## 4. Basic Transcendence Theory

Transcendence theory is usually treated in an independent course at Michigan (777 I believe), but it is closely related to diophantine approximation, and indeed it largely grew out of the subject area. Consider the general form in two variables of degree $k$, $f(x,y) = \sum_{j=0}^{k} a_j x^j y^{k-j}$, and suppose that $f$ is irreducible over $\mathbb{Q}$. When $k = 2$, the solubility of the diophantine equation $f(x,y) = c$, where $c \neq 0$, is (relatively) easy. When the form is definite there are obviously only a finite number of solutions. When it is indefinite, if there is one solution, then the form can be can be related to Pell's equation to show that there are infinitely many solutions. The more interesting question is what happens when $k > 2$ and the form is indefinite. Suppose that there are solutions with $x$ and $y$ arbitrarily large. It follows that $a_0 a_k \neq 0$, for otherwise $x$ or $y$ is a divisor of $c$. The polynomial $F(x) = f(x,1)$ has at least one real root, and we can write $f(x,y) = \prod_\alpha (x - y\alpha) g(x,y)$ where the $\alpha$ are the real roots of $F$ and $g$ has no real roots. Since $f$ is irreducible the $\alpha$ are distinct, and so all of the factors on the right here are large except possibly one in the product over $\alpha$. It follows that for that exceptional $\alpha$ we have $x - y\alpha$ about $y^{1-k}$ in size, and so $\alpha - x/y$ is about $y^{-k}$ in size. On the other hand we know that for most real $\alpha$ we cannot expect such good rational approximations, and it was eventually shown by Thue that the rational approximations to algebraic numbers are never this good. Unfortunately, by Thue's method it is not possible to compute the implicit constants in terms of the given algebraic number, or the coefficients of its minimal polynomial, and although there were important contributions by Siegel, Dyson and Roth this flaw remained.

The modern developments in this have been motivated from a different direction, namely the transcendence of various kinds of general classes of numbers, such as $e^\alpha$ where $\alpha$ is an algebraic number other than 0, or $\alpha^\beta$ where $\alpha$ is an algebraic number other than 0 or 1 and $\beta$ is an algebraic irrational. Actually diophantine approximation, or rather diophantine inequalities are still involved. Consider the special case of Thue's equation, $ax^k - by^k = 1$. Then $(a/b)(x/y)^k = 1 + 1/(by^k)$. Taking logs gives $\log(a/b) + k \log(x/y) \ll 1/(by^k)$. Presumably this is impossible when $y$ is large and $k \geq 3$. Thus one sees that lower bounds for linear forms in logarithms could play a rôle. Although this observation is rather naive, it turns out that nevertheless this is the way forward, in the current state of knowledge. To set ideas we first consider the transcendence of $e$.

**Theorem 4.1.** *$e$ is transcendental.*

*Proof.* For a general polynomial $f(x) = \sum_{j=0}^{m} a_j x^j$ with real coefficients and of degree $m$ for a general complex number $t$ let

$$I(t) = \int_0^t e^{t-u} f(u) du$$

where the integration is along the line segment $[0,t]$. Then, by integration by parts, we have

$$I(t) = \sum_{j=0}^{m} \left( e^t f^{(j)}(0) - f^{(j)}(t) \right).$$

Let $\bar{f}(x) = \sum_{j=0}^{m} |a_j| x^j$. Then using the fact that a complex integral is bounded by the length of the path of integration times the supremum of the modulus of the

integrand on the path, we have

$$|I(t)| \leq |t|e^{|t|}\bar{f}(|t|).$$

Now suppose that $e$ is algebraic and that there are integers $q_j$ such that $q_0 + q_1 e + \cdots + q_n e^n = 0$, $q_0 \neq 0$, $n > 0$. Let $p$ be a sufficiently large prime number and now take

$$f(x) = x^{p-1}(x-1)^p \cdots (x-n)^p,$$

so that the degree $m$ of $f$ is $(n+1)p - 1$. Let $J = q_0 I(0) + q_1 I(1) + \cdots + q_n I(n)$. Then, by our initial formula for $I$ we have

$$J = \sum_{i=0}^{n} q_i \left( e^i \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(i) \right).$$

The first inner sums combine to give 0. Thus

$$J = -\sum_{j=0}^{m} \sum_{i=0}^{n} q_i f^{(j)}(i).$$

Plainly $f^{(j)}(i) = 0$ when $j < p$ and $i > 0$, and when $j < p-1$ and $i = 0$. When $j \geq p$ and $i > 0$ we only get a non-zero contribution to $f^{(j)}(i)$ by differentiating the factor $(x-i)^p$ away. Thus $f^{(j)}(i)$ is an integer divisible by $p!$ in this case. Likewise, when $j \geq p$, $f^{(j)}(0)$ we only get a nonzero contribution by differentiating the factor $x^{p-1}$ away, but in addition at least one of the other factors is differentiated. Thus $f^{(j)}(0)$ is also an integer divisible $p!$. Finally $f^{(p-1)}(0) = (p-1)!(-1)^{np}(n!)^p$.

Since $p$ is sufficiently large, and so $p \nmid q_0$ and $p > n$, $J$ is a non-zero integer divisible by $(p-1)!$. On the other hand we have

$$|J| \leq |q_1|e\bar{f}(1) + \cdots + |q_n|ne^n \bar{f}(n) \leq c^p$$

for some constant depending only on the $q_j$ and $n$. This gives a contradiction when $p$ is sufficiently large.

The same ideas can be used to prove that $\pi$ is transcendental, and to prove Lindemann's theorem to the effect that for any distinct algebraic numbers $\alpha_1, \ldots, \alpha_n$ and any non-zero algebraic numbers $\beta_1, \ldots, \beta_n$ the expression $\beta_1 e^{\alpha_1} + \cdots + \beta_n e^{\alpha_n}$ is not 0. The proofs work in precisely the same way, the main difference being in the choice of polynomial $f$. For Lindemann's theorem one chooses $l$ so that $l\alpha_1, \ldots, l\alpha_n, l\beta_1, \ldots, l\beta_n$ are algebraic integers and defines

$$f_i(x) = l^{np}\big((x-\alpha_1)\cdots(x-\alpha_n)\big)^p/(x-\alpha_i),$$

$I_i(t)$ as in the proof of Theorem 4.1 with $f$ replaced by $f_i$, and

$$J_i = \beta_1 I_i(\alpha_1) + \cdots + \beta_n I_i(\alpha_n) \ \ (1 \leq i \leq n).$$

Then one compares estimates for $J_1 \ldots J_n$.

The transcendence of $\pi$ follows from Lindemann's theorem as follows. Suppose that $\pi$ is algebraic. Then so is $i\pi$, and we also have $1 + e^{i\pi} = 0$, but by Lindemann's theorem $1 + e^{i\pi} \neq 0$.

For the rest of this chapter we will concentrate on the ideas introduced by Baker about thirty years ago and which are concerned basically with linear forms in logarithms.

**Theorem 4.2 (Baker 1966).** *Suppose that $\alpha_1, \dots, \alpha_n$ are non-zero algebraic numbers such that $\log \alpha_1, \dots, \log \alpha_n$ are linearly independent over the rationals. Then $1, \log \alpha_1, \dots, \log \alpha_n$ are linearly independent over the field of all algebraic numbers.*

It has to be remarked that the $\alpha_i$ are allowed to be complex numbers, and that at each occurrence any branch of log is allowed.

*Proof.* The proof will be divided into a number of lemmas. The basic structure is that we argue by contradiction. Thus we suppose that the theorem is false, so that there exist algebraic numbers $\beta_0, \beta_1, \dots, \beta_n$ not all 0 such that

$$\beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \alpha_n = 0$$

and we ultimately derive a contradiction. We may certainly suppose that one of the $\beta_i$ is non-zero, so by relabeling we can suppose that $\beta_n \neq 0$. Then each $\beta_j$ can be replaced by $-\beta_j/\beta_n$, so that $\beta_n$ becomes $-1$. We then have

$$e^{\beta_0} \alpha_1^{\beta_1} \dots \alpha_{n-1}^{\beta_{n-1}} = \alpha_n. \tag{4.1}$$

**Lemma 4.1.** *Suppose that $\alpha$ is an algebraic number satisfying*

$$A_0 \alpha^d + A_1 \alpha^{d-1} + \cdots + A_d = 0,$$

*where $A_0, \dots, A_d$ are rational integers with absolute values at most $A$. Then for each non-negative integer $j$ we have*

$$(A_0 \alpha)^j = A_0^{(j)} + A_1^{(j)} \alpha + \cdots + A_{d-1}^{(j)} \alpha^{d-1}$$

*for some rational integers $A_m^{(j)}$ with absolute values at most $(2A)^j$.*

*Proof of Lemma 4.1.* When $0 \leq j < d$, take $A_j^{(j)} = A_0^j$ and $A_m^{(j)} = 0$ otherwise. When $j = d$ the defining equation for $\alpha$ gives a suitable expression and then for $j > d$ observe that

$$\begin{aligned}(A_0 \alpha)^j &= A_0 \alpha \left( A_0^{(j-1)} + A_1^{(j-1)} \alpha + \cdots + A_{d-1}^{(j-1)} \alpha^{d-1} \right) \\ &= A_0^{(j)} + A_1^{(j)} \alpha + \cdots + A_{d-1}^{(j)} \alpha^{d-1}\end{aligned}$$

where we take

$$A_{-1}^{(j-1)} = 0, \quad A_m^{(j)} = A_0 A_{m-1}^{(j-1)} - A_{d-m} A_{d-1}^{(j-1)} \ \ (0 \leq m < d, j \geq d).$$

It follows from the lemma that if $d$ is the maximum of the degrees of

$$\alpha_1, \dots, \alpha_{n-1}, \ \ \beta_0, \dots, \beta_{n-1}$$

and if $a_1, \dots, a_{n-1}, b_0, \dots, b_{n-1}$ are the leading coefficients in their minimal polynomials, then

$$(a_r \alpha_r)^j = \sum_{s=0}^{d-1} a_{rs}^{(j)} \alpha_r^s, \quad (b_r \beta_r)^j = \sum_{t=0}^{d-1} b_{rt}^{(j)} \beta_r^t, \tag{4.2}$$

where the $a_{rs}^{(j)}$, $b_{rt}^{(j)}$ are rational integers with absolute values at most $c^j$ where $c$ is a positive number depending only on the $\alpha$s and $\beta$s.

**Lemma 4.2.** *Let $M$, $N$ denote integers with $N > M > 0$ and let*

$$u_{ij} \ \ (1 \le i \le M, 1 \le j \le N)$$

*denote integers with absolute values at most $U(\ge 1)$. Then there exist integers $x_1, \dots, x_N$ not all 0, with absolute values at most $(NU)^{M/(N-M)}$ such that*

$$\sum_{j=1}^{N} u_{ij} x_j = 0 \ \ (1 \le i \le M).$$

*Proof of Lemma 4.2.* Let $B = [(NU)^{M/(N-M)}]$. There are $(B+1)^N$ different $N$-dimensional vectors $\mathbf{x}$ with integer entries in the range $0 \le x_j \le N$ and for each such $\mathbf{x}$ the vector $\mathbf{y} = U\mathbf{x}$ has each entry $y_i$ in the range

$$-V_i B \le y_i \le W_i B \ \ (1 \le i \le M),$$

where $-V_i$ is the sum of the negative $u_{ij}$ $(1 \le j \le N)$ and $W_i$ is the sum of the positive $u_{ij}$ $(1 \le j \le N)$. We have $V_i + W_i \le NU$. Thus there are at most $(NUB+1)^M$ different values for $\mathbf{y}$. Moreover

$$(NUB+1)^M \le (NUB+NU)^M = (NU)^M (B+1)^M < (B+1)^N.$$

Thus there are two distinct $\mathbf{x}$ which correspond to identical $\mathbf{y}$ and their difference gives their difference gives the required solution.

We now construct a polynomial of several complex variables which has zeros of high order on its "diagonal". For brevity we use the notation

$$f_{m_0,\dots,m_r}(z_0, \dots, z_r) = (\partial/\partial z_0)^{m_0} \cdots (\partial/\partial z_r)^{m_r} f(z_0, \dots, z_r),$$

**Lemma 4.3.** *For any integers $\lambda_j$ $(0 \le j \le n)$ we use the notation $\gamma_r = \lambda_r + \lambda_n \beta_r$ For each positive integer $h$, let $L = [h^{2-1/(4n)}]$. Then provided that $h$ is sufficiently large there are integers $p(\boldsymbol{\lambda})$ $(0 \le \lambda_0 \le L, \dots, 0 \le \lambda_n \le L)$, not all 0, with absolute values at most $e^{h^3}$, such that the function*

$$\Phi(z_0, \dots, z_{n-1}) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\boldsymbol{\lambda}) z_0^{\lambda_0} e^{\lambda_n \beta_0 z_0} \alpha_1^{\gamma_1 z_1} \cdots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}}$$

*satisfies*

$$\Phi_{m_0,\dots,m_{n-1}}(l, \dots, l) = 0$$

*for all integers $l$ with $1 \le l \le h$ and all non-negative integers $m_0, \dots, m_{n-1}$ with $m_0 + \cdots + m_{n-1} \le h^2$.*

*Proof of Lemma 4.3.* We need to find a suitable set of coefficients $p(\boldsymbol{\lambda})$, and this is essentially a problem in linear algebra. The basic tool at our disposal is the previous lemma. The general monomial term in the sum satisfies

$$(\partial/\partial z_0)^{m_0} \cdots (\partial/\partial z_{n-1})^{m_{n-1}} z_0^{\lambda_0} e^{\lambda_n \beta_0 z_0} \alpha_1^{\gamma_1 z_1} \cdots \alpha_{n-1}^{\gamma_{n-1} z_{n-1}}$$
$$= q(\lambda_0, \lambda_n, z_0) e^{\lambda_n \beta_0 z_0} \prod_{j=1}^{n-1} (\gamma_j \log(\alpha_j))^{m_j} \alpha_j^{\gamma_j z_j}$$

where

$$q(\lambda_0,\lambda_n,z)=\sum_{\mu_0=0}^{m_0}\binom{m_0}{\mu_0}\lambda_0(\lambda_0-1)\dots(\lambda_0-\mu_0+1)z^{\lambda_0-\mu_0}(\lambda_n\beta_0)^{m_0-\mu_0}.$$

Discarding the logarithmic factors which are the same in every term we find that our task is to determine integers $p(\boldsymbol{\lambda})$ such that

$$\sum_{\lambda_0=0}^{L}\cdots\sum_{\lambda_n=0}^{L}p(\boldsymbol{\lambda})q(\lambda_0,\lambda_n,l)e^{\lambda_n\beta_0 l}\prod_{j=1}^{n-1}\gamma_j^{m_j}\alpha_j^{\gamma_j l}=0,$$

and by (4.1) this becomes

$$\sum_{\lambda_0=0}^{L}\cdots\sum_{\lambda_n=0}^{L}p(\boldsymbol{\lambda})q(\lambda_0,\lambda_n,l)e^{\lambda_n\beta_0 l}\prod_{j=1}^{n-1}\gamma_j^{m_j}\prod_{j=1}^{n}\alpha_j^{\lambda_j l}=0,$$

holds for all integers $l$ with $1\le l\le h$ and all non-negative integers $m_0,\dots,m_{n-1}$ with $m_0+\cdots+m_{n-1}\le h^2$. We know from our earlier deliberations, summarised in (4.2) that every power of each algebraic number which occurs can be written in terms of lower powers not exceeding the maximal degree, and in each case we have a reasonable estimate for the size of the coefficients which occur. By the binomial theorem

$$\gamma_r^{m_r}=\sum_{\mu_r=0}^{m_r}\lambda_r^{m_r-\mu_r}(\lambda_n\beta_r)^{\mu_r}.$$

Let

$$Q=(a_1\cdots a_n)^{Ll}b_0^{m_0}\cdots b_{n-1}^{m_{n-1}}.$$

We mutliply our system of equations by $Q$ and substite from (4.2) for the powers of $a_r\alpha_r$ and $b_r\beta_r$ which result. This gives

$$\sum_{s_1=0}^{d-1}\cdots\sum_{s_n=0}^{d-1}\sum_{t_0=0}^{d-1}\cdots\sum_{t_{n-1}=0}^{d-1}A(\mathbf{s},\mathbf{t})\alpha_1^{s_1}\cdots\alpha_n^{s_n}\beta_0^{t_0}\cdots\beta_{n-1}^{t_{n-1}}=0,$$

where

$$A(\mathbf{s},\mathbf{t})=\sum_{\boldsymbol{\lambda}}p(\boldsymbol{\lambda})\sum_{\boldsymbol{\mu}}q_1(\boldsymbol{\lambda})q_2(\boldsymbol{\lambda},\boldsymbol{\mu})q_3(\boldsymbol{\lambda},\boldsymbol{\mu}),$$

and

$$q_1(\boldsymbol{\lambda})=\prod_{r=1}^{n}a_r^{(L-\lambda_r)l}a_{r,s_r}^{(\lambda_r l)},$$

$$q_2(\boldsymbol{\lambda},\boldsymbol{\mu})=\prod_{r=1}^{n-1}\left(\binom{m_r}{\mu_r}(b_r\lambda_r)^{m_r-\mu_r}\lambda_n^{\mu_r}b_{r,t_r}^{(\mu_r)}\right),$$

$$q_3(\boldsymbol{\lambda},\boldsymbol{\mu})=\binom{m_0}{\mu_0}\lambda_0(\lambda_0-1)\cdots(\lambda_0-\mu_0+1)\lambda_n^{m_0-\mu_0}b_n^{\mu_0}l^{\lambda_0-\mu_0}b_{0,t_0}^{(m_0-\mu_0)}.$$

The point of the construction so far is that it suffices that, for each $l,m_0,\dots,m_{n-1}$ under consideration we make each of the $d^{2n}$ equations $A(\mathbf{s},\mathbf{t})=0$. This we can

hope to do because they are linear equations in the $p(\boldsymbol{\lambda})$ with *integer* coefficients. Thus it suffices to obtain suitable bounds for the coefficients and then we can appeal to Lemma 4.2. Clearly

$$|q_1(\boldsymbol{\lambda})| \le \prod_{r=1}^{n} c^{(L-\lambda_r)l} c^{\lambda_r l} \le c_1^{Lh},$$

$$|q_2(\boldsymbol{\lambda}, \boldsymbol{\mu})| \le \prod_{r=1}^{n-1} (c_2 L)^{m_r},$$

$$|q_3(\boldsymbol{\lambda}, \boldsymbol{\mu})| \le 2^{m_0} (\lambda_0 b_n)^{m_0} (c_1 \lambda_n)^{m_0 - \mu_0} l^{\lambda_0 - \mu_0} \le (c_3 L)^{m_0} h^L.$$

The inequalities

$$(m_0 + 1) \cdots (m_{n-1} + 1) \le 2^{m_0 + \cdots + m_{n-1}} \le 2^{h^2}$$

tell us that the coefficient of $p(\boldsymbol{\lambda})$ in the linear form $A(\mathbf{s}, \mathbf{t})$ has absolute value at most

$$U = L^{h^2} (c_4)^{h^2 + hL}.$$

The number $M$ of equations does not exceed the number of choices of $l$, $m_0$, $m_1, \ldots, m_{n-1}$ times the number of choices of $\mathbf{s}$, $\mathbf{t}$, and so $M \le h(h^2+1)^n d^{2n}$. The number of unknowns $N$, i.e. the number of choices for $\boldsymbol{\lambda}$, is $N = (L+1)^{n+1}$. Now

$$N > h^{(2-1/(4n))(n+1)} \ge h^{2n+3/2} > 2d^{2n} h(h^2+1)^n \ge 2M$$

since $h$ is large in terms of $d$ and $n$. Thus, at last, we can appeal to Lemma 4.2 and deduce that the equations are indeed soluble with the solving integers $p(\boldsymbol{\lambda})$ having absolute value at most

$$NU \le h^{2n+2} L^{h^2} (c_4)^{h^2 + hL} \le e^{h^3}.$$

**Lemma 4.4.** *Let $m_0, m_1, \ldots, m_{n-1}$ be any non-negative integers with*

$$m_0 + m_1 + \cdots + m_{n-1} \le h^2,$$

*and let*

$$f(z) = \Phi_{m_0, m_1, \ldots, m_{n-1}}(z, \ldots, z).$$

*Then there is a positive number $c$ depending at most on the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that for any number $z$ we have $|f(z)| \le c^{h^3 + L|z|}$ and for any positive integer $l$ we have either $f(l) = 0$ or $|f(l)| > c^{-h^3 - Ll}$.*

*Proof of Lemma 4.4.* We follow the preliminary manipulations in the proof of the previous lemma. We find that the function $f(z)$ is given by

$$A \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n=0}^{L} p(\boldsymbol{\lambda}) q(\lambda_0, \lambda_n, z) \prod_{j=1}^{n-1} \gamma_j^{m_j} \prod_{j=1}^{n} \alpha_j^{\lambda_j z} = 0,$$

where

$$A = \prod_{j=1}^{n-1} (\log \alpha_j)^{m_j}$$

and

$$q(\lambda_0, \lambda_n, z) = \sum_{\mu_0=0}^{m_0} \binom{m_0}{\mu_0} \lambda_0(\lambda_0 - 1) \dots (\lambda_0 - \mu_0 + 1) z^{\lambda_0 - \mu_0} (\lambda_n \beta_0)^{m_0 - \mu_0}.$$

This latter function is bounded by

$$(c_1 L)^{m_0} (|z| + 1)^L \sum_{\mu=0}^{m_0} \binom{m_0}{\mu} = (2c_1 L)^{m_0} (|z| + 1)^L$$

for a suitable positive number $c_1$. We also have

$$\left| \prod_{j=1}^{n} \alpha_j^{\lambda_j z} \right| \leq c_2^{L|z|}$$

and

$$\left| A \prod_{j=1}^{n-1} \gamma_j^{m_j} \right| \leq (c_3 L)^{m_1 + \dots + m_{n-1}}.$$

The number of terms in the multiple sum is $(1 + L)^{n+1}$, and one has the further inequalities

$$L \leq h^2, \;\; m_0 + \dots + m_{n-1} \leq h^2, \;\; |p(\boldsymbol{\lambda})| \leq e^{h^3}.$$

The desired bound for $|f(z)|$ follows from this.

It remains to deal with the lower bound for $|f(l)|$ when this expression is non-zero. Let $Q$ be as in the proof of the previous lemma. Then $g = QA^{-1} f(l)$ is an algebraic integer with degree at most $d^{2n}$. When we substitute arbitrary conjugates for the $\alpha_r$ and $\beta_s$ we see by the above argument that any conjugate of $f(l)$ has absolute value at most $c_4^{h^3 + Ll}$ and a similar bound pertains for $QA^{-1}$. Now the norm of $g$ on the one hand is a non-zero rational integer when $f(l) \neq 0$ and on the other is a product of $g$ and its conjugates and so

$$1 \leq |QA^{-1}||f(l)| c_4^{(h^3 + Ll)(d^{2n} - 1)} \leq |f(l)| c_4^{(h^3 + Ll) d^{2n}}$$

and this gives the desired conclusion.

There is now a further game which can be played. The range for $l$ in Lemma 4.3 can be extended at the expense of greater restrictions on the range for $m_0 + \dots + m_{n-1}$.

**Lemma 4.5.** *Let $J$ be any integer satisfying $0 \le J \le (8n)^2$. Then for all integers $l$ with $1 \le l \le h^{1+J/(8n)}$ and all non-negative integers $m_0, \dots, m_{n-1}$ with $m_0 + \cdots + m_{n-1} \le h^2/2^J$*

$$\Phi_{m_0,\dots,m_{n-1}}(l, \dots, l) = 0$$

*holds.*

*Proof of Lemma 4.5.* The case $J = 0$ is Lemma 4.3. Let $K$ be an integer with $0 \le K < (8n)^2$ and assume that the lemma is valid for $J = 0, 1, \dots, K$. We deduce the proposition for $J = K + 1$.

Let $N_j = [h^{1+j/(8n)}]$ and $M_j = [h^2/2^j]$. Then it suffices to show that for any integer $l$ with $N_K < l \le N_{K+1}$ and any set of non-negative integers $m_0, \dots, m_{n-1}$ with

$$m_0 + \cdots + m_{n-1} \le M_{K+1}$$

we have $f(l) = 0$ where $f$ is as in Lemma 4.4.

The expression $f_m(r)$ is given by

$$\left(\frac{\partial}{\partial z_0} + \cdots + \frac{\partial}{\partial z_{n-1}}\right)^m \Phi_{m_0,\dots,m_{n-1}}(z_0, \dots, z_{n-1})$$

evaluated at the point $z_0 = \cdots = z_{n-1} = r$ and so by

$$\sum_{\mathbf{j}} \frac{m!}{j_0! \dots j_{n-1}!} \Phi_{m_0+j_0,\dots,m_{n-1}+j_{n-1}}(r, \dots, r)$$

where the sum is over all non-negative integers $j_0, \dots, j_{n-1}$ with $j_0 + \cdots + j_{n-1} = m$. By the inductive hypothesis these derivatives are all 0 for $r$, $m$ with $1 \le r \le N_K$, $0 \le m \le M_{K+1}$ since

$$m_0 + j_0 + \cdots m_{n-1} + j_{n-1} \le 2M_{K+1} \le M_K.$$

Let $F(z) = \prod_{q=1}^{N_K} (z - q)^{M_{K+1}}$. Then $f(z)/F(z)$ is analytic on the closed disc $\mathcal{D}$ centre the origin and radius $\rho = N_{K+1} h^{1/(8n)}$. Thus, by the maximum modulus principle,

$$|f(l)/F(l)| \le \theta/\Theta$$

where $\theta$, $\Theta$ denote respectively the least upper bound for $|f(z)|$ and the greatest lower bound for $|F(z)|$ on the boundary of $\mathcal{D}$. A fairly crude lower bound for $\Theta$ is $(\frac{1}{2}\rho)^{N_K M_{K+1}}$, and Lemma 4.4 gives the upper bound $\theta \le c^{h^3 + L\rho}$. We also have $F(l) \le N_{K+1}^{N_K M_{K+1}}$ and, again by Lemma 4.4, $f(l) = 0$ or $|f(l) > c^{-h^3 - L\rho}$. However the latter possibility would give

$$\left(\frac{1}{2} h^{1/(8n)}\right)^{N_K M_{K+1}} \le (c^2)^{h^3 + L\rho}$$

and as $K < (8n)^2$ and $L\rho \le h^{3+K/(8n)} \le 2^{K+3} N_K M_{K+1}$ this inequality is impossible when $h$ is sufficiently large. Thus $f(l) = 0$ and this completes the proof of the inductive step and hence the lemma.

It is also useful to know something about the size at the origin of higher derivatives of $\Phi$, and the information we require is embodied in the next lemma.

**Lemma 4.6.** *Let $\phi(z) = \Phi(z, \dots, z)$. Then for $j = 0, \dots, h^{8n}$ we have $|\phi_j(0)| < \exp\left(-h^{8n}\right)$.*

*Proof of Lemma 4.6.* Let $X = h^{8n}$ and $Y = [h^2/2^{(8n)^2}]$. By the previous lemma,

$$\Phi_{m_0,\dots,m_{n-1}}(l, \dots, l) = 0$$

for all positive integers $l \leq X$ and all non-negative integers $m_0, \dots, m_{n-1}$ with $m_0 + \cdots + m_{n-1} \leq Y$. Now $\phi_m(r)$ is

$$\left(\frac{\partial}{\partial z_0} + \cdots + \frac{\partial}{\partial z_{n-1}}\right)^m \Phi(z_0, \dots, z_{n-1})$$

evaluated at the point $z_0 = \cdots = z_{n-1} = r$ and so by

$$\sum_{\mathbf{j}} \frac{m!}{j_0! \dots j_{n-1}!} \Phi_{j_0,\dots,j_{n-1}}(r, \dots, r)$$

where the sum is over all non-negative integers $j_0, \dots, j_{n-1}$ with $j_0 + \cdots + j_{n-1} = m$. Thus $\phi_m(r) = 0$ for all positive integers $r$ with $r \leq X$ and all non-negative integers with $m \leq Y$. It follows that $\phi(z)/E(z)$, where $E(z) = ((z-1)\cdots(z-X))^Y$, is analytic within and on the closed disc $\mathcal{D}$ centre the origin and radius $\rho = Xh^{1/(8n)}$. Thus, by the maximum modulus principle, we have, for any $w$ with $|w| < X$, the inequality $|\phi(w)| \leq \xi\Xi^{-1}|E(w)|$ where $\xi$ and $\Xi$ denote the respective upper bound for $|\phi(z)|$ and lower bound for $|E(z)|$ on the boundary $\mathcal{D}$. It is immediate that $|E(w)| \leq (2X)^{XY}$ and $\Xi \geq (\frac{1}{2}h^{1/(8n)})^{XY}$. Moreover, by Lemma 3, we have $\xi \leq c^{h^3+L\rho}$. Hence $|\phi(w)| \leq c^{h^3+L\rho}(\frac{1}{4}h^{1/(8n)})^{-XY}$. We also have $L\rho \leq h^{8n+2} \leq 2^{(8n)^2+1}XY$. Thus $|\phi(w)| < e^{-XY}$. Now we appeal to Cauchy's Integral Formula

$$\phi_j(0) = \frac{j!}{2\pi i}\int_{\mathcal{C}} \frac{\phi(w)}{w^{j+1}}dw,$$

where $\mathcal{C}$ denotes the circle $|w| = 1$ described in the positive sense. Thus the formula on the right is bounded by $j^j e^{-XY}$ from which the desired estimate follows.

We can now begin to say something useful about linear forms in logarithms. The following estimate is comparatively weak, but will be a necessary part of the proof.

**Lemma 4.7.** *There is a positive number $c$ depending at most on the $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that for any integers $t_1, \dots, t_n$, not all 0, and with absolute values at most $T$, we have*

$$|t_1 \log \alpha_1 + \cdots t_n \log \alpha_n| > c^{-T}.$$

*Proof of Lemma 4.7.* Let $a_j$ for $j = 1, \dots, n$ be the leading coefficients of the minimal defining polynomials of $\alpha_j$ or $\alpha_j^{-1}$ according as $t_j \geq 0$ or $t_j < 0$. Then

$$\omega = a_1^{|t_1|} \cdots a_n^{|t_n|}\left(\alpha_1^{t_1} \cdots \alpha_n^{t_n} - 1\right)$$

is an algebraic integer with degrees at most $d^n$, and any conjugate of $\omega$, obtained by substituting arbitrary conjugates of $\alpha_1, \dots, \alpha_n$ has absolute value at most $c'^T$. If $\omega = 0$, then

$$\Omega = t_1 \log \alpha_1 + \cdots + t_n \log \alpha_n$$

is a multiple of $2\pi i$, and in fact is a non-zero multiple since the $\log \alpha_1, \dots, \log \alpha_n$ are, by hypothesis, linearly independent over the rationals. Hence in this case the lemma is immediate. Otherwise, the norm of $\omega$ has absolute value at least 1 and so $|\omega| \geq c'^{-Td^n}$. For any complex $z$ we have $|e^z - 1| \leq |z|e^{|z|}$. Thus $|\omega| \leq |\Omega|e^{|\Omega|} \leq |\Omega|(c'')^T$. Therefore, assuming as we may that $|\Omega| < 1$, the lemma follows once more.

We now begin to collect together the different threads of the proof. There is still an auxiliary result which we will need but we can defer its discussion until we need it. The inequalities of Lemma 4.6 furnish the principal input at this stage. Recall that $\phi(z) = \Phi(z, \dots, z)$ and that

$$\Phi(z_0, \dots, z_{n-1}) = \sum_{\lambda_0=0}^{L} \cdots \sum_{\lambda_n}^{L} p(\lambda_0, \dots, \lambda_n) z_0^{\lambda_0} e^{\lambda_n \beta_0 z_0} \alpha_1^{(\lambda_1+\lambda_n\beta_1)z_1} \cdots \alpha_{n-1}^{(\lambda_{n-1}+\lambda_n\beta_{n-1})z_{n-1}}.$$

We also know that the $p(\boldsymbol{\lambda})$ are integers, not all zero. The idea now is to pick out one which is non-zero, so that it will have modulus at least 1, and to write it as a combination of the other terms and then to use our bounds on $\phi_j$ to obtain a contradiction. From a technical point of view it is simpler to first of all rewrite the terms so that $\phi$ really does look like a function of the single variable $z$. One way of organizing this is as follows. Let $S = L + 1$, $R = S^n$. Then any integer $i$ with $0 \leq i < RS$ can be written uniquely in the form

$$i = \lambda_0 + \lambda_1 S + \cdots + \lambda_n S^n$$

with each $\lambda_j$ an integer in [0,L]. For each such $i$ we define $\nu_i = \lambda_0$ and $\mu_i = \lambda_1 \log \alpha_1 + \cdots + \lambda_n \log \alpha_n$, $p_i = p(\boldsymbol{\lambda})$ and obtain

$$\phi(z) = \sum_{i=0}^{RS-1} p_i z^{\nu_i} e^{\mu_i z}.$$

The idea now is to use an interpolating polynomial to pick out a term with a non-zero coefficient. Suppose that $t$ is a summation index for which $p_t \neq 0$, and let $s$ denote the corresponding exponent $\nu_t$. Now suppose that we can find complex numbers $w_j$ such that the polynomial

$$W(z) = \sum_{j=0}^{RS-1} w_j z^j$$

satisfies $W_s(\mu_t) = 1$ and $W_j(\mu_i) = 0$ for all other $i$ and $j$ with $0 \leq i < R$, $0 \leq j < S$. Then

$$p_t = \sum_{i=0}^{RS-1} p_i W_{\nu_i}(\mu_i)$$

and, by Leibnitz's theorem

$$W_{\nu_I}(\mu_i) = \sum_{j=0}^{RS-1} j(j-1)\cdots(j-\nu_i+1) w_j \mu_i^{j-\nu_i} = \sum_{j=0}^{RS-1} w_j \left[\frac{d^j}{dz^j}(z^{\nu_i} e^{\mu_i z})\right]_{z=0}.$$

When we multiply this expression by $p_i$, sum over the $i$ and take this summation inside the sum over $j$ we obtain

$$p_t = \sum_{j=0}^{RS-1} w_j \phi_j(0).$$

We defer the construction of the polynomial $W$ until later, but we record one further piece of information for use when we construct the polynomial. The numbers $\mu_i$ are the numbers $\lambda_1 \log \alpha_1 + \cdots + \lambda_n \log \alpha_n$. By Lemma 4.7 we have $|t_1 \log \alpha_1 + \cdots t_n \log \alpha_n| > c^{-T}$ whenever the integers $t_i$ are not all 0 and are all bounded by $T$. It follows, therefore, that the $R$ numbers $\lambda_1 \log \alpha_1 + \cdots + \lambda_n \log \alpha_n$ are bounded below by $c^{-L}$ and are all distinct. (Note that the number of indices $i$ is $RS$ and that each $\mu_i$ takes on the same value exactly $S$ times, but, of course, when $\mu_i = \mu_j$ we have $\nu_i \neq \nu_j$). Now we just need a suitable bound for the $w_j$. In fact we will construct a polynomial $W$ in which the $w_j$ are bounded by $(8\mu/\rho)^{RS}$ where $\mu$ is the maximum of 1 and all the $\mu_i$ and $\rho$ is the minumum of 1 and all the $\mu_i - \mu_j$ with $\mu_i \neq \mu_j$. Thus

$$w_j \leq (cLc^L)^{RS} \leq c'^{h^{2n+4}}.$$

Now we can complete the proof of Baker's theorem, modulo the construction of the polynomial $W$. We have

$$1 \leq |p_t| \leq RS \max_j(|w_j \phi_j(0)|) \leq c h^{2n+2} c^{h^{2n+4}} \exp(-h^{8n})$$

and this gives the desired contradiction by taking $h$ sufficently large.

It now remains to construct the interpolating polynomial $W$. The following lemma is what we want. It is farily routine.

**Lemma 4.8.** *Let $R$, $S$ be positive integers and let $\sigma_0, \ldots, \sigma_{R-1}$ be $R$ distinct complex numbers. Define $\sigma$ as the maximum of 1, $|\sigma_0|, \ldots, |\sigma_{R-1}|$ and define $|\rho|$ as the minimum of 1 and the $|\sigma_i - \sigma_j|$ with $0 \leq i < j < R$. Then, for any integers $r$, $s$ with $0 \leq r < R$, $0 \leq s < S$, there exist complex numbers $w_i$ $(0 \leq i < RS)$ with absolute values at most $(8\sigma/\rho)^{RS}$ such that the polynomial*

$$W(z) = \sum_{j=1}^{RS-1} w_j z^j$$

*satisfies $W_j(\sigma_i) = 0$ for all $i$, $j$ with $0 \leq i < R$, $0 \leq j < S$ other than $i = r$, $j = s$, and $W_s(\sigma_r) = 1$.*

*Proof of Lemma 4.8.* The required polynomial is given by

$$W(z) = \left(\frac{-1}{s!}\right) \frac{1}{2\pi i} \int_{\mathcal{C}_r} \frac{(\zeta - \sigma_r)^s U(z)}{(\zeta - z) U(\zeta)} d\zeta,$$

where $U(z) = \big((z - \sigma_0) \cdots (z - \sigma_{R-1}\big)^S$ and $\mathcal{C}_r$ denotes a circle described in the positive sense with centre $\sigma_r$ and sufficiently small radius, less than, say, $\rho$ and $|z - \sigma_r|$ for $z \neq \sigma_r$. The proof depends on two alternative expressions for the

polynomial $W(z)$, one of which establishes the desired values and the other of which enables one to bound its coefficients.

First, since the absolute value of the integrand multiplied by $\zeta$ decreases to 0 as $|\zeta| \to \infty$ Cauchy's residue theorem gives

$$W(z) = \frac{(z-\sigma_r)^s}{s!} + \frac{U(z)}{s!}\frac{1}{2\pi i}\sum_{\substack{j=0\\ j\neq r}}^{R-1}\int_{\mathcal{C}_j}\frac{(\zeta-\sigma_r)^s}{(\zeta-z)U(\zeta)}d\zeta,$$

where $\mathcal{C}_j$, like $\mathcal{C}_r$ above, is a circle about $\sigma_j$ with sufficiently small radius.

The sum over $j$ is a rational function of $z$, analytic at $z = \sigma_r$ and, because $U(z)$ has a zero at $z = \sigma_r$ of order $S$, it follows that $W_j(\sigma_r) = 1$ when $j = s$ and is 0 otherwise.

On the other hand, by Cauchy's Integral Formula,

$$W(z) = \frac{-1}{s!t!}\left(\frac{d^t}{d\zeta^t}\frac{(\zeta-\sigma_r)^S U(z)}{(\zeta-z)U(\zeta)}\right)\Bigg|_{\zeta=\sigma_r},$$

where $t = S - s - 1$. Thus

$$W(z) = (-1)^{t-1}(s!)^{-1}U(z)\sum v(j_0,\dots,j_{R-1})(\sigma_r - z)^{-j_r-1}$$

where the sum is over all non-negative integers $j_0,\dots,j_{R-1}$ with $j_0+\cdots+j_{R-1} = t$, and

$$v(j_0,\dots,j_{R-1}) = \prod_{\substack{i=0\\ i\neq r}}^{R-1}\binom{S+j_i-1}{j_i}(\sigma_r-\sigma_i)^{-S-j_i}.$$

Now $j_r + 1$ lies between 1 and $S$ inclusively and so obviously $W(z)$ is a polynomial of degree at most $RS - 1$. Also $W(z)$, like $U(z)$, has a zero of order $S$ at $\sigma_i$ when $i \neq r$. Thus $W_j(\sigma_i) = 0$ for all $j < S$. Moreover, the typical factor in the product defining $v$ has absolute value not exceeding $2^{S+j_i-1}\rho^{-S-j_i}$, and so

$$|v(j_0,\dots,j_{R-1})| \leq (2/\rho)^{(R-1)+j_0+\cdots+j_{R-1}} \leq (2/\rho)^{RS}.$$

The coefficients of $(\sigma_r - z)^{j_r-1}U(z)$ have absolute values at most $(\sigma+1)^R S$ and the number of terms in the above sum does not exceed $S^R$. Thus the coefficients of $W(z)$ have absolute values at most

$$S^R(\sigma+1)^{RS}(2/\rho)^{RS} \leq (8\sigma/\rho)^{RS},$$

and this completes the proof of the lemma.

The most important aspect of Baker's work is that it can be quantified. By some elaboration (!) and due care to explicit expressions he was able to obtain in 1966 a very useful lower bound for linear forms in logarithms. This was taken up by others and within a short period an essentially best possible bound was abtained by Feldman.

**Theorem 4.3 (Feldman).** *Let $\alpha_1, \dots, \alpha_n$ be non-zero algebaric numbers with degrees at most $d$ and heights at most $A$. Further let $\beta_0, \dots, \beta_n$ be algebraic numbers with degrees at most $d$ and heights at most $B$ where $B \geq 2$. Let*

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n.$$

*Then either $\Lambda = 0$ or there is an effectively computable number $C$ depending at most on $n$, $d$, $A$ and the original determination of the logarithms such that $|\Lambda| > B^{-C}$.*

There are many striking applications of this theorem. For example there is an effective treatment of the Thue equation.

**Theorem 4.4.** *Let $K$ be an algebraic number field with degree $d$, let $\alpha_1, \dots, \alpha_n$ be $n \geq 3$ distinct algebraic integers in $K$, and let $\mu$ be any non-zero algebraic integer in $K$. Then the equation*

$$(x - \alpha_1 y) \cdots (x - \alpha_n y) = \mu$$

*has only a finite numbe of solutions in algebraic integers $x$ and $y$ in $K$ and these can be effectively determined.*

I will not give the proof, but one key ingredient is as follows. Let $\beta_i = x - \alpha_i y$, $\eta_1, \dots, \eta_r$ be a suitable set of units in $K$, let $\gamma_i = \beta_i \eta_1^{a_{i1}} \cdots \eta_r^{a_{ir}}$ where the $a_{ij}$ are suitable integers, and let $\alpha$ be a suitable expression of the form $\frac{(\alpha_j - \alpha_l)\gamma_k}{(\alpha_k - \alpha_l)\gamma_j}$. Then one can apply the previous theorem to give a lower bound for

$$|b_1 \log \eta_1 + \cdots b_r \log \eta_r - \log \alpha - b_0 \log(-1)|$$

where the $b_j$ are rational integers and this leads to an upper bound for the $\beta_j$ and hence the $x$ and $y$.