Dyadic approximation in the Cantor set

Sam Chow

University of Warwick

Joint with Demi Allen (Bristol)
and Han Yu (Cambridge)

Diophantine approximation

By Dirichlet's theorem, or by continued fractions, if $\alpha \in \mathbb{R}$ then

$$\left|\alpha - \frac{a}{n}\right| < \frac{1}{n^2}$$

has infinitely many solutions $(a, n) \in \mathbb{Z} \times \mathbb{N}$. More succinctly

$$\|n\alpha\|_{\mathbb{R}/\mathbb{Z}} < n^{-1} \qquad \text{i.o.}$$

By Hurwitz's theorem, if $\alpha \in \mathbb{R}$ then

$$\|n\alpha\| < \frac{1}{\sqrt{5}n} \qquad \text{i.o.,}$$

and in the case $\alpha = \frac{1+\sqrt{5}}{2}$ the factor $\frac{1}{\sqrt{5}}$ cannot be sharpened.

Metric diophantine approximation

Given $\psi : \mathbb{N} \to [0, \infty)$, does a typical real number α have infinitely many rational approximations a/n such that $\left|\alpha - \frac{a}{n}\right| < \frac{\psi(n)}{n}$?

Theorem 1 (Khintchine 1924)

If $\psi : \mathbb{N} \to [0, \infty)$ is decreasing[a] *then the Lebesgue measure of*

$$W(\psi) = \{\alpha \in [0,1] : \|n\alpha\| < \psi(n) \textit{ i.o.}\}$$

is

$$\begin{cases} 0, & \textit{if } \sum_n \psi(n) < \infty \\ 1, & \textit{if } \sum_n \psi(n) = \infty. \end{cases}$$

[a] The monotonicity condition in the original paper was slightly different.

Borel–Cantelli lemmas

Khintchine's theorem is explained by the Borel–Cantelli lemmas in probability theory. Observe that

$$W(\psi) = \limsup_{n \to \infty} A_n, \qquad A_n = \bigcup_{a=0}^{n} \left(\frac{a - \psi(n)}{n}, \frac{a + \psi(n)}{n} \right) \cap [0,1].$$

For any probability measure ν on $[0,1]$, the Borel–Cantelli lemmas:

I. If $\sum_{n=1}^{\infty} \nu(A_n) < \infty$ then $\nu(W(\psi)) = 0$.

II. If $\sum_{n=1}^{\infty} \nu(A_n) = \infty$ and

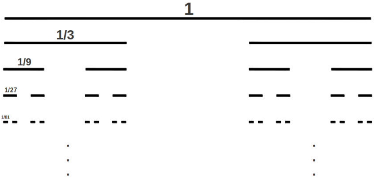
$$\sum_{n,m \leqslant N} \nu(A_n \cap A_m) \ll \left(\sum_{n \leqslant N} \nu(A_n) \right)^2,$$

then $\nu(W(\psi)) > 0$.

The latter can be upgraded to full measure if we can establish these estimates uniformly with $\nu_{\mathcal{I}}$ in place of ν, where $\mathcal{I}$ is an arbitrary subinterval of $[0,1]$.

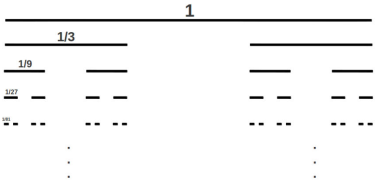
A question of Mahler (1984)

"At the age of 80 I cannot expect to do much more mathematics. I may, however, state a number of questions where perhaps further research might lead to interesting results... How close can irrational elements of Cantor's set be approximated by rational numbers?"



The middle-third Cantor set, K, contains infinitely many rational numbers, since it comprises numbers that can be written using only $0, 2$ in ternary. Its Hausdorff dimension is $\gamma := \frac{\log 2}{\log 3}$.

The Cantor measure



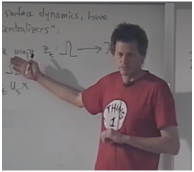
The *Cantor measure*, μ, is a probability measure that assigns a weight of 2^{-N} to each level N subinterval (of length 3^{-N}) in the construction of the Cantor set K. If E is Borel then $\mu(E)$ is the Hausdorff γ-measure of $E \cap K$. All we need is that μ is supported on K and Ahlfors–David regular:

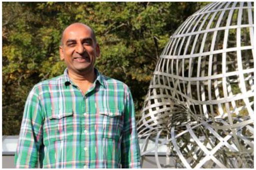
$$\mu(B(z,r)) \asymp r^\gamma \qquad (z \in K, \, 0 < r < 1).$$

Previous work

Theorem 2 (Weiss 2001)

If $\varepsilon > 0$ *then* $\mu(\{\alpha : \|n\alpha\| < n^{-1-\varepsilon} \text{ i.o.}\}) = 0.$





Theorem 3 (Levesley–Salp–Velani 2007)

If $\psi : \mathbb{N} \to [0, \infty)$ *then*

$$\mu(\{\alpha : \|3^n\alpha\| < \psi(3^n) \text{ i.o.}\}) = \begin{cases} 0, & \text{if } \sum_n \psi(3^n)^\gamma < \infty \\ 1, & \text{if } \sum_n \psi(3^n)^\gamma = \infty. \end{cases}$$

Why triadic rationals are easier

The idea in Levesley–Salp–Velani is to estimate[1]

$$\mu\left(\bigcup_{a=0}^{3^n} B\left(\frac{a}{3^n}, \frac{\psi(3^n)}{3^n}\right)\right) = \sum_{a=0}^{3^n} \mu\left(B\left(\frac{a}{3^n}, \frac{\psi(3^n)}{3^n}\right)\right),$$

and then apply the Borel–Cantelli lemmas. Only the balls centred at triadic rationals **in** the Cantor set have measure; the others avoid the Cantor set entirely. Thus, the measure has order of magnitude

$$2^{n+1}(\psi(3^n)/3^n)^\gamma \asymp \psi(3^n)^\gamma.$$

The first Borel–Cantelli lemma now yields the convergence statement. For the divergence statement, they also needed to show that the twofold intersections of the events $\|3^n\alpha\| < \psi(3^n)$ are quasi-independent on average.

[1] Let's assume, as we may, that $\psi(3^n) < 1/2$.

A conjecture of Velani

Conjecture (Velani)

If $\psi : \mathbb{N} \to [0, \infty)$ then

$$\mu(\{\alpha : \|2^n\alpha\| < \psi(2^n) \text{ i.o.}\}) = \begin{cases} 0, & \text{if } \sum_n \psi(2^n) < \infty \\ 1, & \text{if } \sum_n \psi(2^n) = \infty. \end{cases}$$

Let's try to guess what the measure of

$$A_n = \{\alpha : \|2^n\alpha\| < \psi(2^n)\} = \bigcup_a B\left(\frac{a}{2^n}, \frac{\psi(2^n)}{2^n}\right)$$

is. Roughly speaking, if K_N is the N^{th} Cantor level and

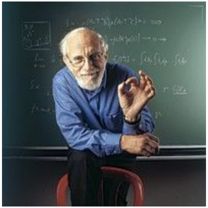
$$\mathrm{dist}\left(\frac{a}{2^n}, K_N\right) < 3^{-N} \asymp \frac{\psi(2^n)}{2^n},$$

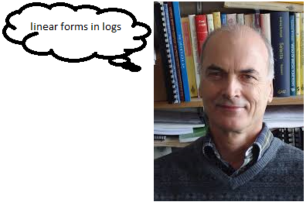
then we get measure $(\psi(2^n)/2^n)^\gamma$ around $a/2^n$. This should have probability $(2/3)^N$, so $\mu(A_n) \approx 2^n(2/3)^N(\psi(2^n)/2^n)^\gamma \asymp \psi(2^n)$?

The x2 x3 phenomenon

Theorem 4 (Furstenberg 1967)

If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ *then* $\overline{\{2^m 3^n \alpha : m, n \in \mathbb{N}\}} = \mathbb{R}/\mathbb{Z}$.





Theorem 5 (Stewart 1980)

With $D_b(\cdot)$ *counting digit changes in base* b, *we have*

$$D_2(n) + D_3(n) \geqslant \frac{(1 - o(1)) \log \log n}{\log \log \log n} \qquad (n \to \infty).$$

Results

Theorem 6 (Allen–C.–Yu 2020, convergence theory)

If

$$\sum_n \psi(2^n), \sum_n 2^{-\frac{\log n}{\log\log n \cdot \log\log\log n}} \psi(2^n)^\gamma < \infty$$

then $\mu(W_2(\psi)) = 0$, *where* $W_2(\psi) = \{\alpha : \|2^n\alpha\| < \psi(2^n) \text{ i.o.}\}$.

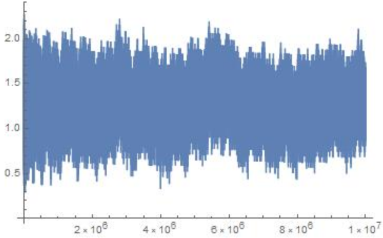
This beats the "elementary benchmark result", which assumes that $\sum_n \psi(2^n)^\gamma < \infty$. For example, the conclusion holds for $\psi(2^n) = n^{-1/\gamma}$.

Theorem 7 (Allen–C.–Yu 2020, divergence theory)

For $\psi(2^n) = 2^{-\log\log n / \log\log\log n}$, *we have* $\mu(W_2(\psi)) = 1$.

The truth about binary and ternary expansions

This is $\frac{D_2(n)+D_3(n)}{\log n}$ against n.



If we assume it's bounded above and below by positive constants, then we get a sharp convergence theory, and an almost-sharp (e.g. $\psi(2^n) = 1/n$) divergence theory. The upper bound is clear, and to motivate the lower bound imagine that n is a power of 2 whose ternary digits change 1/3 of the time.

A consequence of the Lang–Waldschmidt conjecture

Conjecture (Lang–Waldschmidt $\leqslant$ 1978)

Let $a_1, \ldots, a_n, b_1, \ldots, b_n$ be non-zero integers with $\Lambda := \sum_i b_i \log a_i \neq 0$. Then $\log |\Lambda| \geqslant -Cn(\log A + \log B)$, where $A = \max_i |a_i|$ and $B = \max_i |b_i|$. The constant C is effective.

For $c > 0$ small, we prove $D_2(n) + D_3(n) \geqslant c \log \log n$ on this conjecture, stronger than Stewart's unconditional result by a $\log \log \log n$ factor and giving a small n^δ improvement in the approximation results. If c were large then we would have a sharp convergence theory!

The right scale

We wish to show that $\mu(\limsup_{n\to\infty} A_n) = 0$, where

$$A_n = \{\alpha : \|2^n\alpha\| < \psi(2^n)\}.$$

By Borel–Cantelli it suffices for $\sum_n \mu(A_n) < \infty$. What is $\mu(A_n)$?

Lemma 1

Let $N \in \mathbb{N}$ with $3^{-N} \leqslant \psi(2^n)/(5 \cdot 2^n)$, and let C_N be the set of $b/3^N$ in the Cantor set. Then

$$\mu(A_n) \leqslant 2^{-N}|C_N \cap 5A_n|.$$

The idea is that $a/2^n$ and $b/3^N$ need to be close to get positive measure within a 3^{-N}-ball centred at $b/3^N$. Such a ball has Cantor measure $O(2^{-N})$, since $\mu(B(z,r)) \ll r^\gamma$ for all z, and we recall $\gamma = \frac{\log 2}{\log 3} = \dim_H(K)$.

A coarser scale

From the previous slide, it suffices to count dyadic rationals $a/2^n$ within about $\psi(2^n)/2^n$ of the Cantor set. We're currently unable to do this unconditionally, so we examine a coarser scale.

Lemma 2

If $3^{-N} \asymp \frac{\psi(2^n)}{2^n} \leqslant 3^{-M} \ll \frac{1}{2^n}$ then

$$|C_N \cap 5A_n| \ll |C_M \cap 5A_n(M)|,$$

where $A_n(M) = \cup_a B(a/2^n, 3^{-M})$.

To count solutions to

$$\left| \frac{a}{2^n} - \frac{b}{3^N} \right| < \frac{5\psi(2^n)}{2^n},$$

where $b/3^N \in K$, we write $b = b_1 b_2$, where b_1 comprises M ternary digits. The triangle inequality forces $a/2^n$ to lie within about 3^{-M} of $b_1/3^M$.

The benchmark result

Choose $M \leqslant N$ so that $3^{-N} \asymp \psi(2^n)/2^n$ and $3^{-M} \asymp 2^{-n}$, giving

$$\mu(A_n) \ll 2^{-N}|C_N \cap 5A_n| \ll 2^{-N}|C_M \cap 5A_n(M)|$$
$$\ll 2^{M-N} = (3^{M-N})^\gamma \ll (2^n/3^N)^\gamma \ll \psi(2^n)^\gamma.$$

Proposition 8 (Basic convergence assertion)

If $\psi : \mathbb{N} \to [0,\infty)$ *and* $\sum_n \psi(2^n)^\gamma < \infty$ *then* $\mu(W_2(\psi)) = 0$.

We beat this by working at an intermediate scale.

Prelude to Fourier analysis

To estimate $|C_M \cap 5A_n(M)| = 2|L_M \cap 5A_n(M)|$, where L_M is the set of $b/3^M$ with no 1's in ternary, we use a smooth bump function ϕ for $5A_n(M)$ and Dirac masses for L_M.

- Recall $5A_n(M) = \cup_a B(a/2^n, 5/3^M)$. Our $\phi : \mathbb{R} \to [0,\infty)$ is supported on $[-2,2]$, is 1 on $[-1,1]$, is Schwartz (has rapidly decaying derivatives of all orders), and in particular has the rapid Fourier decay property $\hat{\phi}(t) \ll (1+|t|)^{-99}$.
- For $x \in \mathbb{R}$, the Dirac translate δ_x is a distribution (generalised function) with the property that if f is Schwartz then
$$\int_{\mathbb{R}} \delta_x(\alpha) f(\alpha)\,\mathrm{d}\alpha = f(x).$$

Fourier series and Parseval's identity

The upshot is that

$$|C_M \cap 5A_n(M)| \leqslant 2 \int_{\mathbb{R}} f(\alpha) C(\alpha) \, \mathrm{d}\alpha,$$

where

$$C(\alpha) = \sum_{x \in L_M} \delta_x(\alpha), \quad f(\alpha) = \sum_{b=0}^{2^n - 1} \phi(2^{n+k}(\alpha - b2^{-n})), \quad \frac{5}{3^M} \approx \frac{1}{2^{n+k}}.$$

As f, C are supported on a unit interval, we can 1-periodically extend them and expand as Fourier series, applying Parseval to give

$$\sum_{w \in \mathbb{Z}} \hat{f}(w) \hat{C}(w).$$

The passage to digit changes

The rapid decay of $\hat{\phi}$ enables us to restrict attention to a finite sum, and meanwhile (writing $e(z) = e^{2\pi i z}$)

$$\hat{C}(w) = \int_0^1 C(\alpha)e(-w\alpha)\,\mathrm{d}\alpha = \sum_{\varepsilon_1,\ldots,\varepsilon_M \in \{0,2\}} \prod_{j \leqslant M} e(-w\varepsilon_j/3^j)$$

$$= 2^M \prod_{j \leqslant M} \frac{1 + e(-2w/3^j)}{2}$$

is small unless w has few ternary digit changes. We also get a $\sum_{b=0}^{2^n-1} e(-wb/2^n)$ factor from $\hat{f}(w)$, which forces w to be a multiple of 2^n and have fewer binary digit changes, so Stewart's theorem wins the day: the zero frequency dominates and we get the expected count (if our scale is not too fine).

Cheers

Thanks very much for your attention!