FURTHER NUMBER THEORY

SANJU VELANI

1. Dirichlet's Theorem and Continued Fractions

Recall a fundamental theorem in the theory of Diophantine approximation.

Theorem 1.1 (Dirichlet $\sim$ 1842)**.** *For any real number x and $N \in \mathbb{N}$, there exists $p, q \in \mathbb{Z}$ with $1 \leqslant q \leqslant N$ such that*

$$\left| x - \frac{p}{q} \right| \leqslant \frac{1}{qN} \left(\leqslant \frac{1}{q^2} \right)$$

- important consequence.

Theorem 1.2 (Dirichlet)**.** *For any real number x there exists infinitely many rationals $\frac{p}{q}$ $(q > 0)$ such that*

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

Remark. (1) Any real number can be approximated by rationals with 'rate' (denominator)$^{-2}$.

(2) If we impose condition that $\frac{p}{q}$ is reduced (i.e. $(p, q) = 1$) then statement requires x to be irrational; indeed suppose $x = \frac{a}{b} \left(\neq \frac{p}{q} \right)$. Then,

$$\begin{aligned} \left| x - \frac{p}{q} \right| = \left| \frac{a}{b} - \frac{p}{q} \right| &= \frac{|aq - pb|}{bq} \\ \because aq \neq pb \Rightarrow \qquad &\geqslant \frac{1}{bq} \end{aligned}$$

Thus $q < b$ for

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2} \tag{1.1}$$

But there are only a finite number of rationals with $q < b$ of interest.

(3) In view of (2), imposing condition that the rationals are reduced gives a criterion for determining whether a given real number x is rational or irrational.

Key part of the course is to investigate numbes for which the 'rate' of approximation in Dirichlet (i.e. (denominator)$^{-2}$) can be improved by some power. More precisely, for $\tau \geqslant 2$, let $W(\tau)$ denote the set of real numbers x for which there exists infinitely many rationals $\frac{p}{q}$ $(q > 0)$ such that

$$\left|x - \frac{p}{q}\right| < \frac{1}{q^\tau}$$

- $\tau \leqslant 2$ gives us that $W(2) = \mathbb{R}$, which is just the theorem above.
- $\tau > 2$ - studying $W(\tau)$ brings into play 'serious' measure theory - in particular the 'world' of fractals.

Lets rephrase Thereom 2: given any irrational x we can find an infinite sequence of 'good' rationals $\frac{p_n}{q_n}$ which approximate x with 'rate' at least $\frac{1}{q_n^2}$.

Question: Is there a natural way to find these 'good' approximates?

Yes - the convergents $\frac{p_n}{q_n}$ associated with a continued fraction expansion of x satsifies the crucial inequality

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

Let x be irrational. Then x has a unique representation as an infinite continued fraction, i.e. given x there exists a unique sequence of integers $a_0, a_1, a_2, \ldots$ where $a_0 \in \mathbb{Z}$ and $a_i \geqslant 1$ for all $i \geqslant 1$ (called the *partial quotients* of x), such that

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}} = [a_0; a_1, a_2, \ldots].$$

To find the partial quotients use the continued fraction algorithm.

Example 1.1. The Golden Ratio has the simplest continued fraction expansion

$$\frac{\sqrt{5}+1}{2} = [1; 1, 1, 1, \ldots] := [\overline{1}].$$

The above is an example of a quadratic irrational; i.e. a root of an algebraic polynomial of degree 2, such as

$$f(x) = ax^2 + bx + c \qquad (a, b, c \in \mathbb{Z}, a \neq 0).$$

So we have that $\frac{\sqrt{5}+1}{2}$ is a root of $f(x) = x^2 - x - 1$ and it has a *periodic* continued fraction expansion.

Definition: If an infinite continued fraction contains a block $b_1, b_2, \ldots, b_n$ which repeats indefinitely then the continued fraction expansion is called *periodic*.

So, we write

$$[a_0, a_1, \ldots, a_m, \overline{b_1, \ldots, b_n}]$$

for

$$[a_0, a_1, \ldots, a_m, b_1, \ldots, b_n, b_1, \ldots, b_n, \ldots].$$

An elegant statement is

Theorem 1.3. *The continued fraction expansion of irrational x is periodic if an only if x is a quadratic irrational.*

Remark. If x is periodic, i.e. $x = [a_0, \ldots, a_m, \overline{b_1, \ldots, b_n}]$ then the partial quotients in its continued fraction expansion are bounded by a constant

$$K(x) \text{ - namely the } \max\{|a_0|, a_1, \ldots, a_m, b_1, \ldots, b_n\}.$$

Theorem 1.3 $\Rightarrow$ quadratic irrationals have bounded partial quotients - the set Bad.

- Given $x = [a_0, a_1, \ldots]$ consider its nth convergent

$$\frac{p_n}{q_n} = [a_0, a_1, a_2, \ldots, a_n] \quad n \geqslant 0.$$

- p_n and q_n satisfying

$$\begin{aligned} p_0 &= a_0 & q_0 &= 1 \\ p_1 &= a_1 a_0 + 1 & q_1 &= a_1 \\ &\vdots & &\vdots \\ p_n &= a_n p_{n-1} + p_{n-2} & q_n &= a_n q_{n-1} + q_{n-2}. \end{aligned}$$

Lemma 1.1. *$q_{n-1} \leqslant q_n$ for all $n \geqslant 1$ wtih strict inequality when $n > 1$ (Exercise).*

Lemma 1.2. *$(p_n, q_n) = 1$ for all $n \geqslant 1$*

The key properties of convergents are

Theorem 1.4. *If $\frac{p_n}{q_n}$ $(n \geqslant 0)$ is the nth convergent to irrational x then*

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}$$

Corollary 1.1. *The convergents $\frac{p_n}{q_n}$ $(n \geqslant 0)$ satisfy*

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

compare from Theorem 1.2 (just use Lemma 1.1).

Theorem 1.5. *If $\frac{p_n}{q_n}$ $(n \geqslant 0)$ is the nth convergent to irrational x then*

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

UPSHOT: Theorem 1.4 and 1.5 $\Rightarrow$ the approximation by convergents is 'trapped' by the denominators of the convergents.

Theorem 1.6 (Best Approximation)**.** *Let $\frac{p_n}{q_n}$ $(n \geqslant 0)$ be the nth convergent of irrational x. If $1 \leqslant b \leqslant q_n$, the rational $\frac{a}{b}$ satisfies*

$$\left| x - \frac{p_n}{q_n} \right| \leqslant \left| x - \frac{a}{b} \right|.$$

Every other rational with the same or smaller denominator differs from x by a greater amount.

2. Hurwitz and Bad

The theory of continued fractions not only enables us to 'effectively' prove Dirichlet's Theorem (Theorem 1.2) - it enables us to improve it.

Theorem 2.1 (Hurwitz $\sim$ 1891)**.** *For any real number x there exists infinitely many rationals $\frac{p}{q}$ $(q > 0)$ such that*

$$\left| x - \frac{p}{q} \right| < \frac{1}{\sqrt{5}q^2}. \tag{2.1}$$

So the 'rate' of $\frac{1}{q^2}$ is replaced by the 'rate' of $\frac{1}{\sqrt{5}q^2}$.

The continued fraction theory $\Rightarrow$ for x irrational, at least one of the three consecutive convergents

$$\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}, \frac{p_{n+2}}{q_{n+2}}$$

satisfy (2.1).

Question: Can $\frac{1}{\sqrt{5}}$ be replaced by something smaller?

Theorem 2.2. *The constant $\frac{1}{\sqrt{5}}$ in Hurwitz is best possible.*

Proof. Assume that $\frac{1}{\sqrt{5}}$ can be replaced by something smaller, in particular by

$$\frac{1}{\sqrt{5}+\epsilon} \qquad (\epsilon > 0, \text{ arbitrary}).$$

Consider $x_1 = \frac{\sqrt{5}+1}{2}$, root of the polynomial

$$f(t) = t^2 - t - 1 = (t - x_1)(t - x_2)$$

where $x_2 = \frac{1-\sqrt{5}}{2}$. Assume there exists a sequence of rationals $\frac{p_i}{q_i}$ satisfying

$$\left|x - \frac{p_i}{q_i}\right| < \frac{1}{(\sqrt{5}+\epsilon)q_i^2}.$$

Then for i sufficiently large

$$\left|x_2 - \frac{p_i}{q_i}\right| \leqslant \underbrace{|x_2 - x_1|}_{=\sqrt{5}} + \underbrace{\left|x_1 - \frac{p_i}{q_i}\right|}_{<\frac{1}{(\sqrt{5}+\epsilon)q_i^2}<\epsilon}$$

which is just the triangle inequality. It follows that

$$\left|f\left(\frac{p_i}{q_i}\right)\right| < \frac{1}{(\sqrt{5}+\epsilon)q_i^2} \cdot \sqrt{5}+\epsilon$$
$$\Rightarrow \left|q_i^2 f\left(\frac{p_i}{q_i}\right)\right| < 1.$$

However the L.H.S. is a positive integer - a contradiction - there are no integers in $(0,1)$. □

The proof shows that if we replace $\frac{1}{\sqrt{5}}$ by anything smaller (say $\frac{1}{\sqrt{5}+\epsilon}$) then Hurwitz statement is false for the Golden Ratio, i.e. with $x_1 = \frac{\sqrt{5}+1}{2}$ and $\epsilon > 0$ arbitrary there are at most a finite number of rationals $\frac{p}{q}$ such that

$$\left|x - \frac{p}{q}\right| < \frac{1}{(\sqrt{5}+\epsilon)q^2}.$$

Alternatively, there exists a constant say $c(x)$ such that

$$\left|x - \frac{p}{q}\right| > \frac{c(x)}{q^2} \text{ for all } \frac{p}{q} \in \mathbb{Q}.$$

In fact we can take $c(x) = \frac{1}{\sqrt{5}+\epsilon}$ for the Golden Ratio. So, $\frac{\sqrt{5}+1}{2}$ has the simplest continued fraction and is the 'worst' approximable irrational.

2.1. **Badly Approximable Numbers**

Definition: A real number x is said to be *badly approximable* if there exists a constant $c(x) > 0$ such that

$$\left|x - \frac{p}{q}\right| > \frac{c(x)}{q^2} \text{ for all } \frac{p}{q} \in \mathbb{Q}.$$

In view of Theorem 2.2,

$$0 < c(x) < \frac{1}{\sqrt{5}}.$$

Let Bad denote the set of badly approximable numbers. Clearly Bad $\neq \emptyset$ since

$$\gamma = \frac{\sqrt{5}+1}{2} \in \text{Bad}$$

and indeed $t\gamma \in \text{Bad}$ for $t \in \mathbb{Z} \setminus \{0\}$ (Exercise). A connection between Bad and the theory of continued fractions is given by

Theorem 2.3. *An irrational* $x \in \text{Bad}$ *if and only if the partial quotients* a_i *in its continued fraction expansion* $[a_0; a_1, a_2, \dots]$ *are bounded; i.e. there exists a constant* $K(x) > 0$ *such that* $|a_i| \leqslant K(x)$ *for all* i*.*

Corollary 2.1. Bad *is an uncountable set.*

Corollary 2.2. *Let* x *be a quadratic irrational. Then* $x \in \text{Bad}$*.*

Proof. This is just a consequence of Theorem 1.3. □

Remark. A famous conjecture states that the quadratic irrationals are the only algebraic irrationals in Bad. For example does $2^{\frac{1}{3}}$ have bounded partial quotients?

Recall from chapter 1 Theorems 1.4 and 1.5. Then we can look at the proof of Theorem 3.

Proof. $x \in \text{Bad} \Rightarrow x$ has bounded partial quotients. This is an exercise. Now examine and see whether the continued fraction expansion is also bounded. Suppose $x = [a_0; a_1, a_2, \dots]$ with $|a_i| \leqslant K$. Let $\frac{p}{q} \in \mathbb{Q}$ with $q \geqslant 1$. Then there exists $n \geqslant 1$ such that

$$q_{n-1} \leqslant q < q_n.$$

The convergents are the best approximates (Theorem 1.6), so

$$\left| x - \frac{p}{q} \right| \geqslant \left| x - \frac{p_n}{q_n} \right| > \frac{1}{q_n(q_{n+1}+q_n)} = \frac{1}{q_n^2\left(\frac{q_{n+1}}{q_n}+1\right)}.$$

However

$$\begin{aligned} \frac{q_{n+1}}{q_n} + 1 &= \frac{a_{n+1}q_n + q_{n-1}}{1_n} + 1 \\ &= a_{n+1} + \underbrace{\frac{q_{n-1}}{q_n}}_{\leqslant 1} + 1 \qquad \text{(Lemma 1, Chapter 1)} \\ &\leqslant a_{n+1} + 2 \\ &\leqslant K + 2. \end{aligned}$$

Thus

$$\begin{aligned}\left|x - \frac{p}{q}\right| &> \frac{1}{q_n^2(K+2)} \\ &= \frac{1}{q^2(K+2)}\frac{q^2}{q_n^2} \\ &> \frac{1}{q^2(K+2)}\left(\frac{q_{n-1}}{q_n}\right)^2 \quad (q \geqslant q_{n-1}).\end{aligned}$$

If $n = 1$ then

$$\frac{q_0}{q_1} = \frac{1}{a_1} \geqslant \frac{1}{K} > \frac{1}{K+1}.$$

If $n \geqslant 2$,

$$\begin{aligned}\frac{q_{n-1}}{q_n} &= \frac{q_{n-1}}{a_n q_{n-1} + q_{n-2}} \\ &= \frac{1}{a_n + \frac{q_{n-2}}{q_{n-1}}} \geqslant \frac{1}{K+1}\end{aligned}$$

UPSHOT: We have that

$$\left|x - \frac{p}{q}\right| > \frac{1}{q^2(K+2)(K+1)^2}$$

i.e. $x \in$ Bad. □

- Particular goal is to investigate the 'size' of Bad.
- $m(\text{Bad}) = 0$, i.e. zero Lebesgue measure.
- Although it is small in terms of Lebesgue measure it is maximal in terms of dimension as $\dim(\text{Bad}) = 1$

3. Algebraic and Transcendental Numbers

An *algebraic number* is a number x which satisfies an algebraic equation. An equation of the form

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

with each $a_i \in \mathbb{Z}$ and at least one of these a_i is non-zero. A number which is **not** algebraic is called *transcendental*.

EXAMPLE 3.1. We see the following numbers are algebraic:

- Every rational $\frac{a}{b}$ is algebraic because it satisfies $bx - a = 0$.
- $\sqrt{2}$ is algebraic because it satisfies $x^2 - 2 = 0$.
- $i = \sqrt{-1}$ is algebraic because it satisfies $x^2 + 1 = 0$.

An algebraic number satisfies an infinite number of algebraic equations of different degrees. For example $x = \sqrt{2}$ satisfies

$$
\begin{aligned}
x^2 - 2 &= 0 \\
x^4 - 4 &= 0 \\
x^6 - 8 &= 0
\end{aligned}
$$

and so on. If x satisfies an algebraic equation of degree n but none of lower degree then we say that x is of degree n.

EXAMPLE 3.2. $\sqrt{2}$ is of degree 2 and any rational is of degree 1.

The set of algebraic numbers is countable while the set of transcendental numbers is uncountable. It's much harder to prove that a given number is transcendental than to prove that it is irrational. Transcendental $\Rightarrow$ irrational. (If ζ is transcendental and rational then ζ is algebriac, which is a contradiction).

Aim: is to produce as many examples of transcendental numbers as we like.
First a useful notion:

Definition: A real number x is said to be *approximable* by rationals to *order* n if there exists a positive constant c depending only on x such that

$$\left| x - \frac{p}{q} \right| < \frac{c}{q^n}$$

has infinitely many rational solutions $\frac{p}{q}$ $(q > 0)$ with $(p, q) = 1$.

Dirichlet's Theoreom implies that an irrational is approximable to order 2.

EXERCISE: Any rational is approximable to order 1 but not to any higher order.

Liouville used the following to construct transcendental numbers.

Theorem 3.1 (Liouville 1844)**.** *Suppose x is a real algebraic number of degree n. Then there exists a constant $c(x) > 0$ such that*

$$\left| x - \frac{p}{q} \right| > \frac{c(x)}{q^n}$$

for every rational $\frac{p}{q}$ $(q > 0)$ distinct from x.

Corollary 3.1. *A real algebraic number of degree n is* ***not*** *approximable to order $n+1$ or higher.*

Proof. Exercise - indeed $n + 1$ can be replaced by $n + \epsilon$ ($\epsilon > 0$ is arbitrary). $\square$

EXAMPLE 3.3. Let

$$\beta = 10^{-1!} + 10^{-2!} + 10^{-3!} + \cdots$$
$$= 0.\underset{1!2!}{11}000\underset{3!}{1}0\ldots0\underset{4!}{1}0\ldots$$

Now let $N \geqslant 1$ be arbitrary. For $n > N$ let

$$\beta_n = \sum_{i=1}^{n} 10^{-i!} = \frac{p_n}{10^{n!}} = \frac{p_n}{q_n}.$$

Now examine

$$\begin{aligned}
0 < \beta - \frac{p_n}{q_n} &= \beta - \beta_n \\
&= 10^{-(n+1)!} + 10^{-(n+2)!} + \cdots \\
&= 10^{-(n+1)!}\left(1 + 10^{(n+1)!}10^{-(n+2)!} + 10^{(n+1)!}10^{-(n+3)!)} + \cdots\right) \\
&= 10^{-(n+1)!}\left(1 + \underbrace{\frac{1}{10^{(n+2)}}}_{<\frac{1}{10}} + \underbrace{\frac{1}{10^{(n+3)(n+2)}}}_{<\frac{1}{100}} + \underbrace{\frac{1}{10^{(n+4)(n+3)(n+2}}}_{<\frac{1}{1000}} + \cdots\right) \\
&< 10^{-(n+1)!}\underbrace{\sum_{i=0}^{\infty} 10^{-i}}_{\frac{1}{1-\frac{1}{10}}=\frac{10}{9}} < 2 \cdot 10^{-(n+1)!} = 2q_n^{-(n+1)} < 2q_n^{-N}
\end{aligned}$$

because $n > N$.

UPSHOT: For any $n > N$ there exists $\frac{p_n}{q_n} \in \mathbb{Q}$ such that

$$\left|\beta - \frac{p_n}{q_n}\right| < \frac{2}{q_n^N}.$$

Now Liouville implies β is **not** algebraic of degree less than N. But N is arbitrary, so β is not algebraic of any degree, which implies β is transcendental.

Remark. Clearly we could have replaced 10 by other integers to obtain other transcendental numbers. The general principle of the construction is simply that a number defined by a sufficiently rapid sequence of rational approximations is necessarily transcendental.

Definition: A real number ζ is said to be *Liouville number* if for every $m \in \mathbb{N}$ there exists $\frac{p_m}{q_m} \in \mathbb{Q}$ $(q_m > 1)$ such that

$$\left|\zeta - \frac{p_m}{q_m}\right| < q_m^{-m}.$$

In view of Liouville's Theorem

Corollary 3.2. *Any Liouville number is transcendental.*

Indeed the number

$$\beta = \sum_{i=1}^{\infty} 10^{-i!}$$

is a Liouville number.

Remark. Let $\mathcal{L}$ denote the set of Liouville numbers. It follows from definition that if $\zeta \in \mathcal{L}$ then there exists infinitely many rationals $\frac{p}{q}$ such that

$$\left|\zeta - \frac{p}{q}\right| < q^{-\tau}$$

where $\tau > 0$ is arbitrary. i.e. $\mathcal{L} \subset W(\tau)$ - see Chapter 1.

Recall: $W(2) = \mathbb{R}$ and it now follows that for $\tau > 2$ we have $W(\tau) \neq .\emptyset$.

If $\zeta \in \mathcal{L}$, then the rate of approximation in Dirichlet's Theorem (i.e. (denominator)$^{-2}$) can be improved by an arbitrary power. In some sense Liouville numbers are at the opposite end of the 'spectrum' to badly approximable numbers.

A real number x is said to be *very well approximable* if $w \in W(\tau)$ for some $\tau > 2$. Clearly, $\mathcal{L} \subset \mathcal{W}$ where $\mathcal{W}$ is the set of very well approximable numbers.

Theorem 3.2 (Mahler's Conjecture - 1984)**.** *Let K denote the middle third Cantor set.*

$$K := \left\{ x \in [0,1] : x \in \sum_{i=1}^{\infty} a_i 3^{-i},\ \textit{with } a_i \in [0,1] \right\}$$
$$\equiv \textit{set of points in } [0,1] \textit{ whose base 3 expansions are free of the digit 1.}$$

Clearly

$$\zeta = 2\sum_{n=1}^{\infty} 3^{-n!} \in K$$

and is Liouville, i.e. $\zeta \in K \cap 2$. Mahler's Conjecture states that there exists very well approximable numbers other than Liouville numbers in the middle third Cantor set. i.e.

$$(\mathcal{W} \setminus \mathcal{L}) \cap K \neq \emptyset.$$

This was solved in 2007 (Maths Annalen, vol 338) and we will return to the proof later.

Proof. (Liouville) Let ζ be a real, algebraic number of degree n, i.e. ζ is a root of

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

with $a_i \in \mathbb{Z}$ and $a_n \neq 0$.

Note. f is a real continuous function.

Fix an interval around ζ say $I_\lambda \equiv [\zeta-\lambda, \zeta+\lambda]$ (with $\lambda \in (0,1)$) such that I_λ contains no other root of f except ζ ($\lambda = \lambda(\zeta)$). Now there exists $c_1 > 0$ ($c_1(\lambda)$) such that

$$|f'(x)| \leqslant c_1 \quad \text{for all } x \in I_\lambda.$$

Let $c(\zeta) = \min\{\frac{1}{c_1}, \lambda\}$ and consider any rational $\frac{p}{q}$ ($q>0$) then

(1) If $\frac{p}{q} \notin I_\lambda$ then

$$\left|\zeta - \frac{p}{q}\right| \geqslant \lambda \geqslant c(\zeta) \geqslant \frac{c(\zeta)}{q^n}$$

$q \geqslant 1$, $n \geqslant 1$.

(2) If $\frac{p}{q} \in I_\lambda$ and $\frac{p}{q} \neq \zeta$ then

$$0 \neq \left|f\left(\frac{p}{q}\right)\right| = \frac{|a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_1 p q^{n-1} + a_0 q^n|}{q^n} \geqslant \frac{1}{q^n}$$

for $q>0$. By the Mean Value Theorem there exists $x \in (\zeta, \frac{p}{q})$ such that

$$f'(x) = \frac{f\left(\frac{p}{q}\right) - f(\zeta)}{\frac{p}{q} - \zeta}$$

i.e.

$$f\left(\frac{p}{q}\right) = \left(\frac{p}{q} - \zeta\right) f'(x).$$

Hence,

$$\left|\zeta - \frac{p}{q}\right| = \frac{\left|f\left(\frac{p}{q}\right)\right|}{|f'(x)|} \geqslant \frac{1}{c_1 q^n} \geqslant c(\zeta) q^n$$

□

3.1. **Roth's Theorem**

Let ζ be real algebraic number of degree $n \geqslant 2$. Then, Liouville implies that

$$\left|\zeta - \frac{p}{q}\right| < \frac{1}{q^\tau} \tag{3.1}$$

has only finitely many rational solutions if $\tau > n$.

Theorem 3.3 (Thue - 1909)**.** *He showed that* (3.1) *holds if* $\tau > \frac{1}{2}n + 1$

Theorem 3.4 (Siegel - 1921)**.** *He showed that* (3.1) *holds if* $\tau > 2\sqrt{n}$

Theorem 3.5 (Dyson - 1947, Gelfond - 1952)**.** *They showed independently that* (3.1) *holds if* $\tau > \sqrt{2n}$

Theorem 3.6 (K Roth - 1955)**.** *Roth finally showed that* (3.1) *holds if $\tau > 2$, for which he was awarded the fields medal in 1958.*

Roth's result is the best possible.

Note. Liouville's Theorem implies that if $x \in \mathcal{L}$ then x is transcendental. Roth's Theorem implies that if $x \in W(\tau)$ for $\tau > 2$ then x is transcendental.

4. Naturally Occurring Irrationals and Transcendental Numbers

Example 4.1. $\sqrt{2}$ (irrational and algebraic), e, π, $e+\pi$, $\sin\sqrt{2}$.

In Year 3, we proved that e is irrational - this was quite straight forward and made use of the fact that

$$e = \sum_{r=0}^{\infty} \frac{1}{r!}.$$

We need to be slightly more sophisticated to prove that π is irrational.

Lemma 4.1. *For $n \in \mathbb{N}$, let*

$$f(x) = \frac{x^n(1-x)^n}{n!}.$$

Then $f(x)$ and all its derivatives $f^{(m)}(x)$ take integer values at $x = 0$.

Remark. Some remarks about $f(x)$:

(1) Since, $f(1-x) = f(x)$, the conclusion remains true at $x = 1$.
(2) For $0 < x < 1$ we have $0 < f(x) < \frac{1}{n!}$

Proof. To be done as an exercise. □

Theorem 4.1. *π^2 is irrational, which trivially implies that π is irrational.*

Proof. Suppose that $\pi^2 = \frac{a}{b}$ for $a, b \in \mathbb{N}$. With f as in Lemma 4.1, write

$$G(x) = b^n\big(\pi^{2n} f(x) - \pi^{2n-2} f^{(2)}(x) + \pi^{2n-4} f^{(4)}(x) - \cdots + (-1)^n f^{(2n)}(x)\big).$$

In view of the Lemma, we have $G(0)$ and $G(1)$ are integers. Also note that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(G'(x)\sin(\pi x) - \pi G(x)\cos(\pi x)\right) &= G''(x)\sin(\pi x) + \cancel{\pi G'(x)\cos(\pi x)} \\ &\quad - \cancel{\pi G'(x)\cos(\pi x)} + \pi^2 G(x)\sin(\pi x) \\ &= \big(G''(x) + \pi^2 G(x)\big)\sin(\pi x) \\ &= b^n\big(\pi^{2n+2} f(x) + (-1)^n \underbrace{f^{(2n+2)}(x)}_{=0}\big)\sin(\pi x)\end{aligned}$$

because f is a polynomial with leading term x^{2n}.

$$\begin{aligned} &= b^n \pi^{(2n+2)} f(x)\sin(\pi x) \\ &= \pi^2 a^n f(x)\sin(x). \end{aligned}$$

Hence, integrating throughout we obtain

$$\begin{aligned} \pi \int_b^1 a^n f(x)\sin(\pi x)\ \mathrm{d}x &= \left[\frac{G'(x)\sin(\pi x)}{\pi} - G(x)\cos(\pi x)\right]_0^1 \\ &= G(1) + G(0). \end{aligned}$$

But, by our second remark of Lemma 4.1 we have that

$$0 < \pi \int_0^1 a^n \underbrace{f(x)}_{<\frac{1}{n!}} \underbrace{\sin(\pi x)}_{<1}\ \mathrm{d}x < \underbrace{\frac{\pi a^n}{n!}}_{\to 0} < 1$$

the last inequality holds for n large. A contradiction as there are no integers in $(0, 1)$. The proof can be adapted to show that e^y is irrational for every rational $y \neq 0$. □

Hermite in 1873 proved the e is transcendental and π was proved to be transcendental by Lindemann in 1882.

Theorem 4.2 (Lindemann-Weierstrauss, 1885)**.** *$\alpha_1, \dots, \alpha_n$ are algebraic and $\beta_1, \dots, \beta_n$ are algebraic distinct such that not all β_i are zero then*

$$\alpha_1 e^{\beta_1} + \cdots + \alpha_n e^{\beta_n} \neq 0.$$

Equivalently if $\alpha_1, \dots, \alpha_n$ are linearly independent over $\mathbb{Q}$ then for all polynomials $p \in \mathbb{Z}[X_1, \dots, X_n]$ such that $p(\alpha_1, \dots, \alpha_n) \neq 0$.

Corollary 4.1. *π is transcendental. We have that*

$$\begin{aligned} e^{i\pi} &= -1 \\ e^0 &= 1 \\ e^0 + e^{i\pi} &= 0. \end{aligned}$$

Assume that π is algebraic, which implies $i\pi$ is algebraic by the closure of the algebraic numbers. By the Lindermann-Weierstrauss theorem we have that $e^0 + e^{i\pi} \neq 0$ if π is algebraic $\Rightarrow$ contradiction.

From this we can get that $e^\alpha, \sin\alpha, \cos\alpha, \tan\alpha, \sinh\alpha, \cosh\alpha, \tanh\alpha$ are all transcendental for α algebraic. This is because they can all be written as a linear combination of e^α Also, $\log\alpha$ is transcendental for $\alpha \neq 1$ an algebraic number.

Hilbert's 7th Problem: If α is an algebraic number such that $\alpha \neq 0, 1$ and β is algebraic irrational then is α^β transcendental? For example,

$$2^{\sqrt{2}}, (\sqrt{2})^{\sqrt[3]{3}}, \dots, e^{\pi} \in (-1)^{-i}.$$

this was solved around 1934-1935 by Gelfond and Schneider.

5. Metric Diophantine Approximation - the Lebesgue Theory

$\psi : \mathbb{R}^+ \hookleftarrow$ is a real positive function such that $\psi(r) \to 0$ as $r \to \infty$. We say that R approximable functions.

$$W(\psi) = \left\{ x \in I \,\middle|\, \left| x - \frac{p}{q} \right| \leqslant \psi(q) \text{ for inf. many } (p,q) \in \mathbb{Z} \times \mathbb{N} \right\}$$

x is ψ-well approximable $\Rightarrow \{x\} \in W(\psi)$. We aim to determine the size of $W(\psi)$.

We say that *almost all numbers* in an interval (could be $\mathbb{R}$ in which case *almost all real numbers*) have a certain property P if the set of numbers lacking the property P has measure zero.

Example 5.1. Almost all real numbers are non-integers.

$W(\psi)$ is a lim sup set. Let E_i be a sequence of subsets in $\mathbb{R}^n$. The lim sup set denoted by E_∞ associated with this sequence is defined as follows

$$\begin{aligned} E_\infty &= \bigcap_{j=1}^{\infty} \left(\bigcup_{i=j}^{\infty} E_i \right) \\ &= \limsup_{i \to \infty} E_i. \end{aligned}$$

It simply corresponds to the set of points in $\mathbb{R}^n$ that lie in infinitely many of the set E_i ($i \in \mathbb{N}$). lim sup sets are fundamental objects in metric number theory. We now show that $W(\psi)$ is a lim sup set.

For fixed $q \in \mathbb{N}$, let

$$A_q(\psi) = \bigcup_{p=0}^{q} B\left(\frac{p}{q}, \psi(q)\right) \bigcap I$$

Note. We have that

$$m\big(A_q(\psi)\big) \leqslant 2q\psi(q)$$

and we have equality if $\psi(q) < \frac{1}{2q}$.

Then $W(\psi)$ is simply the set of numbers in I which lie in infinitely many sets $A_q(\psi)$ with $q = 1, 2, \ldots$ i.e.

$$W(\psi) = \bigcap_{m=1}^{\infty} \bigcup_{q=m}^{\infty} A_q(\psi)$$

a lim sup set. Now notice that

$$W(\psi) \subset \bigcup_{q=m}^{\infty} A_q(\psi)$$

i.e. the balls (or rather intervals) form a cover for $W(\psi)$. Thus,

$$\begin{aligned} m\big(W(\psi)\big) &\leqslant m\left(\bigcup_{q=m}^{\infty} A_q(\psi)\right) \\ &\leqslant \sum_{q=m}^{\infty} m\big(A_q(\psi)\big) \\ &\leqslant 2\sum_{q=m}^{\infty} q\psi(q) \qquad (\star) \end{aligned}$$

Now suppose

$$\sum_{q=1}^{\infty} q\psi(q) < \infty.$$

Then given $\epsilon > 0$, there exists m_0 such that for all $m \geqslant m_0$ and

$$\sum_{q=m}^{\infty} q\psi(q) < \frac{\epsilon}{2}.$$

It follows from $(\star)$, that $m \geqslant m_0$ and hence

$$m\big(W(\psi)\big) \leqslant \epsilon.$$

But $\epsilon > 0$ is arbitrary whence

$$m\big(W(\psi)\big) = 0$$

Theorem 5.1. *Let $\psi : \mathbb{R}^+ \hookleftarrow$ be a real positive function such that*

$$\sum q\psi(q) < \infty.$$

Then

$$m\big(W(\psi)\big) = 0.$$

This theorem is in fact a simple consequence of a general result in probability theory.

5.1. The convergence Borel-Cantelli Lemma

Let (Ω, μ) be a measure space with $\mu(\Omega) < \infty$ and let E_i $(i \in \mathbb{N})$ be a family of measurable sets in Ω such that

$$\sum_{i=1}^{\infty} \mu(E_i) < \infty.$$

Then

$$\mu(E_\infty) = 0$$

where '$E_\infty = \limsup_{i\to\infty} E_i$', i.e. 'the set of points in Ω which lie in infinitely many E_i $(i \in \mathbb{N})$.

Note. Theorem 1 is a trivial consequence. Let $\Omega = I = [0,1]$, $\mu = m$ and $E_q = A_q(\psi)$. Then, if $\sum m\big(A_q(\psi)\big) < \infty$ then $m\big(W(\psi)\big) = 0$.

Proof. Let $\epsilon > 0$, since $\sum \mu(E_i) < \infty$ then there exists N_0 such that for all $N \geqslant N_0$ we have

$$\mu\left(\bigcup_{i\geqslant N} E_i\right) \leqslant \sum_{i=N}^{\infty} \mu(E_i) < \epsilon.$$

But, for each $N \in \mathbb{N}$ we have

$$E_\infty \subset \bigcup_{i\geqslant N} E_i.$$

Thus

$$\mu(E_\infty) \leqslant \mu\left(\bigcup_{i\geqslant N} E_i\right) < \epsilon.$$

But $\epsilon > 0$ is arbitrary whence

$$\mu(E_\infty) = 0$$ □

A simple consequence of Theorem 1 is the following corollary.

Corollary 5.1. *For $\tau > 2$ we have $m\big(W(\tau)\big) = 0$. Recall that $W(\tau)$ is $W(\psi)$ with $\psi(q) = q^{-\tau}$. Then*

$$\sum q\psi(q) = \sum q^{1-\tau} < \infty$$

for $\tau > 2$.

Also, note that $m\big(W(\psi)\big) = 0$ when

$$\psi(r) = \frac{1}{q^2 \log q^{1+\epsilon}}$$

or

$$\psi(r) = \frac{1}{q^2 \log q (\log\log q)^{1+\epsilon}}.$$

$(\epsilon > 0)$.

In view of Theorem 1, what can we say about $m\big(W(\psi)\big)$ when $\sum q\psi(q) = \infty$? A fundamental thereom in metric number theory is

Theorem 5.2 (Khintchine's Theorem (1924))**.** *Let $\psi : \mathbb{R}^+ \longleftrightarrow$ be a* monotonic *function such that $\psi(r) \to 0$ as $r \to \infty$. Then*

$$m\big(W(\psi)\big) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q\psi(q) < \infty \\ 1 & \text{if } \sum_{q=1}^{\infty} q\psi(q) = \infty \end{cases}$$

Remark. We note that

(1) Khintchine's Thereom is an elegant 'zero-full' law.
(2) $1 = m(I)$
(3) $m\big(W(\psi)\big) = 1$ when

$$\psi(q) = \frac{1}{q^2 \log q} \qquad \text{or} \qquad \psi(q) = \frac{1}{q^2 \log\log q}$$

To prove the 'full measure' result the divergence of the measure sum isn't in general enough. Convergence of the Borel-Cantelli Lemma says that if

$$\sum \mu(E_\infty) < \infty$$

the $\mu(E_{\infty)=0}$. However it's not the case that if $\sum \mu(E_i) = \infty$ then $\mu(E_\infty) = \mu(\Omega)$ or indeed $\mu(E_\infty) > 0$.

Example 5.2. For $i \in \mathbb{N}$ let $E_i = [0, \frac{1}{i}]$, then we have

$$m(E_i) = \frac{1}{i} \Rightarrow \sum m(E_i) = \sum \frac{1}{i} = \infty$$

but clearly $E_\infty = \{0\}$, i.e. $m(E_\infty) = 0$

The problem here is that the sets E_i overlap 'too much' - in fact they are nested. We require that the sets E_i $(i \in \mathbb{N})$ are in 'some sense' independent. If we had independence, i.e. for $i \neq j$

(5.1) $$\mu(E_i \cap E_j) = \mu(E_i)\mu(E_j)$$

and $\sum \mu(E_i) = \infty$, then we could conclude that $\mu(E_\infty) = \mu(\Omega)$. In applications (5.1) rarely holds. More realistic is quasi-independence, i.e. for $i \neq j$ there exists a constant $C \geqslant 1$ such that

(5.2) $$\mu(E_i \cap E_j) \leqslant C\mu(E_i)\mu(E_j).$$

Then if $\sum \mu(E_i) = \infty$ we have

$$\mu(E_\infty) \geqslant \frac{1}{C} > 0.$$

So, by weakening independence to (5.1) we only get positive measure for E_∞ - this is no problem if we already know (by some other means) that $\mu(E_\infty)$ that

$$\mu(E_\infty) = 0 \qquad \text{or} \qquad \mu(\Omega).$$

The following result shows that we can weaken (5.2) and its this result that is often used to establish full measure statements.

5.2. **The Divergence Borel-Cantelli Lemma**

Let (Ω, μ) be a probability measure space (i.e. $\mu(\Omega) = 1$) and let E_i $(i \in \mathbb{N})$ be a family of measurable sets in Ω such that

$$\sum_{i=1}^{\infty} \mu(E_i) = \infty.$$

Then for all Q sufficiently large

$$\sum_{1 \leqslant i,j \leqslant Q} \mu(E_i \cap E_j) \leqslant C \left(\sum_{i=1}^{Q} \mu(E_i) \right)^2$$

'quasi-independence on average', where $C \geqslant 1$ is a constant.

The key to Khintchine's Theorem is establishing that the sets $A_q(\psi)$ are quasi-independent on average; it's here where the monotonicity assumption is required. This is NOT required for the convergence part.

Removing monotonicity from Khintchine's Theorem is a serious problem and is called the Duffin-Schaeffer conjecture. Recall, $W(\psi)$ is invariant under integer/rational translations. Khintchine's Theorem implies

Corollary 5.2. *Let $\psi : \mathbb{R}^+ \hookleftarrow$ be a monotonic function such that $\psi(r) \to 0$ as $r \to \infty$ and*

$$\sum q\psi(q) = \infty.$$

Then almost all real numbers are ψ-well approximable, i.e.

$$m\big(\{x \in \mathbb{R} \mid x \textit{ is not } \psi \textit{ approximable}\}\big) = 0.$$

Corollary 5.3. $m(\text{Bad}) = 0$

Proof. Take the function

$$\psi(q) = \frac{1}{q^2 \log q}$$

$$\text{and } B(\psi) = \left\{ x \in \mathbb{R} \mid \exists\, c(x) > 0 \text{ s.t. } \left| x - \frac{p}{q} \right| > c(x)\psi(q) \text{ for all } \frac{p}{q} (q > 0) \right\}.$$

By Corollary 2 we have $m\big(B(\psi)\big) = 0$ but Bad $\subset B(\psi)$ so $m(\text{Bad}) = 0$. □

In view of Corollary 5.1 we have that $m\big(W(\tau)\big) = 0$ for $\tau > 2$. We'd expect to the 'size' of $W(\tau)$ to decrease as τ increases (i.e. as 'rate of approximation' increases). For example we'd expect $W(300)$ to be 'smaller' than $W(3)$ - clearly $W(300) \subset W(3)$ but Lebesgue measure can't distinguish between these two sets - both are zero.

Upshot: require a 'finer' measurement of size that can distinguish between sets of Lebesgue measure zero.

6. Metric Diophantine Approximation: the Hausdorff Theory

Start by generalising Lebesgue measure and the notion of integer dimension.

6.1. Hausdorff Measure and Dimension

Let F be a non-empty subset in $\mathbb{R}^n$ (no harm in assuming F is uncountable)

!! DIAGRAM !!

We cover our set F by balls B_i of diameter $d_i \leqslant \rho$ (with $\rho > 0$). If $\{B_i\}$ is a countable collection of balls B_i with diameter $d_i \leqslant \rho$ that cover F, i.e.

$$F \subset \bigcup_i B_i$$

then $\{B_i\}$ is said to be a *ρ-cover* of F. For $s \geqslant 0$ and $\rho > 0$ let

$$\mathcal{H}^s_\rho(F) = \inf\left\{ \sum_i d_i^s \;\middle|\; \{B_i\} \text{ is a } \rho\text{-cover of } F \right\} \tag{6.1}$$

As ρ decreases the class of allowed covers of F in (6.1) is reduced. Therefore $\mathcal{H}^s_\rho(F)$ increases as ρ decreases and so approaches a limit as $\rho \to 0$. Then

$$\mathcal{H}^s = \lim_{\rho \to 0} \mathcal{H}^s_\rho(F), \tag{6.2}$$

which exists for any subset F of $\mathbb{R}^n$; can be 0, positive, finite or ∞. The quantity $\mathcal{H}^s(F)$ is the *s-dimensional Hausdorff measure* of the set F. When $s = n$ – an integer then $\mathcal{H}^n$ is comparable to n-dimensional Lebesgue measure m_n, i.e. for $F \subseteq \mathbb{R}^n$

$$c_1 m_n(F) \leqslant \mathcal{F}^n \leqslant c_2 m_n(F). \tag{6.3}$$

where c_1, c_2 are constants (greater than 0) depending only on n.

Some Analysis of (6.3)

Let $\rho < 1$. If $t > s$ and $\{B_i\}$ is a ρ-cover of F then

$$\sum_i d_i^t = \sum_i d_i^s d_i^{t-s} < \rho^{t-s} \sum d_i^s$$

because $d_i \leqslant \rho$. So, only taking infimum we have

$$\mathcal{H}^t_\rho(F) \leqslant \rho^{t-s} \mathcal{H}^s_\rho(F).$$

On letting $\rho \to 0$, we see that if $\mathcal{H}^s(F) < \infty$ then $\mathcal{H}^t(F) = 0$ for $t > s$. Furthermore for $s = 0$ and F uncountable we have that $\mathcal{H}^0(F)$ is infinite. Hence, there exists a critical value of s at which $\mathcal{H}^s(F)$ jumps from ∞ to 0 as s increases.

This critical value of s is called the *Hausdorff dimension* of F – written simply as $\dim F$.

Remarks

- By definition of $\dim F$ we have that

$$\mathcal{H}^s(F) = \begin{cases} 0 & \text{if } s > \dim F \\ \infty & \text{if } s < \dim F. \end{cases}$$

- If $s = \dim F$, then $\mathcal{H}^s(F)$ may be zero or infinite or may satisfy

$$0 < \mathcal{H}^s(F) < \infty$$

  F is said to be an s-set.

- Let $I = [0,1]$. Then $\dim I = 1$ and

$$\mathcal{H}^s(I) = \begin{cases} 0 & \text{if } s > 1 \\ 1 & \text{if } s = 1 \\ \infty & \text{if } s < 1 \end{cases}$$

  I is an s-set.

To calculate $\dim F$ (say $\dim F = \alpha$), its usually the case that we establish upper and lower bounds separately.

$$\begin{aligned} &\dim F \leqslant \alpha && \text{(upper bound)} \\ &\dim F \geqslant \alpha && \text{(lower bound).} \end{aligned}$$

Upper bounds for $\dim F$ are usually easy – we can exploit a 'natural' cover of F.

Example 6.1. Let K be the standard middle Cantor set. Then

$$\dim K \leqslant \frac{\log 2}{\log 3}.$$

Recall $K = \cap_{n=0}^{\infty} K_n$ where K_n is a union of 2^n balls (intervals) of diameter 3^{-n}. Thus, for each $n \in \mathbb{N}$ we have $E_n \supset K$ and so the intervals in E_n form a $\rho = 3^{-n}$ cover of K. It follows that

$$\mathcal{H}^s_\rho(K) \leqslant 2^n (3^{-n})^s = \left(\frac{2}{3^s}\right)^n \to 0$$

as $n \to \infty$ (i.e. $\rho \to 0$) if

$$\frac{2}{3^s} < 1 \Rightarrow s > \frac{\log 2}{\log 3}.$$

In other words

$$\mathcal{H}^s(K) = 0 \text{ if } s > \frac{\log 2}{\log 3}.$$

It follows from the definition of Hausdorff dimension

$$\dim K = \inf\{s \mid \mathcal{H}^s(K) = 0\}$$

that $\dim K \leqslant \frac{\log 2}{\log 3}$. In fact, $\dim K = \frac{\log 2}{\log 3}$ and moreover

$$\mathcal{H}^{\frac{\log 2}{\log 3}}(K) = 1.$$

To prove

$$\dim F \geqslant \frac{\log 2}{\log 3}$$

we need to work with arbitrary covers of K. Let $\{B_i\}$ be an arbitrary ρ-cover with $\rho < 1$. K - is bounded and closed (intersection of closed intervals), i.e. K is compact. Hence without loss of generality we can assume that $\{B_i\}$ is finite. For each B_i, let k be the integer such that

$$3^{-(k+1)} \leqslant d_i < 3^{-k}. \tag{6.4}$$

Then the B_i intersects at most one interval of E_k – separated by at least 3^{-k}. If $j \geqslant k$ then B_i intersects at most

$$2^{j-k} = 2^j 3^{-sk} \leqslant 2^j 3^s d_i^s \tag{6.5}$$

intervals of E_j, where $s = \frac{\log 2}{\log 3}$ and the final inequality holds from (6.4). Now j large enough such that

$$3^{-(j+1)} \leqslant d_i \text{ for all } B_i \in \{B_i\}$$

can be done since the number of $\{B_i\} < \infty$. Then $j \geqslant k$ for each B_i and (6.5) invalid. Furthermore $\{B_i\}$ must intersect all 2^j intervals of E_j – $\{B_i\}$ is a cover for k, i.e.

$$\begin{aligned} 2^j &= \#\{I \in E_j \mid \cup B_i \cap I \neq \emptyset\} \\ &\leqslant \sum_i \#\{I \in E_j \mid B_i \cap I \neq \emptyset\} \\ &\leqslant \sum_i 2^j 3^s d_i^s \end{aligned}$$

by (6.5), i.e. for any arbitrary cover $\{B_i\}$ we have that for $s = \frac{\log 2}{\log 3}$

$$\sum d_i^s \geqslant 3^{-s} = \frac{1}{2}$$

which implies $\mathcal{H}^s(K) \geqslant \frac{1}{2} \Rightarrow \dim K \geqslant \frac{\log 2}{\log 3}$

Even in the simple Cantor set example, the lower bound for $\dim K$ is much more involved than the upper bound – which is easy. This is usually the case and the number theoretic sets $W(\psi)$ and $W(\tau)$ are no exception.

6.2. The Jarnik-Besicovitch Theorem

Recall, that $W(\psi) = \limsup_{q\to\infty} A_q(\psi)$ where

$$A_q(\psi) = \bigcup_{p=0}^{q} B\left(\frac{p}{q}, \psi(q)\right) \bigcap I.$$

Also, for each $m \in \mathbb{N}$

$$W(\psi) \subset \bigcup_{q=m}^{\infty} A_q(\psi)$$

and the balls in here form a cover for $W(\psi)$. $\psi : \mathbb{R}^+ \hookleftarrow$ is a real, positive function such that $\psi(r) \to 0$ as $r \to \infty$. Suppose for the moment that ψ *is monotonic*. Now fix $m \in \mathbb{N}$. Now fix $m \in \mathbb{N}$ and let $\rho = \rho(m) = 2\psi(m)$. Then the balls in $\{A_q(\psi)\}_{q \geqslant m}$ form a ρ cover of $W(\psi)$. Thus,

$$\mathcal{H}^s_\rho\big(W(\psi)\big) \leqslant \sum_{q=m}^{\infty} q\big(2\psi(q)\big)^s$$

The right hand side $\to 0$ as $m \to \infty$ (i.e. $\rho = 2\psi(m) \to 0$) if

$$\sum_{q=1}^{\infty} q\psi^s(q) < \infty$$

i.e. $\mathcal{H}^s\big(W(\psi)\big) = 0$ if the s-volume sum converges. Actually, monotonicity on ψ can be removed (exercise) and we have proved the following theorem.

Theorem 6.1. *Let $\psi : \mathbb{R}^+ \hookleftarrow$ be a function such that $\psi(r) \to 0$ as $r \to \infty$. Then*

$$\mathcal{H}^s\big(W(\psi)\big) = 0 \text{ if } \sum_{q=1}^{\infty} q\psi^s(q) < \infty.$$

When $s = 1$ then $\mathcal{H}^1 \equiv m$ – one dimensional Lebesgue measure, and then the theorem is just the convergent part of Khintchines theorem (Theorem 5.2).

Corollary 6.1.

$$\dim W(\psi) \leqslant \inf\left\{ s \;\middle|\; \sum_{q=1}^{\infty} q\psi^s(q) < \infty \right\}.$$

which just follows from Theorem 6.1 and the definition of Hausdorff dimension.

Corollary 6.2. *For $\tau \geqslant 2$ we have* $\dim W \leqslant \frac{2}{\tau}$.

Proof. We have that

$$\sum_{q=1}^{\infty} q\psi^s(q) = \sum_{q=1}^{\infty} q \cdot q^{-\tau s} < \infty$$

because $\psi(q) = q^{-\tau}$ and $s > \frac{2}{\tau}$ □

The divergence result (Theorem 6.1) and upper bound dimension result (Corollary 6.2) just exploit the natural cover of the lim sup set! The substance of the next result is the lower bound.

Theorem 6.2 (The Janik-Besicovitch Theorem)**.** *For $\tau \geqslant 2$ we have* $\dim\big(W(\tau)\big) = \frac{2}{\tau}$.

From this we see that $W(300)$ is smaller than $W(3)$. Excellent result but gives no information about $\mathcal{H}^s$ as $s = \dim\big(W(\tau)\big)$ – the critical exponent.

!! DIAGRAM !!

Does there exist a Hausdorff measure version of Khintchines Theorem (a Lebesgue measure statement)? Yes!

Theorem 6.3 (Jarnik's Theorem)**.** *Let* $s \in (0,1)$ *and* $\psi : \mathbb{R}^+ \hookleftarrow$ *be a* monotonic *function such that* $\psi(r) \to 0$ *as* $r \to \infty$*. Then*

$$\mathcal{H}^s\big(W(\psi)\big) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q\psi^s(q) < \infty \\ \infty & \text{if } \sum_{q=1}^{\infty} q\psi^s(q) = \infty \end{cases}$$

An Important Remark

In Jarnik's original statement various technical conditions on ψ and indirectly s were imposed, which prevented $s = 1$ – note even as stated the theorem is false when $s = 1$. This is because $\mathcal{H}^s\big(W(\psi)\big) \equiv m\big(W(\psi)\big) = 1$, which is not equal to infinity in the divergence case. However, these technical conditions have recently been removed (2006) and gives the clear cut statement. Moreover it allows us to combine Khintchine and Jarnik to obtain a unifying statement

Theorem 6.4 (Khintchine-Jarnik $\sim$ 2006)**.** *Let* $s \in (0,1]$ *and* $\psi : \mathbb{R}^+ \hookleftarrow$ *be a* monotonic *function such that* $\psi(r) \to 0$ *as* $r \to \infty$*. Then*

$$\mathcal{H}^s\big(W(\psi)\big) = \begin{cases} 0 & \text{if } \sum_{q=1}^{\infty} q\psi^s(q) < \infty \\ \mathcal{H}^s(I) & \text{if } \sum_{q=1}^{\infty} q\psi^s(q) = \infty. \end{cases}$$

Corollary 6.3. *For* $\tau > 2$

$$\mathcal{H}^{\frac{2}{\tau}}\big(W(\tau)\big) = \infty.$$

7. A Mass Transference Principle

The goal is to show that the Hausdorff theory for $W(\psi)$ is in fact a consequence of the Lebesgue theory. In particular – Dirichlet's theorem implies the Jarnik-Besicovitch theorem. Khintchine's Theorem (Lebesgue Statement) implies Jarnik's Theorem (Hausdorff statement). Start by giving the heuristics.

(1) Dirichlet $\Rightarrow$ Jarnik-Besicovitch. We know that

$$m\big(W(2)\big) = \mathcal{H}^1\big(W(2)\big) = \mathcal{H}^1(I).$$

We want to renormalize $\mathcal{H}^1 \to \mathcal{H}^{\frac{2}{\tau}}$ preserves $\mathcal{H}^1$ and ball $B \Rightarrow \dim W(\tau) \geqslant \frac{2}{\tau}$ ($\tau \geqslant 2$) - the hard part of the Jarnik-Besicovitch theorem (note that $\mathcal{H}^{\frac{2}{\tau}}(I) > 0$ and $\mathcal{H}^{\frac{2}{\tau}}(I) = \infty$ if $\tau > 2$).

(2) Khintchine $\to$ Jarnik. This is the divergent part of each theorem in the main subtance. So in Jarnik we're given

$$\sum q\psi^s(q) = \infty \qquad s \in (0,1).$$

Want to show $\mathcal{H}\big(W(\psi)\big) = \infty$. We have that

$$\sum q\psi^s(q) = \infty \Rightarrow \mathcal{H}^1\big(W(\psi)\big) = \mathcal{H}^1(I).$$

We again renormalize $\mathcal{H}^1 \to \mathcal{H}^s$, which preserves $\mathcal{H}^1$ and ball B. Then we obtain a statement such as

$$\mathcal{H}^s\big(W(\psi)\big) = \mathcal{H}^s(I)$$

then for $s < 1$ we have $\mathcal{H}^s(I) = \infty$

7.1. **The Mass Transference Principle (a simplified version)**

Let Ω be a compact subset of $\mathbb{R}^n$ such that

$$B^\delta(B) \asymp r(B)^\delta \qquad (\delta > 0) \tag{7.1}$$

for any ball $B = B(x,r)$ in Ω with $x \in \Omega$ and $r \leqslant r_0$. A consequence of (7.1) in that

$$0 < \mathcal{H}^\delta(\Omega) < \infty$$

and $\dim \Omega = \delta$ – we return to this later. Given a ball B and $s \geqslant 0$, let

$$B^s = B\big(x, r^{\frac{s}{\delta}}\big)$$

and so $B^\delta = B$.

Theorem 7.1 (Mass Transfer Principle (2006))**.** *Let $\{B_i\}_{i\in\mathbb{N}}$ be a sequence of balls in Ω with radii $r(B_i) \to 0$ as $i \to \infty$. If $\mathcal{H}^\delta(\limsup B_i^s) = \mathcal{H}^\delta(\Omega)$ then*

$$\mathcal{H}^s(\limsup B_i^\delta) = \mathcal{H}^s(\Omega)$$

Corollary 7.1. *For $\tau > 2$, $\dim W(\tau) \geqslant \frac{2}{\tau}$.*

Proof. Dirichlet and the Mass Transfer Principle (with $\Omega = [0,1]$, $\delta = 1$, $s = \frac{2}{\tau}$ and $B_{\frac{p}{q}} = B(\frac{p}{q}, q^{-\tau})$ in $[0,1]$) $\square$

Corollary 7.2. *Let $s \in (0,1)$ and $\psi : \mathbb{R}^+ \hookleftarrow$ be a monotonic function such that $\psi(r) \to 0$ as $r \to \infty$. Then*

$$\mathcal{H}^s\big(W(\psi)\big) = 0 \qquad \textit{if} \qquad \sum q\psi^s(q) = \infty.$$

Proof. Khintchine and the Mass Transfer Principle (with $\Omega = [0,1]$, $\delta = 1$, $B_{\frac{p}{q}} = B(\frac{p}{q}, \psi(q))$ $\square$

8. A problem of Kurt Mahler

Recall

Remember, from chapter 3, that a real number is very well approximable if $x \in W(\tau)$ for $\tau > 2$. If $\mathcal{W}$ is the set of very well approximable numbers then

$$m(\mathcal{W}) = 0.$$

Also $\mathcal{L} \subset \mathcal{W}$, the set of Liouville numbers, which have $\dim \mathcal{L} = 0$. We regard K to represent the middle third Cantor set. Clearly

$$\zeta = 2\sum_{n=1}^{\infty} 3^{-n!} \in K \cap \mathcal{L}.$$

Mahler's conjecture

Mahler conjectured that

$$(\mathcal{W} \setminus \mathcal{L}) \cap K \neq \emptyset,$$

i.e. the Liouville numbers are *not* the only well approximable numbers in K.

8.1. Premliminaries - The Cantor set and measure

Rationals of the form $\frac{p}{3^n}$ are special!! – $2 \cdot 2^n$ of all possible rationals $\frac{p}{3^n}$ in I lie in K. Let μ be the natural measure on K – each of the 2^n intervals I_n of length 3^{-n} in E_n has mass 2^{-n}, i.e.

$$\mu(I_n) = 2^{-n} = (3^{-n})^{\gamma}$$

where $\gamma = \frac{\log 2}{\log 3}$. Assume $x \in K$ then $\mu(B) \geqslant \mu(I_{k+1}) \geqslant 3^{-\gamma} r^{\gamma}$. Clearly,

$$\mu(E_n) = \mu\left(\bigcup_{I_n \in \xi_n} I_n\right) = 2^n \cdot 2^{-n} = 1.$$

μ is supported on K and $\mu(K) = 1$. Furthermore, for any subset A of $[0,1]$ we have $A \subset \cup_i I_i$ and $I_i \in e$ where ξ is the collection of all basic intervals in Cantor construction We want to know

$$\mu(A) = \inf\left\{\sum_i \mu(I_i) : A \subset U_i I_i \text{ and } I_i \in \xi\right\}.$$

Let $B = B(x, r)$ be a ball (interval) of radius $r < 1$ and let k be the integer such that $3^{-(k+1)} \leqslant r < 3^{-k}$. Then B intersects at most one interval in ξ_k

!! DIAGRAM !!

Then

$$\mu\big(B(x,r)\big) \leqslant \mu(I_k) = 2^{-k} = (3^{-k})^{\gamma} = 3^{\gamma}(3^{-(k+1)})^{\gamma} \leqslant 3^{\gamma} r^{\gamma}$$

i.e.

$$\mu\big(B(x,r)\big) \leqslant 3^{\gamma} r^{\gamma} \tag{8.1}$$

for arbitrary balls. Now suppose that $x \in K$, then B contains at least one interval in ξ_{k+1} and so

$$\mu\big(B(x,r)\big) \geqslant \mu(I_{k+1}) = 2^{-(k+1)} = (3^{-(k+1)})^{\gamma} \geqslant 3^{-\gamma} r^{\gamma}$$

i.e.

$$3^{-\gamma} r^{\gamma} \leqslant \mu\big(B(x,r)\big) \leqslant 3^{\gamma} r^{\gamma}. \tag{8.2}$$

Indeed (8.1) is all that is required to show that

$$\dim K \geqslant \gamma = \frac{\log 2}{\log 3}.$$

Mass Distribution Principle

Let μ be a measure supported on a subset F of $\mathbb{R}^n$ and suppose for some γ there exists constants $c > 0$ and $\delta > 0$ such that

$$\mu(B) \leqslant c d^{\delta}$$

for all balls B with diameter $d \leqslant \delta$. Then

$$\mathcal{H}^{\gamma}(F) \geqslant \frac{\mu(F)}{c}.$$

In particular, $\dim F \geqslant \gamma$.

Proof. If $\{B_i\}$ is a ρ-cover of F with $\rho < \delta_1$ then

$$\begin{aligned} 0 < \mu(F) = \mu\left(\bigcup_i B_i\right) \\ \leqslant \sum_i \mu(B_i) \\ \leqslant c \sum_i d_i^{\gamma}. \end{aligned}$$

Taking infimum, $\mathcal{H}^{\gamma}_{\rho}(F) \geqslant \frac{\mu(F)}{c}$. Thus $\mathcal{H}^{\gamma}(F) \geqslant \frac{\mu(F)}{c}$ because $\mathcal{H}^{\gamma}_{\rho}$ increases as $\rho \to 0$. □

The Strategy towards Mahler

Rationals with denominator powers of 3 are special for K; let

$$\mathcal{A} = \{3^n \mid n \in 0, 1, 2, \dots\}$$

and consider the set

$$W_{\mathcal{A}}(\psi) = \left\{ x \in I \,\middle|\, \left| x - \frac{p}{q} \right| < \psi(q) \text{for inf. many } (p,q) \in \mathbb{Z} \times \mathcal{A} \right\}.$$

This is the set of 'ψ-well approximable numbers with denominators of rational approximation restrcited to $\mathcal{A}$'.

Here $\psi : \mathbb{R}^+ \hookleftarrow$ is a *monotonic* function such that $\psi(r) \to 0$ as $r \to \infty$ and without loss of generality assume

$$\psi(r) < \frac{1}{2r}$$

otherwise $W_{\mathcal{A}}(\psi) = I$. Now consider the set $W_{\mathcal{A}}(\psi) \cap K$ and for $n \in \mathbb{N}$ let

$$\begin{aligned} A_n(\psi) &= \bigcup_{0 \leqslant p \leqslant 3^n} B\left(\frac{p}{3^n}, \psi(3^n)\right) \cap K \\ &= \bigcup_{0 \leqslant p \leqslant 3^n} B\left(\frac{p}{3^n}, \psi(3^n)\right) \cap K \end{aligned}$$

with $\frac{p}{3^n} \in K$. By definition,

$$W_{\mathcal{A}}(\psi) \cap K = \limsup_{n \to \infty} A_n(\psi).$$

Now, with $\mu =$ to the Cantor measure, we have

$$\mu\big(A_n(\psi)\big) \leqslant 2 \cdot 2^n \cdot 3^\gamma \psi(3^n)^\gamma$$

by †. Indeed since the balls are centred at points of K we have that

$$\mu(A_n(\psi)) \asymp 2^n \psi(3^n)^\gamma = \big(3^n \psi(3^n)\big)^\gamma$$

by †† with $\gamma \equiv \frac{\log 2}{\log 3}$. In view of Borel-Cantelli 'convergence' Lemma (Chapter 5) we have that

$$\mu\big(W_{\mathcal{A}}(\psi) \cap K\big) = 0$$

if $\sum_{n=1}^{\infty} \big(3^n \psi(3^n)\big)^\gamma \leqslant \infty$. The following is a Khintchine type theorem for $W_{\mathcal{A}}(\psi) \cap K$ with respect to μ.

Theorem 8.1. *Let* $\psi : \mathbb{R}^+ \hookleftarrow$ *be a* monotonic *function with* $\psi(r) \to \infty$*as* $r \to 0$*. Then*

$$\mu\big(W_{\mathcal{A}}(\psi) \cap K\big) = \begin{cases} 0 & \text{if } \sum \big(3^n \psi(3^n)\big)^\gamma < \infty \\ 1 = \mu(K) & \text{if } \sum \big(3^n \psi(3^n)\big) = \infty. \end{cases}$$

Essentially this is exercise 19 on the problem sheet.

Now $\mu \equiv \mathcal{H}^\gamma$, which is γ-dimensional Hausdorff measure. Indeed $\mathcal{H}^\gamma(K) = 1 = \mu(K)$. So in view of †† we have $\mathcal{H}^\gamma\big(B(x,r)\big) \asymp r^\gamma$ for $x \in K$. Thus, we are in the position to apply the Mass Transference Principle (Chapter 7) with $\Omega = K$, $\delta = \gamma$ and $B_{\frac{p}{3^n} \in K} = B\left(\frac{p}{3^n}, \psi(3^n)\right)$ - this yields the divergent part of the following (the convergent part is easy).

Theorem 8.2. *Let* $\psi : \mathbb{R}^+ \hookleftarrow$ *be a monotonic function with* $\psi(r) \to 0$ *on* $r \to \infty$ *and* $s \in (0, \gamma)$

$$\mathcal{H}^s\big(W_{\mathcal{A}}(\psi) \cap K\big) = \begin{cases} 0 & \text{if } \sum 3^{n\gamma} \psi(3^n)^s < \infty \\ 1 = \mu(K) & \text{if } \sum 3^{n\gamma} \psi(3^n)^s = \infty. \end{cases}$$

(note $\mathcal{H}^s(K) = \infty$ *if* $s < \gamma$*).*

Theorem 2 'strongly' implies Mahler. With $\psi(r) = r^{-\tau}$ ($\tau > 0$) let's write $W_{\mathcal{A}}(\tau)$ for $W_{\mathcal{A}}(\psi)$. Then theorem 2 implies

$$\dim\big(W_{\mathcal{A}} \cap K\big) = \inf\left\{ s \;\middle|\; \sum 3^{n\tau} 3^{-n\tau s} < \infty \right\} = \sum \big(3^{-\tau s + \gamma}\big)^n$$

Corollary 8.1. *For* $\tau \geqslant 1$ *we have*

$$\dim\big(W_{\mathcal{A}}(\tau) \cap K\big) = \frac{\gamma}{\tau}$$

in particular, for $\tau > 1$

$$\mathcal{H}^{\frac{\gamma}{\tau}}\big(W_{\mathcal{A}}(\tau) \cap K\big) = \infty.$$

For $\tau > 2$, every point in $W_{\mathcal{A}}(\tau)$ is by definition very well approximable, i.e.

$$W_{\mathcal{A}}(\tau) \subset \mathcal{W} \cap K.$$

Thus, for any $\tau > 2$

$$\dim(\mathcal{W} \cap K) \geqslant \dim\big(W_{\mathcal{A}}(\tau) \cap K\big) = \frac{\gamma}{\tau}.$$

We can make τ arbitrarily close to 2, so

$$\dim(\mathcal{W} \cap K) \geqslant \frac{\gamma}{2}.$$

Now, $\dim \mathcal{L} = 0$; then

$$\dim\big((\mathcal{W} \setminus \mathcal{L} \cap K\big) \geqslant \frac{\gamma}{2}$$