An Overview of Diophantine Approximation

Neil Kulkarni

June 2026

Abstract

This paper surveys the theory of rational approximation to irrational numbers, beginning with Dirichlet's theorem, proved both by the pigeonhole principle and by Minkowski's convex body theorem, and extending through Hurwitz's sharp constant, the simultaneous approximation theorem, and the irrationality exponent $\mu(\alpha)$. These results culminate in Roth's theorem, which establishes that $\mu(\alpha) = 2$ for every algebraic irrational, and in Schmidt's Subspace Theorem, its higher-dimensional generalization. The paper concludes by examining a structural feature shared by both results: their proofs are ineffective, yielding sharp exponents while providing no computable bound on the size of solutions, a limitation that has resisted removal.

1 Dirichlet's Approximation Theorem

Theorem 1.1 (Dirichlet's Approximation Theorem)**.** *Let x be a positive real number. For every positive integer N, there is a rational number p/q such that $1 \leq q \leq N$ and $|x - p/q| < \frac{1}{Nq} \leq \frac{1}{q^2}$*

Proof. Let α be an irrational number and N be a positive integer. For every α, there is a constant $k = 0, 1, 2...N$ such that $k\alpha = y_k + x_k$ where y_k is an integer and x_k is a real number in the range [0,1).

As per the Pigeonhole Principle, if there are n items and m containers such that $n > m$, then in one container there will be 2 or more items. Now, since there are N intervals of length $\frac{1}{N}$ and $N + 1$ numbers x_k, then at least two such x_k lie in the same interval.

Now these two x_k be x_i and x_j such that $i < j$, then $|(j - i)\alpha - (y_j - y_i)|$ and rearranging these terms to match our original form to approximate irrational numbers of $k\alpha = y_k + x_k$, we get $|j\alpha - m_j - (i\alpha - m_i)|$. This is simply $|x_j - x_i| < \frac{1}{N} \leq \frac{1}{j-i}$ because both such x lie in the same $\frac{1}{N}$ interval.

Dividing through by $(j-i)$ reveals that

$$|\alpha - \frac{m_j - m_i}{j-i}| < \frac{1}{(j-i)N} \leq \frac{1}{(j-i)^2}$$

, which completes the original proof. □

Theorem 1.2 (Minkowski's Theorem)**.** *Suppose C to be a symmetric, convex, bounded subset $\mathbb{R}$. If $vol(C) > 2^d$, then C contains at least one lattice point other than 0.*

Proof. A subset C of $\mathbb{R}^d$ is bounded if it does not stretch or move towards infinity in any direction, and more formally, if there exists some real number R such that every point x in C satisfies $|x| \leq R$.

- Ex: $x^2 + y^2 \leq 1$ in $\mathbb{R}^2$; bounded
- The real line $\mathbb{R}$; not bounded

A subset C of $\mathbb{R}^d$ is convex if two points x and y in C contain the entire line segment between them also in C.

Finally, a subset C of $\mathbb{R}^d$ is symmetric if whenever x ∈ C, then -x ∈ C, unsurprisingly.

Let $C' = \frac{1}{2}C = \frac{1}{2}x : x \in C$, then $vol(C') > 1$. Since C has volume V in $\mathbb{R}^d$, then $\frac{1}{2}C$ has volume in $\frac{V}{2}^d$. Following this, the claim is that there exists a nonzero integral vector $v \in \mathbb{Z}^d$, such that $v \in C' - C'$. By the symmetric property we defined earlier, if $x \in C'$ and $x + v \in C'$, then $-x \in C$ exists. Hence, $\frac{1}{2}v = \frac{1}{2}(-x) + \frac{1}{2}(v + x) \in C'$ by convexity. Therefore, $C' = \frac{1}{2}C$

Before we continue on to see how Dirichlet's Approximation Theorem extends to Minkowski's Theorem, we need to show how disjoint translates (simply C' shifted by u such that $(C' + u) \cap (C' + v) = \emptyset$) can't all fit in C' and how the set subtraction $C' - C'$ must contain a nonzero integer point.

Proof: By contradiction, suppose that the subtraction set $C' - C' \cap \mathbb{Z} = \{0\}$, then $C' + u$ and $C' + v$ are disjoint for different $u, v \in \mathbb{Z}^d$. Suppose D is the diameter of C', then we have

$$\bigcup_{u \in [-N,N]^d} (C' + u) \subseteq [-(N + D), N + D]^d.$$

Considering the volume on both sides and the fact that the the disjoint translates are contained within the larger sets,

$$(2N+1)^d \cdot \mathrm{vol}(C') \leq 2^d(N+D)^d.$$

This further implies that

$$1 < \mathrm{vol}(C') \leq \left(\frac{2(N+D)}{2N+1}\right)^d.$$

Now, as $N \to \infty$, this inequality becomes false and the contradiction is apparent.

Minkowski's Theorem is essentially that If C is a symmetric, convex, bounded subset of $\mathbb{R}^n$ and $\mathrm{vol}(C) > 2^n$, then C contains at least one lattice point other than 0. Later on, we will see that this number 2 is a specific bound seen everywhere in Diophantine approximation. $\square$

Generalizing Minkowski's Theorem even further, we prove a general result:

Theorem 1.3 (Blichfeldt)**.** *Suppose $A \subset \mathbb{R}^n$ be measurable with vol$(A) > k$, then we can find some $x \in \mathbb{R}^n$ such that $A + x$ contains at least $k+1$ points from $\mathbb{Z}^n$.*

Definition 1.4 (Lebesgue Measure)**.** *The Lebesgue measure μ on $\mathbb{R}^n$ is the unique translation-invariant measure satisfying*

$$\mu([a_1, b_1] \times \cdots \times [a_n, b_n]) = \prod_{i=1}^{n}(b_i - a_i)$$

for all closed boxes in $\mathbb{R}^n$. For a set $A \subset \mathbb{R}^n$, we write vol$(A) = \mu(A)$ and say A is measurable if $\mu(A)$ is well-defined. Lebesgue measure generalizes the notion of length in $\mathbb{R}^1$, area in $\mathbb{R}^2$, and volume in $\mathbb{R}^3$. Recall that translation-invariance means $\mu(A+x) = \mu(A)$ for all $x \in \mathbb{R}^n$. Furthermore, as a measure, μ is countably additive, meaning that for any pairwise disjoint measurable sets $A_1, A_2, \ldots$ we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Proof. Denote $f(x) = \sum_{y \in \mathbb{Z}^n} \mathbf{1}_{A+x}(y)$.

This function essentially counts how many integer lattice points there are using the indicator function, and without generality A is bounded, since the Lebesgue measure is inner regular.

$$\int_{[0,1]^n} f(x)dx = \sum_{y\in\mathbb{Z}^n} \int_{[0,1]^n} \mathbf{1}_{A+x}(y)dx$$

$$= \sum_{y\in\mathbb{Z}^n} \int_{[0,1]^n} \mathbf{1}_{A}(y-x)dx$$

$$= \sum_{y\in\mathbb{Z}^n} \int_{y-[0,1]^n} \mathbf{1}_{A}(t)dt$$

$$= \int_{\mathbb{R}^n} \mathbf{1}_{A}(t)dt$$

$$= \mathrm{vol}(A)$$

$$> k$$

Thus, $\max_{x\in[0,1]^n} f(x) \geq k+1$.

To put it simply, if the average value of $f(x)$ exceeds k, then $f(x)$ can't be $\leq$ k for every single shift, meaning that at least one shift x where $f(x) > k$ exists since $f(x)$ always outputs an integer. $\square$

Theorem 1.5 (Dirichlet's Approximation on Minkowski Convex Body Theorem)**.** *For any $a \in (0,1)$ and $N \in \mathbb{N}$, the pair of natural numbers m, n such that $m \leq N$ exists where*

$$|a - \frac{n}{m}| < \frac{1}{mN}$$

Proof. Given the parallelogram

$$P = \left\{(x,y) : |x| \leq N + \frac{1}{2}, |y - \alpha x| \leq \frac{1}{N}\right\} \subset \mathbb{R}^2$$

, P is bounded and symmetric convex with $|P| = \frac{2(2N+1)}{N} > 2^2$. By Minkowksi's theorem, there exists some $(m,n) \in P \cap \mathbb{Z}^2$.

For this parallelogram, $|m| \leq N + \frac{1}{2}$, so this means m is an integer where $|m| \leq N$ and by symmetry, it means that $|\alpha m - n| < \frac{1}{N}$ Dividing both sides by m brings us to the original

Dirichlet's Approximation Theorem we know: $|\alpha - \frac{n}{m}| < \frac{1}{mN}$. $\square$

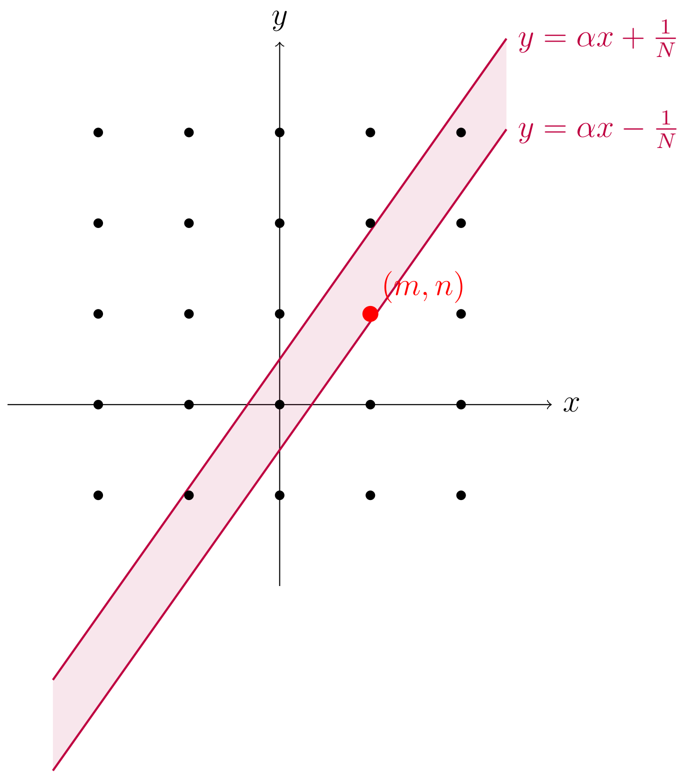

Figure 1: The strip P around the line $y = \alpha x$, with the lattice point (m, n) guaranteed by Minkowski's theorem.

2 Simultaneous Dirichlet Approximation

Before actually diving into the specifics and recent improvement on Dirichlet's Approximation Theorem, the discussion of simultaneous approximation is also an adjacent section that I believe should be covered.

Theorem 2.1 (Simultaneous Dirichlet's Approximation Theorem)**.** *Let $\alpha_1, \ldots, \alpha_n$ be real numbers. For any integer $Q \geq 1$, there exist $p_1, \ldots, p_n \in \mathbb{Z}$ and $q = 1, \ldots, Q$ such that*

$$\left|\alpha_i - \frac{p_i}{q}\right| < \frac{1}{qQ^{1/n}} \quad \textit{for all } i.$$

Proof. Consider the region C in $\mathbb{R}^{n+1}$ defined by

$$-(Q+1) < x_0 < (Q+1), \quad \alpha_i x_0 - Q^{-1/n} < x_i < \alpha_i x_0 + Q^{-1/n}, \quad 1 \leq i \leq n.$$

Since C is defined by $n+1$ independent constraints (n for each x_i and one more for x_0), all these constraints are simply independent intervals of the region C. This means that C is a box in $\mathbb{R}^{n+1}$.

For x_0, the constraint is:

$$-(Q+1) < x_0 < (Q+1)$$

This is an interval of length:

$$(Q+1) - (-(Q+1)) = 2(Q+1)$$

For x_i, the constraint for each x_i is:

$$\alpha_i x_0 - Q^{-1/n} < x_i < \alpha_i x_0 + Q^{-1/n}$$

This is an interval centered at $\alpha_i x_0$ of length:

$$(\alpha_i x_0 + Q^{-1/n}) - (\alpha_i x_0 - Q^{-1/n}) = 2Q^{-1/n}$$

$$\text{Volume} = v(C) = \underbrace{2(Q+1)}_{x_0 \text{ dimension}} \cdot \underbrace{(2Q^{-1/n})^n}_{n \text{ dimensions of } x_i} = 2(Q+1) \cdot \frac{2^n}{Q} = \frac{2^{n+1}(Q+1)}{Q}$$

Since $Q \geq 1$ we have $\frac{Q+1}{Q} > 1$, therefore $v(C) > 2^{n+1}$.

$$v(C) = 2(Q+1)(2Q^{-1/n})^n > 2^{n+1}.$$

□

Lemma 2.2. *If $(x_0, x_1, ..., x_n) \in C$, then $(-x_0, -x_1, ..., -x_n) \in C$ to guarantee symmetry for this new box C.*

Lemma 2.3. *The region C defined by*

$$C = \left\{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} : |x_0| < Q+1,\ |\alpha_i x_0 - x_i| < Q^{-1/n} \textit{ for all } 1 \leq i \leq n\right\}$$

is bounded, symmetric, and convex.

Proof. **Bounded:** The constraint $|x_0| < Q + 1$ bounds x_0. Since $|\alpha_i x_0 - x_i| < Q^{-1/n}$ and x_0 is bounded, each x_i satisfies $|x_i| \le |\alpha_i||x_0| + Q^{-1/n}$, which is also bounded. Hence C is bounded.

Symmetric: Suppose $(x_0, x_1, \ldots, x_n) \in C$. Then $|-x_0| = |x_0| < Q + 1$ and for each i,

$$|\alpha_i(-x_0) - (-x_i)| = |\alpha_i x_0 - x_i| < Q^{-1/n},$$

so $(-x_0, -x_1, \ldots, -x_n) \in C$. Hence C is symmetric.

Convex: Let $(x_0, \ldots, x_n), (y_0, \ldots, y_n) \in C$ and $t \in [0, 1]$. By the triangle inequality:

$$|tx_0 + (1-t)y_0| \le t|x_0| + (1-t)|y_0| < Q + 1,$$

$$|\alpha_i(tx_0 + (1-t)y_0) - (tx_i + (1-t)y_i)| \le t|\alpha_i x_0 - x_i| + (1-t)|\alpha_i y_0 - y_i| < Q^{-1/n}.$$

And finally, hence C is convex. $\square$

Since C is symmetric, convex, bounded, and $v(C) > 2^{n+1}$, Minkowski's theorem guarantees a nonzero integer point $(q, p_1, \ldots, p_n) \in C \cap \mathbb{Z}^{n+1}$.

If $q = 0$ then $|p_i| < Q^{-1/n} \le 1$, forcing $p_i = 0$ for all i, contradicting that $(q, p_1, \ldots, p_n) \ne \mathbf{0}$. Hence $q \ne 0$.

By symmetry of C, we may replace $(q, p_1, \ldots, p_n)$ by $(-q, -p_1, \ldots, -p_n)$ if necessary to arrange $q > 0$. Since $|q| < Q + 1$ and q is a positive integer, we have $1 \le q \le Q$.

The constraint on x_i gives $|\alpha_i q - p_i| < Q^{-1/n}$. Dividing both sides by q:

$$\left|\alpha_i - \frac{p_i}{q}\right| < \frac{Q^{-1/n}}{q} = \frac{1}{qQ^{1/n}} \quad \text{for all } i.$$

Although we took quite a detour from Minkowski to Blichfeldt to discover Simultaneous Dirichlet, the inherent beauty in this reveals all the connections and complexity that arise from the native Dirichlet's Approximation Theorem.

Dirichlet vs Simultaneous Dirichlet

In one dimension, Dirichlet yields $|\alpha - p/q| < 1/q^2$, corresponding to exponent $2 = 1 + 1/1$. In n dimensions, simultaneous Dirichlet gives exponent $1 + 1/n$, which weakens as n grows:

n	Exponent	Bound
1	2	$1/q^2$
2	$3/2$	$1/q^{3/2}$
3	$4/3$	$1/q^{4/3}$
n	$1+1/n$	$1/q^{1+1/n}$

3 Hurwitz Theorem

Theorem 3.1 (Hurwitz Theorem)**.** *For any irrational number ξ, there exists infinitely many rational number $\frac{h}{k}$ such that*

$$\left|\xi - \frac{h}{k}\right| < \frac{1}{\sqrt{5}k^2}$$

has infinitely many rational solutions $\frac{x}{y}$.

Taking the logical next step from Dirichlet's approximation Theorem, Hurwitz's Theorem provides an even sharper error bound on approximating irrationals. In addition, we also see that the golden ratio is the hardest to approximate using this theorem for reasons that will be covered after defining continued fractions.

Definition 3.2 (Continued Fractions)**.** *A finite continued fraction is denoted as*

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

A finite continued fraction can be a rational number, and in accordance, a rational number can be reperesented as a finite continued fraction. The nth convergent is a rational number known as $\frac{h_n}{k_n}$, representing the each continued fraction term of a_n.

In contrast, an infinite continued fraction is an irrational number. Interestingly enough, as the *nth* convergent gets closer to the value of its infinite continued fraction and each term gets more accurate, like numerically evaluating a limit that approaches infinity. Hence, we get that

$$\left|\xi - \frac{h_n}{k_n}\right| < \left|\xi - \frac{h_{n-1}}{k_{n-1}}\right|.$$

This result is important in context of approximating irrational numbers, but also in the sense that it explains why the golden ratio is so difficult to approximate since its growth rate is slow (each a_n is only 1!).

Theorem 3.3 (Legendre's Theorem)**.** *Every rational approximation to an irrational ξ that is accurate to within $\frac{1}{2s^2}$ must be a convergent. More precisely, if $\frac{r}{s}$ is in lowest terms with $s \geq 1$ and satisfies*

$$\left|\xi - \frac{r}{s}\right| < \frac{1}{2s^2}$$

, then $\frac{r}{s}$ is one of the convergents of ξ.

Building off of Legendre's Theorem that gives us the basis to classify a convergent continued fraction, we now introduce how *nth* convergents and reciprocal intertwine.

Lemma 3.4. *For any real number $x > 1$, the nth convergent of $\frac{1}{x}$ is the reciprocal of the $(n-1)$st convergent of x.*

Proof. Let $x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}$ and $\frac{1}{x} = 0 + \cfrac{1}{a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots}}}$. If $\frac{p_n}{q_n}$ and $\frac{h_n}{k_n}$ are the convergents for x and $\frac{1}{x}$ respectively, then:

Convergents of x	**Convergents of $\frac{1}{x}$**
$p_{-1} = 1, p_0 = a_0$	$h_0 = 0, h_1 = 1, h_2 = a_1$
$q_{-1} = 0, q_0 = 1$	$k_0 = 1, k_1 = a_0, k_2 = a_0 a_1 + 1$
$p_n = a_n p_{n-1} + p_{n-2}$	$h_n = a_{n-1} h_{n-1} + h_{n-2}$
$q_n = a_n q_{n-1} + q_{n-2}$	$k_n = a_{n-1} k_{n-1} + k_{n-2}$

Comparing these recurrences, we observe that the sequences (h_n) and (q_{n-1}) satisfy the same recurrence relation $h_n = a_{n-1}h_{n-1} + h_{n-2}$, and similarly (k_n) and (p_{n-1}) satisfy $k_n = a_{n-1}k_{n-1} + k_{n-2}$. Moreover, the initial conditions match: $h_1 = 1 = q_0$ and $k_1 = a_0 = p_0$.

By induction on n, we conclude that $h_n = q_{n-1}$ and $k_n = p_{n-1}$ for all $n \geq 1$. Therefore:

$$\frac{h_n}{k_n} = \frac{q_{n-1}}{p_{n-1}} = \frac{1}{\frac{p_{n-1}}{q_{n-1}}},$$

which is the reciprocal of the $(n-1)$st convergent of x. $\square$

To prove Hurwitz Theorem, we need to show that for any three consecutive convergents, at least one of them satisfies the Hurwitz bound.

Proof. Consider any three consecutive convergents $\frac{h_{n-1}}{k_{n-1}}, \frac{h_n}{k_n}, \frac{h_{n+1}}{k_{n+1}}$ of the continued fraction expansion of ξ. We will demonstrate that at least one satisfies the Hurwitz bound. Define $\gamma_n = k_n^2 \left|\xi - \frac{h_n}{k_n}\right|$.

Our goal is to show that $\min(\gamma_{n-1}, \gamma_n) < \frac{1}{\sqrt{5}}$. Since ξ lies between consecutive convergents, we have

$$\left|\xi - \frac{h_n}{k_n}\right| + \left|\xi - \frac{h_{n-1}}{k_{n-1}}\right| = \left|\frac{h_n}{k_n} - \frac{h_{n-1}}{k_{n-1}}\right| = \frac{1}{k_n k_{n-1}}.$$

Multiply both sides by $k_n^2 k_{n-1}^2$:

$$\gamma_n k_{n-1}^2 + \gamma_{n-1} k_n^2 - k_n k_{n-1} = 0.$$

Replacing n with $n+1$ yields an identical relation:

$$\gamma_{n+1} k_n^2 + \gamma_n k_{n+1}^2 - k_{n+1} k_n = 0.$$

For a finite continued fraction, the recurrence $k_{n+1} = a_{n+1} k_n$ implies $k_{n+1} - k_n = a_{n+1} k_n$. This gives

$$a_{n+1}^2 \gamma_n^2 + 2\gamma_n(\gamma_{n-1} - \gamma_{n+1}) = 1 - a_{n+1}^2(\gamma_n - 1 - \gamma_n + 1)^2.$$

For the right side to remain less than 1, we need $\gamma_{n-1} \neq \gamma_{n+1}$. If they were equal, the previous relation would be violated, implying ξ is rational—a contradiction.

We observe that

$$k_n^2 \left|\xi - \frac{h_{n-1}}{k_{n-1}}\right| = k_n^2 \left|\xi - \frac{h_{n+1}}{k_{n+1}}\right|.$$

This is equivalent to

$$k_{n-1}^2 \left(\xi - \frac{h_{n-1}}{k_{n-1}}\right) = k_{n+1}^2 \left(\xi - \frac{h_{n+1}}{k_{n+1}}\right).$$

Since ξ is irrational, $k_{n-1} \neq k_{n+1}$, thus $\gamma_{n-1} \neq \gamma_{n+1}$. Let $\gamma = \min(\gamma_{n-1}, \gamma_n, \gamma_{n+1})$. Then

$$\gamma^2(k_n^2 + 4) \leq 5\gamma_n^2 + 2\gamma_n(\gamma_{n-1} + \gamma_{n+1}) < 1.$$

Since $a_{n+1} \geq 1$, we have $5\gamma^2 < 1$. Therefore $\gamma < \frac{1}{\sqrt{5}}$, and thus

$$\min(\gamma_{n-1}, \gamma_n, \gamma_{n+1}) < \frac{1}{\sqrt{5}}.$$

The Hurwitz bound $\frac{1}{\sqrt{5}k_n^2}$ is the tightest constant for which infinitely many rational approximations exist. Once approximations exceed $\frac{1}{\sqrt{5}k_n^2}$ in accuracy, only finitely many solutions remain, revealing a whole new area of Diophantine approximation involving narrowing down what constitutes the range of solutions through p-adic analysis (which I will not be covering). □

4 Irrationality Measure

Throughout this paper, we have seen how there have been many different generalizations and new forms of our original problem of approximating irrationals such as the Hurwitz Theorem. As accuracy varies as the degree of the error term increases for specific constraints, we can actually measure and quantify this whole concept through the irrationality measure.

Definition 4.1 (Irrationality Measure)**.** *Let α be a real number, and let R be the set of positive real numbers μ for which*

$$0 < |x - \frac{p}{q}| < \frac{1}{q^\mu}$$

has (at most) finitely many solutions $\frac{p}{q}$ for p and q integers.

Then the irrationality measure is defined as the threshold at which Liouville's approximation theorem kicks in and x is no longer approximable by rational numbers,

$$\mu(x) = \inf_{p \in R} \mu_p.$$

The inf function is the infinum, which is the greatest lower bound of a set C defined as a quantit m such that no member of the set is less than m, but if e is any positive quantity however small, there is always one member less than $m + e$.

If the set R is empty , then $\mu(x)$ is defined to be $\mu(x) = \infty$, and x is a Liouville Number. For a nonempty R, there are three possible options:

$$\mu(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 2 & \text{if } x \text{ is algebraic of degree } > 1 \\ 2 & \text{if } x \text{ is transcendental} \end{cases}$$

The definition of irrationality exponent can be reframed as follows: if x has irrationality exponent μ, then μ is the supremum of all values such that the inequality

$$\left| x - \frac{p}{q} \right| > \frac{C}{q^{\mu+\varepsilon}}$$

holds for any $\varepsilon > 0$ and all sufficiently large q, where C is a constant depending on x and ε.

For irrational numbers with a continued fraction expansion $x = [a_0, a_1, a_2, \ldots]$ and convergents p_n/q_n, the irrationality exponent admits an explicit formula:

$$\mu(x) = 1 + \limsup_{n\to\infty} \frac{\ln q_{n+1}}{\ln q_n}$$

This can be rewritten as

$$\mu(x) = 2 + \limsup_{n\to\infty} \frac{\ln a_{n+1}}{\ln q_n}.$$

A classical example is the golden ratio ϕ, which satisfies

$$\mu(\phi) = 2,$$

a consequence of the equation combined with the continued fraction expansion $\phi = [1; 1, 1, 1, \ldots]$ where all partial quotients equal 1.

Below are some common constants and their respective upper bounds, showing the differences in their irrationality measures.

Constant x	**Upper Bound on** $\mu(x)$	**Reference**
π	7.1032	Hata (1992), Baker's method
e	2	Hermite-Lindemann, transcendence
$\ln 2$	3.891393	Hata (1993), linear forms
$\zeta(3)$	5.513891	Rhin-Viola (1999), Padé approximants
ϕ (golden ratio)	2	Hurwitz, convergents

Table 1: Known upper bounds on the irrationality exponents of special constants.

More interestingly, the Liouville constant $L = \sum_{n=1}^{\infty} 10^{-n!} = 0.110001...$ has $\mu(x) = \infty$.

5 Roth's Theorem / Subspace Theorem

Let α be an irrational, real number. Given a rational number x/y with $x, y \in \mathbb{Z}$, $\gcd(x, y) = 1$ and $y > 0$, we define the quality of x/y to be the real M such that $|\alpha - x/y| = y^{-M}$ if $|\alpha - x/y| < 1$; if $|\alpha - x/y| > 1$ we set $M = 0$. Dirichlet's Theorem implies that every irrational real number has infinitely many rational approximations of quality ≤ 2:

Theorem 5.1. *Let $\alpha \in \mathbb{R}$, $\alpha \notin \mathbb{Q}$. Then there are infinitely many pairs of integers x, y such that*

$$\left|\alpha - \frac{x}{y}\right| \leq y^{-2}, \quad y > 0.$$

The famous theorem of Roth implies that if α is an irrational algebraic number, then for every $\delta > 0$, it has only finitely many rational approximations of quality $> 2 + \delta$.

Theorem 5.2 (Roth's Theorem)**.** *Let $\alpha \in \mathbb{C}$ be an irrational algebraic number and let $\delta > 0$. Then there are only finitely many pairs of integers (x, y) such that*

$$\left|\alpha - \frac{x}{y}\right| \leq y^{-2-\delta}, \quad y > 0.$$

Roth's Theorem is easy to prove if $\alpha \in \mathbb{C} \setminus \mathbb{R}$, or if α is a real quadratic number. For real algebraic numbers α of degree > 3, the proof of Roth's Theorem is very difficult.

In the formulation of the Subspace Theorem, we need some notions from linear algebra, which we recall below. Let n be an integer > 1 and $r \leq n$. We say that linear forms $L_1 = \sum_{j=1}^{n} \alpha_{1j} X_j, \ldots, L_r = \sum_{j=1}^{n} \alpha_{nj} X_j$ with coefficients in $\mathbb{C}$ are linearly dependent if there are $c_1, \ldots, c_r \in \mathbb{C}$, not all 0, such that $c_1 L_1 + \cdots + c_r L_r \equiv 0$. Otherwise, $L_1, \ldots, L_r$ are called linearly independent. If $r = n$, then $L_1, \ldots, L_n$ are linearly independent if and only if their coefficient determinant $\det(L_1, \ldots, L_n) = \det(\alpha_{ij})_{1 \leq i,j \leq n} \neq 0$.

A linear subspace T of $\mathbb{Q}^n$ of dimension r can be described as

$$T = \left\{ \sum_{i=1}^{r} z_i a_i : z_1, \ldots, z_r \in \mathbb{Q} \right\},$$

where $a_1, \ldots, a_r$ are linearly independent vectors from $\mathbb{Q}^n$. Alternatively, T can be described as

$$T = \{x \in \mathbb{Q}^n : L_1(x) = 0, \ldots, L_{n-r}(x) = 0\}$$

where $L_1, \ldots, L_{n-r}$ are linearly independent linear forms in $X_1, \ldots, X_n$ with coefficients in

$\mathbb{Q}$.

The norm of $x = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ is given by

$$\|x\| := \max(|x_1|, \ldots, |x_n|).$$

Theorem 5.3 (Subspace Theorem)**.** *Let $n \geq 2$, let*

$$L_i(X) = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n \quad (i = 1, \ldots, n)$$

be n linearly independent linear forms with algebraic coefficients in $\mathbb{C}$ and let $\delta > 0$. Then the set of solutions of the inequality

$$|L_1(x) \cdots L_n(x)| \leq \|x\|^{-\delta}$$

in $x \in \mathbb{Z}^n$ is contained in a union $T_1 \cup \cdots \cup T_t$ of finitely many proper linear subspaces of $\mathbb{Q}^n$.

We show that the Subspace Theorem implies Roth's Theorem.

Subspace Theorem $\Longrightarrow$ Roth's Theorem. Let (x, y) (with $y > 0$) be a pair of integers satisfying $|\alpha - x/y| \leq y^{-2-\delta}$. Multiplying with y^2 gives

$$|y(x - \alpha y)| \leq y^{-\delta}.$$

Since the linear forms Y and $X - \alpha Y$ are linearly independent, this is an inequality to which the Subspace Theorem is applicable, except that on the right hand side we have y instead of $\|x\| = \max(|x|, |y|)$. However, we certainly have $|x/y| \leq |\alpha| + y^{-2} \leq |\alpha| + 1$. So $|x| \leq (|\alpha| + 1)y =: Cy$ and then also, $\|x\| \leq Cy$. With this observation, our inequality becomes

$$|y(x - \alpha y)| \leq C^{\delta}\|x\|^{-\delta},$$

and now we can apply the Subspace Theorem. We infer that the pairs (x, y) lie in only finitely many proper, i.e., one-dimensional, linear subspaces of $\mathbb{Q}^2$.

Pick one of these subspaces, say T. We show that T contains only finitely many solutions of the stated inequality. We have $T = \{\lambda(x_0, y_0) : \lambda \in \mathbb{Q}\}$ where for (x_0, y_0) we may take a pair of integers with $\gcd(x_0, y_0) = 1$. If $(x, y) \in T \cap \mathbb{Z}^2$ and satisfies the inequality, then

$(x, y) = \lambda(x_0, y_0)$ with $\lambda \in \mathbb{Z}$, and

$$0 < \left|\alpha - \frac{x_0}{y_0}\right| = \left|\alpha - \frac{x}{y}\right| \leq |\lambda|^{-2} y_0^{-2},$$

implying that $|\lambda|$ is bounded in terms of T. So indeed T contains only finitely many solutions. Roth's theorem easily follows. $\square$

The Subspace Theorem states that the set of solutions of the inequality is contained in a finite union of proper linear subspaces of $\mathbb{Q}^n$, but one may wonder whether the inequality doesn't have only finitely many solutions. For instance, it may be that there is a non-zero $x_0 \in \mathbb{Z}^n$ with $L_1(x_0) = 0$. Then for every $d \in \mathbb{Z}$, the point dx_0 is a solution, and this gives infinitely many solutions. To avoid such a construction, let us consider

$$0 < |L_1(x) \cdots L_n(x)| \leq \|x\|^{-\delta}$$

in $x \in \mathbb{Z}^n$.

For instance, let $n = 2$. Then the solutions lie in finitely many one-dimensional subspaces of $\mathbb{Q}^2$, and similarly as in the proof of Roth's Theorem, each of these subspaces contains only finitely many solutions. So altogether, if $n = 2$ then the inequality has only finitely many solutions. We now give an example showing that the inequality may have infinitely many solutions if $n \geq 3$.

Example. Let $0 < \delta < 1$ and consider the inequality

$$0 < |(x_1 + \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 - \sqrt{3}x_3)| \leq \|x\|^{-\delta}$$

to be solved in $x = (x_1, x_2, x_3) \in \mathbb{Z}^3$. The three linear forms on the left-hand side are linearly independent.

We examine the triples of integers $x = (x_1, x_2, x_3) \in \mathbb{Z}^3$ with $x_3 = 0$ and $x_1 x_2 \neq 0$. For these points, $\|x\| = \max(|x_1|, |x_2|, 0)$. By Dirichlet's Theorem, the inequality

$$\left|\sqrt{2} - \frac{x_1}{x_2}\right| \leq |x_2|^{-2}$$

has infinitely many solutions $(x_1, x_2) \in \mathbb{Z}^2$ with $x_2 \neq 0$. For these solutions, $\|x\|$ has the same order of magnitude as $|x_2|$. Indeed,

$$|x_1/x_2| \leq |x_2|^{-2} + \sqrt{2} \leq 1 + \sqrt{2},$$

and so $\|x\| = \max(|x_1|, |x_2|) \leq (1+\sqrt{2})|x_2|$.

For the points under consideration,

$$\begin{aligned}
&0 < |(x_1 + \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 + \sqrt{3}x_3)(x_1 - \sqrt{2}x_2 - \sqrt{3}x_3)| \\
&= |(x_1 + \sqrt{2}x_2)(x_1 - \sqrt{2}x_2)^2| \\
&\leq (1+\sqrt{2})\|x\| \cdot (x_2^{-1})^2 \leq (1+\sqrt{2})^3 \|x\|^{-1} \leq \|x\|^{-\delta},
\end{aligned}$$

provided $\|x\|$ is sufficiently large. It follows that the inequality has infinitely many solutions x lying in the subspace $x_3 = 0$. In a similar manner, one can show that the inequality has infinitely many solutions in the subspaces $x_1 = 0$ and $x_2 = 0$ as well.

By the Subspace Theorem, the solutions of the inequality with $x_1x_2x_3 \neq 0$ lie in finitely many proper linear subspaces of $\mathbb{Q}^3$. With a more refined argument one shows that the inequality has only finitely many solutions with $x_1x_2x_3 \neq 0$.

We give a generalization of the Subspace Theorem which may be useful for certain applications.

We say that a system of $r \geq n$ linear forms $L_1, \ldots, L_r$ in the variables $X_1, \ldots, X_n$ is in general position if each n-tuple of linear forms among $L_1, \ldots, L_r$ is linearly independent.

Theorem 5.4 (Generalized Subspace Theorem)**.** *Let*

$$L_i(X) = \alpha_{i1}X_1 + \cdots + \alpha_{in}X_n \quad (i = 1, \ldots, r,\ r \geq n)$$

be r linear forms with algebraic coefficients in $\mathbb{C}$ in general position and let $\delta > 0$. Then the set of solutions of the inequality

$$|L_1(x) \cdots L_r(x)| \leq \|x\|^{r-n-\delta}$$

in $x \in \mathbb{Z}^n$ is contained in a union $T_1 \cup \cdots \cup T_t$ of finitely many proper linear subspaces of $\mathbb{Q}^n$.

This can be deduced by combining the Subspace Theorem with the following lemma.

Lemma 5.5. *Let $M_1, \ldots, M_n$ be linearly independent linear forms in $X_1, \ldots, X_n$ with complex coefficients. Then there is a constant $C > 0$ such that*

$$\|x\| \leq C \max\{|M_1(x)|, \ldots, |M_n(x)|\}$$

for all $x \in \mathbb{C}^n$.

Proof. Since the linear forms $M_1, \ldots, M_n$ are linearly independent, they span the complex vector space of all linear forms in $X_1, \ldots, X_n$ with complex coefficients. So we can express $X_1, \ldots, X_n$ as linear combinations of $M_1, \ldots, M_n$, i.e.,

$$X_i = \sum_{j=1}^{n} \beta_{ij} M_j \quad \text{with} \quad \beta_{ij} \in \mathbb{C} \quad (i = 1, \ldots, n).$$

Take $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ and put $M := \max_{1 \le i \le n} |M_i(x)|$. Then

$$\max_{1 \le i \le n} |x_i| \le \max_{1 \le i \le n} \sum_{j=1}^{n} |\beta_{ij}| \cdot |M_j(x)| \le C \cdot M$$

with $C := \max_{1 \le i \le n} \sum_{j=1}^{n} |\beta_{ij}|$. □

Proof of the Generalized Subspace Theorem. We partition the solutions x of the inequality into a finite number of subsets according to the ordering of the numbers $|L_1(x)|, \ldots, |L_r(x)|$, and show that each of these subsets lies in at most finitely many proper linear subspaces of $\mathbb{Q}^n$. Consider the solutions $x \in \mathbb{Z}^n$ from one of these subsets, say for which

$$|L_1(x)| \le \cdots \le |L_r(x)|.$$

By the preceding Lemma, for $i = n+1, \ldots, r$, there is a constant C_i such that for all solutions x under consideration,

$$\|x\| \le C_i |L_i(x)|,$$

since $L_1, \ldots, L_{n-1}, L_i$ are linearly independent, and $|L_i(x)|$ has the largest value. Substituting this into the inequality gives

$$|L_1(x) \cdots L_n(x)| \le C\|x\|^{r-n-\delta} \prod_{i=n+1}^{r} |L_i(x)|^{-1} \le C \cdot (C_{n+1} \cdots C_r) \|x\|^{-\delta}.$$

So the solutions x under consideration lie in at most finitely many proper linear subspaces of $\mathbb{Q}^n$. □

We deduce a result on simultaneous approximation.

Theorem 5.6. *Let* $\alpha_1, \ldots, \alpha_n$ *be algebraic numbers in* $\mathbb{C}$ *and let* $C > 0$, $\delta > 0$. *Then the inequality*

$$0 < |\alpha_1 x_1 + \cdots + \alpha_n x_n| \le C\|x\|^{1-n-\delta}$$

in $x = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ has only finitely many solutions.

For $n = 2$ this result is equivalent to Roth's Theorem.

Proof. We proceed by induction on n. For $n = 1$ the assertion is obvious (here we use our assumption $\alpha_1 x_1 \neq 0$). Let $n > 1$ and suppose the theorem is true for linear forms in fewer than n variables.

We apply the Subspace Theorem. We may assume that at least one of the coefficients $\alpha_1, \ldots, \alpha_n$ is non-zero, otherwise there are no solutions. Suppose that $\alpha_1 \neq 0$. Then the inequality implies

$$|(\alpha_1 x_1 + \cdots + \alpha_n x_n) x_2 \cdots x_n| \leq C \|x\|^{-\delta}$$

and by the Subspace Theorem, the solutions of the latter lie in a union of finitely many proper linear subspaces $T_1, \ldots, T_t$ of $\mathbb{Q}^n$. We consider only solutions with $\alpha_1 x_1 + \cdots + \alpha_n x_n \neq 0$. Therefore, without loss of generality we may assume that $\alpha_1 x_1 + \cdots + \alpha_n x_n$ is not identically 0 on any of the spaces $T_1, \ldots, T_t$.

Consider the solutions of the inequality in T_i. Choose a non-trivial linear form vanishing identically on T_i, $a_1 x_1 + \cdots + a_n x_n = 0$. Suppose for instance, that $a_n \neq 0$. Then x_n can be expressed as a linear combination of $x_1, \ldots, x_{n-1}$. By substituting this into the inequality we obtain an inequality

$$0 < |\beta_1 x_1 + \cdots + \beta_{n-1} x_{n-1}| \leq C \|x\|^{1-n-\delta} \leq C \left(\max_{1 \leq i \leq n-1} |x_i| \right)^{2-n-\delta}.$$

By the induction hypothesis, the latter inequality has only finitely many solutions $(x_1, \ldots, x_{n-1})$. So T_i contains only finitely many solutions x of the inequality.

Applying this to $T_1, \ldots, T_t$ we obtain that the inequality has altogether only finitely many solutions. $\square$

6 Acknowledgements

I would like to thank Dr. Simon Rubinstein Salzedo for providing me this opportunity. I would also like to thank my mentor Maxwell Siegel for her guidance and valuable feedback. I also thank my fellow students at Euler Circle.

References

[1] Zichao Dong, *Minkowski's Theorem and Its Applications*, Reading seminar notes, Carnegie Mellon University, 2018–2019.

[2] Jan-Hendrik Evertse, *The Subspace Theorem*, Master Course on Diophantine Equations, Universiteit Leiden, April 2011.

[3] Alexander Gorodnik, *Number Theory, Lecture 12*, Universität Zürich.

[4] Alexander Gorodnik, *Number Theory, Lecture 14: Geometry of Numbers*, Universität Zürich.

[5] Wolfgang M. Schmidt, *Diophantine Approximation*, Lecture Notes in Mathematics 785, Springer-Verlag, Berlin, 1980.

[6] Eric W. Weisstein, *Irrationality Measure*, MathWorld—A Wolfram Web Resource.

[7] Dylan Zukin, *The Farey Sequence and Its Niche(s)*, Senior Project, Whitman College, May 2016.