PURE MATH 944 - DIOPHANTINE APPROXIMATION

Theorem 0.1. (Dirichlet's Theorem) *Let α be a real irrational number, and let $n \in \mathbb{N}$ be a natural number. Then there exist integers p, q with $1 \leq q \leq n$ such that*

$$|q\alpha - p| < \frac{1}{n+1}$$

Proof. Clearly, we may assume that $\alpha > 0$. For $q = 1, \cdots, n$, write $r_q = q\alpha - \lfloor q\alpha \rfloor$. Then the $n+2$ numbers $0, r_1, \cdots, r_n, 1$ (since α is irrational, we have $r_j \neq 0, 1$ for all j) all lie in $[0,1]$ and by the pigeonhole principle, some two of them differ by at most $\dfrac{1}{n+1}$. If there is some r_q such that $|r_q - 1| < 1/(n+1)$ or $|r_q| < 1/(n+1)$ then we are done. Otherwise, there are $1 \leq s, t \leq n$ such that $|r_s - r_t| < 1/(n+1)$. The result follows by noting that $r_s - r_t = r_{s-t}$ if $s > t, r_s > r_t$ and $r_s - r_t = 1 - r_{s-t}$ if $s < t, r_s > r_t$.

□

Theorem 0.2. (Duffin-Schaeffer Theorem) *There exists a sequence of non-negative real numbers $f(1), f(2), \cdots$, such that $\sum_{q=1}^{\infty} f(q) = \infty$, but nonetheless for almost all real α the inequality*

$$\left|\alpha - \frac{p}{q}\right| < \frac{f(q)}{q}$$

has only finitely many solutions for integers p, q.

Proof. Since $\prod_p \left(1 + \frac{1}{p}\right)$ diverges, there exists a strictly increasing sequence $(x_n)_{n=0}^{\infty}$ with $x_0 = 1$ such that $\prod_{x_{i-1} < p \leq x_i} \left(1 + \frac{1}{p}\right) > 2^i + 1$ for all $i \geq 1$. Define $N_i = \prod_{x_{i-1} < p \leq x_i} p$. Note that by construction we have $\gcd(N_i, N_j) = 1$ if $i \neq j$. Now define $f(q)$ to be $2^{-i}\dfrac{q}{N_i}$ if $q | N_i$ and 0 otherwise. Now we define

$$A_q = \left[0, \frac{f(q)}{q}\right] \cup \bigcup_{j=1}^{q-1} \left[\frac{j}{q} - \frac{f(q)}{q}, \frac{j}{q} + \frac{f(q)}{q}\right] \cup \left[1 - \frac{f(q)}{q}, 1\right].$$

Note that the measure of A_q is zero unless $q | N_i$ for some i, and $\mu(A_q) \leq q\left(\dfrac{f(q)}{q}\right)$ otherwise. Also note that $A_q \subset A_{N_i}$ and in fact we have

$$A_{N_i} = \bigcup_{q | N_i} A_q.$$

Therefore, since $\mu(A_{N_i}) \leq 2N_i \left(2^{-i}\frac{q}{N_i q}\right) = 2^{-i}$, it follows that

$$\mu\left(\bigcup_{q|N_i} A_q\right) \leq 2^{-i+1}.$$

Now let A be the set of real numbers $\alpha \in [0,1]$ for which the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{f(q)}{q}$$

has infinitely many solutions in integers p, q. Since only finitely many q's divide N_i for any i, it follows that for any $k_0 \in \mathbb{N}$ we have $A \subset \bigcup_{k=k_0}^{\infty}\left(\bigcup_{q|N_k} A_q\right) = \bigcup_{k=k_0}^{\infty} A_{N_i}$. By sub-additivity of measures it follows that $\mu(A) \leq \sum_{k=k_0}^{\infty} 2^{-k+1} = 2^{-k_0+2}$. In particular, letting $k_0 \to \infty$ we see that $\mu(A) = 0$. On the other hand, we have

$$\sum_{q=1}^{\infty} f(q) = \sum_{i=1}^{\infty} 2^{-i} \sum_{q|N_i, q>1} \frac{q}{N_i}.$$

Note that

$$\sum_{q|N_i, q>1} \frac{q}{N_i} = \frac{1}{N_i}\left(\prod_{p|N_i}(1+p) - 1\right)$$
$$= \prod_{p|N_i}\left(1 + \frac{1}{p}\right) - \frac{1}{N_i} > 2^i + 1 - \frac{1}{N_i} > 2^i$$

by our choice of N_i. Hence, we have

$$\sum_{q=1}^{\infty} f(q) \geq \sum_{i=1}^{\infty} 2^{-i}2^i = \infty.$$

This establishes the existence of a sequence asserted by the theorem. $\square$

If f is as above and we consider the sum $\sum_{q=1}^{\infty}\frac{f(q)\varphi(q)}{q}$ we would obtain

$$\sum_{q=1}^{\infty}\frac{f(q)\varphi(q)}{q} = \sum_{i=1}^{\infty} 2^{-i}\frac{1}{N_i}\sum_{q|N_i, q>1}\varphi(q) = \sum_{i=1}^{\infty} 2^{-i}\frac{N_i - 1}{N_i} < \infty.$$

To investigate the issue further, we will require some further results on the Euler φ function.

Proposition 0.3. *Let $m, n \in \mathbb{Z}^+$. If $\gcd(n,m) = 1$, then $\varphi(nm) = \varphi(n)\varphi(m)$. In other words, φ is a multiplicative function.*

Proof. We have $\varphi(mn) = mn \prod_{p|mn} \left(1 - \frac{1}{p}\right)$, and since $\gcd(m,n) = 1$ it follows that $\prod_{p|mn} \left(1 - \frac{1}{p}\right) = \prod_{p|m} \left(1 - \frac{1}{p}\right) \prod_{p|n} \left(1 - \frac{1}{p}\right)$ and hence $\varphi(mn) = \varphi(m)\varphi(n)$. $\square$

Proposition 0.4. *We have* $\sum_{d|n} \varphi(d) = n$ *for all* $n \in \mathbb{N}$.

Proof. Write C_d to be the subset of $1 \le m \le n$ such that $\gcd(m,n) = d$. Clearly $C_d = \emptyset$ if d does not divide n. Otherwide, if $m \in C_d$, then $\gcd\left(\frac{m}{d}, \frac{n}{d}\right) = 1$, so that $|C_d| = \varphi\left(\frac{n}{d}\right)$. Hence we have

$$n = \sum_{d|n} |C_d| = \sum_{d|n} \varphi\left(\frac{n}{d}\right) = \sum_{d|n} \varphi(d).$$

$\square$

Remark 0.5. By the Mobius inversion formula, we also have $\varphi(n) = n \sum_{d|n} \frac{\mu(d)}{d}$.

Proposition 0.6. *We have* $\sum_{m=1}^{n} \varphi(m) = \frac{3}{\pi^2} n^2 + O(n \log n)$.

Proof. We have

$$\begin{aligned}
\sum_{m=1}^{n} \varphi(m) &= \sum_{m=1}^{n} \sum_{d|m} \frac{m\mu(d)}{d} = \sum_{dd' \le n} d' \mu(d) \\
&= \sum_{d=1}^{n} \mu(d) \sum_{d'=1}^{\lfloor \frac{n}{d} \rfloor} d' \\
&= \sum_{d=1}^{n} \mu(d) \left(\frac{1}{2} \left(\left\lfloor \frac{n}{d} \right\rfloor^2 + \left\lfloor \frac{n}{d} \right\rfloor \right) \right) \\
&= \frac{1}{2} \sum_{d=1}^{n} \mu(d) \left(\frac{n^2}{d^2} + O\left(\frac{n}{d}\right) \right) \\
&= \frac{n^2}{2} \sum_{d=1}^{n} \frac{\mu(d)}{d^2} + O\left(n \sum_{d=1}^{n} \frac{1}{d} \right) \\
&= \frac{n^2}{2} \sum_{d=1}^{n} \frac{\mu(d)}{d^2} + O(n \log n) \\
&= \frac{n^2}{2} \left(\sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} - \sum_{d=n+1}^{\infty} \frac{\mu(d)}{d^2} \right) + O(n \log n) \\
&= \frac{n^2}{2} \prod_{p} \left(1 - \frac{1}{p^2}\right) + O(n \log n) = \frac{3}{\pi^2} n^2 + O(n \log n).
\end{aligned}$$

□

For $n \in \mathbb{N}$, let $\tau(n)$ denote the number of positive divisors of n.

Proposition 0.7. *Let n be a positive integer and let u and v be integers with $v > 0$. Then*

$$\left| \sum_{\substack{u<k\le u+v \\ \gcd(k,n)=1}} 1 - v\frac{\varphi(n)}{n} \right| \le \tau(n).$$

Proof.

$$\begin{aligned}
\left| \sum_{\substack{u<k\le u+v \\ \gcd(k,n)=1}} 1 - v\frac{\varphi(n)}{n} \right| &= \left| \sum_{u<k\le u+v} \sum_{d \mid \gcd(k,n)} \mu(d) - v\frac{\varphi(n)}{n} \right| \\
&= \left| \sum_{u<k\le u+v} \sum_{d \mid (k,n)} \mu(d) - v\sum_{d\mid n} \frac{\mu(d)}{d} \right| \\
&= \left| \sum_{d\mid n} \mu(d) \sum_{\substack{u<k\le u+v \\ d\mid k}} 1 - \sum_{d\mid n} \mu(d)\frac{v}{d} \right| \\
&= \left| \sum_{d\mid n} \mu(d) \left(\sum_{\substack{u<k\le u+v \\ d\mid k}} 1 - \frac{v}{d} \right) \right| \\
&\le \sum_{d\mid n} 1 = \tau(n).
\end{aligned}$$

Now note that $\tau(n) \le 2n^{1/2}$ since if d is a divisor of n then either d or n/d is bounded above by $n^{1/2}$. □

Note that for any $\varepsilon > 0$, we have $n^{1-\varepsilon} < \varphi(n) < n$ for n sufficiently large. In particular, we have the following corollary.

Corollary 0.8. *Let $\varphi_\lambda(n)$ be the number of positive integers m with $m \le \lambda n$ with* $\gcd(m, n) = 1$. *Then*

$$|\varphi_\lambda(n) - \lambda\varphi(n)| \le 2n^{1/2}.$$

In particular, we have $\varphi_\lambda(n) = \varphi(n)(\lambda + \rho)$ with $|\rho| \le cn^{-1/4}$ for some $c > 0$.

Proof. Follows immediately from previous propositions. □

Proposition 0.9. *Let N and M be positive integers and $A > 0$ be a positive real number. The number of positive integer pairs (x, y) with $0 < |Nx - My| \le A$ and with $1 \le x \le M$, $1 \le y \le N$ is at most $2A$.*

Proof. Let $d = \gcd(N, M)$ and write $N_1 = \frac{N}{d}, M_2 = \frac{M}{d}$. It suffices to count the number of pairs of positive pairs of integers (x, y) for which $|N_1 x - M_1 y| \le \frac{A}{d}$ where $1 \le x \le M_1 d$ and $1 \le y \le N_1 d$. Call such a pair (x, y) and admissible pair.

Suppose $x_1 N_1 - y_1 M_1 = x_2 N_1 - y_2 M_1$, with $(x_1, y_1), (x_2, y_2)$ admissible pairs. Then $(x_1 - x_2) N_1 = (y_1 - y_2) M_1$. Since $\gcd(N_1, M_1) = 1$, it follows that $x_1 - x_2, y_1 - y_2$ are multiples of M_1, N_1 respectively.

If h is an integer with $|h| \le \frac{A}{d}$ and $x_1 N_1 - y_1 M_1 = h$, we have that there are at most d solutions in admissible pairs (x, y). To see this, the pair is determined by x, and since for any two distinct solutions x_1, x_2 they must lie in the same congruence class modulo M_1, the number of solutions correspond to the number of such congruence classes in the set $\{1, \cdots, M_1 d\}$, which is at most d. Hence the number of admissible pairs is at most $2d \left\lfloor \frac{A}{d} \right\rfloor \le 2A$ and we are done. □

Theorem 0.10. (Khintchine's Theorem - 1924) *Let $f : \mathbb{R} \to \mathbb{R}^+$ and let $(f(q))_{q=1}^{\infty}$ be a sequence of positive numbers for which*

(i) $\sum_{q=1}^{\infty} f(q) = \infty$,
(ii) The sequence $(qf(q))_{q=1}^{\infty}$ is a decreasing sequence.

Then for all real numbers α with the exception of a set of Lebesgue measure zero, there exist infinitely many rationals p/q for which

$$\left| \alpha - \frac{p}{q} \right| < \frac{f(q)}{q}.$$

For example, the theorem applies to the sequence $f(q) = \frac{1}{q \log q}$ or even $\frac{1}{q \log q \log \log q}$.

Condition (ii) is a stringent one but as the previous Duffin-Schaeffer theorem indicates, some such condition is necessary.

We will derive Khintchine's theorem from the following result, also due to Duffin-Schaeffer.

Theorem 0.11. (Duffin and Schaeffer) *Let $(f(q))_{q=1}^{\infty}$ be a sequence of non-negative real numbers which satisfies*

(i) $\sum_{q=1}^{\infty} f(q) = \infty$,
(ii) $0 \le f(q) \le 1/2$ *for* $q \ge 1$, *and*
(iii) There exist a positive number c such that $\sum_{q=1}^{\infty} f(q) \frac{\varphi(q)}{q} > c \sum_{q=1}^{n} f(q)$ *for infinitely*

many integers n.

Then for all real numbers α, except for a set of Lebesgue measure zero, there exist infinitely many rationals p/q such that $\left|\alpha - \frac{p}{q}\right| < \frac{f(q)}{q}$.

We will introduce some definitions and propositions to establish the theorem.

Definition 0.12. Let θ be a positive real number with $\theta \leq 1/2$, and let $q > 1$ be a positive integer. Denote $E_q^\theta \subset (0,1)$ consisting of $\varphi(q)$ intervals centered at p/q with $\gcd(p,q) = 1$ of radius θ/q.

Proposition 0.13. *Let μ denote the Lebesgue measure of $\mathbb{R}$. Suppose $q, n > 1$ are distinct integers and θ_1, θ_2 are real numbers with $0 \leq \theta_1, \theta_2 \leq 1/2$. Then $\mu(E_n^{\theta_1} \cap E_q^{\theta_2}) \leq 8\theta_1\theta_2$.*

Proof. If an interval $I_1 \subset E_q^{\theta_1}$ overlaps an interval $I_2 \subset E_n^{\theta_2}$ with center m/n, then $0 < \left|\frac{p}{q} - \frac{m}{n}\right| < \frac{\theta_1}{q} + \frac{\theta_2}{n}$ or equivalently, $0 < |np - mq| < \theta_1 n + \theta_2 q$. First suppose that $\theta_1 n \geq \theta_2 q$, so $0 < |np - mq| < 2\theta_1 n$. By proposition 0.9 there are at most $4\theta_1 n$ such solutions.

Therefore $\mu(E_q^{\theta_1} \cap E_n^{\theta_2}) \leq 4\theta_1 n \left(\frac{2\theta_2}{n}\right) = 8\theta_1\theta_2$. Symmetrically, the same arguments hold when $\theta_2 q \geq \theta_1 n$. □

Proposition 0.14. *Let A be a subset of $(0,1)$ consisting of a finite union of intervals. There exists a positive number c, which depends on A, such that if $n > 1$ and $0 < \theta \leq 1/2$, then $\mu(A \cap E_n^\theta) \leq \mu(A)\mu(E_n^\theta)(1 + cn^{-1/4})$.*

Proof. We first prove the result in the case when A is a single interval $(a,b) \subset (0,1)$. The number of intervals in E_n^θ whose centers lie in $(a,b]$ is $\varphi_b(n) - \varphi_a(n)$. Thus the number of intervals of E_n^θ lying entirely in $(a,b]$ is at least $\varphi_b(n) - \varphi_a(n) - 2$. Further, the number of intervals which have some overlap with $(a,b]$ is at most $\varphi_b(n) - \varphi_a(n) + 2$. Thus $\mu(A \cap E_n^\theta) = (\varphi_b(n) - \varphi_a(n) + \gamma)\frac{2\theta}{n}$ where $|\gamma| \leq 2$ is a real number.

By corollary 0.8, we get that

$$\mu(A \cap E_n^\theta) \leq \varphi(n)((b-a) + c_1 n^{-1/4})\frac{2\theta}{n} = \mu(E_n^\theta)\mu(A)(1 + c(A)n^{-1/4}),$$

where $c(A)$ is a constant that depends on A.

Now suppose that A is the union of k disjoint intervals $A_1, \cdots, A_k$. Then put $c = \max(c(A_1), \cdots, c(A_k))$. Then we have

$$\mu(A \cap E_n^\theta) = \mu\left(\left(\bigcup_{i=1}^k A_i\right) \cap E_n^\theta\right) = \mu(E_n^\theta)\mu(A)(1 + cn^{-1/4}).$$

□

Proof. proof of Duffin-Schaeffer theorem Set $f(q) = \theta_q$, for $q = 1, 2, \cdots$. Denote the sets $E_q^{\theta_q}$ by just E_q for brevity. Put $E = \bigcup_{q=2}^{\infty} E_q$. We first prove that the measure of E is 1. If we do this then that will show that for almost all α there exists a rational number p/q with $\left|\alpha - \frac{p}{q}\right| < \frac{\theta_q}{q}$. We shall then show that $\mu\left(\bigcup_{q=k}^{\infty} E_q\right) = 1$ for all $k \geq 3$ and from this we will find infinitely many solutions to the inequality $\left|\alpha - \frac{p}{q}\right| < \frac{f(q)}{q}$.

Suppose to the contrary that $\mu(E) < 1$. Then there exists $\delta > 0$ such that $\mu(E)(1+\delta) < 1$. Suppose there exists a $q_1 > 0$ such that if we put $A = E_2 \cup \cdots \cup E_{q_1}$ then $\mu(A) > \mu(E) - \delta$. Since A is a finite union of intervals, by proposition 0.14 there exists a positive number q_2 such that if $q > q_2$, then

$$\mu(A \cap E_q) \leq \mu(A)\mu(E_q)(1+\delta). \tag{0.1}$$

Let $m > n$ be positive integers larger than $q_1 + q_2$ and put $B = B_{m,n} = E_n \cup \cdots \cup E_m$. We have

$$\sum_{j=n}^{m} \mu(E_j) - \sum_{n \leq j < k \leq m} \mu(E_j \cap E_k) \leq \mu(B) \leq \sum_{j=n}^{m} \mu(E_j).$$

By proposition 0.13, $\mu(E_j \cap E_k) \leq 8\theta_j\theta_k$ and so,

$$\mu(B) \geq \sum_{j=n}^{m} \mu(E_j) - 4\left(\sum_{j=n}^{m} \theta_j\right)^2.$$

By equation (0.1), we have

$$\mu(A \cap B) \leq \sum_{j=n}^{m} \mu(A \cap E_j) \leq \mu(A)\left(\sum_{j=n}^{m} \mu(E_j)\right)(1+\delta).$$

Observe that $\mu(E) \geq \mu(A \cup B) \geq \mu(A) + \mu(B) - \mu(A \cap B)$ and so

$$\mu(E) \geq \mu(A) + \sum_{j=n}^{m} \mu(E_j) - 4\left(\sum_{j=n}^{m} \theta_j\right)^2 - \mu(A)\left(\sum_{j=n}^{m} \mu(E_j)\right)(1+\delta).$$

Hence

$$\mu(E) \geq \mu(A) + \left(\sum_{j=n}^{m} \mu(E_j)\right)(1 - \mu(A)(1+\delta)) - 4\left(\sum_{j=n}^{m} \theta_j\right)^2 \tag{0.2}$$

By assumption, there exists $0 < c \leq 1$ and arbitrarily large integers $m > n > 0$ for which

$$\sum_{j=n}^{m} \theta_j > 1$$

and

$$\sum_{j=n}^{m} \theta_j \frac{\varphi(j)}{j} > \frac{c}{2} \sum_{j=n}^{m} \theta_j.$$

But $\sum_{j=n}^{m} \mu(E_j) = \sum_{j=n}^{m} \theta_j \frac{\varphi(j)}{j} > c \sum_{j=n}^{m} \theta_j$. Thus, by equation (0.2), we have

$$\mu(E) \geq \mu(A) + \left(c \sum_{j=n}^{m} \theta_j \right) (1 - \mu(A)(1+\delta)) - 4 \left(\sum_{j=n}^{m} \theta_j \right)^2 .$$

Put $t = \sum_{j=n}^{m} \theta_j$ and $b = c(1 - \mu(A)(1+\delta))$ so that $\mu(E) \geq \mu(A) + bt - 4t^2$.

Observe that $0 < b < 1$ since $0 < c \leq 1$ and $1 - \mu(A)(1+\delta) < 1$. The maximum of $yb - 4y^2$ for $y \in (0,1)$ occurs when $y = b/8$, at which point $yb - 4y^2 = \frac{b^2}{16}$. We shall now modify the E_j's by replacing θ_j with $2\theta_j$. Denote the set $E_j^{z\theta_j}$ by $E_j^{(1)}$ for $j = 2, 3, \cdots$ where z is chosen so that $\sum_{j=n}^{m} z\theta_j = \frac{b}{8}$. Keep A as before and replace B with B_z where

$$B_z = E_n^{(1)} \cup \cdots \cup E_m^{(1)}.$$

Arguing as before, we obtain

$$\begin{aligned} \mu(E) &\geq \mu(A) + btz - 4(tz)^2 \\ &= \mu(A) + \frac{b^2}{16} \\ &= \mu(A) + \frac{c^2}{16}(1 - \mu(A)(1+\delta))^2 \end{aligned}$$

Notice that as $\delta \to 0$, we have $\mu(A) \to \mu(E)$ and hence

$$\mu(E) \geq \mu(E) + \frac{c^2}{16}(1 - \mu(E)).$$

This inequality is untenable if $\mu(E) < 1$, and hence we must conclude that $\mu(E) = 1$.

Now put $E^{(k)} = \bigcup_{q=k}^{\infty} E_q$ and observe that the same argument holds as before. This implies that $\mu(E^{(k)}) = 1$ for all $k \in \mathbb{N}$. Thus, if we set $E^* = \bigcap_{k=1}^{\infty} E^{(k)}$, then we have

$$\mu(E^*) = 1$$

since E^* is the intersection of countably many sets of full measure. In particular, for each $\alpha \in E^*$, we can find infinitely many rationals p/q such that $\left| \alpha - \frac{p}{q} \right| < \frac{f(q)}{q}$. $\square$

We now show that Khintchine's theorem is a consequence of the Duffin-Schaeffer theorem. In fact, in place of the assumption that $(f(q))_{q \geq 0}$ is a decreasing sequence we will require only $(f(q))_{q \geq 0}$ is decreasing. We will replace $f(q)$ with θ_q for this argument.

Notice that we may suppose that $\theta_q \leq 1/2$ for sufficiently large q since otherwise the inequality $\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q}$ certainly has infinitely many solutions for almost all α. Thus we may replace θ_q with $\min(\theta_q, 1/2)$ to guarantee condition (ii) of the Duffin-Schaeffer theorem.

It remains to show that there exists $c > 0$ such that for infinitely many positive integers n we have

$$\sum_{q=1}^{n} \frac{\theta_q \varphi(q)}{q} > c \sum_{q=1}^{n} \theta_q.$$

Notice that since $(\theta_q)_{q\geq 1}$ is decreasing, we have

$$\begin{aligned}
\sum_{q=1}^{2^n} \theta_q \frac{\varphi(q)}{q} &= \sum_{t=1}^{n} \sum_{q=2^{t-1}+1}^{2^t} \theta_q \frac{\varphi(q)}{q} \\
&\geq \sum_{t=1}^{n} \theta_{2^t} \sum_{q=2^{t-1}+1}^{2^t} \frac{\varphi(q)}{q} \\
&\geq \sum_{t=1}^{n} \frac{\theta_{2^t}}{2^t} \sum_{q=2^{t-1}+1}^{2^t} \varphi(q).
\end{aligned}$$

By proposition 0.6, we get

$$\begin{aligned}
\sum_{q=2^{t-1}+1}^{2^t} \varphi(q) &= \frac{3}{\pi^2}(2^{2t} - 2^{2t-2}) + O(t2^t) \\
&= \frac{9}{4\pi^2} 2^{2t} + O(t2^t) \\
&> c_1 2^{2t},
\end{aligned}$$

for some $c_1 > 0$. Thus

$$\begin{aligned}
\sum_{q=1}^{2^n} \theta_q \frac{\varphi(q)}{q} &\geq \sum_{t=1}^{n} \frac{\theta_{2^t}}{2^t} c_1 2^{2t} \\
&= c_1 \sum_{t=1}^{n} \theta_{2^t} 2^t \\
&\geq c_1 \sum_{q=2}^{2^n+2^n-1} \theta_q \\
&\geq c_1 \sum_{q=2}^{2^n} \theta_q.
\end{aligned}$$

Since $\sum_{q=1}^{\infty} \theta_q = \sum_{q=1}^{\infty} f(q) = \infty$, there exists $c_2 > 0$ such that

$$c_1 \sum_{q=2}^{2^n} \theta_q > c_2 \sum_{q=1}^{2^n} \theta_q.$$

This shows that condition (ii) of the Duffin-Schaeffer theorem is also satisfied, and so Khintchine's Theorem is a corollary of the Duffin-Schaeffer theorem.

Gallagher proved the following result: Let $(f(q))_{q=1}^{\infty}$ be a sequence of non-negative real numbers. Let A be the set of real numbers α in $(0,1)$ for which the inequality $\left|\alpha - \frac{p}{q}\right| < \frac{f(q)}{q}$ has infinitely many solutions in rationals p/q. The measure of A is either 0 or 1.

Duffin and Schaeffer conjectured that for almost all α with respect to Lebesgue measure, the inequality $\left|\alpha - \frac{p}{q}\right| < \frac{f(q)}{q}$ has infinitely many solutions if and only if $\sum_{q=1}^{\infty} \frac{f(q)}{q}\varphi(q)$ diverges. The conjecture is still unsolved, but higher dimensional analogues of it have been proved (Pollington, Vaughan).

Given a real number α, how should we go about finding the good rational approximations p/q to α? We use an algorithm known as the continued fraction algorithm. For any $x \in \mathbb{R}$ recall that $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x. Put $a_0 = \lfloor \alpha \rfloor$. If $\alpha \neq a_0$ then we write $\alpha = a_0 + \frac{1}{\alpha_1}$. Then write $a_1 = \lfloor \alpha_1 \rfloor$. If $a_1 \neq \alpha_1$, we write $\alpha_1 = a_1 + \frac{1}{\alpha_2}$. Continue in this way we generate a sequence of positive integers $a_1, a_2, \cdots$ and real numbers $\alpha_1, \alpha_2, \cdots > 1$. The sequences are finite if $\alpha_i = a_i$ for some $i \in \mathbb{N}$, in which case the algorithm terminates.

If the algorithm terminates, say at $\alpha_n = a_n$, then write

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots + \cfrac{1}{a_n}}}}$$

or more conveniently, $\alpha = [a_0, a_1, \cdots, a_n]$. This expression is called a finite continued fraction.

If the algorithm does not terminate, then we have

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}.$$

Alternatively, we write $\alpha = [a_0, a_1, a_2, \cdots]$. These expressions are known as the continued fraction expression of α.

We will prove that $\alpha = \lim_{n\to\infty}[a_0, a_1, \cdots, a_n]$. The terms $a_0, a_1, \cdots$ are known as the partial quotients of α. Further we will put $[a_0, \cdots, a_n] = \frac{p_n}{q_n}$ where $\gcd(p_n, q_n) = 1$ and $q_n > 0$. The rationals $\frac{p_n}{q_n}$ are known as the *convergents* to α.

We will show that the p_n's and q_n's are generated recursively in the following manner.

Proposition 0.15. *Let α be a real number, and let $\left(\frac{p_n}{q_n}\right)_{n=0}^{\infty}$ be its sequence of convergents and $(a_n)_{n=0}^{\infty}$ be its sequence of partial quotients. Then $(p_n), (q_n)$ both satisfy the recursion*

$$u_n = a_n u_{n-1} + u_{n-2}, n \geq 2 \tag{0.3}$$

with $p_0 = a_0, q_0 = 1, p_1 = a_0 a_1 + 1, q_1 = a_1$.

Proof. We proceed to prove this result by induction. For $n = 2$, we have

$$\begin{aligned}\frac{p_2}{q_2} &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} \\ &= a_0 + \frac{a_2}{a_1 a_2 + 1} \\ &= \frac{a_0 a_1 a_2 + a_0 + a_2}{a_1 a_2 + 1} \\ &= \frac{a_2 p_1 + p_0}{a_2 q_1 + q_0}.\end{aligned}$$

This establishes the base case. Now assume the result holds for $n = k - 1$ with $k \geq 2$ and we will prove it for $n = k$. Consider the associated continued fractions $[a_1, \cdots, a_k]$ and put $[a_1, \cdots, a_{j+1}] = \frac{u_j}{v_j}$ with $\gcd(u_j, v_j) = 1$, $v_j > 0$ for $j = 0, 1, 2, \cdots$. By the inductive hypothesis we have $u_{k-1} = a_k u_{k-2} + u_{k-3}$ and $v_{k-1} = a_k v_{k-2} + v_{k-3}$.

But $\frac{p_j}{q_j} = a_0 + \frac{v_{j-1}}{u_{j-1}}$, for $j = 1, 2, \cdots$. Hence $p_j = a_0 u_{j-1} + v_{j-1}$ and $q_j = u_{j-1}$. Now set $j = k$ to obtain

$$\begin{aligned}p_k &= a_0(a_k u_{k-2} + u_{k-3}) + a_k v_{k-2} + v_{k-3} \\ &= a_k(a_0 u_{k-2} + v_{k-2}) + (a_0 u_{k-3} + v_{k-3}) \\ &= a_k p_{k-1} + p_{k-2},\end{aligned}$$

as desired. Similarly, we have

$$\begin{aligned} q_k &= u_{k-1} \\ &= a_k u_{k-2} + u_{k-3} \\ &= a_k q_{k-1} + q_{k-2}. \end{aligned}$$

This completes the proof. ∎

Recall from the definition of $\alpha_1, \alpha_2, \cdots$ that

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

We also have

$$0 < \frac{1}{\alpha_{n+1}} \leq \frac{1}{a_{n+1}}.$$

Notice that $\alpha \in \left[\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}\right]$.

Proposition 0.16. *If $\left(\frac{p_n}{q_n}\right)_{n=0}^{\infty}$ is the sequence of convergents for a real number α, then*

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

for $n = 0, 1, \cdots$.

Proof. We proceed by induction. For $n = 0$ we have $p_0 q_1 - p_1 q_0 = a_0 a_1 - (a_0 a_1 + 1) = -1$, so the result holds.

Assume that this holds for $n = k - 1$. Then by our recursion for p_k, q_k we have

$$\begin{aligned} p_k q_{k+1} - q_k p_{k+1} &= p_k(a_{k+1} q_k + q_{k-1}) - q_k(a_{k+1} p_k + p_{k-1}) \\ &= p_k q_{k-1} - q_k p_{k-1} \\ &= (-1)^{k+1}, \end{aligned}$$

as required. ∎

Since $\alpha \in \left[\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}\right]$, we see that

$$\left|\alpha - \frac{p_n}{q_n}\right| \leq \left|\frac{q_{n+1} p_n - p_n q_{n+1}}{q_n q_{n+1}}\right| = \frac{1}{q_n q_{n+1}}.$$

We have $q_0 = 1, q_1 = a_1$ and so $q_{n+1} > q_n$ for $n > 0$ and thus

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}$$

for $n = 1, 2, \cdots$. Thus the convergents $\frac{p_n}{q_n}$ are good approximations to α.

Remark 0.17. The continued fraction terminates if and only if α is rational. Further, $\lim_{n \to \infty} [a_0, \cdots, a_n] = \alpha$.

We will now complete the proof of Hurwitz's Theorem by showing that at least one of any three consecutive convergents to α, say $\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}, \frac{p_{n+2}}{q_{n+2}}$ must satisfy

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

Suppose otherwise for the sake of a contradiction. Then we have

$$\left|\alpha - \frac{p_j}{q_j}\right| \geq \frac{1}{\sqrt{5}q_j^2}$$

for $j = n, n+1, n+2$. This implies that

$$\left|\alpha - \frac{p_n}{q_n}\right| + \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right| = \left|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}\right| = \frac{1}{q_n q_{n+1}},$$

and so

$$\frac{1}{\sqrt{5}q_n^2} + \frac{1}{\sqrt{5}q_{n+1}^2} \leq \frac{1}{q_n q_{n+1}} \Rightarrow \frac{1}{\sqrt{5}}\frac{q_{n+1}}{q_n} + \frac{1}{\sqrt{5}}\frac{q_n}{q_{n+1}} \leq 1.$$

Put $\lambda_n = \frac{q_{n+1}}{q_n}$, and hence

$$\begin{aligned}\frac{\lambda_n}{\sqrt{5}} + \frac{1}{\sqrt{5}\lambda_n} \leq 1 &\Rightarrow \lambda_n^2 - \sqrt{5}\lambda_n + 1 \leq 0 \\ &\Rightarrow \left(\lambda_n - \frac{\sqrt{5}}{2}\right)^2 - \frac{1}{4} \leq 0.\end{aligned}$$

Since $\lambda_n \in \mathbb{Q}$, the inequality is strict. Thus $\left(\lambda_n - \frac{\sqrt{5}+1}{2}\right)\left(\lambda_n - \frac{\sqrt{5}-1}{2}\right) < 0$, and so $\frac{\sqrt{5}-1}{2} < \lambda_n < \frac{\sqrt{5}+1}{2}$, in particular $\lambda_n < \frac{1+\sqrt{5}}{2}$. Now, recall that $q_{n+2} = a_{n+2}q_{n+1} + q_n$, so that $\frac{q_{n+2}}{q_{n+1}} = a_{n+2} + \frac{1}{(q_{n+1}/q_n)}$. Observe also that $\lambda_{n+1} < \frac{1+\sqrt{5}}{2}$. But

$$\begin{aligned}\lambda_{n+1} &= a_{n+2} + \frac{1}{\lambda_n} \\ &> 1 + \frac{2}{1+\sqrt{5}} \\ &= \frac{3+\sqrt{5}}{1+\sqrt{5}} \\ &= \frac{1+\sqrt{5}}{2},\end{aligned}$$

a contradiction. This completes the proof of Hurwitz's Theorem.

Proposition 0.18. *For any real number α, the sequence $(|q_1\alpha - p_1|, |q_2\alpha - p_2|, \cdots)$ is a decreasing sequence.*

Proof. The recurrence relations for p_n, q_n hold for any indeterminates and so we may apply them with $\alpha = [a_0, a_1, \cdots, a_n, \alpha_{n+1}]$ to conclude that

$$\alpha = \frac{p_n\alpha_{n+1} + p_{n-1}}{q_n\alpha_{n+1} + q_{n-1}}$$

and so

$$\begin{aligned}\left| q_n\left(\frac{p_n\alpha_{n+1} + p_{n-1}}{q_n\alpha_{n+1} + q_{n-1}}\right) - p_n \right| &= \left| \frac{q_np_n\alpha_{n+1} + q_np_{n-1} - p_nq_n\alpha_{n+1} - p_nq_{n-1}}{q_n\alpha_{n+1} - q_{n-1}} \right| \\ &= \frac{1}{|q_n\alpha_{n+1} + q_{n-1}|}.\end{aligned}$$

But

$$\begin{aligned} q_n\alpha_{n+1} + q_{n-1} &\geq q_n + q_{n-1} \\ &\geq a_nq_{n-1} + q_{n-2} + q_{n-1} \\ &= (a_n + 1)q_{n-1} + q_{n-2} \\ &\geq \alpha_nq_{n-1} + q_{n-2}, \end{aligned}$$

which implies that

$$\left| q_n\left(\frac{p_n\alpha_{n+1} + p_{n-1}}{q_n\alpha_{n+1} + q_{n-1}}\right) - p_n \right| = \frac{1}{q_n\alpha_{n+1} + q_{n-1}} \leq \frac{1}{q_{n-1}\alpha_n + q_{n-2}},$$

and we check that it holds for $n = 1$ also. □

Proposition 0.19. *Let α be a real number. The convergents $\frac{p_n}{q_n}$ to α satisfy*

$$\frac{1}{(a_{n+1} + 2)q_n^2} < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{a_{n+1}q_n^2}.$$

Proof. By proposition 0.18 we have

$$\left|\alpha - \frac{p_n}{q_n}\right| = \frac{1}{q_n(a_n\alpha_{n+1} + q_{n-1})}.$$

Since $a_{n+1} \leq \alpha_{n+1} < a_{n+1} + 1$ and $q_n \geq q_{n-1}$, the result follows. □

The convergents p_n/q_n give the best approximations to α in the sense that if $0 < q < q_{n+1}$, then $|q\alpha - p| \geq |a_n\alpha - p_n|$. To see this, note that since $\det \begin{bmatrix} p_n & q_n \\ p_{n+1} & q_{n+1} \end{bmatrix} = (-1)^{n+1}$, we can find integers u, v such that $p = up_n + vp_{n+1}$ and $q = uq_n + vq_{n+1}$. Note that $u \neq 0$. Further, if $v \neq 0$ then u, v have opposite signs. Thus

$$|q\alpha - p| = |u(q_n\alpha - p_n) + v(q_{n+1}\alpha - p_{n+1})| \geq |q_n\alpha - p_n|.$$

Proposition 0.20. *Let $\alpha \in \mathbb{R}$. If p/q is a rational with $\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q^2}$, then p/q is a convergent to α. In other words, $\frac{p}{q} = \frac{p_n}{q_n}$ for some $n \geq 0$.*

Proof. In fact $\frac{p}{q} = \frac{p_n}{q_n}$ where $q_n \le q < q_{n+1}$, since

$$\begin{aligned} \left|\frac{p}{q} - \frac{p_n}{q_n}\right| &\le \left|\alpha - \frac{p}{q}\right| + \left|\alpha - \frac{p_n}{q_n}\right| \\ &\le \left(\frac{1}{q} + \frac{1}{q_n}\right)|q\alpha - p| \\ &< \frac{2}{q_n}\frac{1}{2q} = \frac{1}{qq_n}. \end{aligned}$$

But if $p/q, p_n/q_n$ are distinct rational numbers, then the absolute value of their difference is at least $\frac{1}{qq_n}$, so the above inequality shows that they must in fact be equal. □

Definition 0.21. The continued fraction $[a_0, a_1, \cdots]$ is said to be ultimately periodic if there exists a non-negative integer n and a positive integer k such that $a_{k+m} = a_m$ for all $m \ge n$.

Theorem 0.22. (Lagrange's Theorem) *A real number α is a quadratic irrational if and only if its continued fraction expansion is ultimately periodic.*

Proof. Suppose that $\alpha = [a_0, \cdots, a_{k-1}, \overline{a_k, \cdots, a_{n+k-1}}]$ where the bar indicates periodicity. Put $\theta = [\overline{a_k, \cdots, a_{n+k-1}}]$ and let $\frac{u_j}{v_j}$ denote the convergents to θ. We have $\theta = [a_k, \cdots, a_{k+n-1}, \theta]$, so that $\theta = \frac{u_{n-1}\theta + u_{n-2}}{v_{n-1}\theta + v_{n-2}}$. Thus $v_{n-1}\theta^2 + (v_{n-2} + u_{n-1})\theta - u_{n-2} = 0$. Further, $\theta \in \mathbb{R} \setminus \mathbb{Q}$ since it has an infinite continued fraction expansion. Thus it is a real quadratic irrational.

But $\alpha = [a_0, \cdots, a_{k-1}, \theta]$ and so $\alpha = \frac{p_{k-1}\theta + p_{k-2}}{q_{k-1}\theta + v_{k-2}}$ and so α is a real quadratic irrational as well.

Suppose now that α is a real quadratic irrational. Let $ax^2 + bx + c$ be the minimal polynomial of α in $\mathbb{Z}[x]$. Then $b^2 - 4ac > 0$ since α is real. Suppose that $\alpha = [a_0, a_1, \cdots]$. Then $\alpha = \frac{p_{n-1}\alpha_n + p_{n-2}}{q_{n-1}\alpha_n + q_{n-2}}$ and so

$$a(p_{n-1}\alpha_n + p_{n-2})^2 + b(p_{n-1}\alpha_n + p_{n-2})(q_{n-1}\alpha_n + q_{n-2}) + c(q_{n-1}\alpha_n + q_{n-2}) = 0.$$

Set $A_n = ap_{n-1}^2 + bp_{n-1}q_{n-1} + cq_{n-1}^2$, $B_n = 2ap_{n-1}p_{n-2} + bp_{n-1}q_{n-2} + bp_{n-2}q_{n-1} + 2cq_{n-1}q_{n-2}$, and $C_n = ap_{n-2}^2 + bp_{n-2}q_{n-2} + cq_{n-2}^2$. In other words, we have

$$A_n\alpha_n^2 + B_n\alpha_n + C_n = 0.$$

Notice that $A_n \ne 0$ since otherwise $ax^2 + bx + c = 0$ has a rational root. Further, $B_n^2 - 4A_nC_n = (b^2 - 4ac)(p_{n-1}q_{n-2} - p_{n-2}q_{n-1})^2 = b^2 - 4ac > 0$.

Now we have $\alpha - \frac{p_n}{q_n} = \frac{\delta_n}{q_n^2}$ with $|\delta_n| \le 1$, for all $n \in \mathbb{N}$. Thus $p_n = q_n\alpha - \frac{\delta_n}{q_n}$, hence

$$\begin{aligned}
A_n &= a\left(q_{n-1}\alpha - \frac{\delta_{n-1}}{q_{n-1}}\right)^2 + b\left(q_{n-1}\alpha - \frac{\delta_{n-1}}{q_{n-1}}\right)q_{n-1} + cq_{n-1}^2 \\
&= (a\alpha^2 + b\alpha + c)q_{n-1}^2 - 2a\alpha\delta_{n-1} + \frac{a\delta_{n-1}^2}{q_{n-1}^2} - b\delta_{n-1} \\
&= -2a\alpha\delta_{n-1} + a\frac{\delta_{n-1}^2}{q_{n-1}^2} - b\delta_{n-1},
\end{aligned}$$

so that $|A_n| \le |2a\alpha| + |a| + |b|$.

Note that $C_n = A_{n-1}$, so $|C_n| \le |2a\alpha| + |a| + |b|$. Finally, we have $|B_n| \le 4|A_nC_n| + |b^2 - 4ac|$. Since $|A_n|, |B_n|, |C_n|$ are bounded, the α_n's are the roots of a finite family of quadratic polynomials, each polynomial has at most two distinct roots (in fact each has exactly two distinct roots since α is irrational). Therefore $\alpha_n = \alpha_{n+k}$ for some $k \in \mathbb{N}$ and $n \ge 1$. Hence the continued fraction expansion is ultimately periodic. □

We say that the continued fraction expansion $[a_0, a_1, \cdots]$ is purely periodic if the period starts at $n = 0$. In other words, for some integer k, we have $a_n = a_{n+k}$ for all $n \ge 0$.

Proposition 0.23. *The continued fraction expansion of a real quadratic irrational α is purely periodic if and only if $\alpha > 1$ and the conjugate β of α satisfies $-1 < \beta < 0$.*

Proof. We claim that the conjugate β_n to α_n also satisfy $-1 < \beta_n < 0$. This follows by induction. Since $\alpha_n = a_n + \dfrac{1}{\alpha_{n+1}}$ we find that $\beta_n = a_n + \frac{1}{\beta_{n+1}}$. But now $a_n \ge 1$ and $-1 < \beta_n < 0$, hence $-1 < \beta_{n+1} < 0$. Observe that since $-1 < \beta_n < 0$, we have $a_n = \left\lfloor \dfrac{-1}{\beta_{n+1}} \right\rfloor$.

Since α is a quadratic irrational we know that there exist distinct integers m, n with $\alpha_m = \alpha_n$. But then $\dfrac{1}{\beta_m} = \dfrac{1}{\beta_n}$ and so $a_{n-1} = a_{m-1}$, which implies that $\alpha_{m-1} = \alpha_{n-1}$. Repeating this argument we find that α has a purely periodic continued fraction expansion.

Suppose that the continued fraction expansion of α is purely periodic. Then $\alpha > a_0 \ge 1$. Further, there is a positive integer n such that $\alpha = \dfrac{p_n\alpha + p_{n-1}}{q_n\alpha + q_{n-1}}$, so $q_n\alpha^2 + (q_{n-1} - p_n)\alpha - p_{n-1} = 0$. Consider the polynomial $f_n(x) = q_nx^2 + (q_{n-1} - p_n)x - p_{n-1}$. We have $f_n(0) = -p_{n-1} < 0$ and $f_n(-1) = (q_n - q_{n-1}) + (p_n - p_{n-1}) > 0$. Thus the polynomial $f_n(x)$ has a root β in $(-1, 0)$, and β is conjugate to α. □

Remark 0.24. Let d be an integer which is positive but not a perfect square. Consider $\alpha = \dfrac{1}{\sqrt{d} - \lfloor\sqrt{d}\rfloor}$. Then $\alpha > 1$ and the conjugate $\dfrac{-1}{\sqrt{d} + \lfloor\sqrt{d}\rfloor}$ satisfies $-1 < \dfrac{-1}{\sqrt{d} + \lfloor\sqrt{d}\rfloor} < 0$. Thus the continued fraction expansion of $\dfrac{1}{\sqrt{d} - \lfloor\sqrt{d}\rfloor}$ is purely periodic.

Consider the rational $\alpha = [a_0, \cdots, a_n]$ and the convergents $p_0/q_0, \cdots, p_n/q_n$ to α. Then $[a_n, \cdots, a_0] = \dfrac{p_n}{p_{n-1}}$ and $[a_n, \cdots, a_1] = \dfrac{q_n}{q_{n-1}}$. To see this, note that $p_n = a_n p_{n-1} + p_{n-2}$ so $\dfrac{p_n}{p_{n-1}} = a_n + \dfrac{1}{\left(\frac{p_{n-1}}{p_{n-2}}\right)}$, hence

$$\frac{p_n}{p_{n-1}} = a_n + \cfrac{1}{a_{n-1} + \cdots + \cfrac{1}{p_1/p_0}},$$

but $\dfrac{p_1}{p_0} = \dfrac{a_1 a_0 + 1}{a_0} = a_1 + \dfrac{1}{a_0}$. This shows that $\dfrac{p_n}{p_{n-1}} = [a_n, \cdots, a_0]$. Similarly, $q_n = a_n q_{n-1} + q_{n-2}$, so $\dfrac{q_n}{q_{n-1}} = a_n + \dfrac{1}{q_{n-1}/q_{n-2}}$ and hence

$$\frac{q_n}{q_{n-1}} = a_n + \cfrac{1}{a + n - 1 + \cdots + \cfrac{1}{q_1/q_0}}.$$

But $q_1/q_0 = a_1/1 = a_1$, and hence $\dfrac{q_n}{q_{n-1}} = [a_n, \cdots, a_1]$.

Proposition 0.25. *Let α be a quadratic irrational with $\alpha > 1$ and conjugate β satisfying $-1 < \beta < 0$. Then $\alpha = [\overline{a_0, \cdots, a_n}]$ and $\dfrac{-1}{\beta} = [\overline{a_n, \cdots, a_0}]$.*

Proof. Let $\theta = [\overline{a_n, \cdots, a_0}]$ so $\theta = [a_n, \cdots, a_0, \theta]$. Let $\dfrac{u_n}{v_n}$ be the convergents to θ. Then $\theta = \dfrac{u_n\theta + u_{n-1}}{v_n\theta + v_{n-1}}$. Now, let $\dfrac{p_n}{q_n}$ be t he nth convergent to α. By the preceding paragraph, it follows that $\dfrac{u_n}{v_n} = \dfrac{p_n}{p_{n-1}}$. By proposition 0.16 we have $\gcd(p_n, p_{n-1}) = 1$, so that $u_n = p_n, v_n = p_{n-1}$. Further, we have $\dfrac{u_{n-1}}{v_{n-1}} = \dfrac{q_n}{q_{n-1}}$ and hence $u_{n-1} = q_n$ and $v_{n-1} = q_{n-1}$, since $\gcd(q_n, q_{n-1}) = 1$. But then $\theta = \dfrac{p_n\theta + q_n}{p_{n-1}\theta + q_{n-1}}$, and therefore

$$p_{n-1}\theta^2 + (q_{n-1} - p_n)\theta - q_n = 0 \Rightarrow -q_n\left(\frac{1}{\theta}\right)^2 + (q_{n-1} - p_n)\left(\frac{1}{\theta}\right) + p_{n-1} = 0.$$

This shows that

$$q_n\left(\frac{-1}{\theta}\right)^2 + (q_{n-1} - p_n)\left(\frac{-1}{\theta}\right) - p_{n-1}.$$

Recall that α is also a root of $q_n x^2 + (q_{n-1} - p_n)x - p_{n-1}$, and therefore $\dfrac{-1}{\theta} = \beta$, as desired. $\square$

Let d be a positive integer which is not a perfect square. Then $\alpha = \sqrt{d} + \lfloor\sqrt{d}\rfloor$ has conjugate $\beta = -\sqrt{d} + \lfloor\sqrt{d}\rfloor$ so $-1 < \beta < 0$. By proposition 0.23, we have

$$\alpha = [\overline{2\lfloor\sqrt{d}\rfloor, a_1, \cdots, a_n}] = [\overline{2a_0, a_1, \cdots, a_n}].$$

By proposition 0.25, we get

$$\frac{-1}{\beta} = [\overline{a_n, \cdots, a_1, 2a_0}].$$

But

$$\sqrt{d} - \lfloor\sqrt{d}\rfloor = 0 + \frac{1}{\frac{1}{\sqrt{d}-\lfloor\sqrt{d}\rfloor}} = [0, \overline{a_n, \cdots, 2a_0}].$$

On the other hand, $\alpha = \sqrt{d} + \lfloor\sqrt{d}\rfloor = [\overline{2a_0, a_1, \cdots, a_n}]$. Thus

$$\sqrt{d} = [a_0, \overline{a_1, \cdots, a_n, 2a_0}] = [a_0, \overline{a_n, \cdots, a_1, 2a_0}].$$

Therefore, $a_n = a_1, a_{n-1} = a_2, \cdots$, so $\sqrt{d} = [a_0, \overline{a_1, a_2, \cdots, a_2, a_1, 2a_0}]$.

We can use this information to find all solutions in integers (x, y) of the equation $x^2 - dy^2 = 1$.

Equations of the form $x^2 - dy^2 = \pm 1, x^2 - dy^2 = \pm 4$ are known as Pell equations.

Fermat had conjectured that for each d with d not a perfect square the equation $x^2 - dy^2 = 1$ has a non-trivial solution, different from $(x, y) = (\pm 1, 0)$. This was established by Lagrange in 1768. Let's consider the equations $x^2 - dy^2 = 1, x^2 - dy^2 = -1$ and suppose that x, y is a non-trivial solution in positive integers to one of them. Then $x \geq \sqrt{dy^2 - 1} \geq y\sqrt{d-1}$. Thus

$$\begin{aligned} |x - \sqrt{d}y| &= \frac{1}{|x + \sqrt{d}y|} \\ &= \frac{1}{x + \sqrt{d}y} \\ &\leq \frac{1}{y(\sqrt{d} + \sqrt{d-1})}. \end{aligned}$$

Now $d \geq 2$ so $\sqrt{d} + \sqrt{d-1} > 2$ hence $|x - \sqrt{d}y| < 1/2y$, so

$$\left|\sqrt{d} - \frac{x}{y}\right| < \frac{1}{2y^2}.$$

By proposition 0.20, x/y is a convergent to $\sqrt{d}$ and $\dfrac{x}{y} = \dfrac{p_n}{q_n}$ for some $n \geq 1$.

Then $\sqrt{d} = \dfrac{p_n\alpha_{n+1} + p_{n-1}}{q_n\alpha_{n+1} + q_{n-1}}$ so

$$q_n\alpha_{n+1}\sqrt{d} + q_{n-1}\sqrt{d} = p_n\alpha_{n+1} + p_{n-1},$$

hence

$$(q_n\sqrt{d} - p_n)\alpha_{n+1} = p_{n-1} - q_{n-1}\sqrt{d} \Rightarrow (p_n - q_n\sqrt{d})\alpha_{n+1} = q_{n-1}\sqrt{d} - p_{n-1}.$$

Therefore,

$$\begin{aligned}(p_n^2 - q_n^2 d) &= (q_{n-1}\sqrt{d} - p_{n-1})(q_n\sqrt{d} + p_n)\\ &= (q_{n-1}q_n d + p_n q_{n-1}\sqrt{d} - p_{n-1}q_n\sqrt{d} - p_n p_{n-1})\\ &= (p_n q_{n-1} - p_{n-1}q_n)\sqrt{d} + (q_n q_{n-1} d - p_n p_{n-1})\\ &= (-1)^{n+1}\sqrt{d} + h, h \in \mathbb{Z}.\end{aligned}$$

Suppose that $p_n^2 - dq_n^2 = \pm 1$. Then $\pm\alpha_{n+1} = (-1)^{n+1}\sqrt{d} + h$ with $h \in \mathbb{Z}$.

The even convergents to $\sqrt{d}$ are smaller than $\sqrt{d}$ and the odd convergents are larger.

Suppose that $p_n^2 - dq_n^2 = 1$. Then from

$$(p_n^2 - dq_n^2)\alpha_{n+1} = (-p_{n-1} + q_{n-1}\sqrt{d})(p_n + q_n\sqrt{d}), \tag{0.4}$$

we see that $-p_{n-1} + q_{n-1}\sqrt{d} > 0$, so we know that $\sqrt{d} > \dfrac{p_{n-1}}{q_{n-1}}$, so $n-1$ is even. Further, if $p_n^2 - dq_n^2 = -1$, then $n-1$ has to be odd.

The convergents of even index are smaller than $\sqrt{d}$ and those of odd index are larger than $\sqrt{d}$, and so by equation (0.4) if $p_n^2 - dq_n^2 = 1$ then $n-1$ has to be even.

Let us consider the case $p_n^2 - dq_n^2 = 1$. Then $\alpha_{n+1} = \sqrt{d} + h$. Thus $\alpha_{n+2} = \alpha_1$. But $\sqrt{d} = [a_0, \overline{a_1, \cdots, a_m}]$ where m is the period, so the minimal positive integer for which $\alpha_1 = \alpha_{m+1} = \alpha_{2m+1} = \cdots$. Therefore $(n+2)-1$ has to be a multiple of m, say $n = lm - 1$ with $l \in \mathbb{N}$. Note that in this case $lm = n+1$ is even.

In the case $p_n^2 - dq_n^2 = -1$ we have $n-1$ is odd and $-\alpha_{n+1} = -\sqrt{d} + h$ so $\alpha_{n+1} = \sqrt{d} - h$, hence $\alpha_{n+2} = \alpha_1$ and we have $n = lm - 1$ as before. Thus lm is odd. This immediately shows that if m is even, then the equation $p_n^2 - dq_n^2 = -1$ has no solutions.

Theorem 0.26. *Let d be a squarefree integer with $d > 1$. Let m be the length of the period of the continued fraction expansion of $\sqrt{d}$. Then*

(i) (x, y) is a solution of the equation $u^2 - dv^2 = 1$ in $\mathbb{N}$ if and only if $x = p_n, y = q_n$ where $\dfrac{p_n}{q_n}$ is a convergent to $\sqrt{d}$ and $n = lm - 1$ where $l \in \mathbb{N}$ and lm is even.

(ii) (x, y) is a solution to $u^2 - dv^2 = -1$ in positive integers x, y if and only if $x = p_n, y = q_n$ where $\dfrac{p_n}{q_n}$ is a convergent to $\sqrt{d}$ and $n = lm - 1$ where l is a positive integer and lm is odd.

Proof. The forward direction in both claims are done already. Hence it suffices to prove the converses.

Suppose that $n = lm - 1$. Then $\alpha_{n+2} = \alpha_1$, by periodicity, and so

$$\sqrt{d} = \frac{p_{n+1}\alpha_{n+2} + p_n}{q_{n+1}\alpha_{n+2} + q_n} = \frac{p_{n+1}\alpha_1 + p_n}{q_{n+1}\alpha_1 q_n}.$$

Recall that $\alpha_1 = \dfrac{1}{\sqrt{d} - a_0}$. We have

$$\begin{aligned}\sqrt{d}(q_{n+1}\alpha_1 + q_n) = (p_{n+1}\alpha_1 + p_n) &\Rightarrow \sqrt{d}(q_{n+1} + q_n(\sqrt{d} - a_0)) = p_{n+1} + p_n(\sqrt{d} - a_0)\\ &\Rightarrow \sqrt{d}(q_{n+1} - a_0 q_n - p_n) + q_n d - p_{n+1} + a_0 p_n = 0.\end{aligned}$$

But $\sqrt{d} \notin \mathbb{Q}$ so $q_{n+1} - a_0 q_n = p_n = 0$, so that $q_{n+1}p_n - p_n^2 = a_0 p_n q_n$ and $p_{n+1}q_n - dq_n^2 = a_0 p_n q_n$. These imply that

$$p_n^2 - dq_n^2 = p_{n+1}q_n - q_{n+1}p_n = (-1)^{n+1},$$

and the result follows from $n + 1 = lm$. □

Are there naturally occurring real numbers with nice continued fractions which do not lie in $\mathbb{Q}(\sqrt{d})$ for any squarefree d?

Yes, for example $e - 1 = [1, 1, 2, 1, 1, 4, 1, 1, 6, \cdots]$.

To see this we introduce the following function. Let $c \in \mathbb{R} \setminus \mathbb{N} \cup \{0\}$ be a real number. Define

$$f_c(x) = \sum_{n=0}^{\infty} \frac{1}{c(c+1)\cdots(c+n-1)} \frac{x^n}{n!},$$

for $x \in \mathbb{R}$. This series converges absolutely for all $x \in \mathbb{R}$.

We can check that $f_c(x) = f_{c+1}(x) + \dfrac{x}{c(c+1)} f_{c+2}(x)$, since

$$\frac{1}{c(c+1)\cdots(c+n-1)} \frac{1}{n!} = \frac{1}{(c+1)\cdots(c+n)} \frac{1}{n!} + \frac{1}{c(c+1)\cdots(c+n)} \frac{1}{(n-1)!}.$$

Thus, for $f_c(x) \neq 0$, we have

$$\begin{aligned}\frac{f_{c+1}(x)}{f_c(x)} &= \frac{f_{c+1}(x)}{f_{c+1}(x) + \frac{x}{c(c+1)} f_{c+2}(x)}\\ &= \frac{1}{1 + \frac{x}{c(c+1)} \frac{f_{c+2}(x)}{f_{c+1}(x)}},\end{aligned}$$

when $f_{c+1}(x) \neq 0$. We put $x = z^2$ to obtain

$$\frac{f_{c+1}(z^2)}{f_c(z^2)} = \frac{1}{\frac{z}{c}\left(\frac{c}{z} + \frac{z}{c+1} \frac{f_{c+2}(z^2)}{f_{c+1}(z^2)}\right)}$$

so

$$\frac{z}{c} \frac{f_{c+1}(z^2)}{f_c(z^2)} = \frac{1}{\frac{c}{z} + \frac{z}{c+1} \frac{f_{c+2}(z^2)}{f_{c+1}(z^2)}}.$$

Therefore we have

$$\frac{z}{c} \frac{f_{c+1}(z^2)}{f_c(z^2)} = \left[0, \frac{c}{z}, \frac{c+1}{z}, \cdots, \frac{c+n}{z}, \alpha_{n+2}\right].$$

Now choose c, z so that $\frac{c}{z}, \frac{c+1}{z}, \cdots$ are positive integers. Then $\alpha_{n+2} \geq 1$ for $n \geq 0$ and

$$\frac{z}{c}\frac{f_{c+1}(z^2)}{f_c(z^2)} = \left[0, \frac{c}{z}, \frac{c+1}{z}, \cdots\right].$$

We observe that if we take $c = 1/2$ abd $z = \frac{1}{2y}, y \in \mathbb{N}$ then the conditions hold. Thus

$$\frac{1}{y}\frac{f_{3/2}(1/4y^2)}{f_{1/2}(1/4y^2)} = [0, y, 3y, 5y, \cdots].$$

Put $w = 1/y$. Then

$$\begin{aligned}
f_{\frac{1}{2}}\left(\frac{w^2}{4}\right) &= w\sum_{n=0}^{\infty}\frac{1}{\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\cdots\left(\frac{2n+1}{2}\right)}\frac{w^{2n}}{n!4^n} \\
&= \sum_{n=0}^{\infty}\frac{4^n n!}{(2n)!}\frac{w^{2n}}{n!4^n} \\
&= w\sum_{n=0}^{\infty}\frac{w^{2n}}{(2n)!} \\
&= \frac{e^w + e^{-w}}{2}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
wf_{\frac{3}{2}}\left(\frac{w^2}{4}\right) &= w\sum_{n=0}^{\infty}\frac{1}{\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\cdots\left(\frac{2n+1}{2}\right)}\frac{w^{2n}}{n!4^n} \\
&= \sum_{n=0}^{\infty}\frac{w^{2n+1}}{(2n+1)!} \\
&= \frac{e^w - e^{-w}}{2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{wf_{3/2}(w^2/4)}{f_{1/2}(w^2/4)} &= \frac{e^w - e^{-w}}{e^w + e^{-w}} \\
&= \frac{e^{1/y} - e^{-1/y}}{e^{1/y} + e^{-1/y}} \\
&= [0, y, 3y, 5y, \cdots].
\end{aligned}$$

If we take $y = 2$ we find that

$$\frac{e-1}{e+1} = [0, 2, 6, 10, 14, \cdots].$$

Theorem 0.27. $e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \cdots]$. *That is,* $a_0 = 2, a_1 = 1$, *and* $a_{3k} = a_{3k+1} = 1$ *for all* $k \geq 1$, *and* $a_{3k+2} = 2(k+1)$.

Proof. Let $\alpha = [2, 1, 2, 1, 1, 4, 1, 1, 6, \cdots]$ and put $\theta = \dfrac{e+1}{e-1} = [2, 6, 10, 14, \cdots]$. Let $\dfrac{r_n}{s_n}$ be the nth convergent to θ and let $\dfrac{p_n}{q_n}$ be the nth convergent to α. Notice that $e = \dfrac{\theta+1}{\theta-1}$ since

$$\frac{\frac{e+1}{e-1}+1}{\frac{e+1}{e-1}-1} = \frac{\frac{2e+1-1}{e-1}}{e+1-e+1}e-1 = e.$$

Since $r_n/s_n \to \theta$ it is enough to show that $p_{3n+1} = r_n + s_n$ and $q_{3n+1} = r_n - s_n$ since then

$$\begin{aligned}\frac{p_{3n+1}}{q_{3n+1}} &= \frac{r_n + s_n}{r_n - s_n}\\ &= \frac{r_n/s_n + 1}{r_n/s_n - 1}\\ &\to e.\end{aligned}$$

This would then show that $\alpha = e$.

We will prove $p_{3n+1} = r_n + s_n$ and $q_{3n+1} = r_n - s_n$ for $n = 0, 1, 2, \cdots$ by induction. For $n = 0$, we have $r_0 = 2, s_0 = 1$ and $p_1 = 3, q_1 = 1$ and for $n = 1$ we have $r_1 = 13, s_1 = 6$, with $p_4 = 19, q_4 = 7$ and so we are done for $n = 0, 1$. For $n \geq 2$ we have $r_n = (4n+2)r_{n-1} + r_{n-2}$ and $s_n = (4n+2)s_{n-1} + s_{n-2}$. In addition, we have $p_{3n-3} = p_{3n-4} + p_{3n-5}, p_{3n-1} = 2np_{3n-2} + p_{3n-3}$, $p_{3n} = p_{3n-1} + p_{3n-2}$, and $p_{3n+1} = p_{3n} + p_{3n-1}$. These imply that

$$\begin{aligned} p_{3n-3} &= p_{3n-4} + p_{3n-5}\\ -p_{3n-2} &= -p_{3n-3} - p_{3n-4}\\ 2p_{3n-1} &= 4np_{3n-2} + 2p_{3n-3}\\ p_{3n} &= p_{3n-1} + p_{3n-2}\\ p_{3n+1} &= p_{3n} + p_{3n-1}\end{aligned}$$

Adding these, we obtain

$$p_{3n+1} = (4n+2)p_{3n-2} + p_{3n-5}$$

and similarly, we obtain

$$q_{3n+1} = (4n+2)q_{3n-2} + q_{3n-5}.$$

It now follows from the recurrence for r_n, s_n and the inductive hypothesis that $p_{3n+1} = r_n + s_n$ and $q_{3n+1} = r_n - s_n$ for $n = 0, 1, 2, \cdots$ □

It is possible to determine the continued fraction expression of $e^{2/y}$ for all $y \in \mathbb{N}$ in this way. Notice that if $\dfrac{p_n}{q_n}$ are the convergents to e then

$$q_{3m-1} \geq \prod_{j=1}^{m}(2j) = 2^m m! \geq \left(\frac{2m}{e}\right)^m.$$

By proposition 0.19, $\left|e - \frac{p_n}{q_n}\right| > \frac{1}{(q_{n+1}+2)q_n^2}$. If $n+1 = 3m-1$ for some $m \in \mathbb{N}$ then $a_{n+1} = 2m = 2\frac{n+2}{3}$ and otherwise $a_{n+1} = 1$. Thus $a_{n+2} + 2 \leq 4n$. But

$$q_n > q_{3\lfloor n/3\rfloor - 1} \geq \left(\frac{2\lfloor n/3\rfloor}{e}\right)^{\lfloor n/3\rfloor}.$$

For $n \geq 3$, $\lfloor n/3\rfloor \geq n/6$ so for $n \geq 3$, we have $q_n \geq \left(\frac{n}{3e}\right)^{n/6}$. Thus there is a positive real number c such that for $n \geq 4$, we have

$$4n < c\frac{\log q_n}{\log\log q_n}.$$

Therefore for $n \geq 4$, we have

$$\left|e - \frac{p_n}{q_n}\right| > \frac{1}{c\frac{\log q_n}{\log\log q_n}q_n^2}.$$

Recall that if p/q is a rational with $\left|e - \frac{p}{q}\right| < \frac{1}{2q^2}$, then p/q is a convergent to e. Therefore there is a positive number c_1 such that if $q > 4$ then

$$\left|e - \frac{p}{q}\right| > \frac{c_1 \log\log q}{(\log q)q^2}.$$

Notice that e cannot be as well approximated by rationals as a typical real number since for almost all reals α the inequality

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

has infinitely many solutions in rationals p/q.

The continued fraction expansion for π is $\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, \cdots]$. No patter has been discerned to date.

Mahler in 1953 proved that there exists $c > 0$ such that $\left|\pi - \frac{p}{q}\right| > \frac{c}{q^{42}}$. In 1993 Hata proved $\left|\pi - \frac{p}{q}\right| > \frac{1}{q^{8.017}}$ for all sufficiently large q. Salikhov proved that $\left|\pi - \frac{p}{q}\right| > \frac{1}{q^{7.6065\cdots}}$.

General question: How do we expect the q_n's to grow and how do we expect the partial quotients to be distributed for a typical real number?

Observe that $q_0 = 1, q_1 = a_1$ and $q_n = a_n q_{n-1} + q_{n-2}$. Note that $q_n \geq u_n$ where $u_0 = 0, u_1 = 1$ and $u_n = u_{n-1} + u_{n-2}$ for $n \geq 2$. Here $u_n = F_n$ is the nth Fibonacci number and

$$u_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Thus $q_n \geq \frac{1}{2\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n$.

Theorem 0.28. *There exists a positive number c such that for all α except a set of Lebesgue measure zero, such that $q_n = q_n(\alpha) < e^{cn}$ for all n sufficiently large with respect to α.*

Proof. (Khintchine) Clearly we may restrict to α in $(0,1)$, since a countable union of sets of measure 0 remains measure 0. Let $g \geq 1$ be a real number and $n \in \mathbb{N}$. We define $E_n(g)$ to be the set of $\alpha \in (0,1)$ for which $a_1 \cdots a_n \geq g$.

Let $(a_1, \cdots, a_n)$ be a sequence of positive integers. We now determine the measure of the set of α's in $(0,1)$ whose first n partial quotients are $a_1, \cdots, a_n$. That is, $\alpha = [0, a_1, \cdots, a_n, \alpha_{n+1}]$.

We have $\alpha = \frac{p_n \alpha_{n+1} + p_{n-1}}{q_n \alpha_{n+1} + q_{n-1}}$ with $\alpha_{n+1} \in [1, \infty)$. Therefore α is in an interval with endpoints $\frac{p_n}{q_n}, \frac{p_n + p_{n-1}}{q_n + q_{n-1}}$. To see this note that

$$\left| \alpha - \frac{p_n}{q_n} \right| = \left| \frac{p_n \alpha_{n+1} + p_{n-1}}{q_n \alpha_{n+1} + q_{n-1}} - \frac{p_n}{q_n} \right| = \frac{1}{q_n(q_n \alpha_{n+1} + q_{n-1})},$$

which is a monotone function of α_{n+1}.

The length of the interval is $\frac{1}{q_n(q_n + q_{n-1})} < \frac{1}{q_n^2}$ and since $q_n > a_n q_{n-1}$ we see that the length is less than $\frac{1}{(a_1 \cdots a_n)^2}$.

Recall that $E_n(g)$ is the set of α in $(0,1)$ for which $a_1 \cdots a_n \geq g$. Thus, $\mu(E_n(g)) < \sum_{a_1 \cdots a_n \geq g} \frac{1}{(a_1 \cdots a_n)^2}$. Note that

$$\prod_{i=1}^n \frac{1}{a_i^2} = \prod_{i=1}^n \left(\frac{a_i + 1}{a_i} \right) \frac{1}{a_i(a_i+1)} \leq 2^n \prod_{i=1}^n \frac{1}{a_i(a_i+1)}.$$

But

$$\begin{aligned} \prod_{i=1}^n \frac{1}{a_i(a_i+1)} &= \prod_{i=1}^n \int_{a_i}^{a_i+1} \frac{dx_i}{x_i^2} \\ &= \int_{a_1}^{a_1+1} \cdots \int_{a_n}^{a_n+1} \frac{dx_1 \cdots dx_n}{x_1^2 \cdots x_n^2}. \end{aligned}$$

Put $J_n(g) = \int_R \frac{dx_1 \cdots dx_n}{x_1^2 \cdots x_n^2}$ where R is the region of $(x_1, \cdots, x_n) \in \mathbb{R}^n$ with $x_i \geq 1$ for $i = 1, 2, \cdots, n$ and $x_1 \cdots x_n \geq g$. We see that $\mu(E_n(g)) \leq 2^n J_n(g)$ and so it remains to estimate $J_n(g)$.

If $g \leq 1$, then $J_n(g) = \left(\int_1^\infty \frac{dx}{dx^2} \right)^n = -1$.

We will prove for $g > 1$ that

$$J_n(g) = \frac{1}{g}\sum_{i=0}^{n-1} \frac{(\log g)^i}{i!}.$$

We will do this by induction on n. For $n = 1$ we have $J_1(g) = \int_g^\infty \frac{dx_1}{x_1^2} = \frac{1}{g}$, as required. Let us assume that the result holds for $n = k$, for some $k \geq 1$. Then

$$J_{k+1}(g) = \int_1^\infty \frac{dx_{k+1}}{x_{k+1}^2} J_k\left(\frac{g}{x_k}\right).$$

Apply the change of variable $u = \frac{g}{x_{k+1}}$ so $du = \frac{-g dx_{k+1}}{x_{k+1}^2}$, hence

$$\begin{aligned}
J_{k+1}(g) &= \int_g^0 \frac{-1}{g} J_k(u)du \\
&= \int_0^g \frac{1}{g} J_k(u)du \\
&= \frac{1}{g}\int_0^1 J_k(u)du + \frac{1}{g}\int_1^g J_k(u)du \\
&= \frac{1}{g} + \frac{1}{g}\int_1^g \frac{1}{u}\left(\sum_{i=0}^{k-1}\frac{(\log u)^i}{i!}\right) du \\
&= \frac{1}{g} + \frac{1}{g}\sum_{i=0}^{k-1}\frac{(\log u)^{i+1}}{(i+1)!}|_1^g \\
&= \frac{1}{g} + \frac{1}{g}\sum_{i=0}^{k-1}\frac{(\log g)^{i+1}}{(i+1)!} \\
&= \frac{1}{g}\sum_{i=0}^{k}\frac{(\log g)^i}{i!},
\end{aligned}$$

as desired.

We see that $\mu(E_n(g)) \leq 2^n J_n(g) = 2^n \frac{1}{g}\sum_{i=0}^{n-1}\frac{(\log g)^i}{i!}$. Now take $g = e^{An}$ for a positive real number $A \geq 1$. Then

$$\begin{aligned}
\mu(E_n(g)) &\leq 2^n \frac{1}{g}\sum_{i=0}^{n}\frac{(An)^i}{i!} \\
&\leq 2^n A^n \frac{1}{e^{An}}\sum_{i=0}^{n}\frac{n^i}{i!} \\
&\leq e^{-An} 2^n A^n e^n \\
&= \exp((\log 2 + \log A + 1 - A)n).
\end{aligned}$$

Choose A so that $\log 2 + \log A + 1 - A < 0$. Then $\sum_{n=0}^{\infty} \mu(E_n(e^{An}))$ converges. By the Borel-Cantelli Lemma, we see that almost all α's in the sense of Lebesgue measure will belong to only finitely many of the $E_n(e^{An})$'s. Thus for almost all α, $a_1 \cdots a_n < e^{An}$ for n sufficiently large in terms of α.

But $q_n = a_n q_{n-1} + q_{n-2}$, hence $q_n \le 2a_n q_{n-1}$ and so $q_n \le 2^n a_1 \cdots a_n$. Therefore $q_n \le 2^n e^{An} = e^{(\log 2 + A)n}$. □

In 1935 Paul Lévy proved by probabilistic arguments that $q_n^{1/n} \to \exp(\pi^2/(12\log 2))$ for almost all real numbers α. To prove results of thsi sort we will use ergodic theory.

Consider a probability space $(\Omega, \Sigma, \mathbb{P})$ consisting of a set Ω, a σ-algebra Σ on Ω, and $\mathbb{P}$ a probability measure on Σ (so that $\mathbb{P}(\Omega) = 1$). We say that $T : \Omega \to \Omega$ is a measure preserving transformation on $(\Omega, \Sigma, \mathbb{P})$ if for $B \in \Sigma$ we have $T^{-1}B \in \Sigma$ and $\mathbb{P}(T^{-1}(B)) = \mathbb{P}(B)$. Let L^1 be the measureable functions of f from Ω to Ω which are integrable. Then if T is measure preserving and $f \in L^1$, we have

$$\int_\Omega f d\mathbb{P} = \int_\Omega (f \circ T) d\mathbb{P}.$$

Definition 0.29. Let T be a measure preserving transformation in a probability space $(\Omega, \Sigma, \mathbb{P})$. Then T is said to be ergodic if whenever $B \in \Sigma$ and $T^{-1}B \subset B$, we have $\mu(B) \in \{0, 1\}$.

Theorem 0.30. (Ergodic Theorem) *Suppose $f \in L^1$ and T is ergodic. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j \alpha) = \int_\Omega f d\mathbb{P}$$

for almost all $\alpha \in \Omega$ with respect to $\mathbb{P}$.

Let $X = (0, 1) \subset \mathbb{R}$ and $\mathcal{B}$ the Borel σ-algebra on $(0, 1)$, and $\mu = \mathbb{P}$ the Lebesgue measure of $(0, 1)$. Let $T : X \to X$ be defined by $T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$. T is not measure preserving with respect to Lebesgue measure, but we can modify μ to give us μ_1, where for all $f \in L^1$ we have

$$\mu_1(f) = \frac{1}{\log 2} \int_0^1 \frac{f(x)}{1+x} dx.$$

Note that μ_1 is still a probability measure.

We claim that T is measure preserving with respect to μ_1. It suffices to check that T is measure preserving on any interval (a, b) with $(a, b) \subset (0, 1)$.

We have $T^{-1}((a, b)) = \bigcup_{n=1}^{\infty} \left(\frac{1}{b+n}, \frac{1}{a+n} \right)$. Since if $a \le \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \le b$, then $\frac{1}{x} = n + \theta$ for some $n \in \mathbb{N}$ with $a \le \theta \le b$. Certainly $\bigcup_{n=1}^{\infty} \left(\frac{1}{b+n}, \frac{1}{a+n} \right)$ is measureable, and since the intervals are disjoint, we have

$$\begin{aligned}
\mu_1\left(\bigcup_{n=1}^{\infty}\left(\frac{1}{b+n},\frac{1}{a+n}\right)\right) &= \sum_{n=1}^{\infty}\mu_1\left(\left(\frac{1}{b+n},\frac{1}{a+n}\right)\right)\\
&= \sum_{n=1}^{\infty}\frac{1}{\log 2}\int_{\frac{1}{b+n}}^{\frac{1}{a+n}}\frac{dx}{1+x}\\
&= \sum_{n=1}^{\infty}\frac{1}{\log 2}\log(1+x)\Big|_{1/(b+n)}^{1/(a+n)}\\
&= \frac{1}{\log 2}\sum_{n=1}^{\infty}\left(\log\left(1+\frac{1}{a+n}\right)-\log\left(1+\frac{1}{b+n}\right)\right)\\
&= \frac{1}{\log 2}\sum_{n=1}^{\infty}\left[\log\left(\frac{a+n+1}{a+n}\right)-\log\left(\frac{b+n+1}{b+n}\right)\right].
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{n=1}^{N}\left[\log\left(\frac{a+n+1}{b+n+1}\right)-\log\left(\frac{a+n}{b+n}\right)\right] &= \log\left(\frac{a+N+1}{b+N+1}\right)-\log\left(\frac{a+1}{b+1}\right)\\
&= \log\left(\frac{b+1}{a+1}\right)-\log\left(\frac{b+N+1}{a+N+1}\right)\\
&\to \log\left(\frac{b+1}{a+1}\right)
\end{aligned}$$

as $N\to\infty$. Hence

$$\mu_1(T^{-1}((a,b))) = \frac{1}{\log 2}\log\left(\frac{b+1}{a+1}\right) = \mu_1((a,b)),$$

so T is a measure preserving transformation with respect to μ_1.

This invariant measure for the transformation T was discovered by Gauss in 1812.

Given $\alpha\in\mathbb{R}$, recall that $\alpha_n = a_n + \frac{1}{\alpha_{n+1}}$ for $n = 0,1,2,\cdots$ or $\alpha_n - a_n = \frac{1}{\alpha_{n+1}}$. This is equivalent to $\left(\frac{1}{\alpha_n}\right)^{-1} - \left(\frac{1}{a_n}\right)^{-1} = \frac{1}{\alpha_{n+1}}$. Note that $\alpha_n \geq 1$ for $n\geq 1$.

Therefore we have that

$$\begin{aligned}
T\left(\frac{1}{\alpha_n}\right) &= \alpha_n - \lfloor\alpha_n\rfloor\\
&= \alpha + n - a_n\\
&= \frac{1}{\alpha_{n+1}}.
\end{aligned}$$

It can be proved that T is ergodic with respect to μ_1. We can take f to be the characteristic function of $\left(\frac{1}{k+1},\frac{1}{k}\right)$ for $k\in\mathbb{N}$ and apply the ergodic theorem to conclude that for almost all α in the sense of the measure μ_1 and hence in the sense of Lebesgue measure,

$$\begin{aligned}\lim_{n\to\infty}\sum_{j=0}^{\infty} f(T^j\alpha) &= \frac{1}{\log 2}\int_X f d\mu_1 \\ &= \frac{1}{\log 2}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{dx}{1+x} \\ &= \frac{1}{\log 2}\log(1+x)\big|_{1/(k+1)}^{1/k} \\ &= \frac{1}{\log 2}\log\left(\frac{(k+1)^2}{k(k+2)}\right)\end{aligned}$$

Therefore for almost all real numbers α, in the sense of Lebesgue measure, the frequency with which k appears as a partial quotient in the continued fraction expansion of α is $\dfrac{1}{\log 2}\log\left(\dfrac{(k+1)^2}{k(k+2)}\right)$. Gauss had conjectured this and it was proved by Kuzman in the 1920s. Thus the expected frequency of 1's is $0.41503\cdots$, of 2's is $0.169925\cdots$, etc.

Observe that if $\alpha = \alpha_0 \in (0,1)$, then $T^n(\alpha) = 1/\alpha_{n+1}$ for $n = 0, 1, \cdots$. Further, $a_n = \lfloor \alpha_n \rfloor$. Thus

$$(a_1\cdots a_n)^{1/n} = (\lfloor T^0(\alpha)^{-1}\rfloor \cdots \lfloor T^{n-1}(\alpha)^{-1}\rfloor)^{1/n}$$

and so

$$\frac{1}{n}\sum_{i=1}^{\infty}\log a_i = \frac{1}{n}\sum_{i=0}^{\infty}\log\left\lfloor\frac{1}{T^i(\alpha)}\right\rfloor.$$

We now take $f(x) = \log\left\lfloor\dfrac{1}{x}\right\rfloor$ and apply the ergodic theorem to deduce that for almost all $\alpha \in (0,1)$, in the sense of Lebesgue measure, that

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\log a_i &= \int_0^1 \frac{1}{\log 2}\frac{\log\lfloor 1/x\rfloor}{1+x}dx \\ &= \frac{1}{\log 2}\sum_{n=1}^{\infty}\log n\int_{1/(n+1)}^{1/n}\frac{dx}{1+x} \\ &= \frac{1}{\log 2}\sum_{n=1}^{\infty}(\log n)\log\left(\frac{1+1/n}{1+1/(n+1)}\right) \\ &= \frac{1}{\log 2}\sum_{n=2}^{\infty}\log n\log\left(\frac{(n+1)^2}{n(n+2)}\right)\end{aligned}$$

Equivalently, we have

$$(a_1\cdots a_n)^{1/n} \to \prod_{n=2}^{\infty}\left(\frac{(n+1)^2}{n(n+2)}\right)^{\frac{\log n}{\log 2}}$$

for almost all α in the sense of Lebesgue measure.

We will deduce the Khintchine-Lévy result about the growth of q_n for almost all α. First we observe that if $[0, a_1, a_2, \cdots] = p_n/q_n$, then

$$q_n = [a_1, \cdots, a_n][a_2, \cdots, a_n] \cdots [a_n] \tag{0.5}$$

since if $[a_j, \cdots, a_n] = \dfrac{x}{b}$ then $[a_{j+1}, \cdots, a_n] = \dfrac{b}{c}$ and so we get equation (0.5) by a telescoping product with first term $\dfrac{q_n}{p_n}$ and last term $a_n/1$.

As an aside, note that $[a_j, \cdots, a_1] = q_j/q_{j-1}$ so $q_n = [a_n, \cdots, a_1] \cdots [a_1]$.

We will first show that if the first $n+1$ partial quotients of α are $[0, a_1, \cdots, a_n]$ then $|\log(T^i(\alpha)) - \log(T^i(p_n/q_n))| < 2^{\frac{-1}{2}(n-1-i)+1}$. We do this by induction on n. It suffices to prove this for $i = 0$. Since α is in an interval with end points $\dfrac{p_n}{q_n}$ and $\dfrac{p_n + p_{n-1}}{q_n + q_{n-1}}$, we have

$$\left|\log\left(\frac{\alpha}{p_n/q_n}\right)\right| \leq \left|\log\left(\frac{\frac{p_n+p_{n-1}}{q_n+q_{n-1}}}{p_n/q_n}\right)\right| = \left|\log\left(\frac{q_n(p_n+p_{n-1})}{p_n(q_n+q_{n-1})}\right)\right|.$$

But

$$\left|\frac{q_n(p_n+p_{n-1})}{p_n(q_n+q_{n-1})} - \frac{p_n(q_n+q_{n-1})}{p_n(q_n+q_{n-1})}\right| = \left|\frac{q_n p_{n-1} - p_n q_{n-1}}{p_n(q_n+q_{n-1})}\right| = \frac{1}{p_n(q_n+q_{n-1})}.$$

Thus $\log\left(\dfrac{\alpha}{p_n/q_n}\right) = \log(1+t)$, with $|t| \leq \dfrac{1}{p_n(q_n+q_{n-1})}$. Now $|\log(1-x)| < 2x$ for $0 < x \leq 1/2$ and $|\log(1+x)| < x$ for the same range. Therefore $\left|\log \alpha - \log\left(\dfrac{p_n}{q_n}\right)\right| < \dfrac{2}{p_n(q_n+q_{n-1})}$ for $n = 1, 2, \cdots$, and since $q_n \geq 2^{\frac{1}{2}(n-1)}$, so

$$\left|\log \alpha - \log\left(\frac{p_n}{q_n}\right)\right| < \frac{2}{2^{\frac{1}{2}(n-1)}},$$

for $n = 1, 2, \cdots$. Therefore

$$\begin{aligned}
\left|\sum_{i=0}^{n-1}\left(\log(T^i(\alpha)) - \log\left(T^i\left(\frac{p_n}{q_n}\right)\right)\right)\right| &< \sum_{i=0}^{n-1} 2^{\frac{-1}{2}(n-1-i)+1} \\
&< 2\sum_{j=0}^{\infty}\left(\frac{1}{\sqrt{2}}\right)^j \\
&= 2\frac{1}{1-1/\sqrt{2}} \\
&= \frac{2\sqrt{2}}{\sqrt{2}-1} < 7.
\end{aligned}$$

Since $-\log q_n = \sum_{i=0}^{n-1} \log\left(T^i\left(\frac{p_n}{q_n}\right)\right)$ we have

$$\left|\sum_{i=0}^{n} \log\left(T^i(\alpha)\right) + \log q_n\right| < 7.$$

Hence, we have

$$\left|\frac{1}{n}\sum_{i=0}^{n} \log(T^i(\alpha))^{-1} - \log q_n\right| < \frac{7}{n}.$$

Therefore for all irrational α we have

$$\lim_{n\to\infty} \frac{1}{n}\left(\sum_{i=0}^{n-1} \log(T^i(\alpha))^{-1} - \log q_n\right) = 0.$$

Thus by the ergodic theorem, with $f(x) = \log(1/x)$, we find that for almost all α, in the sense of Lebesgue measure, we have

$$\lim_{n\to\infty} \frac{\log q_n}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \log(T^i(\alpha))^{-1} = \frac{1}{\log 2}\int_0^1 \frac{\log(1/x)}{1+x}dx.$$

Or equivalently,

$$\lim_{n\to\infty} q_n^{1/n} = \exp\left(\frac{1}{\log 2}\int_0^1 \frac{\log(1/x)}{1+x}dx\right).$$

It remains to show that $\displaystyle\int_0^1 \frac{\log(1/x)dx}{1+x} = \frac{\pi^2}{12}$.

Let $f(x) = \log x$ and $g(x) = \log(1+x)$. Then

$$\int_0^1 \left(\frac{\log(x+1)}{x} + \frac{\log x}{1+x}\right) dx = \log(x)\log(x+1)|_0^1.$$

Since $\lim_{x\to 0^+} \log(x)\log(1+x) = 0$ and $\lim_{x\to 1^-} \log(x)\log(1+x) = 0$, we have

$$\int_0^1 \left(\frac{\log(1+x)}{x} + \frac{\log x}{1+x}\right) dx = 0.$$

Hence

$$\begin{aligned}
\int_0^1 \frac{\log(1/x)}{1+x}dx &= \int_0^1 \frac{\log(1+x)}{x}dx \\
&= \int_0^1 \frac{-1}{x}\left(\sum_{n=1}^{\infty}\frac{(-1)^n x^n}{n}\right)dx \\
&= \sum_{n=1}^{\infty}\int_0^1 \frac{x^{n-1}(-1)^{n-1}}{n}dx \\
&= \sum_{n=1}^{\infty}\frac{x^n(-1)^n}{n^2}\Big|_0^1 \\
&= 1-\frac{1}{4}+\frac{1}{9}-\cdots \\
&= \frac{\pi^2}{8}-\frac{\pi^2}{24}=\frac{\pi^2}{12},
\end{aligned}$$

as required.

Recall the Euclidean algorithm. Given positive integers u and v with $v \geq u$ we compute the gcd of u, v by putting $r_0 = v, r_1 = u$ and $r_{m-1} = a_m r_m + r_{m+1}$ for $m = 1, 2, \cdots$ where a_i's are positive integers and $r_0 \geq r_1 > r_2 > \cdots r_{n+1} = 0$. Thus $\gcd(u, v) = r_n$.

Notice that if $\gcd(u, v) = 1$ then $\frac{v}{u} = [a_m, \cdots, a_1]$. Thus the number of applications of the division algorithm in the Euclidean algorithm for u and v correspond to the length of the continued fraction expression of v/u.

Given two positive numbers u and v with $u \leq v$ let $L(u, v)$ be the number of steps in the Euclidean algorithm to determine $\gcd(u, v)$. In 1970 J. Dixon proved that for $\varepsilon > 0$ there exists $c_0(\varepsilon) > 0$ such that

$$\left|L(u,v)-\frac{12\log 2}{\pi^2}\log v\right| < (\log v)^{\frac{1}{2}+\varepsilon}$$

for all except at most $x^2\exp(-c_0(\log(x))^{\varepsilon/2})$ of the pairs (u, v) with $1 \leq u \leq v \leq x$.

Heilbronn had proved earlier that for each positive integer $v > 10$, we have

$$\frac{1}{\varphi(v)}\sum_{\substack{u=1\\ \gcd(u,v)=1}}^{v} L(u,v)-\frac{12\log 2}{\pi^2}\log v = O((\log\log v)^4).$$

How well can we approximate real algebraic numbers of degree at least 3? The first result of interest was proved by Liouville in 1844.

Theorem 0.31. (Liouville) *Let α be an algebraic number of degree d with $d > 1$. There exists a positive number $c(\alpha)$, which us effectively computable in terms of α,* *such that*

$$\left|\alpha - \frac{p}{q}\right| > \frac{c(\alpha)}{q^d},$$

for every p/q with $q > 0$.

Proof. Let f be the minimal polynomial for α over $\mathbb{Z}$. That is, f is the polynomial in $\mathbb{Z}[x]$ of degree d, with coprime coefficients and positive leading coefficient which has α as a root.

We may assume α is real since if α is not real we may take $c(\alpha) = \frac{1}{2}\min_{\theta\in\mathbb{R}} |\alpha - \theta|$. Since $d > 1$ we have $f\left(\frac{p}{q}\right) \neq 0$. Thus by the mean value theorem, we get

$$\frac{1}{q^d} \le \left|f\left(\frac{p}{q}\right)\right| = \left|f(\alpha) - f\left(\frac{p}{q}\right)\right| = \left|\alpha - \frac{p}{q}\right| |f'(\theta)|$$

where θ is a real number between α and p/q. Note that if $\left|\alpha - \frac{p}{q}\right| \ge 1$ the result holds with $c(\alpha) = 1/2$, and so we may suppose that $\left|\alpha - \frac{p}{q}\right| < 1$.

If $f(x) = a_d x^d + a_{d-1}x^{d-1} + \cdots + a_0$ then $f'(x) = da_d x^{d-1} + \cdots + a_1$ and so

$$|f'(\theta)| \le da_d(|\alpha| + 1)^{d-1} + \cdots + |a_1|.$$

Here we can take $c(\alpha)^{-1} = 2(da_d(|\alpha| + 1)^{d-1} + \cdots + |a_1|)$. □

Liouville constructed the first numbers known to be transcendental with his result.

Theorem 0.32. *The number* $\sum_{n=1}^{\infty} \frac{1}{10^{n!}}$ *is transcendental.*

Proof. Let $\alpha = \sum_{n=1}^{\infty} \frac{1}{10^{n!}}$ and $s_N = \sum_{n=1}^{N} \frac{1}{10^{n!}}$. Clearly, if $s_N = p_N/q_N$ for positive integers p_N, q_N with $\gcd(p_N, q_N) = 1$, then $q_N = 10^{N!}$. Thus we have

$$\begin{aligned}\left|\alpha - \frac{p_N}{q_N}\right| &= \sum_{n=N+1}^{\infty} \frac{1}{10^{n!}} \\ &< \frac{1}{10^{N(N!)}} \\ &= \frac{1}{q_N^N}\end{aligned}$$

Which shows that the conditions of Liouville's theorem do not hold, and hence α is not algebraic. □

Let α be an algebraic number of degree d. The inequality $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\mu}$ has only finitely many solutions in rationals p/q if $\mu > d$ (Liouville's Theorem), if $\mu > d/2 + 1$ (Thue), if $\mu > 2\sqrt{d}$ (Siegel), if $\mu > \sqrt{2d}$ (Dyson), and if $\mu > 2$ (Roth).

Theorem 0.33. (Roth's Theorem) *Let α be an algebraic number and let $\varepsilon > 0$ be a positive real number. Then there exist only finitely many distinct rationals p/q with $q > 0$ for which*

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

Remark 0.34. In light of Dirichlet's Theorem, Roth's Theorem is essentially best possible. In view of Khintchine's Theorem one might expect improvements of Roth's Theorem with $\frac{1}{q^{2+\varepsilon}}$ replaced by $\frac{1}{q^2(\log q)^{1+\varepsilon}}$, but no progress has been made in this direction.

Notice that Roth's Theorem tells us that $a_{n+1} < q_n^{\varepsilon}$ for n sufficiently large. Recall that $q_0 = 1, q_1 = a_1, q_n = a_n q_{n-1} + q_{n-2}$ for $n = 2, 3, \cdots$. Thys $q_n \leq (a_n+1)\cdots(a_1+1)$, whence

$$a_{n+1} < ((a_1+1)\cdots(a_n+1))^{\delta}$$

for n sufficiently large. It follows that $\log\log q_n < c(\alpha)n$ where $c(\alpha)$ is a positive number which depends on α. Davenport and Roth (1955) proved that for each real algebraic number α there is a positive number $c_1(\alpha)$, which depends on α, such that $\log\log q_n < c_1(\alpha)\frac{n}{\sqrt{\log n}}$.

Perhaps the most important applications of Roth's Theorem is to the study of Diophantine equations. Let $m \in \mathbb{N}$. Consider the Diophantine equation $x^3 - 2y^3 = m$, in integers x, y. This equations implies that

$$\left|\frac{x^3}{y^3} - 2\right| = \frac{m}{y^3}$$

which by Roth's Theorem can only be satisfied by at most finitely many pairs of x, y.

Let $F(x,y) = a_n x^n + a_{n-1}x^{n-1}y + \cdots + a_0 y^n \in \mathbb{Z}[x,y]$. Suppose that F is not the zero-form. Then F factors over $\mathbb{C}$ in the form $F(x,y) = L_1(x,y)L_2(x,y)\cdots L_n(x,y)$ where $L_i(x,y) = \gamma_i x + \delta_i y$ for $i = 1, 2, \cdots, n$. Suppose that the discriminant of F is non-zero, or equivalently $i \neq j$ implies that L_i and L_j are linearly independent over $\mathbb{C}$ so F does not have multiple factors. Let (x,y) be an integer point with $F(x,y) \neq 0$. Then by re-ordering the forms we may suppose that

$$0 < |L_1(x,y)| \leq |L_2(x,y)| \leq \cdots \leq |L_n(x,y)|.$$

If $\gamma_1 = 0$ or $\gamma_1 \neq 0$ and $\frac{\delta_1}{\gamma_1} \in \mathbb{Q}$ then $|L_1(x,y)| > c_1$ for some positive number c_1. If $\gamma_1 \neq 0$ and $y = 0$ then $|L_1(x,y)| = |\gamma_1|(|x| + |y|)$. Finally if $\gamma_1 \neq 0$, $\frac{\delta_1}{\gamma_1}$ is irrational, and $y \neq 0$ then $L_1(x,y) = \gamma_1 y\left(\frac{x}{y} - \left(\frac{-\delta_1}{\gamma_1}\right)\right)$. Therefore, by Roth's Theorem, for each $\varepsilon > 0$ there exists a positive number $c_2\left(\varepsilon, \frac{-\delta_1}{\gamma_1}\right)$ such that

$$|L_1(x,y)| \geq c_2|y|^{-1-\varepsilon} \geq \frac{c_2}{(|x|+|y|)^{1+\varepsilon}}.$$

Since L_1 and L_2 are linearly independent over $\mathbb{C}$, we have

$$|L_2(x,y)| \geq \frac{1}{2}\left(|L_2(x,y)| + |L_1(x,y)|\right) > c_3(|x|+|y|)$$

for some $c_3 > 0$. Thus

$$|F(x,y)| > \frac{c_2}{(|x|+|y|)^{1+\varepsilon}} c_3^{n-1}(|x|+|y|)^{n-1} = c_2 c_3^{n-1}(|x|+|y|)^{n-2-\varepsilon}.$$

We conclude that if $F(x,y)$ is a binary form of degree n with non-zero discriminant, then for each $\varepsilon > 0$ there are only finitely many integers x, y for which $|F(x,y)| < (|x|+|y|)^{n-2-\varepsilon}$. In particular, if $n \geq 3$ and $m \in \mathbb{N}$ then the equation $F(x,y) = m$ has only finitely many solutions in integers x, y.

The equation $F(x,y) = m$ is known as Thue equation.

Since the constant in Roth's Theorem is not effectively computable, it is not possible to bound the size of the solutions in Thue equations. However, it is possible to bound then number of solutions. The critical point in showing that $F(x,y) = m$ has only finitely many solutions is that one needs an improvement on Liouville's Theorem. If this can be accomplished effectively then one can 'solve' Thue equations. In fact, there exist effective improvements on Liouville's Theorem. They follow from Baker's estimates for linear forms in the logarithm of algebraic numbers.

We will follow Cassel's version of Roth's Theorem. Thue, Siegel, and Dyson proved their results by examining polynomials in two variables. Roth used polynomials in several variables.

First note that we may assume α is an algebraic integer for the proof of Roth's Theorem, for if α has minimal polynomial $a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$, then $a_n\alpha$ is a root of

$$x^n + a_{n-1}x^{n-1} + a_n a_{n-2}x^{n-2} + \cdots + a_1 a_n^{n-1} x + a_0 a_n^{n-1}.$$

$a_n\alpha$ is thus an algebraic integer. Suppose that

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

for q sufficiently large. Thus we may suppose α is an algebraic integer.

Let α be an algebraic integer with minimal polynomial $x^n + a_{n-1}x^{n-1} + \cdots + a_0$. We denote the height of α by $h = \max\{1, |a_{n-1}|, \cdots, |a_0|\}$. For the proof we will employ polynomials of the form

$$R(x_1, \cdots, x_m) = \sum_{\substack{0 \leq j_i \leq r_i \\ 1 \leq i \leq m}} c(j_1, \cdots, j_m) x_1^{j_1} \cdots x_m^{j_m},$$

where $c(j_1, \cdots, j_m) \in \mathbb{R}$ for all $(j_1, \cdots, j_m)$.

We define $\overline{R}$ by $\overline{R} = \max_{\substack{0 \le j_i \le r_i \\ 1 \le i \le m}} |c(j_1, \cdots, j_m)|$ and we define

$$R_{i_1,\cdots,i_m}(x_1, \cdots, x_m) = \frac{1}{i_1!} \cdots \frac{1}{i_m!} \frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdots \frac{\partial^{i_m}}{\partial x_m^{i_m}} R(x_1, \cdots, x_m).$$

Proposition 0.35. *If R has integer coefficients then $R_{i_1,\cdots,i_m}$ has integer coefficients for any non-negative integers $i_1, \cdots, i_m$. If R has degree r_u in variable x_u for $u = 1, 2, \cdots, m$ then $R_{i_1,\cdots,i_m}$ has degree $r_u - i_u$ for $u = 1, 2, \cdots, m$. Further, we have*

$$\overline{R_{i_1,\cdots,i_u}} \le 2^{r_1+\cdots+r_m} \overline{R}.$$

Proof. Since

$$R_{i_1,\cdots,i_m} = \sum_{i_u \le j_u \le r_u} \binom{j_1}{i_1} \cdots \binom{j_m}{i_m} c(j_1, \cdots, j_m) x_1^{j_1-i_1} \cdots x_m^{j_m-i_m},$$

the result follows on noting that $\binom{j_1}{i_1} \cdots \binom{j_m}{i_m} \le 2^{j_1+\cdots+j_m} \le 2^{r_1+\cdots+r_m}$. □

By Taylor's Theorem in several variables, we have

$$R(x_1 + y_1, \cdots, x_m + y_m) = \sum_{0 \le i_u \le r_u} y_1^{i_1} \cdots y_m^{i_m} R_{i_1,\cdots,i_m}(x_1, \cdots, x_m) \tag{0.6}$$

We shall say that R has index I at $(\alpha_1, \cdots, \alpha_m)$ with respect to $(s_1, \cdots, s_m)$, where $(\alpha_1, \cdots, \alpha_m) \in \mathbb{R}^m$ and $s_1, \cdots, s_m \in \mathbb{N}$, if I is the least value of the sum $\sum_{u=1}^{m} \frac{i_u}{s_u}$ for which $R_{i_1,\cdots,i_m}(\alpha_1, \cdots, \alpha_m)$ does not vanish. Note by equation (0.6) I exists provided that R is not identically zero. If $R \equiv 0$, we put $I = \infty$.

Proposition 0.36. *Let* ind *denote the index of R at $(\alpha_1, \cdots, \alpha_m)$ with respect to $(s_1, \cdots, s_m)$. Then*

(i) $\operatorname{ind} R_{i_1,\cdots,i_m} \ge \operatorname{ind} R - \sum_{u=1}^{m} \frac{i_u}{s_u}$,

(ii) $\operatorname{ind}(R^{(1)} + R^{(2)}) \ge \min\{\operatorname{ind} R^{(1)}, \operatorname{ind} R^{(2)}\}$, *and*

(iii) $\operatorname{ind}(R^{(1)} R^{(2)}) = \operatorname{ind} R^{(1)} + \operatorname{ind} R^{(2)}$.

Proof. (i) is immediate and for (ii) and (iii) put $s = s_1 \cdots s_m$ and $I = \operatorname{ind} R$. Then by (i) t^{sI} is the least power of t occurring in $R(x_1 + t^{s/s_1} y_1, \cdots, x_m + t^{s/s_m} y_m)$ considered as a polynomial in the variable t. □

Proposition 0.37. (Siegel's Lemma) *Let N and M be positive integers with $N > M$. Let $a_{j,k} \in \mathbb{Z}$ for $1 \le j \le M, 1 \le k \le N$ with $|a_{j,k}| \le A$, and $A \ge 1$. Consider the system of linear equations $L_j(x_1, \cdots, x_N) = \sum_{k=1}^{N} a_{j,k} x_k = 0$, for $j = 1, 2, \cdots, M$. There exists a solution in integers $x_1, \cdots, x_N$, not all zero, with*

$$\max_{1 \le i \le N} |x_i| \le \left\lfloor (NA)^{\frac{N}{N-M}} \right\rfloor.$$

Proof. Put $X = \left\lfloor (NA)^{\frac{M}{N-M}} \right\rfloor$. Then $NA < (X+1)^{\frac{N-M}{M}}$, hence

$$NAX \leq (NA)(X+1) < (X+1)^{\frac{N}{M}}.$$

Notice that for any $(z_1, \cdots, z_N) \in \mathbb{Z}^N$ with $0 \leq z_i \leq X$, $i = 1, 2, \cdots, N$ we have

$$-B_j X \leq L_j(z_1, \cdots, z_m) \leq C_j X$$

where $-B_j$ is the sum of the negative coefficients of L_j and C_j is the sum of positive coefficients of L_j. Note that $B_j + C_j \leq NA$ for $j = 1, 2, \cdots, M$. Thus $L_j((z_1, \cdots, z_m))$ takes on at most $NAX + 1$ different values.

Notice that there are $(X+1)^N$ different values of $(z_1, \cdots, z_m)$ but at most $(NAX+1)^M$ different values of $(L_1((z_1, \cdots, z_N)), \cdots, L_M((z_1, \cdots, z_N)))$. Since $(NAX+1)^M < (X+1)^N$ we see that there exist two distinct vectors $\mathbf{z}_1, \mathbf{z}_2$ for which

$$(L_1(\mathbf{z}_1), \cdots, L_M(\mathbf{z}_1)) = (L_1(\mathbf{z}_2), \cdots, L_M(\mathbf{z}_2)).$$

Hence if we put $\mathbf{x} = \mathbf{z}_1 - \mathbf{z}_2$, we obtain

$$(L_1(\mathbf{x}), \cdots, L_M(\mathbf{x})) = (0, \cdots, 0)$$

and the result follows since $\max_i |x_i| \leq X$. $\square$

Proposition 0.38. *For each integer $l \geq 0$, there are rational integers $a_j^{(l)}$ with $0 \leq j \leq n$ such that*

$$\alpha^l = a_{n-1}^{(l)} \alpha^{n-1} + \cdots + a_0^{(l)}$$

with $|a_j^{(l)}| \leq (a+1)^l$.

Proof. This is immediate from the fact that $\alpha^{n-1}, \cdots, \alpha, 1$ form a basis of $\mathbb{Q}(\alpha)$ as a vector space over $\mathbb{Q}$. $\square$

Proposition 0.39. *For any positive integers $r_1, \cdots, r_m$ and real number λ the number of m-tuples of non-negative integers $i_1, \cdots, i_m$ such that $\sum_{u=1}^{m} \frac{i_u}{r_u} \leq \frac{1}{2}(m - \lambda)$ with $0 \leq i_u \leq r_u$, $u = 1, 2, \cdots, m$ is at most $(2m)^{1/2} \lambda^{-1} (r_1 + 1) \cdots (r_m + 1)$.*

Proof. Proof is by induction on m. Note that for $m = 1$ the result is immediate since the number of solutions is at most $r + 1$ and is at most 0 if $\lambda > 1$. Assume the result for $m - 1$. Then for fixed $r = r_m$ and $i = i_m$ the number of $(m-1)$-tuples of integers satisfying

$$\sum_{u=1}^{m-1} \frac{i_u}{r_u} + \frac{i}{r} \leq \frac{1}{2}(m - \lambda)$$

is the same as the number of $(m-1)$-tuples satisfying $\displaystyle\sum_{u=1}^{m-1}\frac{i_u}{r_u} \leq \frac{1}{2}\left(m-\lambda-\frac{2i}{r}\right)$, which is bounded above by $(2(m-1))^{1/2}\dfrac{1}{\left(\lambda+\frac{2i}{r}-1\right)}(r_1+1)\cdots(r_{m-1}+1)$. But

$$\begin{aligned}\sum_{i=0}^{r}\frac{2}{\lambda-1+\frac{2i}{r}} &= \sum_{i=0}^{r}\left(\frac{1}{\lambda-1+\frac{2i}{r}}+\frac{1}{\lambda+1-\frac{2i}{r}}\right)\\ &= \sum_{i=0}^{r}\left(\frac{\lambda+1-\frac{2i}{r}+\lambda-1+\frac{2i}{r}}{\lambda^2-(1-2i/r)^2}\right)\\ &= \sum_{i=0}^{r}\frac{2\lambda}{\lambda^2-(1-2i/r)^2}\\ &\leq 2(r+1)\frac{\lambda}{\lambda^2-r}.\end{aligned}$$

Therefore the total number of m-tuples is at most

$$(2(m-1))^{1/2}\frac{\lambda}{\lambda^2-1}(r_1+1)\cdots(r_{m-1}+1)(r+1) \tag{0.7}$$

If $\lambda \leq (2m)^{1/2}$, then this bound is subsumed by the trivial bound $(r_1+1)\cdots(r_m+1)$. Thus assume $\lambda > (2m)^{1/2}$. We then obtain

$$\lambda^2-1 > \lambda^2\left(1-\frac{1}{2m}\right) > \lambda^2\left(1-\frac{1}{m}\right)^{1/2}$$

and the result follows from (0.7). $\square$

Theorem 0.40. *Let $0<\varepsilon<1$ and let α be an algebraic integer of degree n, with minimal polynomial $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_0$ and put $a=\max(1,|a_{n-1}|,\cdots,|a_0|)$. Let m be an integer with $m>8n^2\varepsilon^{-2}$, and let $r_1,\cdots,r_m$ be positive integers. There exists a polynomial $R(x_1,\cdots,x_m)$ with integer coefficients and degree at most r_u in x_u for $u=1,2,\cdots,m$ which*

(i) does not vanish identically,
(ii) has index at least $\frac{1}{2}m(1-\varepsilon)$ at $(\alpha,\cdots,\alpha)\in\mathbb{R}^m$,
(iii) $\overline{R}\leq 4(a+1)^{r_1+\cdots+r_m}$.

Proof. We write

$$R(x_1,\cdots,x_m)=\sum_{\substack{0\leq j_u\leq r_u\\ 1\leq u\leq m}} c(j_1,\cdots,j_m)x_1^{j_1}\cdots x_m^{j_m},$$

where $c(j_1,\cdots,j_m)$ are $(r_1+1)\cdots(r_m+1)$ integers to be determined. Put $N=(r_1+1)\cdots(r_m+1)$. We want

$$R_{i_1,\cdots,i_m}(\alpha,\cdots,\alpha)=0 \tag{0.8}$$

for all non-negative integers $i_1, \cdots, i_m$ for which $\displaystyle\sum_{u=1}^{m} \frac{i_u}{r_u} \le \frac{1}{2}(m-\varepsilon)$. Plainly (0.8) holds if $i_u > r_u$ for some u with $1 \le u \le m$. In (0.8) we express the powers of α as integer linear combinations of $1, \alpha, \cdots, \alpha^{n-1}$ using proposition 0.38. Then we find that solving (0.8) is the same as solving n-linear equations in the coefficients $c(j_1, \cdots, j_m)$. Since

$$R_{i_1,\cdots,i_m}(\alpha, \cdots, \alpha) = \sum_{\substack{i_u \le j_u \le r_u \\ 1 \le u \le m}} \binom{j_1}{i_1} \cdots \binom{j_m}{i_m} c(j_1, \cdots, j_m) \alpha^{j_1 - i_1} \cdots \alpha^{j_m - i_m},$$

it follows that $R_{i_1,\cdots,i_m}(\alpha, \cdots, \alpha) = 0$ is equivalent to the system of equations

$$\sum_{\substack{i_u \le j_u \le r_u \\ 1 \le u \le m}} \binom{j_1}{i_1} \cdots \binom{j_m}{i_m} a_k^{(j_1 - i_1 + \cdots + j_m - i_m)} c(j_1, \cdots, j_m) = 0$$

for $k = 0, \cdots, n-1$. Since $\displaystyle\binom{j_u}{i_u} \le 2^{j_u} \le 2^{r_u}$ for $u = 1, 2, \cdots, m$ and since $(j_1 - i_1) + \cdots + (j_m - i_m) \le r_1 + \cdots + r_m$, by proposition 0.38 the coefficients are at most $(2(a+1))^{r_1 + \cdots + r_m}$ in absolute value.

Now take $\lambda = m\varepsilon$ in proposition 0.39. The number of m-tuples of non-negative integers is at most $(2m)^{1/2}(m\varepsilon)^{-1}(r_1+1)\cdots(r_m+1)$, hence the number of linear equations with integer coefficients satisfied by the $c(j_1, \cdots, j_m)$'s is at most

$$M \le n(2m)^{1/2}(m\varepsilon)^{-1} N \le N/2,$$

since $m > 8n^2\varepsilon^{-2}$. Thus, by Siegel's Lemma, there exist integers $c(j_1, \cdots, j_m)$, not all zero, such that (0.8) holds for all non-negative integers $i_1, \cdots, i_m$ for which $\displaystyle\sum_{u=1}^{m} \frac{i_u}{r_u} \le \frac{1}{2}m(1-\varepsilon)$ and $A = (2(a+1))^{r_1 + \cdots + r_m}$ with

$$\begin{aligned} \max |c(j_1, \cdots, j_m)| &\le (NA)^{\frac{M}{N-M}} \\ &\le NA \\ &\le (r_1+1)\cdots(r_m+1)(2(a+1))^{r_1+\cdots+r_m} \\ &\le (4(a+1))^{r_1+\cdots+r_m}. \end{aligned}$$

□

Theorem 0.41. *Let $0 < \delta < 1/12$, $0 < \varepsilon < \delta/20$ be positive real numbers. Suppose that $p_u/q_u \in \mathbb{Q}, u = 1, 2, \cdots, m$ are such that $\left|\alpha - \frac{p_u}{q_u}\right| < \frac{1}{q_u^{2+\delta}}$ and $q_u^{\varepsilon} > 64(a+1)\max(1, |\alpha|)$ for $u = 1, 2, \cdots, m$. Let $r_1, \cdots, r_m \in \mathbb{N}$ be such that $r_1 \log q_1 \le r_u \log q_u \le (1+\varepsilon) r_1 \log q_1$ for $u = 1, 2, \cdots, m$. Then the index of the polynomial R constructed in Theorem 0.40 at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $(r_1, \cdots, r_m)$ is at least $\frac{\delta m}{8}$.*

Proof. Let $k_1, \cdots, k_m$ be non-negative integers for which

$$\sum_{u=1}^{m} \frac{k_u}{r_u} < \frac{\delta m}{8}.$$

Put $T(x_1, \cdots, x_m) = R_{k_1,\cdots,k_m}(x_1, \cdots, x_m)$. We must show that $T\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right) = 0$. By Theorem 0.40 and proposition 0.35 we see that T has integer coefficients and $\overline{T} \leq (8(a+1))^{r_1+\cdots+r_m}$. Since T has degree at most r_u in x_u for $u = 1, 2, \cdots, m$, T has at most $(r_1+1)\cdots(r_m+1)$ terms and hence at most $2^{r_1+\cdots+r_m}$ terms. Thus for any non-negative integers $i_1, \cdots, i_m$ we have, by proposition 0.35,

$$|T_{i_1,\cdots,i_m}(\alpha, \cdots, \alpha)| \leq (2 \cdot 2 \cdot (8(a+1)) \max(1, |\alpha|))^{r_1+\cdots+r_m}$$

so that

$$|T_{i_1,\cdots,i_m}(\alpha, \cdots, \alpha)| \leq (32(a+1) \max(1, |\alpha|))^{r_1+\cdots+r_m}. \tag{0.9}$$

By theorem 0.40 the index of R at $(\alpha, \cdots, \alpha)$ with respect to $(r_1, \cdots, r_m)$ is at least $\frac{1}{2}m(1-\varepsilon)$. By proposition 0.36 (ii), the index of T at $(\alpha, \cdots, \alpha)$ with respect to $(r_1, \cdots, r_m)$ is at least

$$\begin{aligned} \frac{1}{2}m(1-\varepsilon) - \sum_{u=1}^{m} \frac{k_u}{q_u} &\geq \frac{1}{2}m(1-\varepsilon) - \frac{\delta m}{8} \\ &= \frac{1}{2}m\left(1 - \varepsilon - \frac{\delta}{4}\right). \end{aligned}$$

Since $0 < \varepsilon < \delta/20$ the index is at least $\frac{1}{2}m\left(1 - \frac{\delta}{3}\right)$. Put $\beta_u = \frac{p_u}{q_u} - \alpha$ for $u = 1, 2, \cdots, m$. By Taylor's Theorem, we have

$$T\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right) = \sum_{\substack{0 \leq i_u \leq r_u \\ 1 \leq u \leq m}} T_{i_1,\cdots .i_m}(\alpha, \cdots, \alpha)\beta_1^{i_1} \cdots \beta_m^{i_m}. \tag{0.10}$$

But $T_{i_1,\cdots,i_m}(\alpha,\cdots,\alpha) = 0$ unless $\displaystyle\sum_{u=1}^{m} \frac{i_u}{r_u} > \frac{1}{2}m\left(1-\frac{\delta}{3}\right)$. For such $i_1,\cdots,i_m$ we have, since $|\beta_u| < \dfrac{1}{q_u^{2+\delta}}$ for $u = 1,2,\cdots,m$ and

$$\begin{aligned}
-\log|\beta_1^{i_1}\cdots\beta_m^{i_m}| &\geq (2+\delta)\sum_{u=1}^{m} i_u \log q_u \\
&= (2+\delta)\sum_{u=1}^{m}\frac{i_u}{r_u}(r_u\log q_u) \\
&\geq (2+\delta)r_1\log q_1 \sum_{u=1}^{m}\frac{i_u}{r_u} \\
&> (2+\delta)r_1\log q_1\left(\frac{1}{2}\left(m-\frac{\delta}{3}\right)\right) \\
&\geq \left(1+\frac{\delta}{2}\right)\left(1-\frac{\delta}{3}\right)\sum_{u=1}^{m} r_u\log q_u\left(\frac{1}{1+\varepsilon}\right).
\end{aligned}$$

Remark 0.42. The coefficient 1/2 in $\dfrac{1}{2}m\left(1-\dfrac{\delta}{3}\right)$ is the exponent 2 in Roth's Theorem.

Observe that $\left(1+\dfrac{\delta}{2}\right)\left(1-\dfrac{\delta}{3}\right) = \left(1+\dfrac{\delta}{6}-\dfrac{\delta^2}{6}\right)$ and that $0 < \delta < 1/12$, so $\left(1+\dfrac{\delta}{2}\right)\left(1-\dfrac{\delta}{3}\right) > 1+\dfrac{\delta}{8}$. Since $0 < \varepsilon < \delta/20$ we have $\left(1+\dfrac{\delta}{8}\right) > (1+\varepsilon)^2$. Thus

$$|\beta_1^{i_1}\cdots\beta_m^{i_m}| < (q^{r_1}\cdots q_m^{r_m})^{-1-\varepsilon}.$$

There are at most $(r_1+1)\cdots(r_m+1) \leq 2^{r_1+\cdots+r_m}$ terms in the sum (0.10). Thus by (0.9), we have

$$\begin{aligned}
\left|q_1^{r_1}\cdots q_m^{r_m}T\left(\frac{p_1}{q_1},\cdots,\frac{p_m}{q_m}\right)\right| &< 2^{r_1+\cdots+r_m}(32(a+1)\max(1,|\alpha|))^{r_1+\cdots+r_m}\left(q_1^{r_1}\cdots q_m^{r_m}\right)^{-\varepsilon} \\
&< 2^{r_1+\cdots+r_m}\left(2^{-(r_1+\cdots+r_m)}\right) < 1
\end{aligned}$$

by the choice of the q_u's.

Since $\left|q_1^{r_1}\cdots q_m^{r_m}T\left(\dfrac{p_1}{q_1},\cdots,\dfrac{p_m}{q_m}\right)\right|$ is an integer less than 1, it must be zero, so we are done. □

We must now extract a contradiction. To this end we introduce Wronskians. Let Δ denote an operator of the form $\dfrac{\partial^{i_1}}{\partial x^{i_1}}\cdots\dfrac{\partial^{i_m}}{\partial x^{i_m}}$. We say that $i_1+\cdots+i_m$ is the order of Δ.

If $\Delta_1,\cdots,\Delta_h$ have orders at most $0,1,\cdots,h-1$ respectively and $\varphi_1,\cdots,\varphi_h$ are functions of $x_1,\cdots,x_m$ we call $\det(\Delta_i\varphi_j)_{1\leq i,j\leq h}$ a (generalized) Wronskian.

If $m = 1$ then there is only one Δ of order i, given by $\dfrac{d^{i-1}}{dx_1^{i-1}}$. As a consequence the only Wronskian that don't vanish identically are of the form $\det\left(\frac{d^{i-1}}{dx_1^{i-1}}\varphi_j\right)$

Proposition 0.43. *Let $\varphi_1, \cdots, \varphi_h$ be rational functions (quotients of polynomials) of variables $x_1, \cdots, x_m$ with coefficients in $\mathbb{Q}$. Suppose that the only rational numbers $c_1, \cdots, c_h$ with $c_1\varphi_1 + \cdots + c_h\varphi_h = 0$ are $c_1 = \cdots = c_h = 0$. Then some Wronskian $\det(\Delta_i\varphi_j)$ does not vanish.*

Remark 0.44. If there is a non-trivial linear combination among the q_i's then all of the Wronskians vanish.

Proof. We shall prove the result by induction on h. When $h = 1$ the only Wronskian is φ_1 itself, and by assumption φ_1 is not identically 0.

Suppose that the result holds for $h - 1$. Note that φ_1 is not identically 0. We put $\varphi_j^* = \varphi_1^{-1}\varphi_j$ for $j = 1, 2, \cdots, h$. By the rule for differentiating products we can express a Wronskian of $\varphi_1^*, \cdots, \varphi_h^*$ as a sum of Wronskians of $\varphi_1, \cdots, \varphi_h$ each multiplied by rational functions of φ_1. It now suffices to look for a non-vanishing Wronskian of $\varphi_1^*, \cdots, \varphi_h^*$. Notice that any non-trivial linear relation over $\mathbb{Q}$ between $\varphi_1^*, \cdots, \varphi_h^*$ gives us such a relation for $\varphi_1, \cdots, \varphi_h$. Thus, without loss of generality we may suppose that $\varphi_1 \equiv 1$.

If φ_h is a constant, say c, then $c\varphi_1 - \varphi_h = 0$, contradicting $\varphi_1, \cdots, \varphi_h$ being linearly independent over $\mathbb{Q}$. Therefore there is some variable, say x_1, for which $\dfrac{\partial\varphi_h}{\partial x_1} \neq 0$.

Suppose there is a non-trivial rational linear combination of $\varphi_2, \cdots, \varphi_h$ which is independent of x_1, say $c_2\varphi_2 + \cdots + c_h\varphi_h$. Then one of $c_2, \cdots, c_{h-1}$ is non-zero and there is no loss of generality in assuming $c_2 \neq 0$, and indeed we may take $c_2 = 1$.

Thus $\dfrac{\partial}{\partial x_1}(c_2\varphi_2 + \cdots + c_h\varphi_h) = 0$. Observe that if we replace φ_2 by $\varphi_2 + c_3\varphi_3 + \cdots + c_h\varphi_h$ we don't change the Wronskians. By doing so we may suppose that $\dfrac{\partial\varphi_2}{\partial x_1} = 0$. We can repeat this argument and in this way we find an integer k with $1 \leq k < h$ for which

$$\frac{\partial\varphi_1}{\partial x_1} = \frac{\partial\varphi_2}{\partial x_1} = \cdots = \frac{\partial\varphi_k}{\partial x_1}$$

and for which there is no non-trivial linear combination of $\varphi_{k+1}, \cdots, \varphi_h$ over $\mathbb{Q}$ which is independent of x_1, or equivalently there is no non-trivial rational linear combination of $\dfrac{\partial\varphi_{k+1}}{\partial x_1}, \cdots, \dfrac{\partial\varphi_h}{\partial x_1}$.

By the inductive hypothesis there exist operators $\tilde{\Delta}_1, \cdots, \tilde{\Delta}_k$ of orders at most $0, 1, \cdots, k - 1$ respectively such that

$$W_1 = \det(\tilde{\Delta}_i\varphi_j)_{1 \leq i,j \leq k} \neq 0.$$

Further, since there are no non-trivial linear relations over $\mathbb{Q}$ between $\frac{\partial \varphi_{k+1}}{\partial x_1}, \cdots, \frac{\partial \varphi_h}{\partial x_1}$ there are operators $\tilde{\Delta}_{k+1}, \cdots, \tilde{\Delta}_h$ of orders at most $0, 1, \cdots, h-(k+1)-1$ respectively for which

$$W_2 = \det\left(\tilde{\Delta}_i \frac{\partial \varphi_j}{\partial x_1}\right)_{k+1 \leq i,j \leq h} \neq 0$$

Put $\Delta_i = \tilde{\Delta}_i$ for $i = 1, 2, \cdots, k$ and $\Delta_i = \tilde{\Delta}_i \frac{\partial}{\partial x_1}$ for $i = k+1, \cdots, h$. Notice that Δ_i is an operator of order at most $i-1$ for $i = 1, 2, \cdots, h$. Then the Wronskian W given by

$$W = \det(\Delta_i \varphi_j)_{1 \leq i,j \leq h}$$

is non-zero since $\frac{\partial \varphi_1}{\partial x_1} = \cdots = \frac{\partial \varphi_k}{\partial x_1} = 0$ and so we have $W = W_1 W_2 \neq 0$. □

Theorem 0.45. *Put $w = w(m, \varepsilon) = 24 \cdot 2^{-m} \left(\frac{\varepsilon}{12}\right)^{2^{m-1}}$, for $m \in \mathbb{N}$ and $0 < \varepsilon < 1/12$. Let $r_1, \cdots, r_m$ be positive integers for which $wr_u \geq r_{u+1}$ for $u = 1, 2, \cdots, m-1$, and let $q_u > 0$ and p_u be co-prime integers such that $q_u^{r_u} \geq q_1^{r_1}$ for $u = 1, 2, \cdots, m$ and $q_u^w \geq 2^{3m}$ for $u = 1, \cdots, m$.*

Suppose that $S(x_1, \cdots, x_m)$ is a polynomial of degree at most r_u in x_u, for $u = 1, \cdots, m$, with integer coefficients and $\overline{S} \leq q_1^{wr_1}$. If S does not vanish identically, then S has index at most ε at the point $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $(r_1, \cdots, r_m)$.

Remark 0.46. Some condition on the r_i's is necessary since for example $S(x_1, x_2) = (x_1 - x_2)^r$ has index 1 at any point $(p/q, p/q)$ with respect to (r, r).

Proof. The proof proceeds by induction on m. We first prove the result when $m = 1$. Suppose that $S\left(\frac{p_1}{q_1}\right) = S'\left(\frac{p_1}{q_1}\right) = \cdots = S^{(t-1)}\left(\frac{p_1}{q_1}\right)$ and $S^{(t)}\left(\frac{p_1}{q_1}\right) \neq 0$. Here we suppose that p_1, q_1 are coprime integers with $q_1 > 0$. Then $S(x) = \left(x - \frac{p_1}{q_1}\right)^t T(x)$ for some $T \in \mathbb{Q}[x]$.

We have $S(x) = (q_1 x - p_1)^t (q_1^{-t} T(x))$, since S has integer coefficients, by Gauss's Lemma, we have $\frac{1}{q_1^t} T(x) \in \mathbb{Z}[x]$. Therefore $q_1^t \leq \overline{S} \leq q_1^{wr_1}$ and hence $t \leq wr_1$. For $m = 1$, $w = w(m, \varepsilon) = 24 \cdot 2^{-1}(\varepsilon/12) = \varepsilon$. Equivalently, $t_1/r_1 \leq \varepsilon$ as required.

We shall now suppose that the result holds for $1 \leq t < m$. We can write S in the form

$$S(x_1, \cdots, x_m) = \sum_{1 \leq j \leq h} \varphi_j(x_1, \cdots, x_{m-1}) \psi(x_m),$$

where φ_j, ψ_j are polynomials with rational coefficients.

In particular, we can take $h = r_m + 1$ and $\psi_j(x_m) = x_m^{j-1}$. We take such a decomposition with h minimal. Then certainly $h \leq r_m + 1$. Suppose there exists a linear relation $c_1\varphi_1 + \cdots + c_h\varphi_h = 0$ with $c_1, \cdots, c_h$ rational and not all zero. Then without loss of generality we may suppose $c_h \neq 0$. Then $\varphi_h = -\frac{c_1}{c_h}\varphi_1 - \cdots - \frac{c_{h-1}}{c_h}\varphi_{h-1}$ and so

$$S = \sum_{j=1}^{h-1} \varphi_j \left(\psi_j - \frac{c_j}{c_h}\psi_h\right)$$

which contradicts the minimality of h. Thus there exists no non-trivial linear relation among $\varphi_1, \cdots, \varphi_h$ over the rationals. Similarly, suppose that there exist rationals $e_1, \cdots, e_h$ not all zero such that $e_1\psi_1 + \cdots + e_h\psi_h = 0$. Without loss of generality we suppose that $e_h \neq 0$. Then

$$S = \sum_{j=1}^{h-1} \psi_j \left(\varphi_j - \frac{e_j}{e_h}\varphi_h\right)$$

which again contradicts the minimality of h. Again, there is no non-trivial rational linear combination among $\psi_1, \cdots, \psi_h$ over $\mathbb{Q}$.

We choose h minimal and conclude there is no non-trivial relation over $\mathbb{Q}$ of $\varphi_1, \cdots, \varphi_h$ and the same holds for $\psi_1, \cdots, \psi_h$. Therefore proposition 0.43,

$$U(x_m) = \det\left(\frac{1}{(i-1)!}\frac{\partial^{i-1}}{\partial x_m^{i-1}}\varphi_j\right)_{1\leq i,j\leq h} \neq 0.$$

Further, by proposition 0.43, there exist operators Δ_i' for $i = 1, \cdots, h$ of the form

$$\Delta_i' = \frac{1}{i_1!}\cdots\frac{1}{i_m!}\frac{\partial^{i_1+\cdots+i_m}}{\partial x_1^{i_1}\cdots\partial^{i_m}}$$

with $i_1 + \cdots + i_m \leq i - 1 \leq h - 1 \leq r_m$ such that

$$V(x_1, \cdots, x_m) = \det(\Delta_i'\varphi_j)_{1\leq i,j\leq h} \neq 0.$$

Next we define $W(x_1, \cdots, x_m)$ by

$$W(x_1, \cdots, x_m) = \det\left(\Delta_i'\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial x_m^{j-1}}S(x_1, \cdots, x_m)\right)_{1\leq i,j\leq h}$$

Thus

$$\begin{aligned} W &= \det\left(\Delta_i'\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial x_m^{j-1}}\left(\sum_{k=1}^{h}\varphi_k\psi_k\right)\right)_{1\leq i,j\leq h} \\ &= \det\left((\Delta_i'\varphi_k)_{1\leq i,j\leq h}\left(\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial x_m^{j-1}}\psi_k\right)_{1\leq j,k\leq h}\right) \\ &= U(x_m)V(x_1, \cdots, x_m) \neq 0 \end{aligned}$$

But

$$\Delta_i'\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial x_m^{j-1}}S(x_1, \cdots x_m) = S_{i_1,\cdots,i_{m-1},j-1}(x_1, \cdots, x_m)$$

and since S has integer coefficients so does $S_{i_1,\cdots,i_{m-1},j-1}$ and therefore W has integer coefficients. By Gauss's Lemma we may write $W = v(x_1, \cdots, x_m)u(x_m)$ where $v(x_1, \cdots, x_m)$ and $u(x_m)$ have integer coefficients. Since $S_{i_1,\cdots,i_{m-1},j-1}$ has degree at most r_u in x_u for $u = 1, 2, \cdots, m$ and since W is given by the determinant of an $h \times h$ matrix, W has degree at most hr_u in x_u for $u = 1, 2, \cdots, m$. In particular, v has degree at most hr_u in x_u for $u = 1, 2, \cdots, m-1$ and $u(x_m)$ has degree at most hr_m in x_m.

Now, by proposition 0.35, we have $\overline{S_{i_1,\cdots,i_{m-1},j-1}} \leq 2^{r_1+\cdots+r_m} q_1^{wr_1}$. There are at most $(r_1+1)\cdots(r_m+1)$ monomials in $S_{i_1,\cdots,i_{m-1},j-1}$ and $(r_1+1)\cdots(r_m+1) \leq 2^{r_1+\cdots+r_m}$. There are at most $h! \leq h^{h-1}$ products in the determinant expansion of W. Since $h \leq r_m + 1$, this is at most $h^{r_m} \leq 2^{hr_m}$. Thus

$$\begin{aligned}
\overline{W} &\leq h!((r_1+1)\cdots(r_m+1))^h \left(2^{r_1+\cdots+r_m} q_1^{wr_1}\right)^h \\
&\leq 2^{hr_m} 2^{h(r_1+\cdots+r_m)} 2^{h(r_1+\cdots+r_m)} q_1^{wr_1 h} \\
&\leq 2^{3h(r_1+\cdots+r_m)} q_1^{wr_1 h} \\
&\leq 2^{3mr_1 h} q_1^{wr_1 h} \\
&\leq (q_1^{2w})^{r_1 h} = q_1^{2wr_1 h},
\end{aligned}$$

by hypothesis. Since $W = uv$ and $v(x_1, \cdots, x_{m-1})$ and $u(x_m)$have integer coefficients and each coefficient of W is obtained as a product of a coefficient of v and a coefficient of u, we see that $\overline{u}, \overline{v} \leq q_1^{2wr_1 h}$. By the definition of w we have $w(m, \varepsilon) = \frac{1}{2} w\left(m-1, \frac{\varepsilon^2}{12}\right)$. We now apply the inductive hypothesis to u and v. First apply it to v with $hr_1, \cdots, hr_{m-1}$ in place of $r_1, \cdots, r_m$ and $\varepsilon^2/12$ for ε and $2w$ for w. Then the hypotheses are satisfied and v has index at most $\varepsilon^2/12$ at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_{m-1}}{q_{m-1}}\right)$ with respect to $hr_1, \cdots, hr_{m-1}$. Thus the index of v as a function of $x_1, \cdots, x_m$ at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $r_1, \cdots, r_m$ is at most $h\varepsilon^2/12$.

Secondly we apply our inductive hypothesis to u with hr_m in place of $r_1, \cdots, r_m$ and $\varepsilon^2/12$ in place of ε and $2w$ in place of w. Since $w = w(m, \varepsilon) \leq \frac{1}{2} w\left(1, \frac{\varepsilon^2}{12}\right)$ and since $q_1^{r_1} \leq q_m^{r_m}$, we have $\overline{u} \leq q_m^{2wr_m h}$.

Thus the index of u at p_m/q_m with respect to hr_m is at most $\varepsilon^2/12$. Thus the index of u as a function of $x_1, \cdots, x_m$ is at most $h\varepsilon^2/12$. Thus, by proposition 0.36, the index I_W of W at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $r_1, \cdots, r_m$ is at most $\frac{h\varepsilon^2}{12} + \frac{h\varepsilon^2}{12} = \frac{\varepsilon^2}{6}$.

We should now estimate I_W in terms of θ where θ is the index of S at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $r_1, \cdots, r_m$. By proposition 0.36 the index of $S_{i_1,\cdots,i_{m-1},j-1}$ is at least

$$\theta - \frac{i_1}{r_1} - \cdots - \frac{i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m} \geq \theta - \frac{i_1 + \cdots + i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m},$$

since $i_1 + \cdots + i_{m-1} \leq i - i \leq h - 1 \leq r_m$, we have

$$\theta - \frac{i_1 + \cdots + i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m} \geq \theta - \frac{r_m}{r_{m-1}} - \frac{j-1}{r_m}.$$

By hypothesis, $r_m/r_{m-1} \leq w$, and thus we obtain

$$\theta - \frac{r_m}{r_{m-1}} - \frac{j-1}{r_m} \geq \theta - w - \frac{j-1}{r_m}.$$

But $m \geq 2$, whence $w \leq 24 \cdot 2^{-2} \left(\frac{\varepsilon}{12}\right)^2 = \frac{\varepsilon^2}{24}$, and so the index of $S_{i_1,\cdots,i_{m-1},j-1}$ at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $r_1, \cdots, r_m$ is at least $\theta - \frac{\varepsilon^2}{24} - \frac{j-1}{r_m}$.

Developing W as a determinant expansion and using the fact that the index is non-negative and proposition 0.36, we find that

$$\begin{aligned} I_W &\geq \sum_{j=1}^{n} \max\left(\theta - \frac{\varepsilon^2}{24} - \frac{j-1}{r_m}, 0\right) \\ &\geq -\frac{h\varepsilon^2}{24} + \sum_{j=1}^{h} \max\left(\theta - \frac{j-1}{r_m}, 0\right). \end{aligned}$$

But $I_W \leq \frac{h\varepsilon^2}{6}$, and therefore

$$\frac{5\varepsilon^2 h}{24} \geq \sum_{j=1}^{h} \max\left(\theta - \frac{j-1}{r_m}, 0\right) \Rightarrow \frac{\varepsilon^2}{4} > \frac{1}{h}\sum_{j=1}^{h} \max\left(\theta - \frac{j-1}{r_m}, 0\right).$$

We have $1 \leq h \leq r_m + 1$, so if $\theta \geq \frac{h-1}{r_m}$ then

$$\begin{aligned} \sum_{j=1}^{h} \max\left(\theta - \frac{j-1}{r_m}, 0\right) &= \frac{1}{h}\sum_{j=1}^{h} \left(\theta - \frac{j-1}{r_m}\right) \\ &= \theta - \frac{h-1}{2r_m} \\ &= \frac{\theta}{2} + \frac{1}{2}\left(\theta - \frac{h-1}{r_m}\right) \\ &\geq \frac{\theta}{2}. \end{aligned}$$

Hence, $\theta/2 < \varepsilon^2/4$ and so $\theta < \varepsilon$. Otherwise, we have $\theta < \frac{h-1}{r_m}$. Then we have

$$\begin{aligned} \frac{1}{h}\sum_{j=1}^{h} \max\left(\theta - \frac{j-1}{r_m}, 0\right) &= \frac{1}{h}\sum_{1 \leq j \leq \theta r_m + 1} \left(\theta - \frac{j-1}{r_m}\right) \\ &\geq \frac{1}{h}\left(\lfloor \theta r_m \rfloor + 1\right)\left(\theta - \frac{\lfloor \theta r_m \rfloor}{2r_m}\right) \\ &\geq \frac{1}{h}\left(\lfloor \theta r_m \rfloor + 1\right)\left(\frac{\theta}{2}\right) \\ &\geq \frac{\theta^2 r_m}{2h}. \end{aligned}$$

Since $h \leq r_m + 1 \leq 2r_m$ we see that

$$\frac{1}{h}\sum_{j=1}^{h}\left(\theta - \frac{j-1}{r_m}, 0\right) \geq \frac{\theta^2}{4}.$$

Hence $\theta^2 < \varepsilon^2$, so $\theta < \varepsilon$ as required. □

Proof. (Roth's Theorem) Suppose that $0 < \delta < 1/12$ and that there are infinitely many solutions in rationals p/q, with $q > 0$, to the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}. \tag{0.11}$$

Choose $\varepsilon > 0$ to be a real number with $0 < \varepsilon < \delta/20$. Next let $m > 8n^2\varepsilon^{-2}$ (here n is the degree of α over $\mathbb{Q}$). Then put $w = w(m, \varepsilon) = 24 \cdot 2^{-m}\left(\frac{\varepsilon}{12}\right)^{2^{m-1}}$. Let p_1/q_1 be a solution to (0.11) with q_1 so large that
(i) $q_1^{\varepsilon} > 64(a+1)\max(1, |\alpha|)$,
(ii) $q_1^w \geq 2^{3m}$, and
(iii) $q_1^w \geq (4(a+1))^m$.

Now choose p_u/q_u for $u = 2, \cdots, m$ to be solutions of (0.11) in co-prime integers p_u, q_u with $q_u > 0$ successively such that
(iv) $\frac{1}{2} w \log q_{u+1} \geq \log q_u$.
Since $q_{u+1} > q_u$ for $u = 1, 2, \cdots, m-1$, we have
(v) $q_u^{\varepsilon} > 64(a+1)\max(1, |\alpha|)$ and also
(vi) $q_u^w \geq 2^{3m}$.
(v), (vi) hold for $u = 2, 3, \cdots, m$.

Next choose r_1 to be an integer so large that $\varepsilon r_1 \log q_1 \geq \log q_m$. Put $r_u = \left\lfloor \frac{r_1 \log q_1}{\log q_u} \right\rfloor + 1$, for $u = 2, 3, \cdots, m$. Then

$$\begin{aligned} r_1 \log q_1 &\leq r_u \log q_u \\ &\leq r_1 \log q_1 + \log q_u \\ &\leq (1+\varepsilon) r_1 \log q_1 \end{aligned}$$

for $u = 1, 2, \cdots, m$.

Then the conditions of theorems 0.40 and 0.41 are satisfied. Further,

$$\frac{r_{u+1}}{r_u} \leq \frac{2\log q_u}{\log q_{u+1}} \leq w.$$

Since

$$\frac{r_1 \log q_1}{\log q_{u+1}} \geq \frac{r_1 \log q_1}{\log q_m} \geq \frac{1}{\varepsilon} \geq 240$$

and

$$r_u \geq \frac{r_1 \log q_1}{\log q_u} \Rightarrow q_u^{r_u} \geq q_1^{r_1}$$

for $u = 1, 2, \cdots, m$, the conditions of Theorem 0.45 are also satisfied.

Next observe that the polynomial R constructed in theorem 0.40 has integer coefficients of size, in absolute value, at most $(4(a+1))^{r_1+\cdots+r_m} \leq (4(a+1))^{mr_1}$ and by (iii) this is at most $q_1^{wr_1}$, hence theorem 0.45 applies with $S = R$. Let I_R be the index of R at $\left(\frac{p_1}{q_1}, \cdots, \frac{p_m}{q_m}\right)$ with respect to $r_1, \cdots, r_m$. By theorem 0.41, I_R is at least $\frac{\delta m}{8}$. By Theorem 0.40 R is not the zero polynomial. Hence by theorem 0.45, I_R is at most ε. Therefore $\frac{\delta m}{8} < \varepsilon$, but $0 < \varepsilon < \delta/20$, and so we have a contradiction. This proves Roth's Theorem. $\square$

Remark 0.47. Roth's Theorem is not effective and it is a very important problem to make the proof effective.

Remark 0.48. Roth's Theorem can be used to prove that numbers of the form $\sum_{n=1}^{\infty} 2^{-3^n}$ are transcendental.

Remark 0.49. In 1959 Cugiani proved that if $\frac{p_1}{q_1}, \frac{p_2}{q_2}, \cdots$ are solutions to $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+20(\log\log\log q)^{-1/2}}}$ with $0 < q_1 < q_2 < \cdots$, then $\limsup_{k\to\infty} \frac{q_{k+1}}{q_k} = \infty$.

For the proof of Roth's Theorem we supposed the existence of several good approximations to α. For the Thue-Siegel approach one can get by with one very good approximation. This is important for effective results. Bombieri used such an approach to improve on the Liouville estimate in some cases. For example, let $r \geq 40$. He proved that there is a positive number $m_0(r)$ which is effectively computable such that if α is a root of $x^r - mx^{r-1} + 1$ and $m > m_0(r)$ then there is an effectively computable positive number $q_0(\alpha)$ such that if $q > q_0(\alpha)$ then $\left|\alpha - \frac{p}{q}\right| > \frac{1}{q^{39.2574}}$.

The first effective and explicit refinement of Liouville's estimate is due to Baker in 1964, although such results are implicit in the works of Thue. For example Baker in 1964 proved that $\left|2^{1/3} - \frac{p}{q}\right| > \frac{10^{-6}}{q^{2.955}}$ for all p, q with $q > 0$. Chudnovsky and Easton refined this. In 1997 Bennett proved $\left|2^{1/3} - \frac{p}{q}\right| > \frac{1}{4}\frac{1}{q^{2.5}}$ for all p, q with $q > 0$.

The first non-trivial effective improvement of the Liouville result which applies to all algebraic numbers α of degree at least 3 is due to Baker and it depends on estimates for linear forms in the logarithm of algebraic numbers. This work in turn builds on earlier work of Gelfond and Schneider who resolved Hilbert's 7th problem. The improvement was small but it sufficed to effectively solve Thue equations. For instance, in 1986 Baker and Stewart proved

Theorem 0.50. *Let a be a positive integer which is not a perfect cube. Let ε be the fundamental unit in the ring of algebraic integers of the field $\mathbb{Q}(a^{1/3})$ (that is, the* *smallest unit larger than 1). Then, for all rationals p/q with $q > 0$, we have*

$$\left|a^{1/3} - \frac{p}{q}\right| > \frac{c}{q^\kappa}$$

where $c = \dfrac{1}{32c_1}$ *and* $\kappa = 3 - \dfrac{1}{c_2}$ *with*

$$c_1 = \varepsilon^{(50 \log\log \varepsilon)^2}, c_2 = 10^{12} \log \varepsilon.$$

This translates into

Theorem 0.51. *Let a and n be positive integers with a not a perfect cube. All solutions in integers x, y of $x^3 - ay^3 = n$ satisfy* $\max(|x|, |y|) < (c_1 n)^{c_2}$ *with c_1, c_2 as in the previous theorem.*

There have been extensions of Roth's Theorem. The first one was to estimate ho well α can be approximated by an element β from a fixed finite extension of $\mathbb{Q}$, say K. We need a measure of the size of β and for this we introduce a height function. Let $f \in \mathbb{Z}[x]$ be of the form $f(x) = a_n x^n + \cdots + a_1 x + a_0$. We put $H(f) = \max_i |a_i|$ and we put $H(\beta) = H(g)$, where g is the minimal polynomial of β over $\mathbb{Z}$. If $\beta = p/q$ is rational, then $H(\beta)$ is simply $\max(|p|, |q|)$.

In 1955, Levesque proved

Theorem 0.52. *Let α be algebraic, let K be a finite extension of $\mathbb{Q}$, and let $\delta > 0$. There are only finitely many elements β of K for which $|\alpha - \beta| < H(\beta)^{-2-\delta}$.*

Notice that we do not insist that α is real.

What happens if instead of fixing the extension field K in which β lies we only require that β is of degree at most d? Siegel, Ramachandra, and Wirsing made progress on this problem.

Theorem 0.53. (Schmidt) *Let $d \in \mathbb{N}$ and let α be a real algebraic number of degree greater than d. Set $\delta > 0$. Then there are only finitely many algebraic numbers β of degree at most d for which $|\alpha - \beta| < H(\beta)^{-d-1-\delta}$.*

Theorem 0.54. (Wirsing) *Let d be a positive integer and suppose that α is a real algebraic number of degree greater than d. Then for every $\delta > 0$ there are infinitely many real β of degree at most d for which $|\alpha - \beta| < H(\beta)^{-d-1+\delta}$.*

Theorem 0.55. (Mahler) *Let α be a real non-zero algebraic number and let $p_1, \cdots, p_r$ be distinct primes. Suppose $\delta > 0$. There are only finitely many rationals p/q with $p = p_1^{a_1} \cdots p_r^{a_r} p'$ and $q = p_1^{b_1} \cdots p_r^{b_r} q'$ where $a_1, \cdots, a_r$ and $b_1, \cdots, b_r$ are non-negative integers and p', q' are co-prime with $p_1, \cdots, p_r$ for which* $\left|\alpha - \dfrac{p}{q}\right| < \dfrac{1}{|p'q'||pq|^\delta}$.

Mahler used such a result to prove thati f $p_1, \cdots, p_r$ are distinct primes and $F(x, y)$ is a binary form with integer coefficients, non-zero discriminant and degree at least 3 then the equation $F(x, y) = p_1^{z_1} \cdots p_r^{z_r}$ has only finitely many solutions in coprime integers x and y and non-negative integers $z_1, \cdots, z_r$. This is known as the Thue-Mahler equation.

Using ideas from the geometry of numbers and building on the work of Roth, Schmidt proved

Theorem 0.56. *For any algebraic numbers $\alpha_1, \cdots, \alpha_n$ with $1, \alpha_1, \cdots, \alpha_n$ linearly independent over $\mathbb{Q}$ and for any $\varepsilon > 0$ there are only finitely many positive integers q for which*

$$q^{1+\varepsilon}\|q\alpha_1\| \cdots \|q\alpha_n\| < 1,$$

where $\|\cdot\|$ is the distance to the nearest integer.

Remark 0.57. It follows from the above theorem that if $1, \alpha_1, \cdots, \alpha_n$ are $\mathbb{Q}$-linearly independent then for each $\varepsilon > 0$ there are only finitely many integers $p_1, \cdots, p_n$ and q with $q > 0$ for which

$$\left|\alpha_i - \frac{p_i}{q}\right| < \frac{1}{q^{1+1/n+\varepsilon}}.$$

The exponent can be shown to be best possible.

The above theorem can be applied to the study of norm form equations - a generalization of the Thue equation. There are also p-adic versions of this work. One consequence is due to Evertse in 1984. Let $p_1, \cdots, p_r$ be distinct prime numbers and let n be a positive integer. There are only finitely many n-tuples of integers $(x_1, \cdots, x_n)$ with the x_i's composed only of primes from $\{p_1, \cdots, p_r\}$ with $x_1 + \cdots + x_n = 0$, and $\gcd(x_1, \cdots, x_n) = 1$ and such that $x_{i_1} + \cdots + x_{i_l} \neq 0$ whenever $\{i_1, \cdots, i_l\}$ is a proper subset of $\{1, \cdots, n\}$. For example, $2^a - 3^b + 5^c + 7^d = 0$ has only finitely many solutions.

Suppose we are given a sequence $(x_n)_{n=1}^{\infty}$ of real numbers in $[0, 1)$. We can ask how well distributed the sequence is in the interval. The first question to ask is whether the sequence is dense. Let α be a real number and consider the sequence $(\{n\alpha\})_{n=1}^{\infty}$ where $\{n\alpha\} = n\alpha - \lfloor n\alpha \rfloor$. If α is rational, then $(\{n\alpha\})_{n=1}^{\infty}$ is finite and hence not dense. Conversely, if α is irrational, then $(\{n\alpha\})_{n=1}^{\infty}$ is dense. To see this, note that all of the terms of the sequence are distinct, since

$$\{n_1\alpha\} = \{n_2\alpha\} \Rightarrow n_1\alpha - n_2\alpha = \lfloor n_1\alpha \rfloor - \lfloor n_2\alpha \rfloor \Rightarrow \alpha = \frac{\lfloor n_1\alpha \rfloor - \lfloor n_2\alpha \rfloor}{n_1 - n_2} \in \mathbb{Q}.$$

Next note that for each $\varepsilon > 0$ we can find distinct positive integers $n_1 > n_2$ such that $|\{n_1\alpha\} - \{n_2\alpha\}| < \varepsilon$. But then $\{(n_1 - n_2)\alpha\} = (n_1 - n_2)\alpha - \lfloor (n_1 - n_2)\alpha \rfloor$. Thus

$$\{(n_1 - n_2)\alpha\} = \{n_1\alpha\} + N_1 + \{n_2\alpha\} + N_2 - N_3,$$

where $N_1 = \lfloor n_1\alpha \rfloor, N_2 = \lfloor n_2\alpha \rfloor, N_3 = \lfloor (n_1 - n_2)\alpha \rfloor$. Thus $\{(n_1 - n_2)\alpha\}$ is either in $(0, \varepsilon)$ or $(1 - \varepsilon, 1)$.

In the former case, $\{m(n_1 - n_2)\alpha\} = m\{(n_1 - n_2)\alpha\}$ for $m = 1, 2, \cdots, k$ where k is the largest integer such that $k\varepsilon < 1$. For every real number $\beta \in [0, 1)$, there is $j, 1 \leq j \leq k$ such that $|\beta - m\{(n_1 - n_2)\alpha\}| < \varepsilon$.

Similarly, in the other case we have $\{m(n_1 - n_2)\alpha\} = 1 - m(1 - \{(n_1 - n_2)\alpha\})$ for $m = 1, 2, \cdots, k$ where $k = \left\lfloor \dfrac{1}{1 - \{(n_1 - n_2)\alpha\}} \right\rfloor$ and again every $\beta \in [0, 1)$ is within ε of one of the multiples. Hence if α is irrational then $(\{n\alpha\})_{n=1}^{\infty}$ is dense in $[0,1)$.

The result is the one-dimensional version of a result of Kronecker. Kronecker proved that if $\alpha_1, \cdots, \alpha_k$ are real numbers with $1, \alpha_1, \cdots, \alpha_k$ linearly independent over $\mathbb{Q}$, then $(\{n\alpha_1\}, \cdots, \{n\alpha_k\})_{n=1}^{\infty}$ is dense in $[0,1]^k$.

A more refined notion than that of being dense is the following:

Definition 0.58. A sequence $(x_n)_{n=1}^{\infty}$ of real numbers is said to be uniformly distributed modulo 1 (u.d. mod 1) if for every pair of real numbers a, b with $0 \leq a < b \leq 1$ we have

$$\lim_{N\to\infty} \frac{A(a,b,N)}{N} = b - a,$$

where $A(a,b,N) = \#\{x_n : n \leq N, a \leq \{x_n\} < b\}$.

Let $\chi_{[a,b)}$ be the characteristic function of $[a,b)$. Then $(x_n)_{n=1}^{\infty}$ is u.d. mod 1 if and only if $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \chi_{[a,b)}(\{x_n\}) = b - a$ for all intervals $[a,b)$ with $0 \leq a < b \leq 1$.

Theorem 0.59. *A sequence $(x_n)_{n=1}^{\infty} \subset \mathbb{R}$ is u.d. mod 1 if and only if for every real valued continuous function f on $[0,1]$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}) = \int_0^1 f(x)dx.$$

Proof. Suppose first that $(x_n)_{n=1}^{\infty}$ is u.d. mod 1 . Let g be a step function on $[0,1]$ so there exist real numbers $0 \leq a_0 < a_1 < \cdots < a_k = 1$ and $s_1, \cdots, s_k$ such that $g = \sum_{i=1}^{k} s_i \chi_[a_i, a_{i+1})$. Then we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g(\{x_n\}) = \sum_{i=1}^{k} s_i(a_i - a_{i-1}) = \int_0^1 g(x)dx.$$

The step functions are uniformly dense in the real valued continuous functions, so there exist step functions f_1, f_2 with $f_1(x) \leq f(x) \leq f_2(x)$ and for which $f_2(x) - f_1(x) < \varepsilon$ for all $x \in [0,1]$. But then

$$\begin{aligned} \int_0^1 f(x)dx - \varepsilon &\leq \int_0^1 f_1(x)dx \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f_1(\{x_n\}) \\ &\leq \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\{x_n\}). \end{aligned}$$

Likewise

$$\begin{aligned}\int_0^1 f(x)dx + \varepsilon &\geq \int_0^1 f_2(x)dx \\ &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f_2(\{x_n\}) \\ &\geq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f(\{x_n\}).\end{aligned}$$

Since ε was arbitrary, it follows that $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f(\{x_n\})$ exists and

$$\int_0^1 f(x)dx = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f(\{x_n\}).$$

Given $\varepsilon > 0$ and $[a,b)$ with $0 \leq a < b \leq 1$ there exist continuous functions g_1, g_2 on $[0,1]$ for which $g_1(x) \leq \chi_{[a,b)}(x) \leq g_2(x)$ and for which $\int_0^1 (g_2(x) - g_1(x))dx < \varepsilon$. Then

$$\begin{aligned}(b-a) - \varepsilon &\leq \int_0^1 g_2(x)dx - \varepsilon \\ &\leq \int_0^1 g_1(x)dx \\ &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N g_1(\{x_n\}) \\ &\leq \liminf_{N\to\infty} \frac{A(a,b,N)}{N} \\ &\leq \limsup_{N\to\infty} \frac{A(a,b,N)}{N} \\ &\leq \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N g_2(\{x_n\}) \\ &= \int_0^1 g_2(x)dx \\ &\leq (b-a) + \varepsilon.\end{aligned}$$

Therefore, $\lim_{N\to\infty} \frac{A(a,b,N)}{N}$ exists and is $b-a$. □

Theorem 0.60. *A sequence $(x_n)_{n=1}^\infty \subset \mathbb{R}$ is u.d. mod 1 if and only if for every complex valued continuous function f on $\mathbb{R}$ with period 1, we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f(x_n) = \int_0^1 f(x)dx.$$

Proof. "⇒" Let f be periodic with period 1. Then we have that $f(x_n) = f(\{x_n\})$. Fur there there exist real valued continuous functions of period 1 f_1, f_2 such that $f = f_1 + if_2$. It then follows that

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(x_n) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_1(\{x_n\}) + \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} if_2(\{x_n\}) \\
&= \int_0^1 f_1(x)dx + i\int_0^1 f_2(x)dx \\
&= \int_0^1 f(x)dx.
\end{aligned}$$

We are done by the previous theorem.

"⇐" For every real valued continuous function f_1 on $\mathbb{R}$ we have $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_1(\{x_n\}) = \int_0^1 f_1(x)dx$ and since f_1 is periodic, we have $f_1(\{x_n\}) = f_1(x_n)$ and the result follows. □

We shall use the above theorem to establish a very useful criterion for a sequence to be u.d. mod 1 due to Herman Weyl in 1916.

Theorem 0.61. (Weyl's Criterion) *A sequence $(x_n)_{n=1}^{\infty}$ of real numbers is u.d. mod 1 if and only if for each non-zero integer h,*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{2\pi i h x_n} = 0.$$

Proof. "⇒" Let $\varepsilon > 0$. Suppose that f is a continuous function which is periodic with period 1 from $\mathbb{R}$ to $\mathbb{C}$. By the Weierstrass approximation theorem, there exists a trigonometric polynomial $g(x)$ such that $\sup_{0\le x\le 1} |f(x) - g(x)| < \varepsilon$. Write

$$g(x) = c_1 e^{2\pi i h_1 x} + \cdots + c_k e^{2\pi i h_k x}$$

with $c_1, \cdots, c_k \in \mathbb{C}$ and $h_1, \cdots, h_k$ are integers. But then

$$\begin{aligned}
\left|\int_0^1 f(x)dx - \frac{1}{N}\sum_{n=1}^{N} f(\{x_n\})\right| &\le \left|\int_0^1 (f(x) - g(x))dx\right| + \left|\int_0^1 g(x)dx - \frac{1}{N}\sum_{n=1}^{N} f(\{x_n\})\right| \\
&\le \int_0^1 |f(x) - g(x)|dx + \left|\int_0^1 g(x)dx - \frac{1}{N}\sum_{n=1}^{N} g(\{x_n\})\right| + \Bigg|\frac{1}{N}\sum_{n=1}^{N} g(\{ \\
&\le \varepsilon + \varepsilon + \varepsilon = 3\varepsilon
\end{aligned}$$

for N sufficiently large. Hence $\frac{1}{N}\sum_{n=1}^{N} f(x_n) = \int_0^1 f(x)dx$. □