University of Waterloo

Lecture Notes

Diophantine Approximations

Prof. Cameron L. Stewart

typed by
Andrej Vuković

January 7, 2020

Contents

Abstract

This is a series of lecture notes for a class on Diophantine approximations and the work of Alan Baker taught by Cam Stewart.

1 Sept. 4, 2019.

Consider the following problems.

(i) Find all integer solutions of $x^3 - 2y^3 = 6$.

(ii) Given $k \in \mathbb{Z}^+$, find all integer solutions of $y^2 = x^3 + k$.

We did not have a general method of tackling these sorts of Diophantine equations until Alan Baker's work in the 1960s. (Baker was a Fields Medalist and Cam's PhD advisor.) This course will focus on that work.

Cam Stewart's office is MC 5016. There is no textbook for this course. The course will consist of three lectures a week, on Monday, Wednesday, and Friday. We will be given research topics related to the material covered in class, and everyone will choose one topic. We will be expected to master the material in the paper, write an essay of about 10 pages, and present to the class a seminar on that paper. These will be scheduled for the end of the term. There will also be an oral exam at the end of the term on the course material.

Definition 1.1. A complex number is *algebraic* if it is the root of a non-zero polynomial with integer coefficients. A complex number that is not algebraic is said to be *transcendental.* The *degree* of an algebraic number is the degree of its minimal (integer) polynomial.

Definition 1.2. In number theory, we say a function $f(x)$ is *effectively computable* in terms of x if given x, there is some Turing machine that computes $f(x)$ in a finite number of steps. (We will not be concerned with Turing machines in this course; you should just think of this terminology as meaning that there is some fairly direct formula for computing the function.)

Baker's work on estimates for linear forms in the logarithms of algebraic numbers, which has implications for the solutions of Diophantine equations, has its origins in the theory of transcendental numbers. In 1844, Liouville proved the following.

Theorem 1.3. *Let α be an algebraic number of degree d with $d > 1$. Then there exists a positive real $C(\alpha)$, which is effectively computable in terms of α, such that*

$$|\alpha - \frac{p}{q}| > \frac{C(\alpha)}{q^d}$$

for any rational number $\frac{p}{q}$ with $p, q \in \mathbb{Z}$ and $q > 0$.

Proof. Notice that the result holds if α is not real since in that case, for any $\frac{p}{q}$ of the prescribed form, we have

$$|\alpha - \frac{p}{q}| \geq |\mathrm{Im}(\alpha)| \geq \frac{|\mathrm{Im}(\alpha)|}{q^d} > \frac{\frac{1}{2}|\mathrm{Im}(\alpha)|}{q^d}.$$

Taking $C(\alpha) := \frac{|\mathrm{Im}(\alpha)|}{2}$, we thus obtain the result.

Now suppose α is real. Let $f \in \mathbb{Z}[x]$ be the minimal polynomial of α, so f is not the zero polynomial, has minimal degree among integer polynomials vanishing at α, has content 1, and has positive leading coefficient. (Recall that the *content* of a polynomial is the gcd of its coefficients.)

Since f is the minimal polynomial of α, it is irreducible over $\mathbb{Q}$, so $f(\frac{p}{q}) \neq 0$. Since f has degree d and has integer coefficients, $f(\frac{p}{q})$ is a rational with denominator at most q^d when written in lowest terms. Therefore,

$$\frac{1}{q^d} \leq |f(\frac{p}{q})|.$$

Now $f(\alpha) = 0$, so by the mean value theorem, when α is real,

$$\frac{1}{q^d} \leq |f(\alpha) - f(\frac{p}{q})| = |\alpha - \frac{p}{q}||f'(y)|, \quad (*)$$

where y is a real number between α and $\frac{p}{q}$.

Notice that if $|\alpha - \frac{p}{q}| > 1$, then we may take $C(\alpha) := 1$ since $1 > \frac{1}{q^d}$. Suppose next that $|\alpha - \frac{p}{q}| \leq 1$. Let $f(x) =: a_d x^d + ... + a_1 x + a_0$. Then $f'(x) = d a_d x^{d-1} + ... + a_1$. The value y is between α and $\frac{p}{q}$, so $|y| \leq |\alpha| + 1$. Thus,

$$|f'(y)| \leq d a_d (|\alpha| + 1)^{d-1} + ... + 2|a_2|(|\alpha| + 1) + |a_1| =: C_1(\alpha).$$

(Note that we have here used the assumption that f has positive leading coefficient.) By (*),

$$|\alpha - \frac{p}{q}| \geq \frac{1/C_1(\alpha)}{q^d} > \frac{1/(2C_1(\alpha))}{q^d},$$

so we can take $C(\alpha) := \frac{1}{2C_1(\alpha)}$.

But what if $C_1(\alpha) = 0$? We claim that this cannot happen because $|f'(y)| > 0$. Indeed, since f is irreducible, it does not have a double root at α, so $f'(\alpha) \neq 0$. For sufficiently large q, if $\frac{p}{q}$ is the best rational approximation (in lowest terms) to α and y is between α and $\frac{p}{q}$, we can make $|f'(y)| \geq \frac{1}{2}|f'(\alpha)|$ by continuity of the derivative f'. Therefore, there is some positive integer N such that for $q \geq N$, the above argument works with $C(\alpha) = \frac{1}{2C_1(\alpha)}$. Moreover, for $1 \leq q < N$, we have only finitely many additional inequalities we want to be satisfied, so we can choose a possibly smaller but positive $C(\alpha)$ so that the same result holds for these finitely many cases. □

Anton remarks that Schmidt's textbook on Diophantine approximations proves the previous result through the use of Taylor's theorem.

Liouville proved that the number

$$\alpha = \sum_{n=1}^{\infty} 10^{-n!}$$

is transcendental. He was able to do this because the sum converges very rapidly. By truncating the sum, you get very good rational approximations, which he showed are too good for α to be an algebraic number. This basic idea is still more or less the main method in proving transcendence results. We now study Liouville's approach in more detail.

Theorem 1.4. *Let* $\alpha = \sum_{n=1}^{\infty} 10^{-n!}$. *Then* α *is transcendental.*

Proof. For each positive integer k, let

$$q_k = 10^{k!} \text{ and } p_k = 10^{k!}(\sum_{n=1}^{k} 10^{-n!}).$$

Then

$$|\alpha - \frac{p_k}{q_k}| = \sum_{n=k+1}^{\infty} 10^{-n!} < \frac{2}{10^{(k+1)!}} = \frac{2}{q_k^{k+1}}. \quad (1)$$

Suppose α is algebraic of degree 1. In this case, it suffices to prove that α is irrational. But this follows from the fact that it does not have a periodic decimal expansion.

Now suppose that α is algebraic of degree $d > 1$. Then by Theorem 1.3, there exists $C(\alpha) > 0$ such that

$$|\alpha - \frac{p_k}{q_k}| > \frac{C(\alpha)}{q_k^d}. \quad (2)$$

Combining (1) and (2), we find that

$$\frac{C(\alpha)}{q_k^d} < \frac{2}{q_k^{k+1}}, \text{ so } q_k^{k+1-d} < \frac{2}{C(\alpha)},$$

which cannot hold for k sufficiently large. □

In 1873, Hermite proved that e is transcendental. In 1874, Cantor proved that the real transcendental numbers are dense in $\mathbb{R}$. He did this by showing that the algebraic numbers are countable. In 1882, Lindemann proved that π is transcendental. More generally, the *Hermite–Lindemann theorem* states that if β is a non-zero complex number, then at least one of $\{\beta, e^{\beta}\}$ is transcendental. Taking $\beta = 2\pi i$ gives that π is transcendental. Lindemann stated and Weierstrass proved in 1885 that if $\beta_1, ..., \beta_n$ are linearly independent over $\mathbb{Q}$, then $e^{\beta_1}, ..., e^{\beta_n}$ are *algebraically independent*, i.e., they do not satisfy any non-trivial polynomial equation over $\mathbb{Q}$. We will see that this is what motivated Baker's work.

2 Sept. 6, 2019

In 1900, Hilbert proposed a list of 23 problems that he felt were fundamental. Cam talks about how Hilbert would master one field of mathematics and then switch gears to another. He would ask naive questions at first, but in a few years he would catch up to the state of the art in the new field.

Hilbert's seventh problem is the following: *Let α be an algebraic number not equal to* 0 *or* 1*, and let β be an irrational algebraic number. Prove that α^β is transcendental.* In 1934, the seventh problem was independently proved by Gelfond and Schneider.

Theorem 2.1. *Let α be an algebraic number not equal to* 0 *or* 1*, and let β be an irrational algebraic number. Then α^β is transcendental.*

This was Schneider's PhD problem! He didn't know that the problem his advisor gave him was famous. There had, however, been some progress in 1929 on this problem that suggested a way to solve it, and Siegel (who was Schneider's advisor) knew there was probably a way to do it. Schneider told Cam that he showed Siegel the first proof he came up with and that Siegel found a mistake. Then Schneider patched it and came back, and Siegel told him, "Congratulations. You'll probably get a PhD."

For an example of what Hilbert's seventh problem is talking about, note that $\sqrt{2}^{\sqrt{2}}$ and $e^\pi = i^{-2i}$ are transcendental. Baker generalized the Gelfond–Schneider theorem in 1967 by proving the following result.

Theorem 2.2. *If $\alpha_1, ..., \alpha_n$ are algebraic numbers different from* 0 *or* 1 *and $\beta_1, ..., \beta_n$ are algebraic numbers such that $\{1, \beta_1, ..., \beta_n\}$ is a $\mathbb{Q}$–linearly independent set, then*

$$\alpha_1^{\beta_1} ... \alpha_n^{\beta_n}$$

is transcendental.

Remark 2.3. We need the condition that that set is $\mathbb{Q}$–linearly independent to avoid, for example, the situation where $\beta_2 = -\beta_1$ and other obviously bad things.

Baker also proved the next theorem.

Theorem 2.4. *If $\beta_0, ..., \beta_n, \alpha_1, ..., \alpha_n$ are any non-zero algebraic numbers, then*

$$e^{\beta_0} \alpha_1^{\beta_1} ... \alpha_n^{\beta_n}$$

is transcendental.

It is natural to ask for a quantitative version of these results. In particular, can we estimate from below the quantity

$$|\beta_0 + \beta_1 \log \alpha_1 + ... + \beta_n \log \alpha_n - \log \alpha_{n+1}|$$

for other algebraic numbers α_{n+1}? It is not *a priori* clear what this means because $\log \alpha_{n+1}$ can be arbitrarily close to the sum of the other terms. So you need to have a lower bound that depends on some information about the other terms. It turns out to depend on a notion of *height* that is a measure of arithmetic complexity.

The above approximation problem motivates the following simpler question. For our purposes, it is enough to consider estimates from below for

$$|\beta_1 \log \alpha_1 + ... + \beta_n \log \alpha_n| \quad (*)$$

where $\beta_1, ..., \beta_n$ are integers. (We are taking $\alpha_{n+1} = 1$ in the earlier expression.) This is a degenerate case of the previous situation. Earlier we wanted $\{1, \beta_1, ..., \beta_n\}$ to be $\mathbb{Q}$–linearly independent, but if the β_i are integers, then that clearly is not the case. Nevertheless, we can still obtain estimates from Baker's proof. We will show that (*) is either 0 or it is bounded away from 0 by an *explicit quantity.* To make the lower bound explicit, we need a measure for the complexity of an algebraic number α.

Definition 2.5. Let

$$f(x) = a_d x^d + ... + a_1 x + a_0$$

be the minimal polynomial of α. We define the *naive height* of α, denoted $H(\alpha)$, by

$$H(\alpha) = \max(|a_d|, |a_{d-1}|, ..., |a_0|).$$

Baker established lower bounds for linear forms in the logarithms of algebraic numbers in 1966. We will state a result from 1993 due to Baker and Wüstholz.

Theorem 2.6. *Let $\alpha_1, ..., \alpha_n$ be algebraic numbers different from 0 and 1, and let $\log \alpha_1, ..., \log \alpha_n$ denote the principal branch of the logarithm function of $\alpha_1, ..., \alpha_n$. Let $K = \mathbb{Q}(\alpha_1, ..., \alpha_n)$, and let $d := [K : \mathbb{Q}]$. Suppose that $A_i := \max(H(\alpha_i), e)$ for $i = 1, ..., n$ and that $b_1, ..., b_n$ are integers, and let $B := \max(|b_1|, ..., |b_n|, e)$. Let*

$$\Lambda := b_1 \log \alpha_1 + ... + b_n \log \alpha_n.$$

If $\Lambda \neq 0$, then

$$|\Lambda| > \exp(-(16nd)^{2(n+2)} \log A_1 ... \log A_n \log B).$$

This seems like a pretty weak lower bound. It's an exponential of a negative exponential! It is not clear immediately that this is a significant result. We also present the following useful proposition.

Proposition 2.7. *Let α be an algebraic number with height H and minimal polynomial $f(x) = a_d x^d + ... + a_0$. Then*

$$|\alpha| < \frac{H}{|a_d|} + 1.$$

Proof. We may assume that $|\alpha| > 1$ since otherwise the result holds. Since α is a root of f,

$$0 = a_d \alpha^d + ... + a_1 \alpha + a_0,$$

so

$$0 = a_d\alpha + a_{d_1} + ... + a_0\alpha^{-d+1},$$

which implies that

$$|a_d\alpha| = |a_{d-1} + ... + a_0\alpha^{-d+1}|.$$

By the triangle inequality, we obtain

$$|a_d\alpha| \leq H \cdot (1 + |\alpha|^{-1} + ... + |\alpha|^{-d+1}) < \frac{H}{(1-|\alpha|^{-1})},$$

so

$$|\alpha| - 1 < \frac{H}{|a_d|}$$

as required. □

Note that $x^d f(\frac{1}{x})$ is the minimal polynomial of α^{-1} provided that $\alpha \neq 0$, so in this case we may apply Proposition 2.7 to conclude that

$$|\alpha| > (\frac{H}{|a_0|} + 1)^{-1}.$$

3 Sept. 9, 2019

The following is a result from Cam's own PhD thesis that gives a weaker bound than the result we saw last class of Baker and Wüstholz.

Proposition 3.1. *Let $b_1, ..., b_n$ be integers with absolute values at most $B \geq 2$. Let $\alpha_1, ..., \alpha_n$ be non-zero algebraic numbers with heights at most $A \geq 2$. Let $d := [\mathbb{Q}(\alpha_1, ..., \alpha_n) : \mathbb{Q}]$. Let*

$$\Lambda := b_1 \log \alpha_1 + ... + b_n \log \alpha_n,$$

where log *denotes the principal branch of the logarithm function. If $\Lambda \neq 0$, then*

$$|\Lambda| > (3A)^{-ndB}.$$

Proof. Let a_j be the leading coefficient in the minimal polynomial of α_j if $b_j > 0$ and α_j^{-1} if $b_j < 0$. Let

$$w := a_1^{|b_1|}...a_n^{|b_n|}(\alpha_1^{b_1}...\alpha_n^{b_n} - 1),$$

which is an algebraic integer of degree at most d. Let σ be an embedding of $\mathbb{Q}(\alpha_1, ..., \alpha_n)$ into $\mathbb{C}$ which fixes $\mathbb{Q}$. Then the conjugate $\sigma(w)$ of w has the form

$$a_1^{|b_1|}...a_n^{|b_n|}(\sigma(\alpha_1^{\epsilon_1})^{|b_1|}...\sigma(\alpha_n^{\epsilon_n})^{|b_n|} - 1)$$

where $\epsilon_i := \frac{b_i}{|b_i|}$ for $i = 1, ..., n$.

Now, by Proposition 2.7,

$$|a_i\sigma(\alpha_i^{\epsilon_i})| < 2A$$

and, as a consequence,

$$|\sigma(w)| < 2(2A)^{nB}.$$

Since $\Lambda = \log(\alpha_1^{b_1}...\alpha_n^{b_n})$, if $w = 0$ and $\Lambda \neq 0$, then Λ is a multiple of $2\pi i$ and the result holds. Suppose now that $w \neq 0$. Let N denote the norm function associated to the field extension $\mathbb{Q}(\alpha_1, ..., \alpha_n)/\mathbb{Q}$. Then $|N(w)| \geq 1$. Since the norm can be written as a product of Galois conjugates and since σ was arbitrary above,

$$1 \leq |N(w)| = |w||\sigma(w)|^{d-1} < |w|(2(2A)^{nB})^{d-1}$$

so

$$|w| > (2(2A)^{nB})^{-d+1}$$

From the inequality $|e^z - 1| \leq |z|e^{|z|} <$ we find on setting $z := \Lambda$ that

$$\frac{|w|}{a_1^{|b_1|}...a_n^{|b_n|}} = |\alpha_1^{b_1}...\alpha_n^{b_n} - 1| \leq |\Lambda|e^{|\Lambda|}.$$

Either $|\Lambda| > \frac{1}{2}$, in which case we are done, or $1 < e^{|\Lambda|} \leq e^{1/2} < 2$. In the latter case,

$$|\Lambda| > \frac{(2(2A)^{nB})^{-d+1}}{e^{|\Lambda|}A^{nB}} > 2^{-d+1}(2A)^{-nBd} > (3A)^{-nBd},$$

as required. (The last step of the last inequality follows from the fact that $n \geq 1$, $B \geq 2$, and $2(\log_2 3 - 1) > 1$ if you work it out.) □

Note that if $\Lambda \neq 0$, then it follows from Proposition 3.1 that

$$|\Lambda| > \exp(-nd(\log 3A)B).$$

Suppose $A \geq e$. Then $3A < A^3$, so

$$|\Lambda| > \exp(-3nd(\log A)B). \quad (1)$$

(Note that although Proposition 3.1 assumes $A \geq 2$ which is already sufficient to obtain $3A < A^3$, we are mimicking the assumption in the hypothesis of Theorem 2.6 in taking $A \geq e$.) Recall that Baker and Wüstholz proved that if $\Lambda \neq 0$, then

$$|\Lambda| > \exp(-(16nd)^{2n+4} \log A_1 ... \log A_n \log B). \quad (2)$$

Without loss of generality, we can suppose that $A = A_n \geq A_i$ for $i = 1, ..., n$. Thus, we see that (2) is an improvement on (1) whenever

$$3nd(\log A)B > (16nd)^{2n+4} \log A_1 ... \log A_n \log B,$$

so whenever

$$B/\log B > \frac{1}{3nd}(16nd)^{2n+4} \log A_1 ... \log A_{n-1}.$$

Thus, we see that we can obtain significant information on linear forms with "large" coefficients. Basically, large powers of algebraic numbers of small height cannot be too close.

We now consider a simpler situation. The next conjecture is due to Lang and Waldschmidt.

Conjecture 3.2. Let $a_1, ..., a_n$ be positive rational numbers, and let $b_1, ..., b_n$ be non-zero integers. Put $B_j := \max(|b_j|, 1)$. Let $B := \max_j B_j$, let $A_j := \max(H(a_j), 1)$, and let

$$\Lambda := b_1 \log a_1 + ... + b_n \log a_n.$$

Let $\epsilon > 0$. Then there exists a positive real number $C(\epsilon)$ such that if $\Lambda \neq 0$, then

$$|\Lambda| > \frac{C(\epsilon)^n B}{(B_1 ... B_n A_1^2 ... A_n^2)^{1+\epsilon}},$$

so

$$|\Lambda| > C(\epsilon)^n B^{-n(1+\epsilon)} A^{-2n(1+\epsilon)}$$

$$> \exp(-3n(\log B + \log A + \log C(\frac{1}{2})^{-1})).$$

Anton asks why Cam is using $C(\frac{1}{2})$ in the previous bound. Cam basically says this is a matter of notational convenience; we just choose C to be sufficiently small at 1/2.

Cam gives some anecdotes. Apparently Serge Lang could write a 200-page textbook in two weeks and would get intensely interested in various causes, and Michel Waldschmidt is an ultramarathon runner and almost single-handedly developed the French school of transcendence theory. Cam thinks that probably nobody in the next millennium will prove the conjecture above. Anton asks whether the result of Baker and Wüstholz has been improved, and Cam says the best currently known result in that direction is due to Matveev.

4 Sept. 11, 2019

We motivate the Lang–Waldschmidt conjecture (Conjecture 3.2) from last class. They considered the set S of linear combinations $b_1 \log a_1 + ... + b_n \log a_n$ with $|b_j| \leq B_j$ and $H(\alpha_j) \leq A_j$ with $\alpha_j > 0$. Then there are at most $(2B_1 + 1)...(2B_n + 1)$ possibilities for the B_j, and since the a_j in Conjecture 3.2 are positive rationals, there are at most $A_1^2 ... A_n^2$ possibilities for the a_j. (The squaring comes from choosing a positive integer numerator and a positive integer denominator.) All these linear combinations lie in the interval $[-nB \log A, nB \log A]$ where $A := \max_j A_j$ and $B := \max_j B_j$. Under the assumptions that (i) the upper bound $(2B_1 + 1)...(2B_n + 1)A_1^2 ... A_n^2$ is asymptotically the same as the actual number of a_j and b_j satisfying those constraints (which is not true in all cases since, for example, the a_j might all be equal) and that (ii) the numbers are independently and uniformly distributed, then the average distance between two of these linear combinations is

$$\frac{2nB \log A}{(2B_1 + 1)...(2B_n + 1)A_1^2 ... A_n^2}.$$

The $1+\epsilon$ in the formula from Conjecture 3.2 is just a fudge factor similar to the one in the statement of the *abc* conjecture. Similarly, the $\log A$ we obtain in the numerator only helps us in comparison to Lang and Waldschmidt's conjecture since the average we get is bigger than the lower bound of the conjecture.

Definition 4.1. Let $f \in \mathbb{Z}[x]$ be non-zero. We define the *Mahler measure*, denoted $M(f)$, of f as follows. Suppose $f(x) = a_d x^d + ... + a_1 x + a_0$ and

$$f(x) = a_d \prod_{i=1}^{d} (x - \alpha_i).$$

Let

$$M(f) := |a_d| \prod_{i=1}^{d} \max(1, |\alpha_i|).$$

For any algebraic number α, we define $M(\alpha)$ to be $M(f)$ where f is the minimal polynomial of α.

The Mahler measure thus defines another height function. The following bound relating the Mahler measure to the naive height is due to Edmund Landau.

Proposition 4.2. *For any algebraic number α of degree d,*

$$M(\alpha) \leq (d+1)^{1/2} H(\alpha).$$

Proof. We will need *Jensen's formula*, which we now recall. Let f be an analytic function on the closed disc of radius r ($r > 0$) centred at the origin in the complex plane. Suppose that $\alpha_1, ..., \alpha_m$ are zeroes of f in that disc, counted with multiplicity. If $f(0) \neq 0$, then

$$\log |f(0)| = -\sum_{k=1}^{m} \log(\frac{r}{|\alpha_k|}) + \frac{1}{2\pi} \int_0^{\pi} \log |f(re^{i\theta})| \, d\theta. \quad (*)$$

We now apply (*) with f the minimal polynomial of α and $r = 1$. (Note that the last coefficient of the minimal polynomial is necessarily non-zero, which implies that $f(0) \neq 0$, so this is alright.) Let $\alpha_1, ..., \alpha_m$ be the roots of f of modulus at most 1. Suppose $f(x) = a_d x^d + ... + a_1 x + a_0 = a_d(x - \alpha_1)...(x - \alpha_d)$. Then $f(0) = a_0$ as mentioned, so

$$\log |a_0| - \log |a_1| - ... - \log |a_m| = \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| \, d\theta.$$

But $\frac{|a_0|}{|\alpha_1 ... \alpha_m|} = |a_d||\alpha_{m+1}|...|\alpha_d| = M(\alpha)$, so

$$\log M(\alpha) = \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| \, d\theta.$$

Therefore,

$$M(\alpha) = \exp(\frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})|\, d\theta),$$

so

$$M(\alpha)^2 = \exp(\frac{1}{2\pi}\int_0^{2\pi} \log|f(e^{i\theta})|^2\, d\theta.$$

By the arithmetic–geometric mean inequality for integrals (which can be found as Thm. 184 of *Inequalities* by Hardy, Littlewood, and Polya),

$$M(\alpha)^2 \le \frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^2\, d\theta = a_0^2 + ... + a_d^2 \le (d+1)H(\alpha)^2$$

as required. □

It can be shown that if $\Lambda = b_1 \log\alpha_1 + ... + b_n \log\alpha_n$ is a linear form in logarithms of algebraic numbers with the b_i some non-zero integer coefficients and if $\Lambda = 0$, then we can find a linear form in $\log\alpha_1$,..., $\log\alpha_n$ with coefficients which are not all zero and are "small". Baker proved this by analytic means, but the proof was quite involved. Stark showed you could do it using the geometry of numbers, and Loxton and van der Poorten gave a nice proof along these lines.

Theorem 4.3. *Let $\alpha_1, ..., \alpha_n$ be non-zero algebraic numbers, and let $\log\alpha_1$,..., $\log\alpha_n$ be linearly dependent over the rationals. Suppose that*

$$A_j := \max(M(\alpha_j), e^{|\log\alpha_j|/d}, e)$$

for $j = 1, ..., n$, where $d := [\mathbb{Q}(\alpha_1, ..., \alpha_n) : \mathbb{Q}]$. Then there exist integers $t_1, ..., t_n$, not all zero, such that

$$t_1 \log\alpha_1 + ... + t_n \log\alpha_n = 0$$

and

$$|t_k| \le (11(n-1)d^3)^{n-1}\frac{\log A_1 ... \log A_n}{\log A_k}$$

for $k = 1, ..., n$.

For the proof, we need some ideas from the geometry of numbers.

Definition 4.4. A set S in $\mathbb{R}^n$ is said to be *symmetric* about the origin $O := (0, 0, ..., 0)$ if whenever $x \in S$, then $-x \in S$ as well. The set S is said to be *convex* if whenever $x \in S$ and $y \in S$, then all points on the line segment $\{(1-t)x + ty \mid 0 \le t \le 1\}$ joining x and y are in S. The *volume* of S is the Riemann integral of the characteristic function of S proved this characteristic function is integrable. (Note that it can be shown that every bounded convex set in $\mathbb{R}^n$ has a volume.) Finally, an *integer point* in $\mathbb{R}^n$ is a point with integer coordinates

5 Sept. 13, 2019

The following theorem is known as *Minkowski's convex body theorem*. It was proved in 1896.

Theorem 5.1. *Let A be a convex set in $\mathbb{R}^n$ that is bounded and symmetric about the origin, and has volume $\mu(A)$. If $\mu(A) > 2^n$, then A contains an integer point different from $0 := (0, 0, ..., 0)$.*

Proof. Let A_m be the set of rational points with denominator m in A, so

$$A_m := \{(\frac{t_1}{m}, ..., \frac{t_n}{m}) \in A \mid (t_1, ..., t_n) \in \mathbb{Z}^n\}.$$

(Note that we do not require the coordinates to be in lowest terms, so in particular $0 \in A_m$ for every m.) We have

$$\lim_{m\to\infty} \frac{|A_m|}{m^n} = \mu(A).$$

Since $\mu(A) > 2^n$ and the limit above holds, for m sufficiently large we have $|A_m| > (2m)^n$. Thus there exist $a = (\frac{a_1}{m}, ..., \frac{a_n}{m})$ and $b = (\frac{b_1}{m}, ..., \frac{b_n}{m})$ in A_m with $a \neq b$ and $a_i \equiv b_i \pmod{2m}$ for $i = 1, ..., n$. Then $\frac{1}{2}(a - b)$ is an integer point different from 0. (Note that it is an integer point because each difference $a_i - b_i$ is congruent to 0 mod $2m$, hence is even.) Since $a, b \in A_m$, they are also in A. Since A is symmetric, $-b$ in A. Since A is convex, $\frac{1}{2}(a-b) \in A$. This completes the proof. □

Remark 5.2. (i) The set $B := \{(x_1, ..., x_n) \in \mathbb{R}^n \mid |x_i| < 1 \text{ for } i = 1, ..., n\}$ is a symmetric convex set in $\mathbb{R}^n$ of volume $\mu(B) = 2^n$, and the only integer point in B is 0. Therefore, the bound Minkowski obtained is sharp.

(ii) Note that if y is a non-zero integer point in A, then $-y$ is also a non-zero integer point in A.

The next theorem is known as *Minkowski's linear forms theorem.*

Theorem 5.3. *Let $B := (\beta_{ij}) \in GL_n(\mathbb{R})$. Let $c_1, ..., c_n \in \mathbb{R}_{>0}$ with $c_1...c_n \geq |\det(B)|$. Then there exists a non-zero point $x = (x_1, ..., x_n) \in \mathbb{Z}^n$ such that*

$$|\beta_{i1}x_1 + ... + \beta_{in}x_n| < c_i \textit{ for } i = 1, ..., n-1$$

and

$$|\beta_{n1}x_1 + ... + \beta_{nn}x_n| \leq c_n.$$

Proof. Write $L_i(x) := \beta_{i1}x_1 + ... + \beta_{in}x_n$ for $i = 1, ..., n$ and $L'_i(x) := \frac{1}{c_i}L_i(x)$ for $i = 1, ..., n$. We wish to find a non-zero $x \in \mathbb{Z}^n$ satifying the following system of inequalities:

$$|L'_i(x)| < 1 \text{ for } i = 1, ..., n-1$$

and

$$|L'_n(x)| \leq 1.$$

Note that the determinant of the matrix associated with $L'_1(x), ..., L'_n(x)$ is at most 1. Thus, we may assume, without loss of generality, that $c_1 = ... = c_n = 1$ and $|\det(B)| \leq 1$.
For each $\epsilon > 0$, we define A_ϵ to be the set of $(x_1, ..., x_n) \in \mathbb{R}^n$ for which

$$|\beta_{i1}x_1 + ... + \beta_{in}x_n| < 1 \text{ for } i = 1, ..., n-1$$

and

$$|\beta_{n1}x_1 + ... + \beta_{nn}x_n| < 1 + \epsilon.$$

Note that A_ϵ is a bounded symmetric subset of $\mathbb{R}^n$. Moreover, it is convex since if $\lambda \in \mathbb{R}$ with $0 \leq \lambda \leq 1$, then

$$|L'_i(\lambda x + (1-\lambda)y)| = |\sum_{j=1}^{n} \beta_{ij}(\lambda x_j + (1-\lambda)y_j)|$$

$$\leq \lambda|L'_i(x)| + (1-\lambda)|L'_i(y)|$$

$$< \begin{cases} \lambda + (1-\lambda) = 1 & \text{for } i = 1, ..., n-1, \\ (\lambda + (1-\lambda))(1+\epsilon) = 1+\epsilon & \text{for } i = n. \end{cases}$$

Since $\mu(A_\epsilon) \geq (1+\epsilon)2^n > 2^n$, by Theorem 5.1, there exists a non-zero integer point $x_\epsilon \in A_\epsilon$. For each positive integer k, we can find a non-zero integer point $x_{\frac{1}{k}} \in A_{\frac{1}{k}}$. The sets $A_1, A_{\frac{1}{2}}, A_{\frac{1}{3}}$, etc. all lie in the bounded set A_1. But there are only finitely many integer points in any bounded set. Thus, there must be some integer point x that is contained in $A_{\frac{1}{k}}$ for infinitely many k. Such an x then satisfies $|L'_i(x)| < 1$ for $i = 1, ..., n-1$ and $L'_n(x)| \leq 1$, which was what we wanted. □

The following result is due to Kronecker from 1857.

Theorem 5.4. *Suppose that α is an algebraic number with $M(\alpha) \leq 1$. Then either $\alpha = 0$ or α is a root of unity.*

Proof. By definition of the Mahler measure, $M(\alpha) = |a_d| \prod_{i=1}^{d} \max(1, |\alpha_i|)$ where a_d is the leading coefficient of the minimal polynomial of α and the α_i are the roots of the minimal polynomial. Thus, $|a_d| \leq 1$, but since the minimal polynomial is an integer polynomial, $a_d = 1$. Thus, α is an algebraic integer. The conjugates of α are $\alpha_1, ..., \alpha_d$. It follows that $M(\alpha^k) \leq 1$ for $k = 1, 2, ..., d$. In particular, $|\alpha_i^k| \leq 1$ for $i = 1, ..., d$.

The elementary symmetric functions in the α_i^k are integers of absolute value at most 2^d. Thus, α^k is a root of one of a finite collection of non-zero polynomials. By the pigeonhole principle, $\alpha^k = \alpha^\ell$ for two distinct positive integers k and ℓ. Thus, either $\alpha = 0$ or α is a root of unity. □

6 Sept. 16, 2019

In 1933, D. H. Lehmer asked the following question.

Question 6.1. Does there exist a positive number ϵ such that if α is an algebraic number with $M(\alpha) < 1+\epsilon$, then $M(\alpha) \leq 1$?

Lehmer showed that if such an ϵ exists, then it is less than 0.17628081.... He gave the example of a polynomial f given by

$$f(x) := x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1.$$

The polynomial f is irreducible with roots $\alpha_1, ..., \alpha_{10}$ and with $\alpha_1 = 1.17628081...$, $\alpha_2 = \alpha_1^{-1}$, and all other roots on the unit circle. Indeed, α_1 is the smallest known *Salem number*, where a Salem number is defined to be a real algebraic integer greater than 1 whose conjugate roots all have absolute value ≤ 1 and at least one of which has absolute value exactly 1. It is conjectured that there are Salem numbers arbitrarily close to 1 and greater than 1, and this seems to be a hard conjecture to prove or disprove.

If we let ϵ vary with the degree d, we can make some progress. In 1979, Dobrowolski proved that if α is a non-zero algebraic number of degree $d \geq 2$ with

$$M(\alpha) < 1 + \frac{1}{1200}(\frac{\log\log d}{\log d})^3,$$

then α is a root of unity. We will prove an easier bound.

Theorem 6.2. *Let d be a positive integer, and let α be a non-zero algebraic number of degree at most d that is not a root of unity. Then*

$$\log M(\alpha) > \frac{1}{11d^2}.$$

To prove Theorem 6.2, we need some preliminary results.

Proposition 6.3. *Let p be a prime, and let $f \in \mathbb{Z}[x_1, ..., x_n]$. Then there exists a polynomial $g \in \mathbb{Z}[x_1, ..., x_n]$ such that*

$$f(x_1^p, ..., x_n^p) - (f(x_1, ..., x_n))^p = pg(x_1, ..., x_n).$$

Proof. Put $x := (x_1, ..., x_n)$, so $x^p = (x_1^p, ..., x_n^p)$. If f is a monomial, say $f(x) := ax_1^{i_1}...x_n^{i_n}$, where the i_j are non-negative integers, then the result holds since

$$f(x^p) - f(x)^p = (a - a^p)x_1^{pi_1}...x_n^{pi_n}$$

and $a - a^p \equiv 0 \pmod{p}$ by Fermat's little theorem.

Suppose that the result holds for monomials f_1 and f_2 so that

$$f_1(x^p) - f_1(x)^p = pg_1(x)$$

and

$$f_2(x^p) - f_2(x)^p = pg_2(x)$$

with $g_1, g_2 \in \mathbb{Z}[x_1, ..., x_n]$. Then

$$(f_1 + f_2)^p - f_1^p - f_2^p = \sum_{k=1}^{p-1} \binom{p}{k} f_1^k f_2^{p-k}.$$

Note that p divides $\binom{p}{k}$ for $k = 1, ..., p-1$. Thus,

$$(f_1 + f_2)(x^p) - f_1(x)^p - f_2(x)^p = p(g_1(x) + g_2(x) - \sum_{k=1}^{p-1} \frac{(p-1)!}{(p-k)!k!} f_1^k f_2^{p-k}).$$

The result follows by induction on the number of monomials in f. □

Proposition 6.4. *Let α be a non-zero algebraic number. Suppose that h and ℓ are two distinct positive integers for which α^h and α^ℓ are conjugates. Then α is a root of unity.*

Proof. Let K be the splitting field of α over $\mathbb{Q}$. Since α^h and α^ℓ are conjugates, there is some element $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ for which $\sigma(\alpha^h) = \alpha^\ell$. We will now show that

$$\sigma^n(\alpha^{h^n}) = \alpha^{\ell^n}$$

for $n = 1, 2, ...$ by induction on n. The claim is plainly true for $n = 1$. Suppose it is true for $1 \leq k \leq n$. Then

$$\begin{aligned}\sigma^{n+1}(\alpha^{h^{n+1}}) &= \sigma(\sigma^n(\alpha^{h^n})^h) \\ &= \sigma((\alpha^{\ell^n})^h) \\ &= (\sigma(\alpha^h))^{\ell^n} \\ &= (\alpha^\ell)^{\ell^n} \\ &= \alpha^{\ell^{n+1}},\end{aligned}$$

as required.

The Galois group $\mathrm{Gal}(K/\mathbb{Q})$ is finite, so σ has finite order, say t. Then $\sigma^t = \mathrm{id}$, so

$$\sigma^t(\alpha^{h^t}) = \alpha^{h^t}.$$

On the other hand,

$$\sigma^t(\alpha^{h^t}) = \alpha^{\ell^t},$$

and so

$$\alpha^{\ell^t} = \alpha^{h^t}.$$

But $\ell \neq h$, so α is a root of unity. □

Definition 6.5. For any algebraic number α of degree d with conjugates $\alpha = \alpha_1, ..., \alpha_d$, we define the *house* of α, denoted $\overline{|\alpha|}$, by

$$\overline{|\alpha|} := \max(|\alpha_1|, ..., |\alpha_d|).$$

7 Sept. 18, 2019

The following theorem from 1978, which we began proving last class, is due to Dobrowolski.

Theorem 7.1. *If α is a non-zero algebraic integer of degree d which is not a root of unity, then*

$$\overline{|\alpha|} > 1 + \frac{1}{4ed^2}.$$

Proof. Let $\alpha =: \alpha_1, ..., \alpha_d$ be the conjugates of α over $\mathbb{Q}$. Let

$$f_h(x_1, ..., x_d) := x_1^h + ... + x_d^h$$

for $h = 1, 2, ...$ Put $S_h := f_h(\alpha_1, ..., \alpha_d)$ for $h = 1, 2, ...$ This is a trace; you can see directly that it is an integer because it is an algebraic integer (being a polynomial in algebraic integers) and is in $\mathbb{Q}$ (being invariant under the action of the Galois group).

By Fermat's little theorem,

$$S_h \equiv S_h^p \pmod{p}.$$

By Proposition 6.3,

$$S_{hp} - S_h^p = pg(\alpha_1, ..., \alpha_d)$$

for some $g \in \mathbb{Z}[x_1, ..., x_d]$. As before, we see that $g(\alpha_1, ..., \alpha_d)$ is an integer. Therefore,

$$S_{hp} \equiv S_h^p \pmod{p}$$

for $h = 1, 2, ...$

Observe that for any positive integer h,

$$|S_h| \leq d\overline{|\alpha|}^h.$$

Suppose that

$$\overline{|\alpha|} \leq 1 + \frac{1}{4ed^2}.$$

By Bertrand's postulate, there is a prime p with

$$2ed < p < 4ed.$$

For $1 \leq h \leq d$,

$$|S_h| \leq d(1 + \frac{1}{4ed^2})^h \leq d(1 + \frac{1}{4ed^2})^d = d\exp(d\log(1 + \frac{1}{4ed^2}))) \leq de,$$

and

$$|S_{hp}| < d(1 + \frac{1}{4ed^2})^{4ed^2} \leq de.$$

Therefore,

$$|S_{hp} - S_h| \leq 2de < p.$$

But since $S_h \equiv S_{hp} \pmod{p}$, we find that $S_h = S_{hp}$ for $h = 1, ..., d$. But the Newton sums determine the elementary symmetric polynomials in $\alpha_1, ..., \alpha_d$ and in $\alpha_1^p, ..., \alpha_d^p$. Therefore,

α and α^p have the same minimal polynomial and are thus conjugate. Therefore, α is a root of unity. □

We finally prove Theorem 6.2.

Proof. Let $c, k \in \mathbb{R}^+$ with $k > c$. Let

$$f(t) := \log(1 + \frac{1}{ct}) - \frac{1}{kt}.$$

Then

$$f'(t) = -\frac{1}{1 + 1/(ct)} \frac{1}{ct^2} + \frac{1}{kt^2},$$

so for $t > 0$ we have $f'(t) > 0$ when $t < \frac{1}{k-c}$. Also, $f'(t) < 0$ for $t > \frac{1}{k-c}$. Since $f(t)$ is positive for t sufficiently large, we see that

$$\log(1 + \frac{1}{ct}) - \frac{1}{kt} > 0$$

for $t > \frac{1}{k-c}$. Take $k = 11$ and $c = 4e$. Then

$$\log(1 + \frac{1}{4et}) > \frac{1}{11t}$$

for $t > \frac{1}{11-4e} = 7.88....$ Thus, for $d \geq 3$,

$$\log(1 + \frac{1}{4ed^2}) > \frac{1}{11d^2}$$

and also for $d = 2$,

$$\log(1 + \frac{1}{16e}) = 0.022732... \geq \frac{1}{44} = 0.02272....$$

The result now follows. □

Remark 7.2. We are about to discuss valuations. However, what Cam calls a "valuation" I call an "absolute value", and what Cam calls an "order" and denotes ord_p I call a "valuation" and denote v_p. Also, there is a notion of equivalence of absolute values (in my terminology), and I call such an equivalence class a "place". What all this means will soon become clearer.

Definition 7.3. Let K be a field. A function $|\ | : K \to \mathbb{R}$ is an *absolute value* if:

(i) For every $a \in K$, $|a| \geq 0$ and $|a| = 0$ if and only if $a = 0$.

(ii) For all $a, b \in K$, $|ab| = |a||b|$.

(iii) For all $a, b \in K$, $|a + b| \leq |a| + |b|$.

Example 7.4. The ordinary absolute value on $\mathbb{C}$ is an absolute value in the sense just defined.

Definition-Example 7.5. Let p be prime. We define the *p-adic absolute value* on $\mathbb{Q}$ by

$$|\frac{a}{b}|_p := p^{-v_p(a/b)}$$

for $a \neq 0$ and $b \neq 0$. The *p-adic valuation* v_p is defined on rationals (in lowest terms) by $v_p(a/b) := v_p(a) - v_p(b)$. If n is a non-zero integer, then $v_p(n)$ is the highest power of p that divides n. (Note that it does not matter whether n is positive or negative: $v_p(n) = v_p(-n)$.) Also, we formally let $v_p(0) := \infty$ where we use the convention that $p^{-\infty} = 0$. Thus, $|0|_p = 0$ for any prime p.

Definition-Example 7.6. Let k be any field and let T be transcendental over k. Let $K := k(T)$ be the field of rational functions in T with coefficients in k. Let λ be a real number with $0 < \lambda < 1$, and let $p(T)$ be an irreducible element of K. Then every $h \in K$ can be written in the form

$$p(T)^q \frac{f(T)}{g(T)}$$

where $f(T)$ and $g(T)$ are not divisible by $p(T)$ and q is an integer. Then we define $|\ |$ on K by

$$|h| := \begin{cases} \lambda^q & \text{if } h \neq 0. \\ 0 & \text{if } h = 0. \end{cases}$$

Notice that this seems to depend on the choice of λ. Later we will see that under the usual notion of equivalence of absolute values, any choice of λ gives an equivalent absolute value.

Definition-Example 7.7. Let K be a field. Define $|\ |_0$ on K by

$$|x|_0 := \begin{cases} 1 & \text{if } x \neq 0. \\ 0 & \text{if } x = 0. \end{cases}$$

This is known as the *trivial valuation* on K.

8 Sept. 20, 2019

We have still not defined when two valuations are equivalent, and we remedy this now.

Definition 8.1. If $|\ |$ and $|\ |_1$ are absolute values on a field K, we say that they are *equivalent* if there exists $\lambda \in \mathbb{R}_{>0}$ such that

$$|a| = |a|_1^\lambda$$

for every $a \in K$. An equivalence class of absolute values is known as a *place*.

Remark 8.2. Note that in Example 7.6, each choice of $0 < \lambda < 1$ defines an equivalent absolute value on $k(T)$ by the definition of equivalence just given.

Definition 8.3. An absolute value on a field K is said to be *non-Archimedean* if

$$|a + b| \leq \max(|a|, |b|)$$

for all $a, b \in K$.

Example 8.4. The absolute value $| \ |_p$ on $\mathbb{Q}_p$ is non-Archimedean for each prime p.

Remark 8.5. Given an absolute value $| \ |$ on a field K, we can define a metric and thus a topology on K by letting $d(a, b) := |a - b|$. One can check that two absolute values induce the same topology on K if and only if they are equivalent.

The following beautiful theorem is due to Ostrowski.

Theorem 8.6. *Every non-trivial absolute value on $\mathbb{Q}$ is equivalent to either the ordinary absolute value $| \ |$ or a p-adic absolute value $| \ |_p$ for some prime p.*

We will write $|a| =: |a|_{p_\infty}$. Therefore, we think of the Euclidean absolute value as the p-adic absolute value where p is an "infinite prime". Let $S(\mathbb{Q}) := \{p_\infty, p \text{ a prime in } \mathbb{Z}\}$. By the uniqueness of prime factorization over $\mathbb{Z}$, we have, for all $a \in \mathbb{Q}$, the product formula

$$\prod_{v \in S(\mathbb{Q})} |a|_v = \begin{cases} 1 & \text{if } a \neq 0. \\ 0 & \text{if } a = 0. \end{cases}$$

Let K be a finite extension of $\mathbb{Q}$. Let O_K be the ring of algebraic integers of K. For each prime p in $\mathbb{Z}$, the ideal (p) in O_K splits as a product

$$(p) = \mathfrak{p}_1^{e_1} ... \mathfrak{p}_t^{e_t}$$

where each $\mathfrak{p}_i$ is a prime ideal in O_K and each e_i is a positive integer for $i = 1, ..., t$.

Definition 8.7. The *residue class degree* of $\mathfrak{p}_i$ is defined to be $f_i := [O_K/\mathfrak{p}_i : O_{\mathbb{Q}}/p]$.

One can show that

$$e_1 f_1 + ... + e_t f_t = [K : \mathbb{Q}].$$

If K is a Galois extension, then $e_1 = ... = e_t$ and $f_1 = ... = f_t$.

We also wish to extend the Euclidean absolute value to the case of an arbitrary number field. Consider the $\mathbb{Q}$-isomorphisms of K into $\mathbb{C}$. Recall that $K = \mathbb{Q}(\alpha)$ for some α by the primitive element theorem and let $\alpha = \alpha_1, ..., \alpha_d$ be the conjugates of α. The $\mathbb{Q}$-isomorphisms σ are then determined by the effect of σ on α.

Thus, we may define $\mathbb{Q}$-isomorphisms $\sigma_i(\alpha) := \alpha_i$ for each $i = 1, ..., d$. Let $\alpha_1, ..., \alpha_{r_1}$ be real and $\alpha_{r_1+1}, ..., \alpha_{r_1+2r_2}$ be complex and not real. (Note that $r_1 + 2r_2 = d$.) We may suppose that $\alpha_{r_1+i} = \overline{\alpha_{r_1+r_2+i}}$ for $i = 1, ..., r_2$.

We now define an absolute value corresponding to each $\mathbb{Q}$-isomorphism $\sigma_1, ..., \sigma_{r_1+r_2}$.

Definition 8.8. The $\mathbb{Q}$-isomorphisms $\sigma_1, ..., \sigma_{r_1+r_2}$ are called *infinite primes* of K. The prime ideals of O_K are called *finite primes* of K. We denote the union of these two sets by

$S(K)$. For each element of $S(K)$, we define an absolute value as follows. For $\beta \in K$ and for $v = \mathfrak{p}$ a prime ideal of O_K, let

$$|\beta|_v := \begin{cases} N_{K/\mathbb{Q}}(\mathfrak{p})^{-\omega_\mathfrak{p}(\beta)/d} & \text{if } \beta \neq 0. \\ 0 & \text{if } \beta = 0. \end{cases}$$

Here $d = [K : \mathbb{Q}]$, $N_{K/\mathbb{Q}}$ denotes the norm, and $\omega_\mathfrak{p}(\beta)$ is the order of $\mathfrak{p}$ in the canonical decomposition of the fractional ideal (β) in K as the product of prime ideals. If $v = \sigma$ is an infinite prime, then

$$|\beta|_v := |\sigma(\beta)|^{g/d}$$

where $g = 1$ if σ is a real embedding and $g = 2$ otherwise.

Once again, it can be shown that the product formula holds because we have unique factorization into prime ideals. That is, if $\alpha \in K$ and $\alpha \neq 0$, then

$$\prod_{v \in S(K)} |\alpha|_v = 1.$$

We now introduce a new height function h on K. (We remark that the notation h often refers to the logarithm of the height we now define.)

Definition 8.9. The *height* of α in K for $\alpha \neq 0$ is defined by

$$h(\alpha) := \prod_{v \in S(K)} \max(1, |\alpha|_v).$$

Remark 8.10. It follows from our definition of the absolute values that if we extend K to L and α is in K, then

$$h(\alpha) = \prod_{v \in S(L)} \max(1, |\alpha|_v).$$

In other words, h is defined on all of the algebraic numbers!

9 Sept. 23, 2019

What is the connection between $h(\alpha)$ and the Mahler measure $M(\alpha)$?

Proposition 9.1. *Suppose that α is a non-zero algebraic number with minimal polynomial $f(x) := a_d x^d + \cdots + a_d = a_d \prod_{i=1}^{d}(x - \alpha_i)$. Then*

$$h(\alpha) = M(\alpha)^{1/d}.$$

Proof. For $\alpha \neq 0$,

$$h(\alpha)^d = \prod_{v \in S(\mathbb{Q}(\alpha))} \max(1, |\alpha|_v)^d$$

$$= \prod_{\mathfrak{p} \in S(\mathbb{Q}(\alpha))} \max(1, N_{K/\mathbb{Q}}(\mathfrak{p})^{-\omega_\mathfrak{p}(\alpha)}) \prod_{\sigma \in S(\mathbb{Q}(\alpha))} \max(1, |\sigma(\alpha)|^{g_\sigma}),$$

where $K := \mathbb{Q}(\alpha)$,

$$= \prod_{\mathfrak{p}\in S(\mathbb{Q}(\alpha))} \max(1, p^{-f\omega_{\mathfrak{p}}(\alpha)}) \prod_{i=1}^{d} \max(1, |\alpha_i|),$$

where p is the rational prime lying under $\mathfrak{p}$ and f is the residue class degree of $\mathfrak{p}$,

$$= |a_d| \prod_{i=1}^{d} \max(1, |\alpha_i|) = M(\alpha).$$

□

It follows that for an algebraic number $\alpha \neq 0$ and a positive integer k,

$$h(\alpha^k) = \prod_{v\in S(\mathbb{Q}(\alpha))} \max(1, |\alpha^k|_v) = \left(\prod_{v\in S(\mathbb{Q}(\alpha))} \max(1, |\alpha|_v) \right)^k = h(\alpha)^k.$$

Recall that if f is the minimal polynomial of α, then $g(x) = x^d f(x^{-1})$ is the minimal polynomial of α^{-1} (where d is the degree of α). But

$$M(\alpha) = \exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log|f(e^{i\theta})| \,\mathrm{d}\theta \right) = \exp\left(\frac{1}{2\pi} \int_0^{2\pi} \log|g(e^{i\theta})| \,\mathrm{d}\theta \right) = M(\alpha^{-1}).$$

Thus,

$$h(\alpha)^{|k|} = h(\alpha^k) \text{ for all } k \in \mathbb{Z}.$$

Furthermore,

$$h(\alpha\beta) = \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} \max(1, |\alpha\beta|_v) \leq \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} \max(1, |\alpha|_v) \prod_{v\in S(\mathbb{Q}(\alpha,\beta))} \max(1, |\beta|_v) \leq h(\alpha)h(\beta).$$

For sums, the bounds you get are not so nice, and Cam does not write them down.

Theorem 9.2. *Let $\alpha_1, \ldots, \alpha_n$ be non-zero algebraic numbers that are multiplicatively dependent. Suppose that*

$$A_j := \max(h(\alpha_j), e)$$

for $i = 1, \ldots, n$ and that $A_1 \leq \ldots A_n$. Let $d := [\mathbb{Q}(\alpha_1, \ldots, \alpha_n) : \mathbb{Q}]$. There exist $t_1, \ldots, t_n \in \mathbb{Z}$, not all zero, such that

$$\alpha_1^{t_1} \cdots \alpha_n^{t_n} = 1$$

with

$$|t_k| \leq \left(11nd^3\right)^n \log A_2 \cdots \log A_n$$

for $k = 1, \ldots, n$.

Remark 9.3. For any positive integer m, *Euler's phi function* $\varphi(m)$ counts the number of integers from $1, 2, \ldots, m$ that are coprime with m. One can show that there exists a positive number c such that foir $m > e^e$, we have

$$\varphi(m) > c\frac{m}{\log\log m}.$$

We will use a weaker estimate.

Lemma 9.4. *Let $m \in \mathbb{N}$. If $\varphi(m) = d$, then $m \leq 2d^2$.*

Proof. Note that

$$\varphi(m)^2 = m^2 \prod_{p|m}\left(1-\frac{1}{p}\right)^2 = m\left[\frac{m}{\prod_{p|m}\left(1-\frac{1}{p}\right)^{-2}}\right].$$

However,

$$\prod_{p|m}\left(1-\frac{1}{p}\right)^{-2} = \prod_{p|m}\left(\frac{p}{p-1}\right)^2 \leq 2\prod_{p|m} p \leq 2m,$$

and so

$$\varphi(m)^2 \geq \frac{m}{2}.$$

Thus, if $\varphi(m) = d$, then

$$m \leq 2d^2.$$

□

10 Sept. 25, 2019

We now prove Theorem 9.2.

Proof. We first suppose that $n = 1$. Thus, α_1 is a root of unity in a field of degree d. Thus, α is an m^{th} root of unity with $m \leq 2d^2$ and $\alpha^m = 1$. This proves the claim for $n = 1$.

Assume now that $n \geq 2$ and that there exist integers $k_1, \ldots, k_n$, not all zero, with

$$\alpha_1^{k_1} \cdots \alpha_n^{k_n} = 1.$$

Fix an integer j for which

$$|k_j| \geq |k_i|$$

for $i = 1, \ldots, n$. Put

$$c_i := \left(11nd^3 \log A_i\right)^{-1}$$

for $i = 1, \ldots, n$, $i \neq j$, and

$$c_j := \left(11nd^3\right)^{n-1} \frac{\log A_1 \cdots \log A_n}{\log A_j}.$$

Notice that

$$c_1 \cdots c_n = 1.$$

Consider the following system of inequalities:

$$(1) \qquad \left| x_i - \frac{k_i}{k_j} x_j \right| \leq c_i \text{ for } i = 1, \dots, n, i \neq j$$

and

$$(2) \qquad |x_j| \leq c_j.$$

Associated to this system is the matrix

$$B = \begin{pmatrix} 1 & 0 & \cdots & -k_1/k_j & \cdots & 0 \\ 0 & & 1 \cdots & -k_2/k_j & \cdots & 0 \\ 0 & 0 & \ddots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & -k_n/k_j & \cdots & 1 \end{pmatrix}$$

One can show that $\det(B) = 1$. (Note that in the 2×2 case, this matrix is $\begin{pmatrix} 1 & -k_1/k_j \\ 0 & 1 \end{pmatrix}$, so it works.) By Minkowksi's linear forms theorem (Theorem 5.1), there exists a non-zero integer point $(b_1, ..., b_n)$ satisfying (1) and (2). Set

$$\alpha := \alpha_1^{b_1} ... \alpha_n^{b_n}.$$

We claim that α is a root of unity. Since $\alpha_1^{k_1} ... \alpha_n^{k_n} = 1$, we can write

$$\alpha^{k_j} = \prod_{i=1}^{n} \alpha_i^{b_i k_j - b_j k_i},$$

so by (1),

$$h(\alpha)^{|k_j|} \leq \prod_{i=1, i \neq j}^{n} h(\alpha_i)^{c_i |k_j|}.$$

It follows that $h(\alpha) \leq \prod_{i=1, i\neq j}^{n} h(\alpha_i)^{c_i}$, so

$$\frac{1}{d} \log M(\alpha) \leq \sum_{i=1, i \neq j}^{n} c_i \log h(\alpha_i).$$

Therefore,

$$\log M(\alpha) \leq d \sum_{i=1, i \neq j}^{n} \frac{\log h(\alpha_i)}{11nd^3 \max(\log h(\alpha_i), 1)}$$

$$= \frac{1}{11d^2}.$$

By Theorem 6.2, α is a root of unity. As before, α is a root of unity in a field of degree at most d and is thus an m^{th} root of unity with $m \leq 2d^2$. Thus,

$$1 = \alpha^m = \alpha_1^{b_1 m} ... \alpha_n^{b_n m}.$$

Furthermore, by (2),

$$|b_j m| \leq c_j m \leq (2d^2)(11nd^3)^{n-1} \frac{\log A_1 \cdots \log A_n}{\log A_j}$$

$$\leq (11nd^3)^n \log A_2 \cdots \log A_n$$

since $A_1 \leq A_2 \leq \cdots \leq A_n$.
For $i \neq j$, we use the fact that $|k_i| \leq k_j$ for $i \neq j$ to deduce from (1) and (2) that

$$|b_i| \leq c_i + |b_j| \leq 1 + c_j \leq \frac{3}{2} c_j,$$

so

$$|b_i m| \leq 3d^2 (11nd^3)^{n-1} \frac{\log A_1 \cdots \log A_n}{\log A_j}$$

$$\leq (11nd^3)^n \log A_2 \cdots \log A_n$$

for $i = 1, \ldots, n, i \neq j$. The result follows. □

Cam talks about how Harold Stark was given the PhD problem of pushing the bound up on non-existence of imaginary quadratic fields with class number 1. He proved a very strong bound. Then he was reading a math review by Morgan Ward of a paper by Heegner where Heegner proved that the class number problem was solved. This scared Stark, but his advisor made some calls and it turned out that Heegner's proof was flawed. Stark got a job at the University of Michigan and continued to work on the problem. He was then denied tenure because his file did not look very impressive. That summer, he solved the problem and was hired as a full professor at MIT. He then discovered, along with several others, that Heegner's argument was essentially sound and that the proof could be patched.

We will show some applications of Theorem 2.6 and Theorem 9.2. Let Let N be an integer, and let $S_{a,b}(N)$ be the sum of the digit sum of N in base a with the digit sum of N in base b. In 1970, Senge and Strauss proved that if a and b are integers larger than 1 with $\log a / \log b$ irrational, then the number of integers n for which $S_{a,b}(n)$ lies below a given bound is finite. (Here Cam points out that Strauss had a joint paper with Einstein and another one with Erdős.) The proof was not effective because given the bound, the proof does not tell us the size of n needed for the sum to exceed the bound. We can overcome this issue by means of Theorems 2.6 and 9.2.

Let α and β be integers with $0 \leq \alpha < a$ and $0 \leq \beta < b$. Denote the number of digits in the expansion of a positive integer n in base a all of which are different from α by $L_{\alpha,a}(n)$. Define $L_{\beta,b}(n)$ similarly, and put

$$L_{\alpha,a,\beta,b}(n) := L_{\alpha,a}(n) + L_{\beta,b}(n).$$

Then the sum of the digits of n in base a plus the sum of the digits of n in base b is greater than or equal to $L_{0,a,0,b}(n)$.

11 Sept. 30, 2019

Example 11.1. Since $33 = 2^5 + 1$ and $33 = 3^3 + 2 \cdot 3$,

$$L_{0,2,0,3}(33) = L_{0,2} + L_{0,3} = 2 + 2 = 4.$$

The sum of the digits appearing in the two expansions is 5.

Example 11.2. Since $63 = 2^5 + 2^4 + 2^3 + \cdots 2 + 1 = 2 \cdot 5^2 + 2 \cdot 5 + 3$, we have

$$L_{1,2,2,5}(63) = 1.$$

Notice that the condition of Senge and Strauss that $\log a / \log b$ is irrational is necessary because if $a^r = b^s$ with $r, s \in \mathbb{N}$, then when $n = a^{rk} = b^{sk}$ for $k = 1, 2, \ldots$, we have that the sum of the digits in base a plus the sum of the digits in base b is the same as $L_{0,a,0,b}(n) = 2$.

The following is a 1980 result of Stewart. Cam explains that he proved it during a Dutch PhD defense. Apparently Dutch PhD defenses are rather formal affairs and anyone who attempts one gets their thesis: you have effectively already been awarded your thesis by the time you get to the defense, unlike in the British or Canadian systems.

12 Oct. 2, 2019

Theorem 12.1. *Let a and b be integers larger than one, and suppose that $\log a / \log b$ is irrational. Let α and β be integers with $0 \leq \alpha < a$ and $0 \leq \beta < b$. There is a positive number C, which is effectively computable in terms of a and b, such that if n is an integer larger than 25 then*

$$L_{\alpha,a,\beta,b}(n) > \frac{\log \log n}{\log \log \log n + C} - 1.$$

Proof. Suppose $n > a + b$. We can form the expansions

$$n = a_1 a^{m_1} + \alpha \left(\frac{a^{m_1} - 1}{a - 1} \right) + a_2 a^{m_2} + \cdots + a_{n'} a^{m_r}$$

and

$$n = b_1 b^{\ell_1} + \beta \left(\frac{b^{\ell_1} - 1}{b - 1} \right) + b_2 b^{\ell_2} + \cdots + b_t b^{\ell_t}$$

where $0 < a_1 < a$, $-\alpha \leq a_i < a - \alpha$ with $a_i \neq 0$ for $i = 2, \ldots, r$ and $0 < b_1 < b$, $-\beta \leq b_i < b - \beta$ with $b_i \neq 0$ for $i = 2, \ldots, t$, and $m_1 > m_2 > \cdots > m_r \geq 0$, $\ell_1 > \ell_2 > \cdots > \ell_t \geq 0$. We put

$$\theta := c_1 \log \log n \quad (1)$$

where c_1 is a positive number larger than 4 which is effectively computable in terms of a and b alone.

We shall assume that $c_2, c_3, \ldots$ are positive numbers which are effectively computable in terms of a and b alone and may be determined independently of c_1. We first assume that

$$n > c_2 > 25.$$

Define k to be the unique integer for which

$$\theta^k \leq \frac{\log n}{4 \log a} < \theta^{k+1}, \quad (2)$$

and put

$$\Theta_1 := (0, \theta], \Theta_2 := (\theta, \theta^2], \ldots, \Theta_k := (\theta^{k-1}, \theta^k].$$

If each interval Θ_s for $s = 1, \ldots, k$ contains at least one term either of the form $m_1 - m_i$ or $\ell_1 - \ell_j$, then the theorem holds since then

$$L_{\alpha,a,\beta,b}(n) \geq r + t - 2 \geq k, \quad (3)$$

while from (2),

$$(k+1) \log \theta > \log \log n - \log(4 \log a),$$

hence

$$k > \frac{\log \log n}{\log \theta} - \log(4 \log a) - 1.$$

Thus,

$$k > \frac{\log \log n}{\log \log \log n + \log a} - \log(4 \log a) - 1 \quad (4).$$

Our result now follows from (3) and (4) since $L_{\alpha,a,\beta,b}(n) \geq 0$. Thus we may assume that there exists an integer s with $1 \leq s \leq k$ for which Θ_s contains no term of the form $m_1 - m_i$ or $\ell_1 - \ell_j$. Define p and q by the inequalities

$$m_i - m_p \leq \theta^{s-1}, m_1 - m_{p+1} \geq \theta^s, \quad (5)$$

$$\ell_1 - \ell_q \leq \theta^{s-1}, \ell_1 - \ell_{q+1} \geq \theta^s \quad (6)$$

with the convention that m_{r+1} and ℓ_{t+1} are zero. We now write

$$(b-1)(a-1)n = ((b-1)(a-1)a_1 + (b-1)\alpha)\, a^{m_1} + (b-1)(a-1)a_2 a^{m_2} + \cdots + (b-1)(a-1)a_r a^{m_r} - (b-1)\alpha$$

$$= A_1 a^{m_p} + A_2$$

where A_1 and A_2 are integers and

$$A_1 := ((b-1)(a-1)a_1 + (b-1)\alpha)\, a^{m_1 - m_p} + \cdots + (b-1)(a-1)a_p.$$

We have

$$0 < A_1 < (b-1)(a-1)a^{m_1-m_p+1} + (b-1)\alpha a^{m_1-m_p},$$

so

$$0 < A_1 < 2(b-1)(a-1)a^{m_1-m_p+1} \quad (7).$$

Then

$$A_2 = (b-1)(a-1)a_{m_{p+1}}a^{m_{p+1}} + \cdots + (b-1)(a-1)a_r a^{m_r} - (b-1)\alpha,$$

so

$$0 \le |A_2| \le 2(b-1)(a-1)a^{m_{p+1}+1}. \quad (8)$$

Similarly, we have

$$(b-1)(a-1)n = B_1 b^{\ell_q} + B_2,$$

where

$$0 < B_1 < 2(b-1)(a-1)b^{\ell_1-\ell_q+1} \quad (9)$$

and

$$0 \le |B_2| < 2(b-1)(a-1)b^{\ell_{q+1}+1}. \quad (10)$$

We have

$$\frac{(b-1)(a-1)n}{(b-1)(a-1)n} = 1 = \frac{A_1 a^{m_p}}{B_1 b^{\ell_q}}\left(1 + \frac{A_2}{A_1 a^{m_p}}\right)\left(1 + \frac{B_2}{B_1 b^{\ell_q}}\right)^1.$$

If x and y are real numbers with absolute value at most $1/2$, then

$$\max\left\{\frac{1+x}{1+y}, \frac{1+y}{1+x}\right\} \le 1 + 4\max(|x|,|y|). \quad (11)$$

Notice that

$$\frac{|A_2|}{A_1 a^{m_p}} < \frac{2(b-1)(a-1)a^{m_{p+1}+1}}{(b-1)(a-1)a^{m_1}} \le 2a^{-m_1+m_{p+1}+1},$$

and by (5),

$$m_1 - m_{p+1} \ge \theta^s \ge \theta = c_1 \log\log n,$$

and thus for n sufficiently large,

$$\frac{|A_2|}{A_1 a^{m_p}} < \frac{1}{2} \text{ and similarly } \frac{|B_2|}{B_1 b^{\ell_q}} < \frac{1}{2}.$$

Thus on putting $R := \frac{A_1 a^{m_p}}{B_1 b^{\ell_q}}$, we see from (11) that

$$1 \le \max(R, R^{-1}) \le 1 + 4\max\left\{\frac{|A_2|}{A_1 a^{m_p}}, \frac{|B_2|}{B_1 b^{\ell_q}}\right\}$$

$$\le 1 + 8\max\{a^{-m_1+m_{p+1}+1}, b^{-\ell_1+\ell_{q+1}+1}\}.$$

Furthermore, since $\log(1+x) < x$ for $x > 0$,

$$0 < |\log R| \le 8ab\max\left\{a^{-m_1+m_{p+1}}, b^{-\ell_1+\ell_{q+1}}\right\}.$$

Therefore, if $\log R \neq 0$, then by (5) and (6),

$$\log|\log R| < c_3 - c_4\theta^s. \quad (12)$$

On the other hand, we have

$$|\log R| = \left|\log\left(\frac{A_1}{B_1}\right) + m_p \log a - \ell_q \log b\right|,$$

and we may apply Theorem 2.6 to ive a lower bound for $|\log R|$. We take $n = 3$, $d = 1$, and b_1, b_2, b_3 to be $\frac{A_1}{B_1}$, a, and b, respectively, in Theorem 2.6. Note that m_p and ℓ_q are at most $\frac{\log n}{\log 2}$. The height of $\frac{A_1}{B_1}$ is at most the maximum of A_1 and B_1. Then by Theorem 2, if $\log R \neq 0$, then

$$|\log R| \geq \exp(-c_5 \log(4 \max(A_1, B_1)) \log\log n).$$

Now, from (7) and (9),

$$\log|\log R| > -c_6 \max(1, m_1 - m_p, \ell_1 - \ell_q) \log\log n,$$

which, from (5) and (6), yields

$$\log|\log R| > -c_7\theta^{s-1} \log\log n.$$

Comparing this estimate with (12), we find that

$$c_4\theta^s < c_7\theta^{s-1}\log\log n + c_3,$$

hence

$$\theta < c_8 \log\log n + c_9.$$

However, this contradicts (1) if c_1 is chosen to be larger than $c_8 + c_9$. This is possible since c_8 and c_9 are determined independently of c_1. Therefore, the assumption that $\log R \neq 0$ must be false, so we have $\log R = 0$. In particular,

$$\log\frac{A_1}{B_1} + m_p \log a - \ell_q \log b = 0.$$

By Theorem 9.2, there exists a relation of the form

$$x_1 \log\frac{A_1}{B_1} + x_2 \log a + x_3 \log b = 0$$

with integer coefficients x_1, x_2, x_3, not all zero, satisfying

$$\max\{|x_1|, |x_2|, |x_3|\} \leq c_{10} \log(\max(|A|, |B|)).$$

[TO BE CONTINUED NEXT CLASS] □

13 Oct. 4, 2019

We continue the proof from last class. We had just arrived at $\log R = 0$. In particular,

$$\log \frac{A_1}{B_1} + m_p \log a - \ell_q \log b = 0.$$

By Theorem 9.2, there exists a relation

$$x_1 \log \frac{A_1}{B_1} + x_2 \log a + x_3 \log b = 0$$

with x_1, x_2, x_3 integers, not all 0, satisfying

$$\max(|x_1|, |x_2|, |x_3|) \leq c_{10} \log(\max |A_1|, |B_1|).$$

By (5) and (7), $\log |A_1| < c_{11}\theta^{s-1}$. By (6) and (9), $\log B_1 < c_{12}\theta^{s-1}$. Thus, by (2),

$$|x_2| \leq c_{13}\theta^{s-1} \leq c_{13}\theta^{k-1} < \frac{\log n}{4 \log a}$$

for n sufficiently large. Now, by (5),

$$m_p \geq m_1 - \theta^{s-1},$$

and since $m_1 \geq \frac{\log n}{2 \log a}$ and $\theta^{s-1} < \frac{\log n}{4 \log a}$, we see that

$$m_p > \frac{\log n}{4 \log a} > |x_2|.$$

If $x_1 = 0$, then $\frac{\log a}{\log b}$ is rational. If $x_1 \neq 0$, then we may eliminate $\log \frac{A_1}{B_1}$ from the two equations to get

$$(x_1 m_p - x_2) \log a - (x_1 \ell_q - x_3) \log b = 0,$$

and since $m_p > |x_2|$, we see that $m_p - x_2 \neq 0$, hence again $\frac{\log a}{\log b}$ is rational, as required. □

Let p be a prime with $p \geq 3$. Let $1 = n_1 < n_2 < \ldots$ be the increasing sequence of positive integers all of whose prime factors are at most p. In 1898, Størmer proved that $\liminf_{i\to\infty}(n_{i+1} - n_i) > 2$. In 1908, Thue proved that $\lim_{i\to\infty}(n_{i+1} - n_i) = \infty$. In 1965, Erdős gave the following improvement of Thue's result, by means of a theorem of Mahler (which used a p-adic version of Thue's work): "Let $0 < \epsilon < 1$. Then there exists a positive number $N(\epsilon)$ such that for $n_i > N(\epsilon)$, $n_{i+1} - n_i > n_i^{1-\epsilon}$." In 1973, Tijdeman applied estimates for linear forms in logarithms to prove the following theorem.

Theorem 13.1. *Let p be a prime number, and let $n_1 < n_2 < \ldots$ be the sequence of positive integers all of whose prime factors are at most p. There exists a positive number C, which* *is effectively computable in terms of p, such that*

$$n_{i+1} - n_i > \frac{n_i}{(\log n_i)^C} \textit{ for } n_i \geq 3.$$

Tijdeman also showed that there's an upper bound (with a different constant C), so apart from adjusting constants, you cannot do better than Tijdeman's bound.

Proof. Let $p_1, \ldots, p_k$ be the primes of size at most p. Let us consider the prime decomposition of n_i and n_{i+1}:

$$n_{i+1} = p_1^{a_1} \cdots p_k^{a_k}$$

where $a_i \geq 0$ for $i = 1, \ldots, k$, and

$$n_i = p_1^{b_1} \cdots p_k^{b_k}$$

where $b_j \geq 0$ for $j = 1, \ldots, k$. Note that

$$\log \frac{n_{i+1}}{n_i} = (a_1 - b_1) \log p_1 + \cdots + (a_k - b_k) \log p_k.$$

This is a linear form in logarithms, and it is small when n_i and n_{i+1} are near each other. We invoke Theorem 2.6 with $K := \mathbb{Q}$, $d := 1$, and $\alpha_1, \ldots, \alpha_k$ given by $p_1, \ldots, p_k$, respectively, and $n := k$. The coefficients are $(a_1 - b_1), \ldots, (a_k - b_k)$. We can estimate their size as follows:

$$\max_{i=1,\ldots,k} |a_i - b_i| \leq \frac{\log n_{i+1}}{\log n_i} \leq 1 + \frac{\log n_i}{\log 2}.$$

By Theorem 2.6,

$$\left| \log \left(\frac{n_{i+1}}{n_i} \right) \right| > \exp\left(-k^{ck} \log p_1 \cdots \log p_k \log \log n_i \right).$$

Thus, since $\log p_i < c_1 \log k$ for $i = 1, \ldots, k$ by the prime number theorem,

$$\log \left(\frac{n_{i+1}}{n_i} \right) > \exp\left(-k^{c_2 k} \log \log n_i \right).$$

Hence,

$$\log \left(\frac{n_{i+1}}{n_i} \right) > \exp\left(-e^{c_3 p} \log \log n_i \right),$$

so

$$\log \left(\frac{n_{i+1}}{n_i} \right) > \frac{1}{(\log n_i)^{e^{c_3 p}}}.$$

On the other hand,

$$\log \left(\frac{n_{i+1}}{n_i} \right) = \log \left(1 + \frac{n_{i+1} - n_i}{n_i} \right) < \frac{n_{i+1} - n_i}{n_i}.$$

The result follows by comparing the upper and lower bounds established above for $\log(n_{i+1}/n_i)$.

□

14 Oct. 7, 2019

Tijdeman also proved the following theorem.

Theorem 14.1. *Let p and q be distinct primes, and let $1 = n_1 < n_2 < \dots$ be the sequence of positive integers whose prime factors are all p or q. There exist positive numbers c and N_0, which are effectively computable in terms of p and q, such that*

$$n_{i+1} - n_i < \frac{n_i}{(\log n_i)^c}$$

for $n_i > N_0$.

To prove Theorem 14, we need some basic results from Diophantine approximation.

For any real number α, we define a sequence of real numbers $\alpha_0, \alpha_1, \dots$ by putting $\alpha_0 := \alpha$ and

$$\alpha_k := \frac{1}{\alpha_{k+1} - [\alpha_{k-1}]}$$

provided $\alpha_{k-1} - [\alpha_{k-1}] \neq 0$ for $k = 1, 2, \dots$ (where $[\cdot]$ denotes the integer part). Next, we put $a_k := [\alpha_k]$ for $k = 0, 1, 2, \dots$. Then

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{k-1} + \cfrac{1}{\alpha_k}}}}.$$

We put, for $k = 0, 1, 2, \dots$,

$$\frac{p_k}{q_k} := a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_{k-1} + \cfrac{1}{a_k}}}}$$

where $q_k > 0$ and $\gcd(p_k, q_k) = 1$. The numbers a_i are known as the *partial quotients* of α, and the rationals p_k/q_k are known as the *convergents* of α. We have the following facts: If p_k/q_k is a convergent to α with $k \geq 1$, then

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k q_{k+1}} < \frac{1}{q_k^2}.$$

A finite continued fraction

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}$$

is denoted by $[a_0, a_1, \dots, a_n]$. Note that given α, we find p_k/q_k for $k = 1, 2, \dots$ to be given by $p_k/q_k = [a_0, \dots, a_k]$.

We will now show that p_k and q_k are generated by the following recursive rule:

$$p_n = a_n p_{n-1} + p_{n-2}, q_n = a_n q_{n-1} + q_{n-2} \quad (1)$$

for $n \geq 2$, where $p_0 = a_0$, $q_0 = 1$ and $p_1 = a_0 a_1 + 1$, $q_1 = a_1$. We will do so by induction. We check that the result holds for $n = 2$. Assume it holds for $n = k - 1 \geq 2$, and we will prove it for $n = k$. Define coprime integers p'_j and q'_j with $q_j > 0$ for $j = 0, 1, 2, \dots$ by

$$\frac{p'_j}{q'_j} := [a_1, \dots, a_{j+1}]$$

and apply our inductive hypothesis with $j = k - 1$ to get

$$p'_{k-1} = a_k p'_{k-2} + p'_{k-3},$$

$$q'_{k-1} = a_k q'_{k-2} + q'_{k-3}.$$

However,

$$\frac{p_j}{q_j} = a_0 + \frac{q'_{j-1}}{p'_{j-1}} = \frac{a_0 p'_{j-1} + q'_{j-1}}{p'_{j-1}},$$

hence $p_j = a_0 p'_{j-1} + q'_{j-1}$ and $q_j = p'_{j-1}$. Thus, on taking $j = k$, we get

$$\begin{aligned} p_k &= a_0(a_k p'_{k-2} + p'_{k-3}) + a_k q'_{k-2} + q'_{k-3} \\ &= a_k(a_0 p'_{k-2} + q'_{k-2}) + a_0 p'_{k-3} + q'_{k-3} \\ &= a_k p_{k-1} + p_{k-2} \end{aligned}$$

and

$$q_k = p'_{k-1} = a_k p'_{k-2} + p'_{k-3} = a_k q_{k-1} + q_{k-2}$$

as required. By definition of $\alpha_1, \alpha_2, \dots$, we have

$$\alpha = [a_0, a_1, \dots, a_n, \alpha_{n+1}],$$

where $0 < 1/\alpha_{n+1} \leq 1/a_{n+1}$ and α lies between p_n/q_n and p_{n+1}/q_{n+1}.

Proposition 14.2. *With the above notation, we have*

$$p_n q_{n+1} - p_{n+1} q_n = (-1)^{n+1}$$

for $n = 0, 1, 2, \dots$.

Proof. We prove this by induction. For $n = 0$ it holds since

$$p_0 q_1 - p_1 q_0 = a_0 a_1 - (a_1 a_0 + 1) = -1 = (-1)^{0+1}.$$

Assume it holds for $n = k - 1$ and then apply the recurrence relation (1). We get

$$p_k q_{k+1} - p_{k+1} q_k = p_k(a_{k+1} q_k + q_{k-1}) - (a_{k+1} p_k + p_{k-1}) q_k$$

$$= o_k q_{k-1} - p_{k-1} q_k$$
$$= (-1)^{k+1}$$

as required. □

One can show that if

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q^2},$$

then $p/q = p_k/q_k$ for some $k \in \mathbb{N}$. Finally, we have

$$p_0/q_0 < p_2/q_2 < \cdots < \alpha < \cdots < p_5/q_5 < p_3/q_3 < p_1/q_1,$$

and by Proposition 14.2,

$$\frac{p_{k-1}}{q_{k-1}} - \frac{p_k}{q_k} = \frac{(-1)^k}{q_{k-1}q_k}.$$

15 Oct. 9, 2019

The following lemma is due to Tijdeman.

Lemma 15.1. *Let p and q be distinct primes. Let $h_0/k_0, h_1/k_1, \ldots$ be the sequence of convergents to $\log p/\log q$. There exists a positive number c, which is effectively computable in terms of p and q, such that*

$$k_{j+1} < k_j^c \log q \text{ for } j = 2, 3, \ldots.$$

Proof. One has $k_j \geq 2$ for $j \geq 2$. Since

$$\left|\frac{\log p}{\log q} - \frac{h_j}{k_j}\right| < \frac{1}{k_j k_{j+1}} \text{ for } j = 0, 1, 2, \ldots,$$

we obtain

$$|k_j \log p - h_j \log q| < \frac{\log q}{k_{j+1}}. \quad (1)$$

On the other hand, by Theorem 2.6 with $\alpha_1 := p$ and $\alpha_2 := q$, we see that

$$|k_j \log p - h_j \log q| > \exp(-c_1 \log \max(k_j, h_j)).$$

Here $c_1, c_2, \ldots$ are positive numbers that are effectively computable in terms of p and q (although we have not used c_i for $i \neq 1$ yet!). Since h_j/k_j is approximately $\log p/\log q$, we see that

$$|k_j \log p - h_j \log q| > \exp(-c_2 \log k_j) = k_j^{-c_2} \quad (2)$$

Our result then follows from (1) and (2). □

We are now ready to prove Theorem 14.1.

Proof. Put $n := n_i =: p^u q^v$ where u and v are non-negative integers. We assume, without loss of generality, that $p^u \geq \sqrt{n}$, so

$$u \geq \frac{\log n}{2\log p}. \quad (*)$$

Let $h_0/k_0, h_1/k_1, \ldots$ be the sequence of convergents to $\log p/\log q$. Then $k_1 < k_2 < \ldots$, and so for some index j we have $k_j \leq u < k_{j+1}$. We may suppose that $n \geq 3$ and $j \geq 2$. There are two cases to consider depending on whether h_j/k_j is larger than $\log p/\log q$ or it is not.

Case 1: $\frac{h_j}{k_j} > \frac{\log p}{\log q}$.

Put $n' := p^{u-k_j} q^{v+h_j}$, and note that $n' \in \mathbb{Z}$ and $n' > n$. We have

$$\frac{h_j}{k_j} - \frac{\log p}{\log q} < \frac{1}{k_j k_{j+1}}.$$

Thus,

$$\log \frac{n'}{n} = \log \frac{q^{h_j}}{p^{k_j}} = h_j \log q - k_j \log p < \frac{\log q}{k_{j+1}}.$$

Recall that

$$k_{j+1} > u \geq \frac{\log n}{2\log p}.$$

Thus,

$$\log \frac{n'}{n} < \frac{2\log p \log q}{\log a}, \quad (1)$$

so for n sufficiently large in terms of p and q,

$$\frac{n'}{n} - 1 < \exp\left(\frac{2\log p \log q}{\log n}\right) - 1 < \frac{1}{2}.$$

For $x \in \mathbb{R}^+$,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots > x - \frac{x^2}{2}$$

and

$$x - \frac{x^2}{2} > \frac{x}{2} \text{ for } 0 < x < \frac{1}{2}.$$

Thus,

$$\log \frac{n'}{n} = \log\left(1 + \left(\frac{n'}{n} - 1\right)\right) > \frac{1}{2}\left(\frac{n'}{n} - 1\right)$$

for n sufficiently large. Therefore, by (1),

$$\frac{n'}{n} < \frac{4\log p \log q}{\log n},$$

so $n' < n + c_1 \frac{n}{\log n}$, where $c_1, c_2, \ldots$ are positive numbers that are effectively computable in terms of p and q. Since $n_{i+1} \leq n'$,

$$n_{i+1} < n_i + c_1 \frac{n_i}{\log n_i}.$$

Case 2: $\frac{h_j}{k_j} < \frac{\log p}{\log q}$.

Then

$$\frac{h_{j-1}}{k_{j-1}} > \frac{\log p}{\log q}.$$

Put $n' := p^{u-k_{j-1}} q^{v+h_{j-1}}$. Again $n' \in \mathbb{Z}$ and $n' > n$, so

$$n' \geq n_{i+1}.$$

We have

$$\frac{h_{j-1}}{k_{j-1}} - \frac{\log p}{\log q} < \frac{1}{k_{j-1} k_j}.$$

Therefore,

$$\log \frac{n'}{n} \log \frac{q^{h_{j-1}}}{p^{k_{j-1}}} = h_{j-1} \log q - k_{j-1} \log p < \frac{\log q}{k_j}.$$

We find from Lemma 16 that

$$k_j > \left(\frac{k_{j+1}}{\log q} \right)^{1/c}.$$

Since $u \geq \frac{\log n}{2 \log p}$ and since $k_{j+1} > u$, we see that

$$\log \frac{n'}{n} < \frac{\log q}{k_j} < \frac{(\log q)^{1+1/c}}{(k_{j+1})^{1/c}} < \frac{(2 \log p)^{1/c} (\log q)^{1+1/c}}{(\log n)^{1/c}}. \quad (2)$$

Thus for n sufficiently large,

$$\log \frac{n'}{n} > \frac{1}{2} \left(\frac{n'}{n} - 1 \right) \quad (3)$$

and so from (2) and (3),

$$n' < n + c_2 \frac{n}{(\log n)^{1/c}}.$$

Since $n' \geq n_{i+1}$,

$$n_{i+1} < n_i + c_2 \frac{n_i}{(\log n_i)^{1/c}} < n_i + \frac{n_i}{(\log n_i)^{c_4}},$$

as required. □

16 Oct. 11, 2019

Today we give some background information about algebraic number theory.

Let K be a finite extension of $\mathbb{Q}$. Then there exists a monic polynomial $f \in \mathbb{Q}[x]$ that is irreducible and for which K is given by $\mathbb{Q}[x]/f\mathbb{Q}[x]$.

Suppose that f is of degree n and factors over $\mathbb{C}$ as $f(x) = (x-\alpha_1)\cdots(x-\alpha_n)$. By the primitive element theorem, there exists an element θ such that $K = \mathbb{Q}(\theta)$. Note that $\alpha_1, \ldots, \alpha_n$ are distinct since f is irreducible over $\mathbb{Q}$.

There are n distinct embeddings of K in $\mathbb{C}$. They are given by mapping θ to α_i for $i = 1, 2, \ldots, n$ and fixing elements of $\mathbb{Q}$. The images $\mathbb{Q}(\alpha_1), \ldots, \mathbb{Q}(\alpha_n)$ are conjugate fields in $\mathbb{C}$. We may order the α_i so that $\alpha_1, \ldots, \alpha_{r_1}$ are real and $\alpha_{r_1+1}, \ldots, \alpha_n$ are not real. Then we have

$$\alpha_{r_1+i} = \overline{\alpha_{r_1+r_2+i}}$$

for $i = 1, 2, \ldots, r_2$ where $r_1 + 2r_2 = n$.

The ring of algebraic integers of K, denoted by O_K, consists of the elements of K which are roots of a monic polynomial with integer coefficients.

Example 16.1. We have $O_{\mathbb{Q}} = \mathbb{Z}$.

Definition 16.2. An integral domain O is said to be a *Dedekind domain* if the following three conditions are satisfied:

(i) O is a Noetherian ring.
(ii) O is integrally closed in its field of fractions.
(iii) All non-zero prime ideals of O are maximal ideals.

We will not prove the following fact.

Proposition 16.3. *The ring of algebraic integers of K when K is a finite extension of $\mathbb{Q}$ is a Dedekind domain.*

In a Dedekind domain, we have unique factorization into prime ideals. However, we need not have unique factorization into irreducible elements of O_K. (There was a failed attempt to prove FLT in the 1800s that assumed this.)

Let K be a finite extension of $\mathbb{Q}$ with ring of algebraic integers O_K. There exist elements $w_1, \ldots, w_n$ such that each element in O_K can be written as an integral linear combination of $w_1, \ldots, w_n$. Further, the representation is unique. The set $\{w_1, \ldots, w_n\}$ is known as an *integral basis* for O_K. Any two integral bases are related by a matrix of determinant ± 1. The *discriminant* D of K is defined by

$$D := \det((\sigma_i(w_j))_{i,j})^2$$

where $\sigma_1, \ldots, \sigma_n$ are the embeddings of K into $\mathbb{C}$. Notice that the $\sigma_i(w_j)$ are algebraic integers and D is the same as $\sigma_i(D)$ for $i = 1, \ldots, n$, so D is an integer and $D \neq 0$.

Consider the set S of non-zero ideals of O_K. We define a relation $\sim$ on S by saying that $a \sim b$ if there exist $\alpha, \beta \in O_K$ with $\alpha\beta \neq 0$ such that

$$a(\alpha) = b(\beta),$$

where (α) is the principal ideal in O_K generated by α and similarly for (β). Then $\sim$ is an equivalence relation. If $a \sim (1)$, then a is a principal ideal. We may define a multiplication on equivalence classes by multiplication on representatives of the classes, and this is well-defined. This turns the set of equivalence classes into a group. The group is a finite abelian group called the *ideal class group* of K. The order of the group is called the *class number* of K and is denoted by h_K or just h when K is understood. In particular, a^h is a principal ideal whenever a is a non-zero ideal of O_K. The class number h measures, in some sense, how far O_K is from having unique factorization (i.e., being a UFD).

17 Oct. 21, 2019

Last week was reading week.

Let K be a finite extension of $\mathbb{Q}$, and let O_K denote the ring of algebraic integers. Let $U(K)$ denote the group of units (i.e., invertible elements) of O_K. Note that $U(K)$ contains all roots of unity in K.

In 1846, Dirichlet proved that $U(K)$ is a finitely-generated abelian gorup of rank $r_1 + r_2 - 1$ where r_1 is the number of real embeddings of K in $\mathbb{C}$ and $r - 2$ is the number of pairs of complex embeddings. In particular, $U(K)$ is isomorphic (under addition, not multiplication, as Jason points out) to

$$\mu(K) \times \mathbb{Z}^r$$

where $\mu(K)$ is a finite torsion subgroup of $U(K)$ containing the roots of unity in K and $r := r_1 + r_2 - 1$. As always, let

$$\sigma_1, \ldots, \sigma_{r_1}$$

be the real embeddings and

$$\sigma_{r_1+1}, \ldots, \sigma_{r_1+r_2}, \overline{\sigma_{r_1+1}}, \ldots, \overline{\sigma_{r_1+r_2}}$$

be the complex embeddings. By Dirichlet's theorem, there exist units $u_1, \ldots, u_r$ such that if x is a unit in O_K, then there exists a root of unity ζ and integers $a_1, \ldots, a_n$ such that

$$x = \zeta u_1^{a_1} \cdots u_r^{a_r}.$$

In this case, $\{u_1, \ldots, u_r\}$ is known as a *fundamental system of units*. In general, it is not unique. However, we can attach a unique volume to a system of fundamental units.

We define the logarithmic embedding L of $K^\times$ into $\mathbb{R}^{r_1+r_2}$ given by

$$L(x) := (\log|\sigma_1(x)|, \ldots, \log|\sigma_{r_1}(x)|, 2\log|\sigma_{r_1+1}(x)|, \ldots, 2\log|\sigma_{r_1+r_2}(x)|).$$

Then L is an abelian group homomorphism. Furthermore, for any $\alpha \in K$ we define the *norm* from K to $\mathbb{Q}$ of α by

$$N_{K/\mathbb{Q}}(\alpha) := \prod_{i=1}^{d} \sigma_i(\alpha)$$

where $d := r_1 + 2r_2$. The norm is multiplicative, and the norm of an algebraic integer is an integer (a *rational* integer, i.e., an element of $\mathbb{Z}$). The norm of a unit is ±1. Therefore, $L : U(K) \to H$, where H is the hyperplane in $\mathbb{R}^{r_1+r_2}$ given by

$$H := \{(x_1, \ldots, x_{r_1+r_2}) \in \mathbb{R}^{r_1+r_2} \mid x_1 + \cdots + x_{r_1+r_2} = 0\}.$$

Furthermore, the image of $U(K)$ under L is a lattice in H, and the kernel is $\mu(K)$. The volume of a fundamental region of the lattice is called the *regulator* of K and is denoted by R_K. The regular does not depend on the choice of fundamental units.

By definition, we have

$$R_K = |\det(\delta_j \log|\sigma_j(u_i)|)_{i,j=1,\ldots,r}|$$

where

$$\delta_j := \begin{cases} 1 & \text{if } 1 \le j \le r_1, \\ 2 & \text{if } r_1 < j < r_1 + r_2. \end{cases}$$

In 1989, Friedman proved that

$$R_K > 0.2052.$$

In 1918, Landau proved that there is a positive number C, which depends on $d = r_1 + 2r_2$, such that

$$h_K R_K < C|D|^{1/2}(\log|D|)^{d-1}.$$

Furthermore, there exists a fundamental system of units $\{u_1, \ldots, u_r\}$ such that

$$\max_{1 \le i \le r} |\log|u_j^{(i)}|| < C(d) R_K$$

where $C(d)$ is a positive number which depends on d. The norm of an ideal I in O_K is defined by $N(I) := |O_K/I|$. For any $\alpha \in O_K$, we let (α) denote the principal ideal generated by α. We have $N((\alpha)) = |N_{K/\mathbb{Q}}(\alpha)|$.

For any finite extension $K/\mathbb{Q}$, we define the *Dedekind zeta function* $\zeta_K(s)$ for $s \in \mathbb{C}$, $\mathrm{Re}(s) > 1$, by

$$\zeta_K(s) := \sum_a \frac{1}{Na^s}$$

where the sum is over all non-zero ideals a of O_K. Just as for the Riemann zeta function, we have an Euler product given by

$$\zeta_K(s) = \prod_p \frac{1}{1 - \frac{1}{Np^s}}$$

where the product is over all prime ideals p of O_K. The function $\zeta_K(s)$ can be analytically continued to all of $\mathbb{C}$ except for $s = 1$ where there is a simple pole. There is a functional equation relating $\zeta_K(s)$ to $\zeta_K(1-s)$, and there is a generalized Riemann hypothesis stating that all of the zeroes of $\zeta_K(s)$ with $0 \leq \mathrm{Re}(s) \leq 1$ have $\mathrm{Re}(s) = \frac{1}{2}$.

Let $w(K)$ be the number of roots of unity in K. Then the residue at $s = 1$ of $\zeta_K(s)$ is significant. We have

$$\lim_{s \to 1} (s-1)\zeta_K(s) = 2^{r_1}(2\pi)^{r-2} \frac{h(K)R_K}{w(K)\sqrt{|D|}}.$$

18 Oct. 23, 2019

Let $F(x,y) = a_n x^n + \cdots + a_1 x y^{n-1} + a_0 y^n$ be a binary form with integer coefficients, $n \geq 3$, and with non-zero discriminant. Let m be a non-zero integer. The equation

$$F(x,y) = m \quad (*)$$

in integers x and y is known as a *Thue equation.* It is given this name because in 1909, Thue proved that (*) has only finitely many solutions in integers x and y. His proof was ineffective in that it did not yield an upper bound for the size of solutions. For example,

$$x^3 - 2y^3 = 6$$

has only one solution $(x,y) = (2,1)$.

In 1968, Baker, by means of estimates for linear forms in logarithms of algebraic numbers, gave a method for finding all solutions of Thue equations. This follows from the following theorem.

Theorem 18.1. *Let $F \in \mathbb{Z}[x,y]$ be an irreducible binary form of degree at least 3, and let m be a non-zero integer. All solutions in integers (x,y) to $F(x,y) = m$ satisfy*

$$\max(|x|,|y|) < |2m|^{C \log\log|3m|}$$

where $C > 0$ is effectively computable in terms of F.

Proof. Let $C_1, C_2, \dots$ denote positive numbers which are effectively computable in terms of F. Let

$$F(x,y); = a_n x^n + \cdots + a_1 x y^{n-1} + a_0 y^n.$$

We may suppose without loss of generality that $a_n = 1$ since if not we replace $F(X,y)$ by $a_n^{n-1}F(x,y) = F(X,y)$ where $X := a_n x$ and replace m by $a_n^{n-1}m$.

We write

$$F(x,1) = (x-\alpha^{(1)})\cdots(x-\alpha^{(n)})$$

where $\alpha^{(1)},\dots,\alpha^{(n)}$ are real and $\alpha^{(r_1+i)} = \overline{\alpha^{(r_1+r_2+i)}}$ with $r_1+2r_2=n$. Then if $F(x,y)=m$, we have

$$(x-\alpha^{(1)}y)\cdots(x-\alpha^{(n)}y) = m.$$

Put

$$\beta^{(i)} := x - \alpha^{(i)}y$$

for $i=1,\dots,n$. Let $K := \mathbb{Q}(\alpha^{(1)})$, and let $\eta_1,\dots,\eta_r$ be a fundamental syste mof units for K chosen so that

$$|\log|\eta_i^{(j)}|| < C_1.$$

Every point P in $\mathbb{R}^r$ is within C_2 of some point of the lattice with basis

$$(\log|\eta_i^{(1)}|,\dots,\log|\eta_i^{(r)}|)$$

for $i=1,\dots,r$. Take

$$P := (\log||m|^{-1/n}\beta^{(1)}|,\dots,\log||m|^{-1/n}\beta^{(r)}|).$$

Then we see that there are integers $b_1,\dots,b_r$ such that

$$|b_1\log|\eta_1^{j)}| + \cdots + b_r\log|\eta_r^{(j)}| + \log||m|^{-1/n}\beta^{(j)}|| < C_2$$

for $j=1,\dots,r$. Thus, if we set

$$\gamma^{(j)} := \beta^{(j)}(\eta_1^{(j)})^{b_1}\cdots(\eta_r^{(j)})^{b_r} \quad (**)$$

for $j=1,\dots,n$, then

$$\left|\log\left(|m|^{-1/n}|\gamma^{(j)}|\right)\right| < C_2 \quad (1)$$

for $j=1,\dots,r$. Since $|\gamma^{(r_1+i)}| = |\gamma^{(r_1+r_2+i}|$ for $i=1,\dots,r_2$, we see that (1) holds for $j=1,\dots,n$ with the exception of $j=n$ when $r_2=0$ and of $j=r_1+r_2$ and $j=r_1+2r_2$ when $r_2\neq 0$. But

$$|\gamma^{(1)}\cdots\gamma^{(n)}| = |m|,$$

so

$$\sum_{j=1}^{n}\log||m|^{-1/n}|\gamma^{(j)}|| = 0.$$

Therefore, we see that (1) holds for $j=1,\dots,n$. Observe that $\gamma=\gamma^{(1)}$ is an algebraic integer and is a root of a monic polynomial whose coefficients are elementary symmetric polynomials in $\gamma^{(1)},\dots,\gamma^{(n)}$ and so have size at most $C_3|m|$ in absolute value. This proof will be continued next class!

19 Oct. 25, 2019

I was away for convocation. The following is basically this portion of the proof from Dan Wolczuk's notes.

Taking logarithms of (**), we obtain

$$b_1 \log|\eta_1^{(j)}| + \cdots + b_r \log|\eta_r^{(j)}| = \log\left|\frac{\gamma^{(j)}}{\beta^{(j)}}\right|$$

for $j = 1, \ldots, r$. By Cramer's rule, for $i = 1, \ldots, r$ we have

$$b_i = \frac{\det\begin{pmatrix} \log|\eta_1^{(1)}| & \cdots & \log|\frac{\gamma^{(1)}}{\beta^{(1)}}| & \cdots & \log|\eta_r^{(1)}| \\ \vdots & & \vdots & & \vdots \\ \log|\eta_1^{(r)}| & \cdots & \log|\frac{\gamma^{(r)}}{\beta^{(r)}} & \cdots & \log|\eta_r^{(r)}| \end{pmatrix}}{\det\begin{pmatrix} \log|\eta_1^{(1)}| & \cdots & \log|\eta_r^{(1)}| \\ \vdots & & \vdots \\ \log|\eta_1^{(r)}| & \cdots & \log|\eta_r^{(r)}| \end{pmatrix}} \quad (2).$$

We want to bound $|b_i|$ in terms of m, hence to bound $|x|$ and $|y|$ in terms of m. Let $B := \max(|b_1|, \ldots, |b_r|)$ and suppose $|b_i| = B$. Let Δ be the denominator of (2). Then

$$\Delta = R_K 2^{-r_2}.$$

Expanding the numerator of (2) along the i^{th} column, we get

$$\max_{1 \le j \le r} \left|\log\left|\frac{\gamma^{(j)}}{\beta^{(j)}}\right|\right| > C_4 B.$$

Suppose the maximum occurs for $j = J$. Then

$$\begin{aligned} |\log|m|^{-1/n}\log|\beta^{(J)}|| &= \left|\log\frac{|\beta^{(J)}}{\gamma^{(J)}} + \log|m|^{-1/n}|\gamma^{(J)}|\right| \\ &> C_4 B - C_2. \end{aligned}$$

We have

$$\sum_{j=1}^{n} \log|m|^{-1/n}|\beta^{(j)}| = 0,$$

so it follows that for some ℓ with $1 \le \ell \le n$,

$$\log(|m|^{-1/n}|\beta^{(\ell)}|) < \frac{C_2 - C_4 B}{n-1}.$$

In particular,

$$|\beta^{(\ell)}| < |m|^{1/n} C_6 e^{-C_5 B}.$$

Now, since $|\beta^{(1)} \cdots \beta^{(n)}| = |m|$, there exists an integer k with $k \neq \ell$ such that

$$|\beta^{(k)} > |m|^{1/n} C_6^{-1/(n-1)} e^{C_5 B/(n-1)}. \quad (4).$$

Let j be an integer satisfying $1 \leq j \leq n$, $j \neq k$, and $j \neq \ell$. (Since $n \geq 3$, such a j exists.) Then we have

$$(\alpha^{(k)} - \alpha^{(\ell)})\beta^{(j)} - (\alpha^{(j)} - \alpha^{(\ell)})\beta^{(k)} = (\alpha^{(k)} - \alpha^{(j)})\beta^{(\ell)}.$$

Note that

$$(\alpha^{(k)} - \alpha^{(\ell)})(x - \alpha^{(j)}y) - (\alpha^{(j)} - \alpha^{(\ell)})(x - \alpha^{(k)}y) = (\alpha^{(k)} - \alpha^{(j)})(x - \alpha^{(\ell)}y).$$

The coefficient of y in this expression is

$$-(\alpha^{(k)} - \alpha^{(\ell)})\alpha^{(j)} + (\alpha^{(j)} - \alpha^{(\ell)})\alpha^{(k)} = -(\alpha^{(\ell)}\alpha^{(k)} - \alpha^{(\ell)}\alpha^{(j)})$$

$$= -(\alpha^{(k)} - \alpha^{(j)})\alpha^{(\ell)}.$$

Dividing by $(\alpha^{(k)} - \alpha^{(\ell)})\beta^{(k)}\gamma^{(j)}/\gamma^{(k)}$, we get

$$\frac{\beta^{(j)}\gamma^{(k)}}{\gamma^{(j)}\beta^{(k)}} - \frac{\alpha^{(j)} - \alpha^{(\ell)}\gamma^{(k)}}{\alpha^{(k)} - \alpha^{(\ell)}\gamma^{(j)}} = \frac{\alpha^{(k)} - \alpha^{(j)}}{\alpha^{(k)} - \alpha^{(\ell)}} \frac{\beta^{(\ell)}\gamma^{(k)}}{\beta^{(k)}\gamma^{(j)}}.$$

Thus,

$$\left(\frac{\eta_1^{(k)}}{\eta_1^{(j)}}\right)^{b_1} \cdots \left(\frac{\eta_r^{(k)}}{\eta_r^{(j)}}\right)^{b_r} - \alpha_{r+1} = \lambda$$

where

$$\alpha_{r+1} := \frac{\alpha^{(j)} - \alpha^{(\ell)}}{\alpha^{(k)} - \alpha^{(\ell)}} \frac{\gamma^{(k)}}{\gamma^{(j)}}$$

and

$$\lambda := \left(\frac{\alpha^{(k)} - \alpha^{(j)}}{\alpha^{(k)} - \alpha^{(\ell)}}\right) \frac{\beta^{(\ell)}\gamma^{(k)}}{\beta^{(k)}\gamma^{(j)}}.$$

Let $\alpha_i := \eta_i^{(k)}/\eta_i^{(j)}$. Then

$$\alpha_1^{b_1} \cdots \alpha_r^{b_r} \alpha_{r+1}^{-1} = 1 + \lambda.$$

20 Oct. 28, 2019

Today we finish the proof of the upper bound on solutions to the Thue equation. In what follows, when you see log, you can assume we are taking the principal value of the logarithm.

Recall that B is the maximum of $|b_1|, \ldots, |b_n|$. Note that if the maximum occurs for $|b_i|$, then we may suppose that b_i is positive by replacing the unit η_i by η_i^{-1}. We still have a fundamental system of units after making this replacement, so the regulator does not change.

Now,

$$\log(\alpha_1^{b_1}\cdots\alpha_r^{b_r}\alpha_{r+1}^{-1}) = \log\left(1+\frac{\lambda}{\alpha_{r+1}}\right).$$

Set

$$\Lambda := \log(\alpha_1^{b_1}\cdots\alpha_r^{b_r}\alpha_{r+1}^{-1}) = b_1\log\alpha_1+\cdots+b_r\log\alpha_r-\log\alpha_{r+1}+b_{r+2}\log(-1),$$

where b_{r+2} is chosen so that all the logarithms take their principal value. Thus,

$$|b_{r+2}| \le |b_1|+\cdots+|b_r|+1+1 \le (r+2)B.$$

We put $K := \mathbb{Q}(\alpha^{(j)},\alpha^{(k)},\alpha^{(\ell)})$ and $d := [K:\mathbb{Q}]$. Then $d \le n^3$. We may apply Thm. 2 since $\lambda \neq 0$, hence $\Lambda \neq 0$, to get

$$|\Lambda| > \exp(-c_6\log A_{r+1}\log((r+2)B))$$

where $A_{r+1} := \max(H(\alpha_{r+1}), e)$. Put

$$\theta_1 := \frac{\alpha^{(j)}-\alpha^{(\ell)}}{\alpha^{(k)}-\alpha^{(\ell)}},$$

$$\theta_2 := \gamma^{(k)}, \text{ and}$$

$$\theta_3 := 1/\gamma^{(j)},$$

so $\alpha_{r+1} = \theta_1\theta_2\theta_3$. Then

$$H(\alpha_{r+1}) = H(\theta_1\theta_2\theta_3) \le 2^d M(\theta_1\theta_2\theta_3)$$

$$= (2h(\theta_1\theta_2\theta_3))^d \le (2h(\theta_1)h(\theta_2)h(\theta_3))^d$$

$$\le (2M(\theta_1)M(\theta_2)M(\theta_3)^{-1})^d$$

$$\le (2d^{3/2}H(\theta_1)H(\theta_2)H(\theta_3^{-1}))^d$$

by Proposition 4.2,

$$\le (C_7m^2)^d.$$

Thus,

$$\log A_{r+1} \le C_8\log 2|m|.$$

Thus, by (6), we have

$$\log|\Lambda| > -c_9\log 2|m|\log B. \quad (7)$$

But

$$|\Lambda| = |\log(1+\lambda/(\alpha_{r+1}))| = |\log\left(1+\frac{\alpha^{(k)}-\alpha^{(j)}}{\alpha^{(j)}-\alpha^{(\ell)}}\right)\frac{\beta^{(\ell)}\beta^{(k)}}{|}.$$

By (4) and (5),

$$\left|\frac{\alpha^{(k)}-\alpha^{(j)}}{\alpha^{(j)}-\alpha^{(\ell)}}\frac{\beta^{(\ell)}}{\beta^{(k)}}\right| \leq c_{10}e^{-c_{11}B}.$$

Thus, there exists c_{12} such that either $B < c_{12}$ or $c_{10}e^{-c_{11}B} < \frac{1}{2}$. Since $|\log(1+z)| < 2|z|$ for $|z| < 1/2$ and $z \in \mathbb{C}$, we see that

$$|\Lambda| < 2c_{10}e^{-c_{11}B},$$

hence

$$\log|\Lambda| < c_{13} - c_{11}B.$$

Therefore, $B < c_{14}$ or $\log|\Lambda| < -c_{15}B$. Thus, by (7),

$$\frac{B}{\log B} < c_{16}\log|2m|.$$

Therefore,

$$B < c_{17}\log|2m|\log\log|4m|.$$

But now

$$x = \frac{\alpha^{(2)}\beta^{(1)}-\alpha^{(1)}\beta^{(2)}}{\alpha^{(2)}-\alpha^{(1)}},$$

$$y = \frac{\beta^{(1)}-\beta^{(2)}}{\alpha^{(2)}-\alpha^{(1)}},$$

so

$$\max(|x|,|y|) < c_{18}\max(|\beta^{(1)}|,|\beta^{(2)}|).$$

We have

$$|\beta^{(i)}| = |\gamma^{(i)}(\eta_1^{(i)})^{-b_1}\cdots(\eta_r^{(i)})^{-b_r}|$$

for $i = 1, 2$, hence

$$\max(|x|,|y|) < c_{19}|m|^{1/n}e^{c_{20}B} < c_{21}|2m|^{c_{22}\log\log|4m|}.\Box$$

Since Baker's original proof, the upper bound on $\max(|x|,|y|)$ has been improved to $|2m|^C$ by work of Feldman. Feldman's invented the *Feldman polynomials* to this end.

In fact, one can prove that if K is a finite extension of $\mathbb{Q}$; if F is a binary form of degree at least 3, non-zero discriminant, and coefficients in O_K; and if μ is a non-zero element of O_K, then the Diophantine equation

$$F(x,y) = \mu$$

has only finitely many solutions in elements $x, y \in O_K$. Furthermore,

$$\max(\overline{|x|},\overline{|y|}) < C$$

where C is a positive number depending only on F, K, and μ. (This holds even if F is not irreducible.)

21 Oct. 30, 2019

We are given the list of essay topics for the course and are told to sign into PMATH 940 on Learn, click on "Connect > Groups > View Available Groups", and then join the group you want by number. Cam is not looking for a verbatim repetition of the paper, but a 10-page article discussing the paper, key arguments in it, and related work. Each talk will be one hour long.

Let $f \in \mathbb{Z}[x]$ be a polynomial of degree $d \geq 2$. Let n be a positive integer with $m \geq 3$. We can consider the superelliptic equation

$$y^m = f(x)$$

in integers x and y. Suppose $f(x) = (x - \alpha_1) \cdots (x - \alpha_d)$. If a factor of y divides $x - \alpha_i$ and $x - \alpha_j$, then it divides the difference $\alpha_i - \alpha_j$, and roughly speaking, each $x - \alpha_i$ should be "almost an m^{th} power". By the theory of the generalized Thue equation, we should therefore get finitely many solutions.

To make the argument rigorous, we need the following facts from algebraic number theory:

If K is a finite extension of $\mathbb{Q}$, there is a fundamental system of units of O_K, say $\eta_1, \ldots, \eta_r$, such that

$$\max_{i,j=1,\ldots,r} |\log|\eta_j^{(i)}|| < CR \quad \text{(I)}$$

where R is the regulator and C is a positive number that depends on $[K : \mathbb{Q}]$.

When we have such a fundamental system, then every unit η in $U(K)$ can be written as

$$\eta = \eta^1 \eta_1^{b_1} \cdots \eta_r^{b_r}$$

where η^1 is a root of unity, $b_1, \ldots, b_r$ are integers, and

$$\overline{|\eta^1|} = 1. \quad \text{(II)}$$

Next, suppose that α is a non-zero element of O_K for which $|N_{K/\mathbb{Q}}(\alpha)| \leq M$. Then there exists a positive number C_2, which is effectively computable in terms of d, R, and M, and a unit ϵ such that

$$\overline{|\epsilon\alpha|} < C_2. \quad \text{(III)}$$

Finally, let $\mathfrak{a}$ be a non-zero ideal of O_K. There exists a non-zero ideal $\mathfrak{b}$ such that $\mathfrak{a}\mathfrak{b}$ is a principal ideal and

$$N(\mathfrak{b}) \leq \sqrt{|D|} \quad \text{(IV)}$$

where D is the discriminant.

Theorem 21.1. *Let f be a monic polynomial with integer coefficients and at least 2 simple roots, and let m be an integer with $m \geq 3$. There exists a positive number C, which is* *effectively computable in terms of f and m, such that if x and y are integers with*

$$y^m = f(x),$$

then

$$\max(|x|, |y|) < C.$$

To prove this result, we need the following lemma.

Lemma 21.2. *Let $m \geq 2$. Let f be a monic polynomial with integer coefficients and a simple root α. Let K be the splitting field of f. If x and y are integers for which*

$$y^m = f(x),$$

then there exist algebraic integers γ, ϕ, and δ such that

$$x - \alpha = \left(\frac{\gamma}{\phi}\right)\delta^m$$

and $\max(\overline{|\gamma|}, \overline{|\phi|}) < C$ *where C is a positive number depending only on m and f.*

Proof. Let $f(x) =: (x - \alpha_1)\cdots(x - \alpha_n)$, and suppose that $\alpha := \alpha_1$. Let $C_1, C_2, \ldots$ denote positive numbers that are effectively computable in terms of m and f. Let $\eta_1, \ldots, \eta_r$ be a fundamental system of units satisfying (I). Put

$$\Delta := \prod_{i=2}^{n} [\alpha - \alpha_i]$$

where $[\alpha - \alpha_i]$ is the principal ideal generated by $\alpha - \alpha_i$ for $i = 2, \ldots, n$. Since α is a simple root, Δ is a non-zero ideal.

Suppose that x and y are integers for which $y^m = f(x)$. If $x = \alpha$, the result holds with $\delta = 0$ and $\gamma = \phi = 1$. Suppose now that $x \neq \alpha$, and consider the ideal equation

$$[y]^m = [x - \alpha][x - \alpha_2]\cdots[x - \alpha_n].$$

Let $\mathfrak{p}$ be a prime ideal which divides $[x - \alpha]$. Let ℓ_i be the exact power of $\mathfrak{p}$ dividing $[x - \alpha_i]$ for $i = 1, \ldots, n$.

We continue this proof next class.

22 Nov. 1, 2019

We continue the proof from last class.

Recall that $\Delta = \prod_{i=2}^{n}[\alpha - \alpha_i]$. We considered the ideal equation

$$[y]^m = [x - \alpha_1]\cdots[x - \alpha_n]. \quad (*)$$

For any prime ideal $\mathfrak{p}$, let $\mathfrak{p}^{\ell_i}$ be the exact power of $\mathfrak{p}$ dividing $[x-\alpha_i]$, $i = 1, \ldots, n$.

First, suppose that $\ell_1 \geq \ell_i$ for $i = 2, \ldots, n$. Then

$$\mathfrak{p}^{\ell_i} \mid [x-\alpha] - [x-\alpha_i] \supseteq [\alpha - \alpha_i]$$

for $i = 2, \ldots, n$. Thus,

$$\mathfrak{p}^{\ell_2+\cdots+\ell_n} \mid \Delta.$$

Therefore by (*),

$$\ell_1 + \ell_2 + \cdots + \ell_n \equiv 0 \pmod{m},$$

hence

$$\ell_1 \equiv -(\ell_2 + \cdots + \ell_n) \equiv (m-1)(\ell_2 + \cdots + \ell_n) \pmod{m}.$$

Next, suppose that $\ell_1 < \ell_j$ for some j with $2 \leq j \leq n$. Then

$$\mathfrak{p}^{\ell_1} \mid [\alpha - \alpha_j],$$

so $\mathfrak{p}^{\ell_1} \mid \Delta$. In both cases, the exponent of $\mathfrak{p}$ dividing $[x-\alpha]$ is congruent to $a \pmod{m}$ where a is at most $(m-1)$ times the power of $\mathfrak{p}$ dividing Δ.

In particular,

$$[x-\alpha] = \mathfrak{a}\mathfrak{b}^m \quad (1)$$

where $\mathfrak{a}$ and $\mathfrak{b}$ are non-zero ideals in O_K and $\mathfrak{a} \mid \Delta^{m-1}$. By (IV) from last class, there eixst non-zero ideals $\mathfrak{a}_1$ and $\mathfrak{b}_1$ with

$$\max(N(\mathfrak{a}_1), N(\mathfrak{b}_1)) < C_1$$

for which the ideals $\mathfrak{a}\mathfrak{a}_1$ and $\mathfrak{b}\mathfrak{b}_1$ are principal, say $\mathfrak{a}\mathfrak{a}_1 = [\gamma_1]$ and $\mathfrak{b}\mathfrak{b}_1 = [\delta_1]$. Multiplying both sides of (1) by $\mathfrak{a}_1\mathfrak{b}_1^m$, we get

$$\mathfrak{a}_1\mathfrak{b}_1^m[x-\alpha] = [\gamma_1][\delta_1]^m.$$

Notice that $\mathfrak{a}_1\mathfrak{b}_1^m$ is a principal ideal, say $\mathfrak{a}_1\mathfrak{b}_1^m = [\phi_1]$. Furthermore,

$$N([\phi_1]) = N(\mathfrak{a}_1)N(\mathfrak{b}_1)^m < C_2.$$

Also,

$$N([\gamma_1]) = N(\mathfrak{a})N(\mathfrak{a}_1) \leq N(\Delta)^{m-1}N(\mathfrak{a}_1) < C_3.$$

By (III) from last class, we can find associates γ_2 and ϕ_2 of γ_1 and ϕ_1, respectively, such that

$$\max(\overline{|\gamma_2|}, \overline{|\phi_2|}) < C_4.$$

Therefore, since

$$\mathfrak{a}_1\mathfrak{b}_1^m[x-\alpha] = \mathfrak{a}\mathfrak{a}_1(\mathfrak{b}\mathfrak{b}_1)^m,$$

we deduce that

$$x - \alpha = \epsilon(\gamma_2/\phi_2)\delta_1^m$$

where ϵ is a unit in O_K. By (I) and (II) from last class, we can write ϵ as

$$\epsilon_1\epsilon_2^m$$

where ϵ_1 and ϵ_2 are units and

$$\overline{|\epsilon_1|} < C_5.$$

Thus,

$$x - \alpha = \left(\frac{\epsilon_1\gamma_2}{\phi_2}\right)(\epsilon_2\delta_1)^m.$$

Set $\gamma := \epsilon_1\gamma_2$, $\phi := \phi_2$, and $\delta := \epsilon_2\delta_1$. We then have

$$\overline{|\gamma|} = \overline{|\epsilon_1\gamma_2|} < C_6$$

and

$$\overline{|\phi|} = \overline{|\phi_2|} < C_4$$

which implies the result. □

Proof of Theorem 21.1. Let $C_1, C_2, \ldots$ be positive numbers which are effectively computable in terms of f and m. Let

$$f(x) =: (x - \alpha_1)\cdots(x - \alpha_n),$$

and suppose that α_1 and α_2 are simple roots of f.

Let K be the splitting field of f. By Lemma 21.2, there exist

$$\gamma_1, \gamma_2, \phi_1, \phi_2, \delta_1, \delta_2 \in O_K$$

with $\gamma_1\phi_1 \neq 0$ and $\gamma_2\phi_2 \neq 0$ such that

$$x - \alpha_i = \left(\frac{\gamma_i}{\phi_i}\right)\delta_i^m \quad (1)$$

for $i = 1, 2$, and

$$\max(\overline{|\gamma_1|}, \overline{|\gamma_2|}, \overline{|\phi_1|}, \overline{|\phi_2|}) < C_1. \quad (2)$$

Therefore,

$$\gamma_1\phi_2\delta_1^m - \gamma_2\phi_1\delta_2^m = (\alpha_2 - \alpha_1)\phi_1\phi_2.$$

Thus, (δ_1, δ_2) is a solution of $g(x, y) = \mu$ where

$$\mu := (\alpha_1 - \alpha_2\phi_1\phi_2)$$

and

$$g(x, y) := \gamma_1\phi_2 x^m - \gamma_2\phi_1 y^m.$$

Since $\gamma_1\phi_2\gamma_2\phi_1 \neq 0$ and $\alpha_1 \neq \alpha_2$ by the generalization to algebraic number fields of our result on the Thue equation,

$$\max(\overline{|\delta_1|}, \overline{|\delta_2|}) < C_2. \quad (3)$$

By (1), (2), and (3), $|x| < C_3$ and $|y| < C_4$, which proves the claim. □

Remark 22.1. If $m = 2$ and f has at least 3 simple zeroes, then a slightly more complicated argument yields an effective upper bound for $\max(|x|, |y|)$.

We will soon begin our study of the Catalan equation. Apparently Tijdeman proved there are only finitely many solutions but the upper bound was huge. People tried to reduce this bound, but meanwhile Mihailescu gave a different approach that allowed him to solve it completely. Tijdeman's argument generalizes to arbitrary algebraic number fields, but Mihailescu's does not. We will cover the former in this class, not the latter.

23 Nov. 4, 2019

In 1844, Catalan conjectured that the only two consecutive powers of positive integers are 8 and 9. This comes down to solving the Diophantine equation

$$x^m - y^n = 1 \quad (1)$$

in integers x, y, m, and n all larger than 1.

In 1976, Tijdeman proved that all solutions of (1) are less than some effectively computable positive number. In 2002, Mihailescu gave a complete proof of the conjecture using properties of cyclotomic fields. Mihailescu's proof, developed in the process of refining Tijdeman's result, actually does not use results about linear forms in logarithms, relying instead on more algebraic methods. We will follow Tijdeman's argument.

The following result is due to Tijdeman.

Theorem 23.1. *There exists an effectively computable positive number C such that all solutions of (1) in integers $x, y, m, n > 1$ satisfy*

$$\max(x, y, m, n) < C.$$

Proof. Let $C_1, C_2, \ldots$ denote effectively computable positive numbers. We may suppose without loss of generality that m and n are distinct primes, say p and q. We now may consider the equivalent equation

$$x^p - y^q = \epsilon \quad (2)$$

with $p > q$, $x, y > 1$, and $\epsilon \in \{1, -1\}$. We first assume that $p, q, x, y > C_1$. Note that p and q are both odd. Furthermore, since $p > q$, by (2) $x < y$. Also by (2), $(x, y) = 1$.

Moreover, from (2) we have

$$x^p = y^q + \epsilon = (y+\epsilon)(y^{q-1} - \epsilon y^{q-2} + \cdots + \epsilon^{q-1}).$$

Let d be the gcd of $y+\epsilon$ and $y^{q-1} - \epsilon y^{q-2} + \cdots + \epsilon^{q-1}$. Note that $y = \epsilon + (y-\epsilon)$. Thus,

$$y^q = \epsilon^q + \binom{q}{1}\epsilon^{q-1}(y-\epsilon) + \cdots + (y-\epsilon)^q,$$

so

$$\frac{y^q - \epsilon^q}{y-\epsilon} = \binom{q}{1}\epsilon^{q-1} + \binom{q}{2}\epsilon^{q-1}(y-\epsilon) + \cdots + (y-\epsilon)^{q-1}.$$

Suppose $h \mid y-\epsilon$ and $h \mid \frac{y^q-\epsilon^q}{y-\epsilon}$. Thus, $h \mid q$. But q is an odd prime, so $h=1$ or $h=q$, and q divides $\frac{y^q-\epsilon^q}{y-\epsilon}$ to the first power only. In a similar fashion, we can write $y = -\epsilon + (y+\epsilon)$, so $d=1$ or $d=q$, and q divides $\frac{y^q+\epsilon}{y+\epsilon}$ to exactly the first power.

Thus, there are integers $\delta \in \{0,-1\}$ and $s>0$ such that

$$y + \epsilon = q^\delta s^p. \quad (3)$$

In a similar way, we have

$$y^q = x^p - \epsilon = (x-\epsilon)(x^{p-1} + \cdots + \epsilon^{p-1}),$$

and so there are integers $\gamma \in \{0,-1\}$ and $r>0$ such that

$$x - \epsilon = p^\gamma r^q. \quad (4)$$

In fact, if $\gamma = -1$, then $p \mid r$, and if $\delta = -1$, then $q \mid s$. Further, since x and y exceed C_1, we see that r and s are both larger than 1. Thus, we see that

$$p^\gamma r^q > 2^{q-1} \text{ and } q^\delta s^p > 2^{p-1}.$$

Now by (3), $y = q^\delta s^p - \epsilon$, and by (4), $x = p^\gamma r^q + \epsilon$, so we have

$$(p^\gamma r^q + \epsilon)^p - (q^\delta s^p - \epsilon)^q = \epsilon. \quad (5)$$

From (2), (3), and (4) we deduce that

$$2^p r^{pq} \geq (r^q+1)^p + 1 \geq x^p + 1 \geq y^q \geq (q^\delta s^p - 1)^q \geq \frac{s^{pq}}{(2q)^q}.$$

Similarly,

$$2^q s^{pq} \geq (s^p+1)^q + 1 \geq y^q + 1 \geq x^p \geq (p^\gamma r^q - 1)^p \geq \frac{r^{pq}}{(2p)^p}.$$

Thus, since $p > q > C_1$,

$$s^{pq} \leq 4^p q^q r^{pq}.$$

Hence

$$s \leq 4^{1/q} q^{1/p} r \leq 2r. \quad (6)$$

Similarly,

$$r \leq (4p)^{1/q} s. \quad (7)$$

We will continue this proof next class.

24 Nov. 6, 2019

We continue the proof from last class.

Recall that we consider

$$x^p - y^q = \epsilon$$

with p, q odd primes, $p > q$, and $\epsilon \in \{-1, 1\}$. Also, recall Equations (3) and (4):

$$y + \epsilon = q^\delta s^p,$$

$$x - \epsilon = p^\gamma r^q$$

where $\delta, \gamma \in \{-1, 0\}$. Thus, we obtained Equation (5):

$$(p^\gamma r^q + \epsilon)^p - (q^\delta s^p - \epsilon)^q = \epsilon.$$

Furthermore, we proved Inequalities (6) and (7):

$$s \leq 2r,$$

$$r \leq (4p)^{1/q} s.$$

We now note that

$$\max((x-1)^p, (y-1)^q) < x^p = y^q + \epsilon < \min((x+1)^p, (y+1)^q). \quad (8)$$

Therefore,

$$(x - \epsilon)^p - (y + \epsilon)^q \neq 0,$$

so

$$(p^\gamma r^q)^p - (q^\delta s^p)^q \neq 0.$$

Thus, if we set

$$\Lambda_1 := p \log(p^\gamma r^q) - q \log(q^\delta s^p),$$

we see that

$$\Lambda_1 \neq 0.$$

Plainly

$$p^\gamma r^q \geq 2^{q-1}$$

and either $2^{q-1} \geq 12p^3$ or $12p^3 > 2^{q_1}$. In the latter case,

$$\log 12 + 3\log p > (q-1)\log 2,$$

so

$$q < C_2 \log p. \quad (9)$$

We now assume that $2^{q_1} > 12p^3$. (Note that we cannot have $2^{q-1} = 12p^3$ because p is an odd prime.) Then in particular, $p^\gamma r^q > 12p^3$. Now, $x - \epsilon = p^\gamma r^q$, so since $|\epsilon| = 1$,

$$\left|\frac{x}{p^\gamma r^q} - 1\right| = \frac{1}{p^\gamma r^q}, \quad (10)$$

and $x^p - y^q = \epsilon$, so

$$\left|\frac{y^q}{x^p} - 1\right| = \frac{1}{x^p}. \quad (11)$$

Also, $y + \epsilon = q^\delta s^p$, so

$$\left|\frac{y}{q^\delta s^p} - 1\right| = \frac{1}{q^\delta s^p}. \quad (12)$$

Since $|\log(1+z)| \leq 2|z|$ for $|z| \leq 1/2$, we deduce from (10), (11), and (12) and from the inequalities $-1 \leq \gamma \leq 0$, $-1 \leq \delta \leq 0$, and $p > q$ that

$$\left|\log\left(\frac{x}{p^\gamma r^q}\right)\right| = |\log x - \log(p^\gamma r^q)|,$$

and so

$$\left|\log\left(\frac{x}{p^\gamma r^q}\right)\right| = \left|\log\left(1 + \left(\frac{x}{p^\gamma r^q} - 1\right)\right)\right| \leq 2\left|\frac{x}{p^\gamma r^q} - 1\right| \leq \frac{2}{p^\gamma r^q}.$$

Thus,

$$|p\log x - p\log(p^\gamma r^q)| \leq 2p^{1-\gamma} r^{-q} \leq 2p^2 r^{-q}. \quad (13)$$

Similarly we find that

$$|p\log x - q\log y| \leq \frac{2}{x^p} \leq 2pr^{-q} \quad (14)$$

and

$$|q\log y - q\log(q^\delta s^p)| \leq 2q^{1-\delta} s^{-p} \leq 2q^2 s^{-q},$$

which since $p > q$ and (7) holds gives

$$|q\log y - q\log(q^\delta s^p)| \leq 8p^3 r^{-q}. \quad (15)$$

Therefore,

$$|p\log(p^\gamma r^q) - q\log(q^\delta s^p)| \leq 12p^3 r^{-q}. \quad (16)$$

Recall that $\Lambda_1 = p\log(p^\gamma r^q) - q\log(q^\delta s^p)$, and so

$$\Lambda_1 = p\log(p^\gamma) - q\log(q^\delta) + pq\log(r/s).$$

We now apply Theorem 2.6 with $A_1 := p$, $A_2 := q < p$, $A_3 := 2r$, since $s \leq 2r$, $n := 3$, $d := 1$, and $B := p^2$. (These variable names do not agree with the ones from the statement of the theorem, but you get the idea.) Recall that $\Lambda_1 \neq 0$, so

$$|\Lambda_1| > \exp(-C_3(\log p)^3 \log r). \quad (17)$$

Comparing (16) and (17), we find that

$$r^q \leq 12p^3 r^{C_3(\log p)^3} < r^{C_4(\log p)^3}.$$

Thus, we see that

$$q < C_4(\log p)^3$$

and by our assumption (9) we see that

$$q < C_5(\log p)^3.$$

We will now show that p is bounded. It follows from (3), (4), and (8) that

$$(p^\gamma r^q + \epsilon)^p - q^{\delta q} s^{pq} = x^p - (y+\epsilon)^q \neq 0.$$

Thus,

$$\Lambda_2 = p\log(p^\gamma r^q + \epsilon) - q\log(q^\delta s^p) \neq 0.$$

Also, $\Lambda_2 = -q\log(q^\delta) + p\log\left(\frac{p^\gamma r^q + \epsilon}{s^q}\right)$. We have, by (14) and (15), that

$$|p\log x - q\log(q^\delta s^p)| \leq \frac{2}{x^p} + 2q^2 s^{-p}.$$

Furthermore, since $x^p = y^q + \epsilon$ and $y + \epsilon = q^\delta s^p$,

$$x^p = y^q + \epsilon > y + \epsilon = q^\delta s^p \geq \frac{s^p}{q}.$$

Thus,

$$\left|-q^\delta \log q + p\log\left(\frac{p^\gamma r^q + \epsilon}{s^q}\right)\right| \leq 4q^2 s^{-p},$$

i.e., $|\Lambda_2| \leq 4q^2 s^{-p}$.

We will continue this proof next class.

25 Nov. 8, 2019

Recall that

$$\Lambda_2 = -q\delta\log q + p\log\left(\frac{p^\gamma r^q + \epsilon}{s^q}\right)$$

and

$$|\Lambda_2| < 4q^2 s^{-p}. \quad (21)$$

Thus, by Theorem 2.6 with $n := 2$, $A_1 := q$, $A_2 := 5ps^q$, $B := p$. Since $\Lambda_2 \neq 0$, we obtain

$$|\Lambda_2| > \exp(-C_6(\log p)^2 \log(5ps^q)). \quad (22)$$

Comparing (21) and (22), we find that

$$s^p \leq 4q^2(5ps^q)^{C_6(\log p)^2},$$

and by (19) we see that $s^p \leq s^{C_7(\log p)^5}$. Thus, $p \leq C_7(\log p)^5$, hence $p \leq C_8$ and also $q \leq C_8$. Therefore, we may assume that p and q are fixed, and then by Theorem 21.1, we see that $|x|$ and $|y|$ are bounded. This completes the argument subject to our original hypothesis that $|x|, |y|, p$, and q all exceed C_1.

To handle the remaining cases, we can again use arguments based on estimates for linear forms in logarithms. However, we can also appeal to a result of Hyyrö from 1964. He proved that if $x^m - y^n = 1$ with $x, y, m, n > 1$ and $(x, y, m, n) \neq (3, 2, 2, 3)$, then $x, y > 10^{11}$.

In 1850, V.-A. Lebesgue proved that there are no solutions of

$$x^m - y^2 = 1$$

in integers $x, y, m > 1$. In 1965, Chao Ko proved that there is only one solution of

$$x^2 - y^n = 1$$

in integers $x, y, n > 1$. These three results combined with our argument suffice to complete the proof. □

One might wonder whether one could prove there are only finitely many solutions to the other famous equation

$$x^n + y^n = z^n$$

using a linear forms approach. In 1977, Cam proved that there are only finitely many solutions in integers x, y, z, n with $n > 2$ and $|x-y|$ bounded to this equation. This is a nice result!

We now give an estimate for linear forms in two logarithms based on a zero estimate due to Nesternko. Cam describes Nesterenko as an excellent Russian mathematician who is also a very proficient ballroom dancer.

Lemma 25.1. *Let L and $K_1, \ldots, K_n$ be integers with $0 \leq K_1 < K_2 < \cdots < K_n < L$, and let $\mathcal{E}$ be a set of at least L non-zero complex numbers. There exist $a_1, \ldots, a_n$ in $\mathcal{E}$ such that*

$$\det(a_j^{K_i})_{i,j=1,\ldots,n} \neq 0.$$

Proof. We proceed by induction on n. Certainly the result holds for $n = 1$ since the elements of $\mathcal{E}$ are non-zero. Let $n > 1$, and suppose that the result holds for $n-1$. Thus there exist $a_1, \ldots, a_{n-1}$ in $\mathcal{E}$ such that

$$\det(a_i^{K_j})_{i,j=1,\ldots,n-1} \neq 0.$$

Consider the polynomial $P(z)$ given by

$$P(z) := \det \begin{pmatrix} a_1^{K_1} & \cdots & a_{n-1}^{K_1} & z^{K_1} \\ \vdots & \ddots & \vdots & \vdots \\ a_1^{K_n} & \cdots & a_{n-1}^{K_n} & z^{K_n} \end{pmatrix}$$

$$= Az^{K_n} + \cdots,$$

where $A := \det(a_i^{K_j})_{i,j=1,\ldots,n-1} \neq 0$. Note that $P(a_i) = 0$ for $1 \leq i \leq n-1$, so

$$P(z) = (z - a_1) \cdots (z - a_{n-1}) Q(z)$$

for some $Q(z)$. Since the cardinality $|\mathcal{E} \setminus \{z_1, \ldots, z_{n-1}\})|$ is greater than or equal to

$$L - n - 1 > K_n - (n-1) = \deg Q(z),$$

we can find a_n in $\mathcal{E}$ such that $P(a_n) \neq 0$. Our result follows. □

Proposition 25.2. *Let α_1, α_2, and β be complex numbers with $\alpha_1\alpha_2 \neq 0$. Let K, L, R_1, R_2, S_1, and S_2 be positive integers. [At this point, I have to leave to proctor a test. But I got the rest of the Proposition later from Pranabesh's notes.] Let $P \in \mathbb{C}[x,y]$ be a non-zero polynomial with degree at most $K-1$ in x and degree $L-1$ in y. Put $R := R_1 + R_2 - 1$ and $S := S_1 + S_2 - 1$. Suppose that the cardinality of*

$$\{\alpha_1^r \alpha_2^s \mid 0 \leq r < R_1, 0 \leq s < S_1\}$$

is $\geq L$ and that the cardinality of

$$\{r + s\beta \mid 0 \leq r < R_2, 0 \leq s < S_2\}$$

is $\geq (K-1)L$. Then at least one of the numbers

$$p(r + s\beta, \alpha_1^r \alpha_2^s)$$

for $0 \leq r < R$ and $0 \leq s < S$ is non-zero.

26 Nov. 11, 2019

We now prove Proposition 25.2.

Proof. We may suppose that $P(x, 0) \neq 0$ since otherwise $P(x,y) = y^m Q(x,y)$ with $m \geq 1$ and we could replace P by Q because $\alpha_1\alpha_2 \neq 0$. Suppose that

$$P(r + s\beta, \alpha_1^r \alpha_2^s) = 0 \quad (1)$$

for $0 \leq r < R$ and $0 \leq s < S$. Let us write

$$P(x,y) = \sum_{i=1}^{n} Q_i(x) y^{K_i}$$

with the Q_i's non-zero and

$$0 = K_1 < K_2 < \cdots < K_n < L.$$

Put

$$\mathcal{E} := \{\alpha_1^r \alpha_2^s \mid 0 \leq r < R_1, 0 \leq s < S_1\}.$$

By assumption, the cardinality of $\mathcal{E}$ is at least L. By Lemma 25.1, we have that there is a subset $\mathcal{L}$ of $\{(r,s) \mid 0 \leq r < R_1, 0 \leq s < S_1\}$ of size n such that

$$B := \det((\alpha_1^r \alpha_2^s)^{K_i})_{\substack{i=1,\ldots,n \\ (r,s)\in\mathcal{L}}} \neq 0.$$

We now define for each pair $(r,s) \in \mathcal{L}$ the polynomial $P_{r,s}(x,y)$ by

$$P_{r,s}(x,y) := P(x + r + s\beta, \alpha_1^r \alpha_2^s y).$$

Thus,

$$P_{r,s}(x,y) = \sum_{i=1}^{n} Q_i(x + r + s\beta)(\alpha_1^r \alpha_2^s)^{K_i} y^{K_i}. \quad (2)$$

We now put

$$\Delta(x) := \det\left(Q_i(x + r + s\beta)(\alpha_1^r \alpha_2^s)^{K_i}\right)_{\substack{i=1,\ldots,n \\ (r,s)\in\mathcal{L}}}.$$

Write

$$Q_i(x) = b_i x^{m_i} + \cdots$$

with $b_i \neq 0$ for $i = 1, \ldots, n$. We see that

$$\Delta(x) = b_1 \cdots b_n B X^{m_1 + \cdots + m_n} + \cdots.$$

Notice that $b_1 \cdots b_n B \neq 0$.

Consider the system of equations (2) as a system of linear equations in $z_1, \ldots, z_n$ where $z_i := y^{K_i}$ for $i = 1, \ldots, n$. Thus by Cramer's rule and the fact that $K_1 = 0$, there exist polynomials $S_{r,s} \in \mathbb{C}[x,y]$ for $(r,s) \in \mathcal{L}$ for which

$$z_1 \Delta(x) = y^{K_1} \Delta(x) = \Delta(x) = \sum_{(r,s)\in\mathcal{L}} P_{r,s}(x,y) S_{r,s}(x). \quad (3)$$

Note that by (1),

$$P_{r,s}(r_0 + s_0\beta, \alpha_1^{r_0} \alpha_2^{s_0}) = P(r + r_0 + (s + s_0)\beta, \alpha_1^{r+r_0} \alpha_2^{s+s_0}) = 0$$

for $0 \leq r_0 < R_2$ and $0 \leq s_0 < S_2$. By (3),

$$\Delta(r_0 + s_0\beta) = 0$$

for $0 \leq r_0 < R_2$ and $0 \leq s_0 < S_2$. By assumption, the cardinality of

$$\{r_0 + s_0\beta \mid 0 \leq r_0 < R_2, 0 \leq s_0 < S_2\}$$

is greater than $(K-1)L$. Thus,

$$(K-1)L \leq m_1 + \cdots + m_n \leq n(L-1) < L(K-1),$$

which is a contradiction. Therefore, (1) does not hold, and the result follows. □

We will use these results to establish a version of Theorem 2.6 with $n = 2$.

Theorem 26.1. *Let α_1 and α_2 be non-zero algebraic numbers, and let $K := \mathbb{Q}(\alpha_1, \alpha_2)$, $d := [K : \mathbb{Q}]$. Let $\log \alpha_1$ and $\log \alpha_2$ be some branch of the logarithm function at α_1 and α_2, respectively. Let b_1 and b_2 be non-zero integers. Put*

$$\Lambda := b_1 \log \alpha_1 + b_2 \log \alpha_2.$$

Put

$$A_i := \max\{h(\alpha_i)^d, \exp(|\log \alpha_i|), e\}$$

for $i = 1, 2$.

There exists a positive number C, which is effectively computable in terms of d, such that if $\Lambda \neq 0$, then

$$|\Lambda| > \exp(-C \log A_1 \log A_2 (\log B')^2)$$

where

$$B' := \max\left(3, \frac{|b_1|}{\log A_2} + \frac{|b_2|}{\log A_1}\right).$$

Presentations will be during the last week of classes. We will arrange oral exam dates in the period after December 10.

27 Nov. 13, 2019

We now prove Theorem 26.1.

Proof. We may assume without loss of generality that α_1 and α_2 have absolute value at least 1. We may also assume that $b_1 > 0$ and that $b_2 < 0$. Replacing b_2 with $-b_2$, we can write

$$\Lambda = b_2 \log \alpha_2 - b_1 \log \alpha_1$$

with $b_1, b_2 > 0$.

Let $K \geq 3$, $L \geq 2$, R_1, R_2, S_1, and S_2 be positive integers. Put $N := KL$, $R := R_1 + R_2 - 1$, and $S := S_1 + S_2 - 1$. (Use $|\mathcal{S}|$ to denote the cardinality of a set $\mathcal{S}$.) If the conditions

$$|\{\alpha_1^r \alpha_2^s \mid 0 \leq r < R_1, 0 \leq s < S_1\}| \geq L \quad (1)$$

and

$$|\{rb_2 + sb_1 \mid 0 \leq r < R_2, 0 \leq s < S_2\}| \geq (K-1)L \quad (2)$$

hold, then the $(KL \times RS)$ matrix

$$\left(\binom{rb_2 + sb_1}{k} \alpha_1^{\ell r} \alpha_2^{\ell s}\right)_{(r,s),(k,\ell)},$$

where $0 \leq r < R$, $0 \leq s < S$, $k = 1, \ldots, K$, and $\ell = 1, \ldots, L$ is of maximal rank since otherwise there exist complex numbers $c_{k,\ell}$ for $k = 0, \ldots, K-1$, $\ell = 0, \ldots, L-1$ such that the polynomial

$$P(X, Y) := \sum c_{k,\ell} \binom{X}{k} Y^\ell = 0$$

for $X = rb_2 + sb_1$, $0 \leq r < R_2$ and $0 \leq s < S_2$, and $Y = \alpha_1^r \alpha_2^s$ for $0 \leq r < R_1$ and $0 \leq s < S_1$. However, this contradicts Proposition 25.2.

Suppose that (1) and (2) hold. We can then extract an $N \times N$ minor of the matrix with non-zero determinant, say

$$\Delta := \det\left(\binom{r_j b_2 + s_j b_1}{k_i} \alpha_1^{\ell_i r_j} \alpha_2^{\ell_i s_j}\right)_{i,j=1,\ldots,N} \neq 0. \quad (3)$$

We first observe that

$$\sum_{i=1}^{N} k_i = \frac{N}{K} \sum_{i=0}^{K-1} i = \frac{N}{K} \frac{K(K-1)}{2} = \frac{N(K-1)}{2}.$$

We also introduce the quantity b where

$$b := ((R-1)b_2 + (S-1)b_1) \left(\prod_{k=1}^{K-1} k!\right)^{-\frac{2}{K^2-K}}.$$

If we expand the determinant in (3), we get $N!$ terms, and each term is a product $E\alpha_1^{E_1}\alpha_2^{E_2}$ where E is a product of binomial coefficients

$$\binom{r_j b_2 + s_j b_1}{k_i}$$

while E_1 is a sum of terms $\ell_i r_j$ and E_2 is a sum of terms $\ell_i s_j$. We have

$$1 \leq E \leq \frac{((R-1)b_2 + (S-1)b_1)^{\sum_{i=1}^N k_i}}{\prod_{i=1}^N (k_i!)} = \frac{((R-1)b_2 + (S-1)b_1)^{\frac{N(K-1)}{2}}}{\left(\prod_{k=1}^{K-1} k!\right)^L}$$

Also we have $E_1 \leq NLR$ and $E_2 \leq NLS$. Therefore, we have

$$h(\Delta) \leq \frac{N!((R-1)b_2 + (S-1)b_1)^{\frac{N(K-1)}{2}}}{\left(\prod_{k=1}^{K-1} k!\right)^L} h(\alpha_1)^{NLR} h(\alpha_2)^{NLS}.$$

Thus,

$$h(\Delta) \leq N^N b^{\frac{KN}{2}} h(\alpha_1)^{NLR} h(\alpha_2)^{NLS}.$$

From this we will deduce that

$$\log|\Delta| \geq -dN\log N - dKN\frac{\log b}{2} - dLN(R\log h(\alpha_1) + S\log h(\alpha_2)).$$

The proof will be continued next class.

28 Nov. 15, 2019

Recall that the essays are due Dec. 6. They may be submitted to Cam electronically. Cam gives out a sign-up sheet for our talks. Final exams will probably be half an hour each sometime in the range December 10–12. (My talk is from 10:30 to 11:30 on Wednesday, Dec. 4.)

Recall also that

$$0 \neq \Delta = \det\left(\binom{r_j b_2 + s_j b_1}{k_i} \alpha_1^{\ell_i r_j} \alpha_2^{\ell_i s_j}\right)_{i,j=1,\ldots,N},$$

that we defined b by

$$b := ((R-1)b_2 + (S-1)b_1)\left(\prod_{k=1}^{K-1} k!\right)^{-2/(K^2-K)},$$

and that we defined E_0 by

$$E_0 := \frac{(((R-1)b_2 + (S-1)b_1)^{N(K-1)/2}}{\left(\prod_{k=1}^{K-1} k!\right)^L}.$$

The terms in the expansion of the determinant Δ have the form $E\alpha_1^{E_1}\alpha_2^{E_2}$ where $E \leq E_0$, $E_1 \leq NLR$, and $E_2 \leq NLS$.

Suppose that $M(\alpha_i) = |a_i| \prod_\sigma \max(1, |\sigma(\alpha_i)|)$ for $i = 1, 2$ where $a_1, a_2 \in \mathbb{N}$. Then $a_1\alpha_1$ and $a_2\alpha_2$ are algebraic integers, so $a_1^{NLR} a_2^{NLS}\Delta$ is an algebraic integer. Since $\Delta \neq 0$, it follows that

$$|\prod_\sigma \sigma(a_1{}^N LRa_2^{NLS}\Delta)| \geq 1.$$

Notice that

$$|\sigma\Delta| \leq N!E_0 \max(1, |\sigma(\alpha_1)|)^{NLR} \max(1, |\sigma(\alpha_2)|)^{NLS},$$

so

$$|\Delta| \geq (N!E_0)^{-d} \left(\prod_{\sigma\neq \mathrm{id}} \max(1, |\sigma(\alpha_1)|)\right)^{-NLR} \left(\prod_{\sigma\neq \mathrm{id}} \max(1, |\sigma(\alpha_2)|)\right)^{-NLS} (a_1^{NLR} a_2^{NLS})^{-d}.$$

Since $E_0 \leq b^{KN/2}$, we have

$$|\Delta| \geq N^{-Nd} b^{-KNd/2} M(\alpha_1)^{-dNLR} M(\alpha_2)^{-dNLS}.$$

Therefore,

$$\log|\Delta| \geq -dN \log N - (dKN/2) \log b - d^2 LN(R \log h(\alpha_1) + S \log h(\alpha_2)). \quad (4)$$

We now remark that

$$\begin{aligned} b &\leq (Rb_2 + Sb_1) \left(\prod_{k=1}^{K-1} k^{k-K}\right)^{2/(K^2-K)} \\ &\leq (Rb_2 + Sb_1) \exp\left(2/(K^2-K) \sum_{k=1}^{K-1} (k-K) \log k\right). \end{aligned}$$

But notice that

$$\begin{aligned} \sum_{k=1}^{K-1} k \log k &\leq 2 + \int_1^{K-1} (x \log x + (1/2)x)\, \mathrm{d}x \\ &\leq 2 + [\frac{1}{2} x^2 \log x]_1^{K-1} \\ &\leq 2 + \frac{1}{2}(K-1)^2 \log(K-1). \end{aligned}$$

and

$$\begin{aligned} \sum_{k=1}^{K-1} \log k &= \log((K-1)!) \\ &> \log\left(\left(\frac{K-1}{e}\right)^{K-1}\right) \\ &= (K-1) \log\left(\frac{K-1}{e}\right), \end{aligned}$$

so

$$\begin{aligned}
&\exp\left(\frac{2}{K(K-1)}\sum_{k=1}^{K-1}(k-K)\log k\right)\\
&< \exp\left(\frac{2}{K(K-1)}\left(2+\frac{1}{2}(K-1)^2\log(K-1)-K(K-1)\log\left(\frac{K-1}{e}\right)\right)\right)\\
&\leq \exp\left(\frac{4}{K(K-1)}+\frac{K-1}{K}\log(K-1)-2\log\left(\frac{K-1}{e}\right)\right)\\
&\leq \exp\left(1+\log(K-1)-2\log(K-1)+2\right)\\
&\leq \exp(3-\log(K-1))\\
&\leq e^3/(K-1) < e^4/K.
\end{aligned}$$

Thus,

$$b \leq \frac{Rb-2+Sb_1}{K}e^4. \quad (5)$$

We now introduce the quantity Λ' given by

$$\Lambda' := \Lambda\max\{e^{|\Lambda|LS/b_2}LS/b_2, e^{|\Lambda|LR/b_1}LR/b_1\}.$$

Next, let ρ be a real number larger than 1. We will now prove a lemma.

Lemma 28.1. *If* $|\Lambda'| \leq \rho^{-N+1/2}$, *then*

$$\log|\Delta| \leq -N^2\log(\rho/2)+N\log N+\frac{KN\log\rho}{2}+\frac{KN\log b}{2}+\rho NLR|\log\alpha_1|+\rho NLS|\log\alpha_2|.$$

Proof. We may assume without loss of generality that

$$b_1|\log\alpha_1| \leq b_2|\log\alpha_2|.$$

We have $\Lambda = b_2\log\alpha_2 - b_1\log\alpha_1$, so

$$\log\alpha_2 = \frac{\Lambda}{b_2}+\frac{b_1}{b_2}\log\alpha_1.$$

Put $\beta := b_1/b_2$. Then

$$\log\alpha_2 = \beta\log\alpha_1+\Lambda/b_2,$$

so

$$\alpha_2 = \alpha_1^\beta e^{\Lambda/b_2}.$$

By the multilinearity of the determinant,

$$\Delta = \det\left(\frac{b_2^{k_i}}{k_i!}(r_j+s_j\beta)^{k_i}\alpha_1^{\ell_i r_j}\alpha_2^{\ell_i s_j}\right)_{1\leq i,j\leq N}.$$

Notice that

$$\alpha_1^{\ell_i r_j}\alpha_2^{\ell_i s_j} = \alpha_1^{\ell_i(r_j+s_j\beta)}e^{\Lambda/b_2\ell_i s_j}.$$

Now we put

$$e^{\frac{\Lambda}{b_2}\ell_i s_j} = 1+\Lambda'\theta_{i,j}$$

where

$$\theta_{i,j} = \frac{\exp(\ell_i s_j\Lambda/b_2)-1}{\Lambda'},$$

so that

$$|\theta_{i,j}| \leq \frac{b_2(\exp(|\ell_i|s_j\Lambda/b_2)-1)}{LS\Lambda e^{|\Lambda|LS/b_2}}$$

Since $e^x-1\leq xe^x$ for $x>0$,

$$|\theta_{i,j}| \leq \frac{b_2(|\ell_i|s_j\Lambda/b_2)e^{\ell_i s_j\Lambda/b_2}}{LS\Lambda e^{|\Lambda|LS/b_2}} \leq 1.$$

□

29 Nov. 18, 2019

Thus,

$$\Delta = \det\left(\frac{b_i^{k_i}}{k_i!}(r_j+s_j\beta)^{k_i}\alpha_1^{\ell_i(r_j+s_j\beta)}(1+\theta_{ij}\Lambda')\right)_{i,j=1,\ldots,n}.$$

Therefore, Δ can be written as

$$\Delta = \sum_{I\subseteq\{1,\ldots,n\}}(\Lambda')^{N-|I|}\Delta_I \quad (6)$$

where

$$\Delta_I := \det\begin{pmatrix}\varphi_i(z_1) & \cdots & \varphi_i(z_N)\\ \theta_{i1}\varphi_i(z_1) & \cdots & \theta_{iN}\varphi_i(z_N)\end{pmatrix}.$$

We mean by this notation that a row of the matrix is of the form $(\varphi_i(z_1),\ldots,\varphi_i(z_N))$ when $i\in I$ and of the form $(\theta_{i1}\varphi_i(z_1),\ldots,\theta_{iN}\varphi_i(z_N))$ when $i\notin I$. Also,

$$\varphi_i(z) := b_2^{k_i}\frac{z^{k_i}}{k_i!}\alpha_i^{\ell_i z}$$

and

$$z_j := r_j+s_j\beta$$

for $1\leq i,j\leq n$. We now define, for each $I\subseteq\{1,\ldots,n\}$, the function $\Phi_I(x)$ by

$$\Phi_I(x) := \det\begin{pmatrix}\varphi_i(xz_1) & \cdots & \varphi_i(xz_N)\\ \theta_{i1}\varphi_i(xz_1) & \cdots & \theta_{iN}(xz_N)\end{pmatrix}.$$

Here we use the same notational convention as before, so the first row is the form taken when $i \in I$ and the second row is the form taken otherwise.

Observe that $\Phi_I(1) = \Delta_I$. Our next step is to show that $\Phi_I(x)$ has a zero of multiplicity $(\nu^2 - \nu)/2$ where $\nu := |I|$. Notice that we certainly have that $x^{\sum_{i \in I} k_i}$ divides $\Phi_I(x)$, and so $\Phi_I(x)$ has multiplicity at least

$$(0 + \cdots + 0)(L \text{ times}) + (1 + \cdot + 1)(L \text{ times}) + \cdots + \left[\frac{\nu}{L}\right] + \cdots + \left[\frac{\nu}{L}\right] (L \text{ times})$$

$$= L\left(\frac{\left[\frac{\nu}{L}\right]^2 - \left[\frac{\nu}{L}\right]}{2}\right).$$

In fact, Laurent showed by examining the Taylor expansion of the $\varphi_i(xz)$'s that the multiplicity is at least

$$\frac{\nu^2 - \nu}{2}.$$

By the maximum modulus principle applied to $\Phi_I(x)/x^{(\nu^2-\nu)/2}$, we find that

$$|\Delta_I| = |\Phi_I(1)| \leq \rho^{-(\nu^2-\nu)/2} \max_{|x|=\rho} |\Phi_I(x)|.$$

By (6),

$$|\Delta| \leq 2^N \max_{0 \leq \nu \leq N} \left(\rho^{-(N-1/2)(N-\nu)-(\nu^2-\nu)/2}\right) \max_I \max_{|x|=\rho} |\Phi_I(x)|.$$

Notice that

$$\begin{aligned} &\min_{0 \leq \nu \leq N} ((N-1/2)(N-\nu) + (\nu^2-\nu)/2) \\ &= \min_{0 \leq \nu \leq N} (N^2 - N/2 - N\nu + \nu^2/2) \\ &= N^2 - N/2 - N^2 + N^2/2 \\ &= N(N-1)/2. \end{aligned}$$

Thus,

$$|\Delta| \leq 2^N \rho^{-(N^2-N)/2} \max_I \max_{|x|=\rho} |\Phi_I(x)|. \quad (7)$$

Furthermore,

$$|\Phi_I(x)| \leq N! \left(\prod_{i=1}^{N} \frac{(b_2|x|(R+\beta S))^{k_i}}{k_i!}\right) \exp(\sum_{i=1}^{N} \ell_i(R+S\beta)|x|| \log \alpha_1|).$$

Since $\beta|\log \alpha_1| \leq |\log \alpha_2|$, we have

$$|\Phi_I(x)| \leq N!(|x|b)^{(K-1)N/2} \exp(|x|(NLR|\log \alpha_1| + NLS|\log \alpha_2|)). \quad (8)$$

For $N \geq 6$, we have $2^N N! \leq N^N$, and so by (7) and (8),

$$\log|\Delta| \leq \frac{-N^2 \log\rho}{2} + \frac{N\log\rho}{2} + N\log N + \frac{(K-1)N\log\rho}{2} + \frac{(K-1)N\log b}{2}$$

$$+\rho NLR|\log\alpha_1| + \rho NLS|\log\alpha_2|,$$

and the lemma follows. □

We now continue the *Proof of Theorem 26.1.*

We compare our upper bound for $\log|\Delta|$ with our lower bound, which we established under the assumption that Λ' is small. We get

$$-2d\log N - 2dK\log b - 2d^2 LR\log h(\alpha_1) - 2d^2 LS\log h(\alpha_2)$$

$$\leq -N\log\rho + \log\rho + 2\log N + (K-1)\log\rho + (K-1)\log b + 2\rho LR|\log\alpha_1| + 2\rho LS|\log\alpha_2|,$$

so

$$N\log\rho \leq 2(d+1)\log N + 3dK\log b + K\log\rho +$$
$$2LR(\rho|\log\alpha_1| + d^2\log h(\alpha_1)) + 2LS(\rho|\log\alpha_2| + d^2\log h(\alpha_2)).$$

30 Nov. 20, 2019

Cam says that for the oral exam, we should know the strategy of every proof, but we aren't expected to know every detail. The exams will be half an hour long and in MC 5479. Mine is from 11 to 11:30 on Wednesday, Dec. 11.

Cam also says that the essay should be a longer expository result of the paper we write about (in particular, don't just rewrite the paper), situating it in historical context by relating it to any relevant problems.

From where we left off last class, we get

$$N\log\rho \leq 2(d+1)\log N + 3dK\log b + K\log\rho + 2L(\rho+d)(R\log A_1 + S\log A_2). \quad (9)$$

Set

$$B := \max\{\log\rho, d(7+\log B')\},$$

and, letting $[\cdot]$ denote the integer part, set

$$K := [c^2 B\log A_1\log A_2], L := [B], R_1 := [B\log A_1],$$

$$S_1 := [B\log A_2], R_2 := [cB\log A_1] + 1, S_2 := [cB\log A_2] + 1$$

where $c \geq 1$ will be chosen later.

Notice that $R_1, S_1 \geq L$, so if α_1 and α_2 are not both roots of unity, then condition (1) holds. On the other hand, if they are both roots of unity our result follows immediately.

The argument now splits into two cases.

Case (i): The set $\{rb_2 + sb_1 \mid 0 \leq r < R_2, 0 \leq s < S_2\}$ has R_2S_2 elements.

Case (ii): The set $\{rb_2 + sb_1 \mid 0 \leq r < R_2, 0 \leq s < S_2\}$ has fewer than R_2S_2 elements.

We first deal with Case (i). In Case (1),

$$R_2S_2 > (cB \log A_1)(cB \log A_2)$$

and $(K-1)L \leq c^2 B \log A_1 \log A_2$, so $R_2S_2 > (K-1)L$. The conditions (1) and (2) hold. Now, by (5),

$$d \log B \leq B.$$

Thus the right hand side of (9) is bounded from above by

$$2(d+1)\log(c^2B^2 \log A_1 \log A_2) + 3c^2B^2 \log A_1 \log A_2$$

$$+c^2 \log \rho B \log A_1 \log A_2 + (2(\rho + d)B)(2(1+c))B \log A_1 \log A_2,$$

which in turn is less than

$$(2(d+1)\log(c^2) + 4c^2 + 8(\rho + d)c)B^2 \log A_1 \log A_2.$$

But

$$N \log \rho = KL \log \rho \geq \frac{1}{2}c^2B^2 \log A_1 \log A_2 \log \rho,$$

so

$$\frac{1}{2}c^2 \log \rho \leq 4c^2 + 2(d+1)\log(c^2) + 8(\rho + d)c.$$

Take $\rho := c$. Then, taking c sufficiently large in terms of d, we obtain a contradiction. Therefore, in Case (i), our assumption that $|\Lambda'| < \rho^{-N+1/2}$ is false. Thus, $|\Lambda'| \geq \rho^{-N+1/2}$.

We get

$$|\Lambda| > \exp(-(N+1/2)\log c)(\max\{\exp(|\Lambda|LS/b_2)(LS/b_2), \exp(|\Lambda|LR/b_1)(LR/b_1)\})^{-1}.$$

We may assume $|\Lambda| < 1/(LS)$ and $|\Lambda| < 1/(LR)$ since otherwise the result holds. Then

$$1 + \log R + \log S + \log L < B^2 \log A_1 \log A - 2,$$

so

$$|\Lambda| > \exp(-2N \log c).$$

But $N = KL > (1/2)c^2B^2 \log A_1 \log A_2$, so the result follows in Case (i).

Finally, we consider Case (ii). Then there exist integers r and s with $|r| \leq R-1$ and $|s| \leq S-1$ for which

$$rb_2 + sb_1 = 0.$$

Accordingly, $b_2 = (-s/r)b_1$. Then

$$|\Lambda| = |b_1 \log \alpha_1 + b_2 \log \alpha_2|$$

$$= |b_1 \log \alpha_1 - (s/r)b_1 \log \alpha_2|$$

$$= |b_1/r||r \log \alpha_1 - s \log \alpha_2|,$$

and by Proposition 3.1 and the fact that $H(\alpha) \leq 2^d h(\alpha)^d$, the result follows. This completes the proof of Theorem 26.1. □

Cam explains that having proved this bound for a linear form of two logarithms, there is a cheap trick that allows us to get a bound for linear forms of n logarithms strong enough to solve the Thue equation. However, it is not as strong as Theorem 2.6. To get that, you basically need to redo this whole proof for the n logarithm case. This involves showing the linear form is small and then constructing a complicated polynomial in n variables that vanishes at too many points in a small volume.

31 Nov. 22, 2019

Cam asks us to email our essays to cstewart@uwaterloo.ca on or before December 6. He says he will not hold us to technical details on the proof of the result about linear forms in two logarithms during our oral exams but wants us to understand all the results in the course. He says he will ask us some questions at the blackbord, he'll be happy, we'll be happy, and everyone will walk away delighted.

Let $\alpha_1, \ldots, \alpha_n$ be non-zero algebraic numbers of heights $A_1, \ldots, A_n$ (with $A_i \geq 2$). Put $K := \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$ and $d := [K : \mathbb{Q}]$. Let $\mathfrak{p}$ be a prime ideal in O_K lying over the rational prime p. Let $\text{ord}_{\mathfrak{p}}(\theta)$ for $\theta \in K$ be the order of $\mathfrak{p}$ in the fractional ideal generated by θ.

In 1977, van der Poorten proved the following.

Theorem 31.1. *There exist positive numbers C_0 and C_1 such that*

$$ord_{\mathfrak{p}}(\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1) < (C_0 nd)^{C_1 n} \frac{p^d}{\log p} \log A_1 \cdots \log A_n (\log B)^2$$

for all integers $b_1, \ldots, b_n$ of absolute value at most B (≥ 2) for which $\alpha_1^{b_1} \cdots \alpha_n^{b_n} \neq 1$.

There were some technical points overlooked by van der Poorten. These were fixed by Kunrui Yu in 1989.

For any integer n, let $P(n)$ denote the greatest prime factor of n with the convention that $P(0) = P(\pm 1) = 1$. In 1977, Shorey, van der Poorten, Tijdeman, and Schinzel used both complex and p-adic estimates for linear forms to prove the following.

Theorem 31.2. *Let F be a binary form with integer coefficients, degree at least 3, and non-zero discriminant. Let x and y be coprime integers for which $F(x,y) \neq 0$, and put*

$$z := \max(|x|, |y|, 3).$$

There is a positive number C, which is effectively computable in terms of F, such that

$$P(F(x,y)) > C \log\log z.$$

In 1984/85, Oesterlé and Masser made the following conjecture, known as the *abc conjecture.*

Conjecture 31.3. Let $\epsilon > 0$. Then there exists a positive number $C(\epsilon)$ such that if a, b, and c are positive integers with $(a,b,c) = 1$ and

$$a + b = c,$$

then

$$c < C(\epsilon) G^{1+\epsilon}$$

where

$$G = G(a,b,c) := \prod_{p|abc} p.$$

With Kunrui Yu, Cam showed in 2001 by means of estimates for p-adic and complex linear forms that there exists a positive number C_1 such that, with the previous assumptions,

$$c < \exp(C_1 G^{1/3} (\log G)^3).$$

Cam now tells us the story of Mochizuki's attempt to prove the *abc* conjecture.

The remainder of the course will be spent giving talks. I will not take notes on these.