LECTURE 1: INTRODUCTION AND BAKER’S THEOREM

Samit Dasgupta
Duke University

Math 790, 2/23/2021
Introduction to Transcendence Theory
A complex number $\alpha \in \mathbb{C}$ is called **algebraic** if it is the root of a nonzero polynomial $P(x) \in \mathbb{Q}[x]$.

The set of all algebraic numbers is a field $\overline{\mathbb{Q}}$.

**Hilbert’s 7th Problem**

Let $a \in \overline{\mathbb{Q}}$, $a \neq 0,1$, $b \in \overline{\mathbb{Q}}$, $b \notin \mathbb{Q}$. Is $a^b$ necessarily transcendental?

Example: Gelfond’s constant $\sqrt{2}^{\sqrt{2}} \notin \overline{\mathbb{Q}}$.

$a^b = c \iff b \log a = \log c$. 
DAVID HILBERT
Gelfond-Schneider Theorem (1934, 1935)

Yes. Equivalently, if \( a, c \in \overline{Q} \) and \( \log a, \log c \) are linearly dependent over \( \overline{Q} \), then they are linearly dependent over \( Q \).

Note, here \( \log a \) means “any logarithm of \( a \)”, i.e. a number \( x \) such that \( e^x = a \).

Baker’s Theorem (1966)

Suppose \( \alpha_1, \ldots, \alpha_n \in \overline{Q}^* \) such that \( \log \alpha_1, \ldots, \log \alpha_n \) are linearly independent over \( Q \). Then \( \log \alpha_1, \ldots, \log \alpha_n \) are linearly independent over \( \overline{Q} \).
**Four Exponentials Conjecture**

Let \( a, b, c, d \in \overline{Q}^* \) such that

\[
\det(M) = 0, \quad \text{where} \quad M = \begin{pmatrix} \log a & \log b \\ \log c & \log d \end{pmatrix}.
\]

Then either the rows or columns of \( M \) are linearly dependent over \( Q \).

**Six Exponentials Theorem**

Let \( a, b, c, d, e, f \in \overline{Q}^* \) such that

\[
\text{rank}(M) < 2, \quad \text{where} \quad M = \begin{pmatrix} \log a & \log b & \log c \\ \log d & \log e & \log f \end{pmatrix}.
\]

Then either the rows or columns of \( M \) are linearly dependent over \( Q \).
Baker’s Theorem (1966)

Suppose $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}^*$ such that $\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over $\mathbb{Q}$. Then $\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over $\overline{\mathbb{Q}}$.

Full theorem is stronger in two ways:

- Get 1, $\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over $\overline{\mathbb{Q}}$.
- Get effective version, giving a lower bound on

$$\left| \beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n \right|$$

for algebraic numbers $\beta_i$. 
Theorem
Suppose $\alpha_1, \ldots, \alpha_n \in \overline{Q}^*$ such that $\log \alpha_1, \ldots, \log \alpha_n, 2\pi i$ are linearly independent over $Q$. Then $\log \alpha_1, \ldots, \log \alpha_n$ are linearly independent over $\overline{Q}$.

Reason for including $2\pi i$

We’ll show $\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n = 0$ with $\beta_1, \ldots, \beta_n \in \overline{Q}$ not all zero implies $\alpha_1^{\lambda_1} \alpha_2^{\lambda_2} \cdots \alpha_n^{\lambda_n} = 1$ for integers $\lambda_i$ that are not all zero. This implies that $\log \alpha_1, \ldots, \log \alpha_n, 2\pi i$ are linearly dependent over $Q$.

The method of proof of the full theorem is exactly the same, just slightly more complicated in details.
KEY IDEAS IN PROOF

The proof of Baker’s Theorem is both beautiful and brilliant, but writing down the details can seem technical, so let’s first give an overview of what I view as the 4 key ideas/stratagems.

(1) How to show $\alpha_1^{\lambda_1} \alpha_2^{\lambda_2} \cdots \alpha_n^{\lambda_n} = 1$.

We’ll show $\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n = 0$ with $\beta_1, \ldots, \beta_n \in \overline{\mathbb{Q}}$ not all zero implies $\alpha_1^{\lambda_1} \alpha_2^{\lambda_2} \cdots \alpha_n^{\lambda_n} = 1$ for integers $\lambda_i$ that are not all zero.

Key concept: auxiliary polynomials.

We will need to construct polynomials in several variables with many algebraic roots.
(2) How to construct polynomials with many roots (method 1).

Pigeonhole/Box Principle, Siegel’s Lemma.

(3) How to construct polynomials with many roots (method 2).

Discreteness of algebraic integers.

(4) Extrapolation Method. Baker’s amazing insight using a generalized Schwarz’ Lemma: bootstrap from functions $f$ with many zeroes (but not enough) such that many derivatives of $f$ also have zeroes, to show (using part (2)) that $f$ has more zeroes than we thought.
How to show $\alpha_1^{\lambda_1}\alpha_2^{\lambda_2}\cdots\alpha_n^{\lambda_n} = 1$ for integers $\lambda_i$ that are not all zero.

We will construct a nonzero polynomial

$$f(x_1, \ldots, x_n) = \sum_{\lambda_i=0}^{L} p_{\lambda_1,\ldots,\lambda_n} x_1^{\lambda_1} \cdots x_n^{\lambda_n}$$

such that $f(\alpha_1^z, \ldots, \alpha_n^z) = 0$ for $z = 0, 1, 2, \ldots, (L + 1)^n - 1$. 
There are \((L + 1)^n\) numbers \(\alpha^\lambda = \alpha^\lambda_1 \cdots \alpha^\lambda_n\). Let \(M\) be the matrix whose rows are indexed by \(z = 0, 1, 2, \ldots, (L + 1)^n - 1\) and columns are indexed by the \(\alpha^\lambda\), with corresponding entry \(\alpha^{z\lambda}\).

The equation \(f(\alpha^z_1, \ldots, \alpha^z_n) = 0\) for \(z = 0, 1, \ldots, (L + 1)^n - 1\) says precisely that \(Mv = 0\), where \(v\) is the column vector of the \(p_\lambda\).

Therefore \(\det(M) = 0\). But \(M\) is a Vandermonde matrix. It follows that \(\alpha^\lambda = \alpha^{\lambda'}\) for some distinct \(\lambda, \lambda'\), whence \(\alpha^{\lambda - \lambda'} = 1\).
We are left with finding a polynomial in $n$ variables

$$f(x) = \sum_{\lambda} p_{\lambda} x^{\lambda}$$

such that $\alpha^z$ is a root for lots of integers $z$.

**Step 1:** Express this as a system of linear equations in the $p_{\lambda}$ with integral coefficients.

For each $\alpha_i$, let $c_i$ denote the leading coefficient of the integral minimal polynomial of $\alpha_i$. Then an easy induction shows that for each $j \geq 0$, there exist integers $a_{i,j,s}$ such that

$$\left(c_i \alpha_i\right)^j = \sum_{s=0}^{d-1} a_{i,j,s} \alpha_i^s$$
We then calculate

\[(c_1 \cdots c_n)^L z f(\alpha z) = (c_1 \cdots c_n)^L z \sum_{\lambda} p_\lambda \alpha^\lambda z\]

\[= \sum_{\lambda} p_\lambda c^{Lz - \lambda z} (c\alpha)^\lambda z\]

\[= \sum_{\lambda} p_\lambda c^{Lz - \lambda z} \prod_{i=1}^{n} \sum_{s=0}^{d-1} (\alpha_i)^s a_{i,\lambda_i z,s}\]

\[= \sum_{s_1,s_2,\ldots,s_n=0}^{d-1} \alpha^s \sum_{\lambda} p_\lambda c^{Lz - \lambda z} \prod_{i=1}^{n} a_{i,\lambda_i z,s}\]

Therefore we can force \(f(\alpha z) = 0\) by imposing \(d^n\) linear equations on the \(p_\lambda\) with integer coefficients.
Step 2: Siegel’s Lemma. Let $N > 2M > 0$ be integers, and let $A = (a_{i,j})$ be an $M \times N$ matrix of integers such that $|a_{i,j}| < H$. There is a nonzero vector $b \in \mathbb{Z}^N$ such that $Ab = 0$ and each coordinate of $b$ has absolute value less than $2NH$. 

Proof. Consider all vectors $b \in \mathbb{Z}^N$ with coordinates of absolute value $\leq NH$. There are $(2NH)^N$ such vectors. For each such $b$, each coordinate of $Ab$ has size at most $(NH)^2$.

The total number of possible vectors $Ab$ is $(2(NH)^2)^M$. By Pigeonhole, two distinct $b$ must give the same value of $Ab$. Their difference gives the desired vector.
Basic Principle (Form 1): If $a \in \mathbb{Z}$ and $|a| < 1$ then $a = 0$.

Basic Principle (Form 2): If $a \in \mathbb{Q}$, $da \in \mathbb{Z}$, and $|a| < 1/d$, then $a = 0$.

Basic Principle (Form 3): If $a \in \overline{\mathbb{Q}}$, $da \in \overline{\mathbb{Z}}$, such that

- $|a| < \epsilon$,
- $|\sigma(a)| < M$ for all $\sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$,
- $[\mathbb{Q}(a) : \mathbb{Q}] = n$,
- $\epsilon M^{n-1}d^n < 1$,

then $a = 0$.

**Punch line:** to show that an algebraic number is 0, it suffices to show that it is tiny and that its conjugates are not too big.
(4) EXTRAPOLATION

**Schwarz’ Lemma.** Let \( f: \mathbb{C} \rightarrow \mathbb{C} \) be an entire function with a zero of order \( \geq T \) at \( z = 0 \). Let \( 0 < r < R \). Then

\[
|f|_r \leq \left( \frac{r}{R} \right)^T |f|_R,
\]

where \( |f|_r = \max \{ |f(z)| : |z| < r \} \).

**Punch line:** Roots of the derivative of a function can force good upper bounds. If these are strong enough at certain algebraic arguments, they can force the function to vanish there by the discussion of the previous slide.

If we get more roots of the function and its derivative, we get improved upper bounds, which can then in turn lead to more zeroes.

Bootstrapping in this way, we can hope to repeat the argument enough times to yield enough roots of the function.
**EXTRAPOLATION: BAKER’S LEMMA**

**Lemma.** Let \( f : \mathbb{C} \rightarrow \mathbb{C} \) be an entire function, let \( \epsilon > 0 \), and let \( A, B, C, T, U \) be large positive integers such that

\[
\frac{\epsilon}{4} C > \frac{2T + UAB}{A(\log A)^{1/2}} + \frac{UBA^\epsilon}{\log A}.
\]

Suppose that

- \( |f(z)| \leq e^{T+U}|z| \) for \( z \in \mathbb{C} \).
- \( f^{(t)}(z) = 0 \) for \( t = 0, \ldots, C \) and \( z = 1, 2, \ldots, A \).

Then

\[
|f(z)| \leq e^{-(T+Uz)(\log A)^{1/2}} \quad \text{for} \quad z = 1, \ldots, AB.
\]
The function
\[ h(z) = \frac{g(z)}{(z - 1)^C \cdots (z - A)^C} \]
is entire by the second assumption.

By the maximum modulus principle on the circle of radius \( A^{1+\epsilon}B \) around the origin, we have for \( |z| \leq AB \):
\[ |h(z)| \leq \max_{|w|=A^{1+\epsilon}B} |h(w)|, \]
hence
\[ |f(z)| \leq \max_{|w|=A^{1+\epsilon}B} |f(w)| \cdot \max_{|w|=A^{1+\epsilon}B} \left| \frac{(z - 1)(z - 2) \cdots (z - A)}{(w - 1)(w - 2) \cdots (w - A)} \right|^C, \]
PROOF OF BAKER’S LEMMA

Now
\[
\left| \frac{(z-1)(z-2) \cdots (z-A)}{(w-1)(w-2) \cdots (w-A)} \right| \leq \frac{(AB)^A}{(A^{1+\epsilon/2}B)^A} = e^{-(\epsilon/2)A \log A}.
\]

Meanwhile
\[
|f(w)| \leq e^{T+UA^{1+\epsilon}B}.
\]

Our goal is to show that \( |f(z)| \leq e^{-(T+UA)\log A}^{1/2} \), so it suffices to show that
\[
-(\epsilon/2)AC \log A + (T + UA^{1+\epsilon}B) \leq -(T + UA)\log A^{1/2}.
\]

It is easy to see that this is implied by the assumption of the lemma,
\[
\frac{\epsilon}{2} C > \frac{2T + UAB}{A(\log A)^{1/2}} + \frac{UBA^{\epsilon}}{\log A}.
\]