p$ (resp. $(a, b)_\infty$) to equal 1 if $ax^2 + by^2 - z^2 = 0$ has a non-trivial solution in F and to equal -1 otherwise. In this case the Steinberg relation is clear, since $(x, y, z) = (1, 1, 1)$ is a non-trivial solution when $a + b = 1$. Skew-symmetry is also obvious, but proving bimultiplicativity takes some effort. When $F = \mathbb{R}$ we see that $(a, b)_\infty = -1$ if and only if $a, b < 0$, correlating to the fact that $\{-1, -1\} = -1$ in $K_2(\mathbb{R})$.

Now we will outline the work of Quillen in defining higher K-groups, for which he was awarded the Fields medal. His idea was to define $K_n(A)$ for $n \geq 2$ in terms of the fundamental groups of a certain extended classifying space of $GL(A)$. This yields a definition which is both compatible with Milnor's construction of $K_2(A)$ and possess all the general properties one would hope for, such as the existence of long exact sequences. There is also good computational evidence that Quillen's definition is a "correct" way to define higher K-groups.

A similar (and related, we shall see) situation arose with group cohomology. For a group G, satisfactory definitions for $H_0(G)$, $H_1(G)$, and $H_2(G)$ had been developed and the question naturally arose as to the best way to extend these to a whole sequence of homology groups $H_n(G)$. One satisfactory approach implements the fundamental groups $\pi_n(X)$ of algebraic topology. Recall that $\pi_n(X)$ as a set consists of homotopy classes of continuous maps $f : S^n \to X$ with a fixed basepoint. For any group G there is a classifying space BG with the property that $\pi_1(BG) = G$ and $\pi_n(G) = 0$ for $n > 1$. We then define $H_n(G)$ to equal $H_n(BG)$.

For a topological space X we know that $\pi_1(X)^{\mathrm{ab}} = H_1(X)$. We have already seen that $K_1(A) = GL(A)/E(A) = GL(A)^{\mathrm{ab}}$, so taking $G = GL(A)$ we have

$$K_1(A) = G^{\mathrm{ab}} = \pi_1(BG)^{\mathrm{ab}} = H_1(BG).$$

If we have a presentation $1 \to R \to F \to \pi_1(X) \to 1$ for $\pi_1(X)$ then Hopf showed that

$$(R \cap [F, F])/[R, F] \cong H_2(X)/\mathrm{im}(h_2),$$

where $h_n : \pi_n(X) \to H_n(X)$ is the Hurewicz map. Applying these ideas to $X = BG$ we find $\pi_1(BG) \cong G$, $\pi_2(BG) = 0$, and

$$(R \cap [F, F])/[R, F] \cong H_2(BG) = H_2(G).$$

But when G is a perfect group the left-hand side is exactly $\ker(\phi)$ in our construction of a universal central extenstion of G, so $H_2(G)$ is a Schur multiplier. In particular, using $G = E(A)$ we conclude that $K_2(A) = H_2(E(A))$.

We now discuss Quillen's "+-Construction" for defining higher K-groups. As above we consider the classifying space BG of $G = GL(A)$, except we extend it to a topological space BG^+ by adjoining certain 2-cells, then some other 3-cells to compensate. The addition of the 2-cells is designed so that

$$\pi_1(BG^+) = K_1(A) = G^{\mathrm{ab}}.$$

The addition of the extra cells do not affect the homology groups, i.e. $H_n(BG^+) = H_n(BG) = H_n(G)$. However, the fundamental groups are no longer trivial, and we declare that $K_n(A) = \pi_n(BG^+)$. This definition does in fact agree with our earlier description of $K_2(A)$.

Although the fundamental groups $\pi_n(X)$ are notoriously difficult to compute, there is computational evidence to recommend this approach. For example, Quillen has shown that $K_{2m}(\mathbb{F}_q) = 0$ while $K_{2m-1}(\mathbb{F}_q) = \mathbb{Z}/(q^m - 1)\mathbb{Z}$. This fits in with Borel's theorem, which implies that

$$-\zeta_{\mathbb{F}_q}(-m)^{-1} = \#K_{2m-1}(\mathbb{F}_q).$$

Here $\#G$ means the order of the finite group G, and $\zeta_{\mathbb{F}_q}(s) = (1 - q^{-s})^{-1}$. In the 1970's it was shown that $K_3(\mathbb{Z}) \cong \mathbb{Z}/48\mathbb{Z}$. More recently the computations $K_4(\mathbb{Z}) = 0$ and $K_5(\mathbb{Z}) \cong \mathbb{Z}$ were made. It is conjectured that $K_{22}(\mathbb{Z}) \cong \mathbb{Z}/691\mathbb{Z}$; the relevance of 691 being that it is the twelfth Bernoulli number.

The real usefulness of Quillen's +-construction can be seen, for instance, in the long exact localization sequence of a number field F:

$$\begin{aligned} 0 \to K_{2m}(\mathcal{O}_F) \to K_{2m}(F) \to \bigoplus_{\mathfrak{p}} & K_{2m-1}(\mathcal{O}_F/\mathfrak{p}) \to \\ & \to K_{2m-1}(\mathcal{O}_F) \to K_{2m-1}(F) \to 0. \end{aligned}$$

The zeros appear because of the fact that $K_{2m}(F) = 0$ when F is a finite field. Soulé showed that the final two groups are isomorphic for all $m \geq 1$, so the sequence may be shortened to

$$0 \to K_{2m}(\mathcal{O}_F) \to K_{2m}(F) \to \bigoplus_{\mathfrak{p}} K_{2m-1}(\mathcal{O}_F/\mathfrak{p}) \to 0.$$

For $m = 1$ the final map is given by tame symbols. More precisely, let ν be the valuation corresponding to a given prime $\mathfrak{p} \subset \mathcal{O}_F$. Then the homomorphism $(\cdot,\cdot)_\nu : F^* \times F^* \to K_1(\mathcal{O}_F/\mathfrak{p}) = (\mathcal{O}_F/\mathfrak{p})^*$ factors through $K_2(F)$ by its universal property. The resulting map λ_ν is given by $\{a, b\} \mapsto (a, b)_\nu$ for $a, b \in F^*$. Due to this result we can describe $K_2(\mathcal{O}_F)$ as the intersection of the subgroups $\ker(\lambda_\nu)$ in $K_2(F)$ over all valuations ν. The upshot is that $K_2(\mathcal{O}_F)$ can be computed in many cases. For example, Tate handled $F = \mathbb{Q}(\sqrt{-d})$ for several small positive values of d using this method.

Belabas and Gangl have adapted Tate's method to the computer, producing an algorithm which gives explicit generators for the group $K_2(\mathcal{O}_F)$ along with bounds on their orders. In many cases this algorithm yields an effective proof of the structure of $K_2(\mathcal{O}_F)$. In particular, they have shown that when $F = \mathbb{Q}(\sqrt{-303})$, then $K_2(\mathcal{O}_F) \cong \mathbb{Z}/22\mathbb{Z}$ is generated by

$$\{\tfrac{1}{2}(-37 - 3\sqrt{-303}), \tfrac{1}{2}(-73 + \sqrt{-303})\}^5.$$

The exponent is needed to guarantee that the resulting element of $K_2(F)$ is in the kernel of every map λ_ν. The algorithm mimics a standard approach to computing class groups, in that it finds a small set of generators, then produces relations among them, yielding a group $\tilde{K}_2(\mathcal{O}_F)$ of which $K_2(\mathcal{O}_F)$ is a quotient. If enough relations are found then certain bounds demonstrate that these two groups are the same, accomplishing the task.

Lectures 9–10: Mahler Measure

Let $P \in \mathbb{C}[x_1, x_1^{-1}, \dots, x_n, x_n^{-1}]$ be a non-zero Laurent polynomial in n variables. We define the (logarithmic) Mahler measure of P as

$$m(P) = \frac{1}{(2\pi i)^n} \int_{\mathbb{T}^n} \log |P(x_1, \dots, x_n)| \, \frac{dx_1}{x_1} \cdots \frac{dx_n}{x_n},$$

where $\mathbb{T}^n$ is the n-torus; that is, the set of points $(x_1, \dots, x_n) \in \mathbb{C}^n$ for which $|x_j| = 1$, $1 \le j \le n$. Since we may have $P = 0$ on the n-torus it is not immediately clear that this integral necessarily converges; however, we shall soon show that it does, giving a well defined function on Laurent polynomials. By making the change of variables $x_j = e^{i\theta_j}$ we may rewrite our definition in the form

$$m(P) = \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} \log |P(e^{i\theta_1}, \dots, e^{i\theta_n})| \, d\theta_1 \cdots d\theta_n.$$

Thus $m(P)$ is the average of $\log |P|$ over the n-torus, and the exponential Mahler measure $M(P) = e^{m(P)}$ is the geometric mean of $|P|$ over the n-torus.

Historically, Mahler measure played an important role in Lehmer's optimization of a technique for finding large primes first introduced by Pierce. In a paper in 1918 Pierce suggested restricting the search for large primes to numbers of a special form, as follows. Consider a polynomial $P(x) \in \mathbb{Z}[x]$ of degree d which factors over $\mathbb{C}$ as $P(x) = (x - \alpha_1) \cdots (x - \alpha_d)$. We claim that the quantity

$$\Delta_n = \prod_{j=1}^{d} (\alpha_j^n - 1)$$

is an integer. One way to see this is to let $\xi = e^{2\pi i/n}$ and write

$$\Delta_n = \prod_{j=1}^{d} \prod_{k=1}^{n} (\alpha_j - \xi^k) = \prod_{k=1}^{n} (-1)^d P(\xi^k).$$

Each factor $P(\xi^k)$ is an algebraic integer in a cyclotomic extension of $\mathbb{Q}$, and the entire expression splits neatly into a product of norms, each of which must be an integer. For example, when $n = 6$ we know $F = \mathbb{Q}(\xi) = \mathbb{Q}(\sqrt{-3})$ is a quadratic extension; let $\sigma : \sqrt{-3} \mapsto -\sqrt{-3}$ be the non-trivial automorphism. Then we have

$$\begin{aligned} \Delta_6 &= \prod_{k=1}^{6} (-1)^d P(\xi^k) \\ &= P(1) \cdot P(-1) \cdot (P(\xi)\sigma(P(\xi)) \cdot (P(\xi^2)\sigma(P(\xi^2)) \\ &= P(1)P(-1)N_{F/\mathbb{Q}}(P(\xi))N_{F/\mathbb{Q}}(P(\xi^2)) \in \mathbb{Z}. \end{aligned}$$

Another approach is to argue that Δ_n equals the resultant of $P(x)$ and $x^n - 1$, up to sign, and hence must be an integer.

The advantage of testing Δ_n for primality is that prime divisors of such integers must satisfy a number of congruence conditions. For instance, taking $P(x) = x - 2$ yields the familiar case of $\Delta_n = 2^n - 1$, the Mersene primes. Here we must have $n = q$, a prime, and $p \equiv 1 \bmod q$, $p \equiv \pm 1 \bmod 8$. Using a modification of the above argument one can show that Δ_n/Δ_m is an integer when $m|n$, since the quotient will also be a product of norms. For example,

$$\Delta_6/\Delta_3 = P(-1)N_{F/\mathbb{Q}}(P(\xi)) \in \mathbb{Z}.$$

This observation limits our search for primes to the integers Δ_q/Δ_1, q prime. Using these ideas applied to the polynomial $P(x) = x^3+x+1$ Pierce was able to demonstrate that $\Delta_{61}/\Delta_1 = 4459734401$ was prime with only 169 trial divisions, compared to the 6655 ordinarily needed for a number of that magnitude.

In order to maximize the gain in the reduction of trial divisions needed one wants to find $P(x)$ for which the integers Δ_n grow as slowly as possible. Lehmer observed that as long as $P(x)$ does not vanish at any roots of unity then

$$\begin{aligned}\frac{1}{n}\log|\Delta_n| &= \frac{1}{n}\sum_{k=1}^{n}\log|P(\xi^k)| \\ &\sim \frac{1}{2\pi}\int_0^{2\pi}\log|P(e^{i\theta})|\,d\theta \\ &= m(P).\end{aligned}$$

Thus $\Delta_n \sim M(P)^n$, which shifts the focus to finding polynomials with small Mahler measure. An indispensable tool for analyzing the integrals involved is due to Jensen.

Theorem: (Jensen's formula) *Let $f(x)$ be a holomorphic function on an open set containing the unit disc, with $f(0) \neq 0$. If $\alpha_1, \ldots, \alpha_t$ are the roots of $f(x)$ inside or on the unit disc, listed according to their multiplicity, then*

$$\frac{1}{2\pi}\int_0^{2\pi}\log|f(e^{i\theta})|\,d\theta = \log|f(0)| - \sum_{j=1}^{t}\log|\alpha_j|.$$

If $f(x) = P(x) = a(x-\alpha_1)\cdots(x-\alpha_d)$ is a polynomial with $\alpha_j \neq 0$ then $\log|f(0)| = \log|a| + \log|\alpha_1| + \cdots + \log|\alpha_d|$, so Jensen's formula implies that

$$m(P) = \log|a| + \sum_{j=1}^{d}\log^+|\alpha_j|,$$

where we define $\log^+|x|$ to equal $\log|x|$ for $|x| \geq 1$, while $\log^+|x| = 0$ otherwise. (Note that this formula is valid even if $P(0) = 0$.) Equivalently, this result may be written as

$$M(P) = |a|\prod_{j=1}^{d}\max\{1, |\alpha_j|\}.$$

We remark that $m(P_1P_2) = m(P_1)+m(P_2)$ and $M(P_1P_2) = M(P_1)M(P_2)$. It is now clear that $M(P) \geq 1$ for $P(x) \in \mathbb{Z}[x]$. For purposes of searching for primes the case $M(P) = 1$ is not interesting; as we shall see the values of Δ_n repeat periodically and hence do not grow. Lehmer's goal was to find $P(x)$ for which $M(P)$ was greater than 1 but as close as possible, so that the Δ_n grew slowly. Evidently one should look for monic polynomials with all but one root inside or on the unit circle, with the remaining root is as near the unit circle as possible. The table below lists the optimal such polynomials by degree for $1 \leq d \leq 5$.

degree	optimal $P(x)$	$M(P) \approx$
1	$x-2$	2.000
2	x^2-x-1	1.618
3	x^3-x-1	1.324
4	x^4-x^3+1	1.380
5	$x^5-x^4+x^3-x+1$	1.349

Lehmer's best result, which remains unbroken even in the computer age, was $P(x) = x^{10}+x^9-x^7-x^6-x^5-x^4-x^3+x+1$, for which $M(P) \approx 1.176$. One marvels at how he unearthed this polynomial, which has eight complex roots on the unit circle, and two real roots located just above and below 1.

The interested reader may wonder how these polynomials fare at producing large primes. The degree five polynomial listed has $P(1) = 1$, so $\Delta_1 = -1$ and our candidates for primes are $|\Delta_q|$ for q prime. For $2 \leq q \leq 71$ (the first twenty primes), $|\Delta_q|$ is a prime in twelve cases, including $|\Delta_{71}| = 1762072889$. However, beginning with $q = 127$, the thirty-first prime, $|\Delta_q|$ is divisible by steadily (but not quite monotonically) increasing powers of two. By the time $q = 541$, the hundredth prime, $|\Delta_q|$ is divisible by 2^{182}.

The degree ten polynomial found by Lehmer behaves even more curiously. To begin with, $\Delta_1 = -1$ as before. In fact, $|\Delta_n| = 1$ twenty-two times, including $|\Delta_q| = 1$ for ten of the first twelve prime values of q. More suprising, perhaps, is the fact that $|\Delta_n|$ is a perfect square for all $1 \leq n \leq 162$. Furthermore, for primes q in this range (the first thirty-seven primes), $|\Delta_q|$ is either equal to 1 or the square of a prime. So in a sense the $|\Delta_q|$ are extremely reliable in producing primes. However, $|\Delta_{163}| = 2 \cdot 5^2 \cdot 1313590206173$ interrupts this pattern. (The appearance of 163 here seems remarkable.) For primes $q \geq 163$, $|\Delta_q|$ appears to be either a perfect square or a product of several smaller factors and a single large prime to the first power.

We asserted above that for $P(x) \in \mathbb{Z}[x]$ the values of Δ_n were periodic when $M(P) = 1$. This behavior is explained by a result of Kronecher, for which we will need the following bound.

Lemma: *If $P(x) = a_d x^d + \cdots + a_1 x + a_0$ then $|a_m| \leq \binom{d}{m} M(P)$.*

Proof: We know that a_m/a_d equals a symmetric sum of the roots α_j of $P(x)$, up to sign. This sum involves $\binom{d}{m}$ terms, each of which is a product of $d-m$ of the roots. Clearly the absolute value of each term is less than $\prod \max\{1, |\alpha_j|\}$, the product over all the roots. Therefore by the triangle inequality

$$|a_m| = |a_d| \cdot \left|\frac{a_m}{a_d}\right| \leq |a_d| \binom{d}{m} \prod_{j=1}^{d} \max\{1, |\alpha_j|\} = \binom{d}{m} M(P),$$

according to the product formula for $M(P)$. Equality can be obtained by the polynomial $P(x) = (x+1)^d$, for instance.

Proposition: (Kronecher) *A polynomial $P(x)$ with integral coefficients satisfies $M(P) = 1$ if and only if $P(x)$ is cyclotomic; that is, $P(x)$ has leading coefficient ± 1 and roots which are either zero or roots of unity.*

Proof: Let $P(x) = a \prod (x - \alpha_j)$ have degree d. If $P(x)$ is cyclotomic then $|a| = 1$ and $|\alpha_j| = 0$ or 1 for all j, so clearly $M(P) = 1$. Conversely, if $M(P) = 1$ then $a = \pm 1$, and after factoring out the largest power of x dividing $P(x)$ it is not hard to see that $|\alpha_j| = 1$ for all roots $\alpha_j \neq 0$. However, we do not know a priori that the α_j are roots of unity.

To establish this fact, let $P_k(x) = a^k \prod (x - \alpha_j^k) = \sum a_{km} x^m$. By definition $M(P_k) = M(P)^k = 1$, so the lemma gives $|a_{km}| \leq \binom{d}{m}$ for all k. (In fact, $P_k(x) \in \mathbb{Z}[x]$, but our lemma applies to any polynomial with complex coefficients.) Hence there are a finite number of choices for each coefficient of $P_k(x)$, so $P_k(x) = P_l(x)$ for some positive integers k and l. In particular, these polynomials have the same roots, so $\alpha_j^k = \alpha_{\sigma(j)}^l$ for some permutation $\sigma \in \Sigma_d$. From here it is easy to argue that each root α_j must satisfy an equation of the form $\alpha_j^M = \alpha_j^N$ for positive integers $M < N$, so is either 0 or a root of unity.

In general let $P(x_1, \ldots, x_n) = \sum a_{m_1 \cdots m_n} x_1^{m_1} \cdots x_n^{m_n}$ be a polynomial in n variables such that P is of degree d_j in the variable x_j. Then the lemma above can be generalized via an induction to show that

$$|a_{m_1\cdots m_n}| \le \binom{d_1}{m_1} \cdots \binom{d_n}{m_n} M(P).$$

Assuming that $P \neq 0$, this gives a positive lower bound for $M(P)$, which implies that $m(P) > -\infty$. Thus to conclude that $m(P)$ exists for polynomials, we need only prove that $m(P)$ is bounded above. Using the fact that $\log|x| < |x|$ and the triangle inequality we find that

$$\begin{aligned} m(P) &= \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} \log|P(e^{i\theta_1}, \ldots, e^{i\theta_n})| \, d\theta_1 \cdots d\theta_n \\ &< \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} |P(e^{i\theta_1}, \ldots, e^{i\theta_n})| \, d\theta_1 \cdots d\theta_n \\ &\le \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} \left(\sum |a_{m_1\cdots m_n}|\right) \, d\theta_1 \cdots d\theta_n \\ &= \sum |a_{m_1\cdots m_n}|. \end{aligned}$$

Hence $m(P) < \infty$, so Mahler measure exists for all non-zero polynomials P. We shall soon see that for any Laurent polynomial Q we have $m(Q) = m(P)$ for some polynomial P, ensuring that $m(Q)$ exists in general.

Mahler was interested in relating the "size" of polynomials to the "size" of their product, and he used the expression $\|P\| = \sum |a_{m_1\cdots m_n}|$ to define "size." We have seen that $m(P) < \|P\|$; Mahler showed that in fact $M(P) \le \|P\|$. (The reader may wish to attempt a proof in the one variable case using the product formula for $M(P)$. The Hardy-Littlewood-Polya inequality may be implemented to obtain a general proof.) Next, using the bound on $|a_{m_1\cdots m_n}|$ stated above and summing over all coefficients yields

$$\|P\| \le \sum \binom{d_1}{m_1} \cdots \binom{d_n}{m_n} M(P) = 2^{d_1+\cdots+d_n} M(P),$$

with equality when $P(x_1, \ldots, x_n) = (x_1+1)^{d_1} \cdots (x_n+1)^{d_n}$.

We can now demonstrate Mahler's main result. Let $P_1, \ldots, P_N$ each be polynomials in n variables, and denote the degree of x_j in P_k by d_{jk}. Also set $D_j = \sum d_{jk}$, the degree of x_j in the product of the P_k. The above inequalities imply that

$$\begin{aligned} \prod_{k=1}^{N} \|P_k\| &\le \prod_{k=1}^{N} 2^{d_{1k}+\cdots+d_{nk}} M(P_k) \\ &= 2^{D_1+\cdots+D_n} M(\textstyle\prod P_k) \\ &\le 2^{D_1+\cdots+D_n} \|\textstyle\prod P_k\|, \end{aligned}$$

since $M(P)$ is multiplicative. The final result is a statement only about $\|P\|$, although its proof uses Mahler measure in a crucial way.

To avoid cumbersome notation, we implement the multi-index notation in our discussion of Laurent polynomials $P(x_1, \ldots, x_n) \in \mathbb{C}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]$. For $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{Z}^n$ we write $x^{\mathbf{m}}$ to mean $x_1^{m_1} \cdots x_n^{m_n}$. Thus we can write

$$P(x) = \sum_{\mathbf{m} \in \mathbb{Z}^n} a_{\mathbf{m}} x^{\mathbf{m}}, \qquad \text{almost all } a_{\mathbf{m}} = 0.$$

To each Laurent Polynomial P we can associate its Newton polytope Δ_P, the convex hull of the set of points $\{\mathbf{m} \in \mathbb{Z}^n | a_{\mathbf{m}} \neq 0\}$. When $n = 2$ we will often present P by simply listing each non-zero coefficient $a_{\mathbf{m}}$ at the point $\mathbf{m}$ in $\mathbb{Z}^2$. Thus $P(x,y) = 3y - xy + 2x^2y - 2x^{-1} + 4 + 5xy^{-1}$ would be written

$$\begin{array}{cccccc}
 & & | & & & \\
 & & 3 & -1 & 2 & \\
— & -2 & 4 & & & — \\
 & & & 5 & & \\
 & & | & & &
\end{array}$$

with the axes indicated by the line segments. In this case Δ_P is a quadrilateral with two interior points and one point along the boundary. Notice that for a vector $\mathbf{m}$ and polynomial $P(x)$, multiplication by $x^{\mathbf{m}}$ corresponds to a translation by $\mathbf{m}$ of the Newton polytope. In other words, $\Delta_{x^{\mathbf{m}}P} = \Delta_P + \mathbf{m}$. Mahler measure is clearly unaffected by such multiplication (i.e. $m(x^{\mathbf{m}}P) = m(P)$), since $\log|x^{\mathbf{m}}| = 0$ on $\mathbb{T}^n$. Therefore we will often neglect to specify the origin when presenting a polynomial via its Newton polytope.

The action of $GL_n(\mathbb{Z})$ on $\mathbb{Z}^n$ also preserves the combinatorial aspects of Δ_P and the value of $m(P)$. To be precise, for $\mathbf{A} \in GL_n(\mathbb{Z})$ we let $\mathbf{A}$ act on Laurent polynomials of n variables via $\mathbf{A} : x^{\mathbf{m}} \mapsto x^{\mathbf{Am}}$. This corresponds to the usual action of $\mathbf{A}$ on $\mathbb{Z}^n$ when we consider Newton polytopes. Thus $\Delta_{\mathbf{A}(P)} = \mathbf{A}(\Delta_P)$, and $\mathbf{A}(\Delta_P)$ has the same number of interior points, boundary points, and vertices as Δ_P by properties of general linear transformations.

We claim that Mahler measure is preserved by $\mathbf{A}$, so that $m(\mathbf{A}(P)) = m(P)$. To see why, let $\mathbf{B} = \mathbf{A}^{-1}$ and declare the change of variables $\phi : \mathbb{T}^n \to \mathbb{T}^n$ defined by

$$\phi(x_1, \dots, x_n) = (x_1^{b_{11}} \cdots x_n^{b_{n1}}, \dots, x_1^{b_{1n}} \cdots x_n^{b_{nn}}) = (y_1, \dots, y_n).$$

This change of variables is set up precisely so that $\mathbf{A}(P(\phi(x))) = P(x)$. Computing the matrix of partials reveals that

$$\left|\frac{\partial y}{\partial x}\right| = x_1^{b_{11}+\cdots+b_{1n}-1} \cdots x_n^{b_{n1}+\cdots+b_{nn}-1}|\mathbf{B}|.$$

Note that $\mathbf{B} \in GL_n(\mathbb{Z})$ implies that $|\mathbf{B}| = 1$. Since ϕ is a diffeomorphism of $\mathbb{T}^n$ we conclude by the change of variable theorem that

$$\begin{aligned}
&(2\pi i)^n m(\mathbf{A}(P)) = \\
&\quad = \int_{\mathbb{T}^n} \log|\mathbf{A}(P(y))| \frac{dy_1}{y_1} \cdots \frac{dy_n}{y_n} \\
&\quad = \int_{\mathbb{T}^n} \log|\mathbf{A}(P(\phi(x)))| \frac{x_1^{b_{11}+\cdots+b_{1n}-1} \cdots x_n^{b_{n1}+\cdots+b_{nn}-1}|\mathbf{B}|}{(x_1^{b_{11}} \cdots x_n^{b_{n1}}) \cdots (x_1^{b_{1n}} \cdots x_n^{b_{nn}})} \, dx_1 \cdots dx_n \\
&\quad = \int_{\mathbb{T}^n} \log|P(x)| \frac{dx_1}{x_1} \cdots \frac{dx_n}{x_n} \\
&\quad = (2\pi i)^n m(P).
\end{aligned}$$

The result $m(\mathbf{A}(P)) = m(P)$ is actually valid for any $\mathbf{A} \in M_n(\mathbb{Z})$ with $|\mathbf{A}| \neq 0$. For example, in the one-variable case consider $m(P(x^3))$. Making the change of variables $\phi(x) = x^3 = y$ we find that

$$3\int_{\mathbb{T}^1} \log|P(y)| \frac{dy}{y} = \int_{\mathbb{T}^1} \log|P(x^3)| \cdot 3\frac{dx}{x},$$

which implies that $m(P(x)) = m(P(x^3))$. The initial factor of three stems from the fact that $\phi : \mathbb{T}^1 \to \mathbb{T}^1$ is no longer a diffeomorphism, but a three-fold cover. The general proof requires more intricate bookkeeping, so we omit it.

Let Δ be the Newton polytope of a Laurent polynomial P and let τ be a facet (face of codimension 1) of Δ. Choose a linear embedding $\psi : \mathbb{Z}^{n-1} \hookrightarrow \mathbb{Z}^n$ whose image is the set of lattice points lying on the hyperplane containing τ. Then we can construct a Laurent polynomial P_τ in $n-1$ variables via

$$P_\tau = \sum_{\mathbf{m}' \in \mathbb{Z}^{n-1}} a_{\psi(\mathbf{m}')} x^{\mathbf{m}'}.$$

This formal construction simply gives a polynomial P_τ whose array of coefficients essentially matches P on τ. Of course, P_τ depends on ψ, but any two such embeddings will differ only by a translation and a general linear transformation, so $m(P_\tau)$ is well-defined. In our example above let τ be the top facet of Δ. Thus for a suitable embedding ψ we would have $P_\tau = 3t^2 - t + 2$. Continuing with the same orientation around Δ the other polynomials associated with the facets would be $-2t+3$, $5t-2$, and $2t+5$.

The relevance of the P_τ is underscored by the following result.

Theorem: (Smyth) *Let P have Newton polytope Δ. If τ is any facet of Δ, then $m(P) \geq m(P_\tau)$.*

Before giving the proof we explore a few of the consequences of this theorem. Applying it repeatedly, we eventually reach the zero-dimensional faces η of Δ, which are the vertices of Δ. The polynomials P_η are constants equal to the coefficients of P at these vertices. Therefore $m(P) \geq m(P_\eta) = \log|a_{\mathbf{m}}|$ for any $\mathbf{m}$ which is a vertex of Δ. In the one-dimensional case we conclude that $m(P(x)) \geq \max\{\log|a_d|, \log|a_0|\}$, where a_d and a_0 are the leading and final coefficients. These must both equal ±1 in order to have $m(P) < \log 2$. More generally, to find polynomials $P \in \mathbb{Z}[x,y]$ with small Mahler measure one should ensure that $P_\tau = 0$ for all faces τ, which is equivalent to asking that all P_τ are cyclotomic, by Kronecher's theorem.

Let us return now to the proof that $m(P) \geq m(P_\tau)$.

Proof: We first simplify matters by assuming that $P(x) \in \mathbb{C}[x_1, \ldots, x_n]$ is a polynomial. Furthermore, suppose that the facet τ is contained in the hyperplane $m_n = d$, where $m_1, \ldots, m_n$ are coordinates for $\mathbb{Z}^n$. The upshot of these hypotheses is that if we think of $P(x)$ as a polynomial in x_n,

$$P(x) = P_d(x_1, \ldots, x_{n-1})x_n^d + \cdots + P_1(x_1, \ldots, x_{n-1})x_n + P_0(x_1, \ldots, x_{n-1}),$$

then we can take $P_d(x_1, \ldots, x_{n-1})$ as P_τ. Factoring out P_d we are left with a monic polynomial in x_n whose roots $\alpha_j(x_1, \ldots, x_{n-1})$ depend continuously on the first $n-1$ variables, so that

$$\log|P(x)| = \log|P_d(x_1, \ldots, x_{n-1})| + \sum_{j=1}^{d} \log|(x_n - \alpha_j(x_1, \ldots, x_{n-1}))|.$$

Integrating with respect to x_n in the expression for $m(P)$ and using Jensen's formula yields

$$m(P) = m(P_d) + \frac{1}{(2\pi i)^{n-1}} \int_{\mathbb{T}^{n-1}} \sum_{j=1}^{d} \log^+ |\alpha_j(x_1, \ldots, x_{n-1})| \, \frac{dx_1}{x_1} \cdots \frac{dx_{n-1}}{x_{n-1}}.$$

But $\log^+$ is non-negative, so we conclude that $m(P) \geq m(P_d) = m(P_\tau)$.

To finish the proof, we reduce the general case to the one considered above.

Lemma: *Let Δ be a convex polytope with vertices in $\mathbb{Z}^n$ and let τ be any facet. Then there exists an affine transformation (i.e. a translation followed by a linear* *map) such that the image of τ lies in the hyperplane $m_n = 0$ and the image of Δ lies in the region $m_n \le 0$.*

Proof: By translating Δ if necessary, we may assume that the hyperplane containing τ passes through the origin. This hyperplane has equation $\mathbf{m} \cdot \mathbf{u} = 0$ for some primitive vector $\mathbf{u} \in \mathbb{Z}^n$. (The coordinates of a primitive vector have no common prime divisor.) According to a theorem of Hermite, there exists a matrix $\mathbf{A} \in GL_n(\mathbb{Z})$ whose bottom row is the vector $\mathbf{u}$. If $\mathbf{m}_0$ is any point on the hyperplane, then $\mathbf{A}\mathbf{m}_0$ satisfies $m_n = 0$ by construction. Since $\mathbf{A}(\Delta)$ is convex, it will lie entirely within the region $m_n \le 0$ or $m_n \ge 0$. In the latter case, replace $\mathbf{u}$ by $-\mathbf{u}$ to obtain the desired result.

Therefore given any Laurent polynomial Q with Newton polytope Δ_Q, facet τ, and associated polynomial Q_τ, we first apply the lemma to Δ_Q, then translate the resulting polytope so that it lies entirely within the region $m_1 \ge 0, \dots,$ $m_n \ge 0$. The net effect of this sequence of translations and linear maps is to obtain a new polynomial P with $m(Q) = m(P)$ as well as $m(Q_\tau) = m(P_\tau)$. We can now use the above argument to deduce that $m(Q) \ge m(Q_\tau)$ for any Laurent polynomial Q.

The ideas just presented may be used to present an alternate proof of the convergence of $m(Q)$. We have seen that $m(Q) \ge m(Q_\eta) = \log |a_{\mathbf{m}}|$ for the non-zero coefficient $a_{\mathbf{m}}$ at any vertex η of Δ_Q. Thus $m(Q)$ is bounded below. On the other hand, if we choose $\mathbf{A} \in GL_n(\mathbb{Z})$ so that no face of $\mathbf{A}(\Delta_Q)$ (of any dimension) is parallel to the plane $m_n = 0$ then the polynomial $P = \mathbf{A}(Q)$ will be monic with respect to x_n, i.e. $P_d(x_1, \dots, x_{n-1}) = 1$. Therefore

$$m(Q) = m(P) = \frac{1}{(2\pi i)^{n-1}} \int_{\mathbb{T}^{n-1}} \sum_{j=1}^{d} \log^+ |\alpha_j(x_1, \dots, x_{n-1})| \frac{dx_1}{x_1} \cdots \frac{dx_{n-1}}{x_{n-1}},$$

where as before $\alpha_j(x_1, \dots, x_{n-1})$ are the roots of $P(x)$ viewed as a polynomial in x_n. Since the leading coefficient never vanishes and $\mathbb{T}^{n-1}$ is compact, the functions $\alpha_j(x_1, \dots, x_{n-1})$ are bounded, hence $m(Q)$ is bounded above as well. (A concrete example may clarify this reasoning. Consider $P(x, y) = (x-1)y^2 - 3xy + (x+1)$. The roots y as a function of x are

$$\alpha_1(x) = \frac{3x + \sqrt{5x^2 + 4}}{2(x-1)}, \quad \alpha_2(x) = \frac{3x - \sqrt{5x^2 + 4}}{2(x-1)},$$

which are not bounded near $x = 1$. If instead we had $P(x, y) = y^2 - 3xy + (x+1)$ then $\alpha_j(x) = \frac{1}{2}(3x \pm \sqrt{9x^2 - 4x - 4})$, which are bounded for $|x| = 1$.) Finally, we mention in passing a third approach, which involves defining $\mu_P(t)$ to be the normalized Haar measure on $\mathbb{T}^n$ of the set $\{x \in \mathbb{T}^n | \; |P(x)| \le t\}$. Then we have $m(P) = \int_0^\infty \log t \; d\mu_P(t)$, so the behavior of $\mu_P(t)$ for t small determines the convergence. It has been shown that $\mu_P(t) \sim Ct^\delta$ for positive constants C, δ depending on P, which effectively counteracts the blowing up of $\log t$ as $t \to 0$.

The subject of Mahler measure achieved new relevance during the 1980's with the work of Smyth in which he proved identities such as $m(x + y + z) = L'(\chi_{-3}, -1)$ and $m(x + y + z + w) = \frac{7}{2\pi^2}\zeta(3)$. We will give an elementary proof of the former to illustrate the techniques involved and provide a contrast to the more powerful methods which will be developed in subsequent lectures.

Proposition: (Smyth) *Let $\chi_{-3}(n) = \left(\frac{-3}{n}\right)$ be the quadratic Dirichlet character with conductor 3 given by $\chi_{-3}(n) = 0,\ 1,\ -1$ when $n \equiv 0,\ 1,\ 2 \mod 3$, respectively. Then $m(x + y + z) = L'(\chi_{-3}, -1)$.*

Proof: We first make the simplifying observation that for a given z with $|z| = 1$, $|x + y + z| = |\frac{x}{z} + \frac{y}{z} + 1|$ and $\frac{x}{z}$ traces out $\mathbb{T}^1$ as x does. Thus

$$\begin{aligned} m(x+y+z) &= \frac{1}{(2\pi i)^3} \int_{\mathbb{T}^3} \log \left| \frac{x}{z} + \frac{y}{z} + 1 \right| \frac{dx}{x} \frac{dy}{y} \frac{dz}{z} \\ &= \frac{1}{(2\pi i)^2} \int_{\mathbb{T}^2} \log |x + y + 1| \frac{dx}{x} \frac{dy}{y}, \end{aligned}$$

which is just $m(x + y + 1)$. By Jensen's formula the latter expression equals

$$\frac{1}{2\pi i} \int_{\mathbb{T}^1} \log^+ | - x - 1| \frac{dx}{x}.$$

Letting $x = e^{i\theta}$ leads us to evaluate

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log^+ |e^{i\theta} + 1| \, d\theta = \frac{1}{2\pi} \int_{-2\pi/3}^{2\pi/3} \log |e^{i\theta} + 1| \, d\theta,$$

since for θ outside this range, $|e^{i\theta}+1| < 1$ so $\log^+$ vanishes there. (The geometry of the situation comes into play here.) Using symmetry, the fact that $|e^{i\theta}+1| = 2\cos(\theta/2)$, and the substitution $\theta/2 \mapsto \theta$ we obtain

$$m(x + y + 1) = \frac{2}{3} \log 2 + \frac{2}{\pi} \int_0^{\pi/3} \log(\cos\theta) \, d\theta.$$

By taking derivatives one can verify that

$$\int \log(\cos\theta) \, d\theta = \theta \log(\cos\theta) - \theta \log(1 + e^{2i\theta}) + \frac{i}{2} \mathrm{Li}_2(-e^{2i\theta}) + \frac{i\theta^2}{2}.$$

Since the integral will be real we may ignore the imaginary parts, yielding

$$\begin{aligned} m(x+y+1) &= \Re\left(\tfrac{i}{\pi}\mathrm{Li}_2(\xi^{-1})\right) \\ &= \frac{\sqrt{3}}{4\pi} \cdot 2 \left(\frac{1}{1^2} + \frac{1}{2^2} - \frac{1}{4^2} - \frac{1}{5^2} + \cdots \right) \\ &= \frac{\sqrt{3}}{4\pi} \cdot 3 \left(\frac{1}{1^2} - \frac{1}{2^2} + \frac{1}{4^2} - \frac{1}{5^2} + \cdots \right) \\ &= \frac{3\sqrt{3}}{4\pi} L(\chi_{-3}, 2). \end{aligned}$$

Here $\xi = e^{2\pi i/6}$ and the coefficients of the series repeat mod 6.

Finally, we use the functional equation to relate this value to $L'(\chi_{-3}, -1)$. Recall that we defined

$$L^\star(\chi, s) = |\Delta_F|^{s/2} \Gamma_-(s) L(\chi, s)$$

for odd quadratic Dirichlet characters χ, which satisfies the functional equation $L^\star(\chi, s) = L^\star(\overline{\chi}, 1 - s)$. In our case χ_{-3} is the Dirichlet character associated with $F = \mathbb{Q}(\sqrt{-3})$, so $\chi_{-3}(n) = \left(\frac{-3}{n}\right)$ is odd, real-valued, and the discriminant of F is $\Delta_F = -3$. Using $\Gamma_-(s) = \pi^{-(s+1)/2}\Gamma(\frac{s+1}{2})$ and solving the functional equation for $L(\chi_{-3}, 1 - s)$ we obtain

$$\begin{aligned} L(\chi_{-3}, 1-s) &= 3^{s-\frac{1}{2}} \pi^{\frac{1}{2}-s} \frac{\Gamma\left(\frac{s+1}{2}\right)}{\Gamma\left(1 - \frac{s}{2}\right)} L(\chi_{-3}, s) \\ &= 2^{1-s} 3^{s-\frac{1}{2}} \pi^{-s} \sin(\tfrac{\pi s}{2}) \Gamma(s) L(\chi_{-3}, s). \end{aligned}$$

Here we employed the duplication formula $\Gamma(\frac{s}{2})\Gamma(\frac{s+1}{2}) = 2^{1-s}\pi^{-1/2}\Gamma(s)$ and the identity $\sin(\pi z) = \pi/(\Gamma(z)\Gamma(1-z))$. Substituting $s = 2$, we find that $L(\chi_{-3}, -1) = 0$, but taking the derivative of both sides with respect to s and then evaluating at $s = 2$ yields

$$-L'(\chi_{-3}, -1) = 2^{-1}3^{\frac{3}{2}}\pi^{-2}\,\frac{d}{ds}\left(\sin\left(\frac{\pi s}{2}\right)\right)\bigg|_{s=2}\Gamma(2)L(\chi_{-3}, 2),$$

or $L'(\chi_{-3}, -1) = \frac{3\sqrt{3}}{4\pi}L(\chi_{-3}, 2)$, which brings us at long last to the conclusion that $m(x + y + 1) = L'(\chi_{-3}, -1)$.

In general the exact value of $m(P)$ for $P = x_1 + \cdots + x_n$ is unknown. Determining numerical values to high precision is difficult because of the vanishing of P on the n-torus. The measure $\mu_P(t)$ is closely related to Pearson's random walk, a phenomenon which appears in a wide variety of scientific papers. The question Pearson asks is, beginning at the origin of the complex plane and taking n steps of magnitude 1 in random directions, what is the probability of ending up within a distance t of the origin? Ironically, it is easier to compute $m(P)$ for larger values of n (say $n = 100$ as opposed to $n = 10$), due to an asymptotic estimate of the form

$$m(x_1 + \cdots + x_n) \sim \frac{\log n}{2} - \gamma - \sum_{j=1}^{\infty}\frac{c_j}{n^j}.$$

The series does not actually converge, but the partial sums seem to approach a limit before blowing up; this apparent limit is a good approximation of the Mahler measure.

We conclude this section with a brief discussion of some of the analysis related to Mahler measure. To begin, $m(P)$ is continuous as a function of its coefficients. Secondly, Boyd has shown that for Laurent polynomials P in two variables,

$$\lim_{n\to\infty} m(P(x, x^n)) = m(P).$$

In particular, if $m(P(x, y))$ is small then we obtain a sequence of polynomials $P(x, x^n)$ in one variable with small Mahler measure. (This may have been the original motivation for the computation of $m(x + y + 1)$.) To highlight the amount of care needed to prove such a statement we point out that for $\phi : \mathbb{T}^2 \to \mathbb{C}$ continuous it is true that

$$\lim_{n\to\infty}\frac{1}{2\pi i}\int_{\mathbb{T}^1}\phi(x, x^n)\,\frac{dx}{x} = \frac{1}{(2\pi i)^2}\int_{\mathbb{T}^2}\phi(x, y)\,\frac{dx}{x}\frac{dy}{y}.$$

Intuitively, the map $x \mapsto (x, x^n)$ from $\mathbb{T}^1$ to $\mathbb{T}^2$ fills out the 2-torus more and more completely as $n \to \infty$. However, $\log|P|$ is usually not continuous. The key step in Boyd's argument is to show that

$$\int_{|P|<\epsilon}\log|P(x, x^n)|\,\frac{dx}{x} \longrightarrow 0 \qquad \text{as } \epsilon \to 0$$

using a finer estimate of $\mu_P(t)$ for t small than was quoted before.

To illustrate the dependence of his result on the fact that P is a Laurent polynomial, Boyd provides the following (contrived) counterexample. Define $G : \mathbb{T}^1 \to \mathbb{T}^1$ as follows. Let $x = e^{i\theta} \in \mathbb{T}^1$, and set

$$G(x) = x^{2^k}, \quad 1 - \frac{2}{2^k} \leq \theta < 1 - \frac{1}{2^k},\ k = 1, 2, \ldots$$

By construction, $G(x)$ wraps around $\mathbb{T}^1$ once for each subinterval. Then define $F : \mathbb{T}^2 \to \mathbb{C}$ by $F(x, y) = G(x) - y$. (Note that F is not a polynomial, and is discontinuous at $(1, y)$.) We claim that

$$\lim_{n\to\infty} \frac{1}{2\pi i} \int_{\mathbb{T}^1} \log |F(x, x^n)| \, \frac{dx}{x} \neq \frac{1}{(2\pi i)^2} \int_{\mathbb{T}^2} \log |F(x, y)| \, \frac{dx}{x} \frac{dy}{y}.$$

Taking $n = 2^k$ we see that $F(x, x^n)$ is identically equal to zero on the subinterval $1-\frac{2}{n} \leq \theta < 1-\frac{1}{n}$, so the integral diverges to $-\infty$ there. (It is left as an exercise to verify that the integral equals zero on the other intervals.) Hence the limit on the left-hand side does not exist. On the other hand, the right-hand side equals

$$\frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \log |G(e^{i\theta}) - e^{i\phi}| \, d\theta d\phi = \frac{1}{2\pi} \int_0^{2\pi} \log^+ |G(e^{i\theta})| \, d\theta = 0,$$

which is the counterexample.