On the Duffin-Schaeffer conjecture

Dimitris Koukoulopoulos[1]

joint work with James Maynard[2]

[1]Université de Montréal
[2]University of Oxford

Stony Brook University
October 3, 2019

Diophantine approximation

Given an **irrational** number α, we seek **rational approximations**

$$\frac{a}{q} \approx \alpha$$

Two things to look for:

- the **complexity** of the approximation, i.e. how big q is
- the **quality** of the approximation, i.e. how close a/q is to α

Optimal balance of complexity vs. quality?

i.e. for which choices of $(\Delta_q)_{q=1}^{\infty}$ do we have ∞-many solutions to

$$\left|\alpha - \frac{a}{q}\right| \leqslant \Delta_q \ ?$$

Continued fractions

Let $\alpha \in [0,1] \setminus \mathbb{Q}$ and define:

$$n_1 = \lfloor 1/\alpha \rfloor \quad \rightsquigarrow \quad \alpha = \frac{1}{n_1 + \alpha_1}, \quad 0 < \alpha_1 < 1$$

$$n_2 = \lfloor 1/\alpha_1 \rfloor \quad \rightsquigarrow \quad \alpha = \cfrac{1}{n_1 + \cfrac{1}{n_2 + \alpha_2}}, \quad 0 < \alpha_2 < 1$$

$$\alpha \approx \frac{a_j}{q_j} := \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{\cdots + \cfrac{1}{n_j}}}} = j\text{-th convergent}$$

We have the recurrence formula $\begin{cases} a_j = n_j a_{j-1} + a_{j-2} \\ q_j = n_j q_{j-1} + q_{j-2} \end{cases}$

CF as best approximations

$$\left|\alpha - \frac{a_j}{q_j}\right| = \min\left\{\left|\alpha - \frac{a}{q}\right| : 1 \leqslant q \leqslant q_j\right\}$$

$$\frac{1}{2q_j q_{j+1}} \leqslant \left|\alpha - \frac{a_j}{q_j}\right| \leqslant \frac{1}{q_j q_{j+1}} < \frac{1}{q_j^2}$$

$$\left|\alpha - \frac{a}{q}\right| < \frac{1}{2q^2} \quad \& \quad (a, q) = 1 \qquad \Longrightarrow \qquad \frac{a}{q} \in \left\{\frac{a_1}{q_1}, \frac{a_2}{q_2}, \dots\right\}$$

Metric diophantine approximation

$$\lambda = \text{Lebesgue measure}$$

Question: What is the **typical** quality of approximation of α by its convergents (i.e. what happens λ-almost everywhere)?

• Example: it is known that the sequence $n_1, n_2, \ldots$ is typically unbounded.

• Given errors $(\Delta_q)_{q=1}^{\infty}$, let

$$\mathcal{K} := \{\alpha \in [0,1] : |\alpha - a/q| \leqslant \Delta_q \text{ for } \infty\text{-many } a, q\}$$

Khinchin (1924) proved that if $q^2\Delta_q \searrow$, then:

$$\textstyle\sum_q q\Delta_q < \infty \quad \Longrightarrow \quad \lambda(\mathcal{K}) = 0$$
$$\textstyle\sum_q q\Delta_q = \infty \quad \Longrightarrow \quad \lambda(\mathcal{K}) = 1$$

Corollary: for a typical α, we have $|\alpha - a/q| \leqslant 1/(q^2 \log q)$ ∞-often (and a/q must be a convergent as soon as $q \geqslant 10$)

Why is Khinchin correct?

$$\mathcal{K}_q := \bigcup_{0 \leqslant a \leqslant q} \left[\frac{a}{q} - \Delta_q, \frac{a}{q} + \Delta_q \right]$$

$$N(\alpha) = \#\{q : \alpha \in \mathcal{K}_q\}$$

$$\mathbb{E}_{\alpha \in [0,1]}[N(\alpha)] = \sum_q \lambda(\mathcal{K}_q) = 2 \sum_q q\Delta_q$$

$$\mathcal{K} := \limsup_{q \to \infty} \mathcal{K}_q = \{\alpha \in [0,1] : \alpha \in \mathcal{K}_q \text{ for } \infty\text{-many } q\}$$

• 'easy' direction of Borel-Cantelli : $\sum_{q \in \mathcal{S}} q\Delta_q < \infty \quad \Rightarrow \quad \lambda(\mathcal{K}) = 0.$

• Khinchin's theorem establishes the 'hard' direction of Borel-Cantelli when $q^2 \Delta_q \searrow$

Note: must show the sets $\mathcal{K}_q$ are sufficiently quasi-independent.

The Duffin-Schaeffer conjecture

Question: What is the most general Khinchin-type result?

i.e. for which sequences $(\Delta_q)_{q=1}^{\infty}$ are there ∞-many solutions to

$$\left|\alpha - \frac{a}{q}\right| \leqslant \Delta_q \ ?$$

- If $\Delta_q q^2 \searrow$, then $\Delta_q = O(1/q^2)$.

  What about larger Δ_q? (We are moving away from the theory of continued fractions.)

- If $\Delta_q q^2 \searrow$, then either $\Delta_q > 0$ for all q, or $\Delta_q = 0$ for all large enough q.

  What about sequences supported on sparser sets? e.g. using denominators that are primes, powers of 10, or perfect squares?

$\rightsquigarrow$ must focus on **reduced** fractions (avoids overcounting; deals with non-multiplicative structure of support of Δ_q)

The Duffin-Schaeffer conjecture

$$\mathcal{A}_q := \bigcup_{\substack{1 \leqslant a \leqslant q \\ \gcd(a,q)=1}} \left[\frac{a}{q} - \Delta_q, \frac{a}{q} + \Delta_q\right], \qquad \mathcal{A} = \limsup_{q \to \infty} \mathcal{A}_q$$

- Here $\lambda(\mathcal{A}_q) = 2\varphi(q)\Delta_q$, where

$$\varphi(q) = \#(\mathbb{Z}/q\mathbb{Z})^* = q \prod_{p|q}(1 - 1/p) = \text{Euler's totient function}$$

- Hence, the 'easy' Borel-Cantelli lemma yields:

$$\sum_q \varphi(q)\Delta_q < \infty \qquad \Rightarrow \qquad \lambda(\mathcal{A}) = 0$$

- Duffin and Schaeffer (1941) conjecture a strong converse is also true:

$$\sum_q \varphi(q)\Delta_q = \infty \qquad \Rightarrow \qquad \lambda(\mathcal{A}) = 1.$$

- Gallagher (1961) proved there is 0-1 law: $\qquad \lambda(\mathcal{A}) \in \{0, 1\}$

A key difference

$$\mathcal{S} := \text{supp}(\Delta_q) = \{q : \Delta_q > 0\}$$

$\mathcal{S}$ could be a very sparse/irregular set, which also forces Δ_q to be large (can no longer use continued fractions)

We can think of the Duffin-Schaeffer Conjecture (DSC) as follows:

We are given:

- $\mathcal{S}$ a set of *admissible denominators*
- for each $q \in \mathcal{S}$, an *admissible error* $0 < \Delta_q \leqslant \frac{1}{2q}$

$$\mathcal{A} := \left\{ \alpha \in [0,1] : \left| \alpha - \frac{a}{q} \right| \leqslant \Delta_q \quad \text{for } \infty\text{-many } q \in \mathcal{S}, \gcd(a,q) = 1 \right\}$$

Question: $\lambda(\mathcal{A}) = 0$ or $\lambda(\mathcal{A}) = 1$?

Previous results on DSC

- Duffin-Schaeffer (1941): DSC is true when $\varphi(q) \asymp q$ on average when weighted with $(\Delta_q)_{q\in\mathcal{S}}$

  Example: $\mathcal{S} = \{\text{primes}\}$
- Erdős (1970) & Vaaler (1978): DSC is true when $\Delta_q = O(1/q^2)$ (useful when $\mathcal{S}$ is relatively large so that $\sum_{q\in\mathcal{S}} \varphi(q)/q^2 = \infty$)
- Pollington-Vaughan (1990): DSC is true in $\mathbb{R}^d$ for $d > 1$
- Many results establishing DSC when there is 'extra divergence', i.e. when $\sum_{q\in\mathcal{S}} \frac{\varphi(q)\Delta_q}{L_q} = \infty$;

  Aistleitner (2019): can take $L_q = (\log\log q)^{\varepsilon}$

New results

Theorem (K.-Maynard (2019))

The Duffin-Schaeffer conjecture is true

Corollary (Catlin's conjecture)

$\mathcal{K} := \{\alpha \in [0,1] : |\alpha - a/q| \leqslant \Delta_q$ *for* ∞*-many* $a, q\}$

$C := \sum_q \varphi(q) \max\{\Delta_q, \Delta_{2q}, \dots\}$

We then have $\lambda(\mathcal{K}) = 1$ *when* $C = \infty$, *whereas* $\lambda(\mathcal{K}) = 0$ *when* $C < \infty$.

Using a theorem of Beresnevich-Velani we also obtain:

Corollary

$\mathcal{A} := \{\alpha \in [0,1] : |\alpha - a/q| \leqslant \Delta_q$ *for inf. many* **coprime** $a, q\}$

Assume $\sum_q \varphi(q)\Delta_q < \infty$, *so that* $\lambda(\mathcal{A}) = 0$. *Then*

$$\dim_{Hausdorff}(\mathcal{A}) = \min\left\{\beta \geqslant 0 : \sum_q \varphi(q)\Delta_q^\beta < \infty\right\}$$

Inverting Borel-Cantelli

$$\text{Set-up}: \quad \mathcal{A}_q = \bigcup_{\substack{1 \leqslant a \leqslant q \\ \gcd(a,q)=1}} \left[\frac{a}{q} - \Delta_q, \frac{a}{q} + \Delta_q\right], \qquad \mathcal{A} = \limsup_{\substack{q \to \infty \\ q \in \mathcal{S}}} \mathcal{A}_q,$$

$$\lambda(\mathcal{A}_q) = 2\varphi(q)\Delta_q, \qquad \sum_{q \in \mathcal{S}} \lambda(\mathcal{A}_q) = \infty.$$

Working heuristic: the sets $\mathcal{A}_q$ are quasi-independent events of the probability space $[0, 1]$ and should thus have limited overlap if the sum of their measures is $\leqslant 1$.

$$\textbf{Goal}: \quad \sum_{q \in [x,y] \cap \mathcal{S}} \lambda(\mathcal{A}_q) \approx 1 \quad \Longrightarrow \quad \lambda\Big(\bigcup_{q \in [x,y] \cap \mathcal{S}} \mathcal{A}_q\Big) \approx 1.$$

This is enough because it implies $\lambda(\mathcal{A}) > 0$, and thus $\lambda(\mathcal{A}) = 1$ by Gallagher's 0-1 law.

Cauchy-Schwarz

- $N(\alpha) = \#\{q \in [x,y] \cap \mathcal{S} : \alpha \in \mathcal{A}_q\} \quad \rightsquigarrow \quad \bigcup_{q \in [x,y] \cap \mathcal{S}} \mathcal{A}_q = \mathrm{supp}(N)$

- $\int N(\alpha) \mathrm{d}\alpha = \sum_{q \in [x,y] \cap \mathcal{S}} \int 1_{\mathcal{A}_q}(\alpha) \mathrm{d}\alpha = \sum_{q \in [x,y] \cap \mathcal{S}} \lambda(\mathcal{A}_q)$

- $\left(\int N(\alpha) \mathrm{d}\alpha \right)^2 \leqslant \lambda\big(\mathrm{supp}(N)\big) \int N(\alpha)^2 \mathrm{d}\alpha$

$$\Leftrightarrow \quad \sum_{q \in [x,y] \cap \mathcal{S}} \lambda(\mathcal{A}_q) \leqslant \lambda\Big(\bigcup_{q \in [x,y] \cap \mathcal{S}} \mathcal{A}_q \Big) \sum_{q,r \in [x,y] \cap \mathcal{S}} \lambda(\mathcal{A}_q \cap \mathcal{A}_r).$$

Revised goal: $\sum_{q \in [x,y] \cap \mathcal{S}} \lambda(\mathcal{A}_q) \approx 1 \quad \Longrightarrow \quad \sum_{q,r \in [x,y] \cap \mathcal{S}} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \lesssim 1$

The Erdős-Vaaler argument

Assume $\Delta_q = 1/q^2$ for $q \in \mathcal{S}$, and that $y = 2x$ (to fix size of q)

$$\sum_{q \in [x,2x] \cap \mathcal{S}} \lambda(\mathcal{A}_q) \approx 1 \qquad \iff \qquad \sum_{q \in [x,2x] \cap \mathcal{S}} \frac{\varphi(q)}{q} \approx x$$

For simplicity: ignore the weights $\varphi(q)/q$ and think of $\mathcal{S}$ as an arbitrary set of $\asymp x$ integers in $[x, 2x]$

Pollington-Vaughan: for $q, r \in \mathcal{S}$, we have

$$\frac{\lambda(\mathcal{A}_q \cap \mathcal{A}_r)}{\lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)} \geqslant \log t \qquad \Longrightarrow \qquad L_t(q,r) := \sum_{\substack{p \mid \frac{qr}{\gcd(q,r)^2} \\ p \geqslant t}} \frac{1}{p} \geqslant 1.$$

$$\rightsquigarrow \qquad \sum_{q,r \in [x,2x] \cap \mathcal{S}} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \lesssim \int_1^\infty \frac{\#\{q,r \in [x,2x] : L_t(q,r) \geqslant 1\}}{x^2} \cdot \frac{\mathrm{d}t}{t}$$

Anatomical statistics

$$\mathbb{E}_{q,r\in[x,2x]}\Big[L_t(q,r)\Big] \leqslant \mathbb{E}_{q,r\in[x,2x]}\Bigg[\sum_{p|q,\,p\geqslant t}\frac{1}{p} + \sum_{p|r,\,p\geqslant t}\frac{1}{p}\Bigg]$$

$$= 2\sum_{p\geqslant t}\frac{1}{p}\cdot\mathbb{P}_{q\in[x,2x]}(p|q)$$

$$\approx 2\sum_{p\geqslant t}\frac{1}{p^2} \lesssim \frac{2}{t\log t}$$

In fact, using Chernoff's inequality we find:

$$\frac{\#\{q,r\in[x,2x] : L_t(q,r)\geqslant 1\}}{x^2} = O(e^{-t})$$

$$\rightsquigarrow \qquad \sum_{q,r\in[x,2x]\cap\mathcal{S}}\lambda(\mathcal{A}_q\cap\mathcal{A}_r) \lesssim \int_1^\infty O(e^{-t})\mathrm{d}t = O(1).$$

Generalizing Erdős-Vaaler

Assume $\exists c \in (0,1)$ such that $\Delta_q = 1/q^{1+c}$ for $q \in \mathcal{S}$.

$$\sum_{q \in [x,2x] \cap \mathcal{S}} \lambda(\mathcal{A}_q) \approx 1 \qquad \iff \qquad \sum_{q \in [x,2x] \cap \mathcal{S}} \frac{\varphi(q)}{q} \approx x^c$$

For simplicity: ignore the weights $\varphi(q)/q$ and think of $\mathcal{S}$ as an arbitrary set of x^c integers in $[x, 2x]$

Pollington-Vaughan: for $q, r \in \mathcal{S}$, we have

$$\frac{\lambda(\mathcal{A}_q \cap \mathcal{A}_r)}{\lambda(\mathcal{A}_q)\lambda(\mathcal{A}_r)} \geqslant \log t \qquad \Longrightarrow \qquad \left\{ \begin{array}{ll} (\mathbf{1}) & L_t(q,r) \geqslant 1 \\ (\mathbf{2}) & x^{1-c}/t \leqslant \gcd(q,r) \leqslant x^{1-c} \end{array} \right\}$$

(Think of t as large but much smaller than x.)

$$\rightsquigarrow \qquad \sum_{q,r \in [x,2x] \cap \mathcal{S}} \lambda(\mathcal{A}_q \cap \mathcal{A}_r) \lesssim \int_1^\infty \frac{\#\left\{ q, r \in \mathcal{S} : \begin{array}{l} L_t(q,r) \geqslant 1 \\ t^{-1} \leqslant \frac{\gcd(q,r)}{x^{1-c}} \leqslant 1 \end{array} \right\}}{x^{2c}} \cdot \frac{\mathrm{d}t}{t}$$

Two conditions

Goal: if $\mathcal{S} \subset [x, 2x]$ is a set of x^c integers, show that

$$\#\left\{q, r \in \mathcal{S} : \begin{array}{l} L_t(q,r) \geqslant 1 \\ t^{-1} \leqslant \frac{\gcd(q,r)}{x^{1-c}} \leqslant 1 \end{array} \right\} \leqslant \frac{x^{2c}}{t}.$$

(1) The **anatomical condition** $L_t(q,r) \geqslant 1$ offers exponential gains in t when q, r are sampled over a *dense* subset of $[x, 2x]$

(2) $x^{1-c} \geqslant \gcd(q,r) \geqslant x^{1-c}/t$ is a **structural condition.** The heart of the proof is understanding how often it occurs.

Analysis of the structural condition $\gcd(q, r) \approx x^{1-c}$

$$\sum_{\substack{x \leqslant q \leqslant 2x \\ \gcd(q,r) \geqslant x^{1-c}/t}} 1 \leqslant \sum_{\substack{d|r \\ d \geqslant x^{1-c}/t}} \sum_{\substack{x \leqslant q \leqslant 2x \\ d|q}} 1$$

$$\leqslant \sum_{\substack{d|r \\ d \geqslant x^{1-c}/t}} \frac{x}{d}$$

$$\leqslant tx^c \cdot \#\{d|r\}$$

$$\rightsquigarrow \qquad \#\left\{q, r \in \mathcal{S} : \begin{array}{l} L_t(q,r) \geqslant 1 \\ \gcd(q,r) \geqslant \frac{x^{1-c}}{t} \end{array} \right\} \lesssim tx^{2c+o(1)} = t^2 \cdot x^{o(1)} \cdot \frac{x^{2c}}{t}$$

- Hope to remove t^2 by exploiting the condition $L_t(q, r) \geqslant 1$.
- But how to remove the factor $x^{o(1)}$?

One divisor to rule them all

The guiding model problem

Let $\mathcal{S} \subset [x, 2x]$ be a set of x^c integers. Assume there are $\geqslant |\mathcal{S}|^2/t$ pairs $(q, r) \in \mathcal{S} \times \mathcal{S}$ with $\gcd(q, r) \geqslant x^{1-c}/t$. Must it be the case that there is an integer $d \geqslant x^{1-c}/t$ that divides $\gg |\mathcal{S}'|t^{-O(1)}$ elements of $\mathcal{S}$?

If yes, we are done: replace $\mathcal{S}$ by $d\mathcal{S}' = \{dq : q \in \mathcal{S}'\}$.

We then have:

- $\mathcal{S}' \subset [1, 2x/d] \subset [1, 2tx^c]$
- $\#\mathcal{S}' \geqslant x^c t^{-O(1)}$ (almost positive proportion)

$\rightsquigarrow$ Use the anatomical condition $L_t(q, r) \geqslant 1$ to annihilate $t^{O(1)}$

The graph of dependencies

Consider the graph $G = (\mathcal{S}, \mathcal{E})$, where:

- $\mathcal{S} \subset [x, 2x] \cap \mathbb{Z}$ with $\#\mathcal{S} = x^c$
- $\mathcal{E} = \{(v, w) \in \mathcal{S} \times \mathcal{S} : \gcd(v, w) \geqslant x^{1-c}/t,\ L_t(v, w) \geqslant 1\}$

Assuming that the edge density is $\geqslant 1/t$, must it be the case that a positive proportion of the edges arise from a fixed divisor $d \geqslant x^{1-c}/t$?

Compressing GCD graphs

The tuple $G = (\mathcal{V}, \mathcal{W}, \mathcal{E}, M, N, D, u)$ is called a *CGD graph* if:

- $(\mathcal{V}, \mathcal{W}, \mathcal{E})$ is a bipartite graph;
- $\mathcal{V} \subset [M, 2M]$ and $\mathcal{W} \subset [N, 2N]$;
- $\mathcal{E} \subset \{(v, w) \in \mathcal{V} \times \mathcal{W} : \gcd(v, w) \geqslant D,\ L_t(v, w) \geqslant u\}$;

Goal: start with $G^{\mathrm{start}} = (\mathcal{S}, \mathcal{S}, \mathcal{E}^{\mathrm{start}}, x, x^{1-c}/t, 1)$ where $\mathcal{E}^{\mathrm{start}} = \{(v, w) \in \mathcal{S} \times \mathcal{S} : \gcd(v, w) \geqslant x^{1-c}/t,\ L_t(v, w) \geqslant 1\}$.

Arrive at $G^{\mathrm{end}} = (\mathcal{V}^{\mathrm{end}}, \mathcal{W}^{\mathrm{end}}, \mathcal{E}^{\mathrm{end}}, M^{\mathrm{end}}, N^{\mathrm{end}}, D^{\mathrm{end}}, 1/2)$, where:

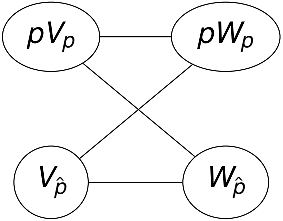
- $D^{\mathrm{end}} = 1$ (i.e. no more GCD conditions);
- $M^{\mathrm{end}} N^{\mathrm{end}} \leqslant \left(\frac{x}{x^{1-c}/t}\right)^2 = t^2 x^{2c}$ (because we have factored out one fixed divisor of size $\geqslant x^{1-c}/t$).

Also need: $\#\mathcal{E}^{\mathrm{end}} \geqslant x^{2c} t^{-O(1)}$.

Working prime by prime

For simplicity: $\mathcal{S}$ contains only square-frees

- $V_p = \{v/p : v \in \mathcal{V},\ p|v\} \subset [x/p, 2x/p]$
- $\mathcal{V}_{\hat{p}} = \{v \in \mathcal{V} : p \nmid v\} \subset [x, 2x]$



"subgraph"	M	N	D	MN/D^2
$(\mathcal{V}, \mathcal{W})$	x	x	x^{1-c}	x^{2c}
$(\mathcal{V}_p, \mathcal{W}_p)$	x/p	x/p	x^{1-c}/p	x^{2c}
$(\mathcal{V}_{\hat{p}}, \mathcal{W}_{\hat{p}})$	x	x	x^{1-c}	x^{2c}
$(\mathcal{V}_{\hat{p}}, \mathcal{W}_p)$	x	x/p	x^{1-c}	x^{2c}/p
$(\mathcal{V}_p, \mathcal{W}_{\hat{p}})$	x/p	x	x^{1-c}	x^{2c}/p

A quality-increment argument

"subgraph"	M	N	D	MN/D^2
$(\mathcal{V}, \mathcal{W})$	x	x	x^{1-c}	x^{2c}
$(\mathcal{V}_p, \mathcal{W}_p)$	x/p	x/p	x^{1-c}/p	x^{2c}
$(\mathcal{V}_{\hat{p}}, \mathcal{W}_{\hat{p}})$	x	x	x^{1-c}	x^{2c}
$(\mathcal{V}_{\hat{p}}, \mathcal{W}_p)$	x	x/p	x^{1-c}	x^{2c}/p
$(\mathcal{V}_p, \mathcal{W}_{\hat{p}})$	x/p	x	x^{1-c}	x^{2c}/p

quality of a GCD graph: $q(G) = \delta(G)^{10} \cdot |\mathcal{V}| \cdot |\mathcal{W}| \cdot \dfrac{D^2}{MN}$

Hard cases : $\dfrac{|\mathcal{V}_p|}{|\mathcal{V}|}, \dfrac{|\mathcal{W}_p|}{|\mathcal{W}|} = 1 - O(1/p)$ or $\dfrac{|\mathcal{V}_{\hat{p}}|}{|\mathcal{V}|}, \dfrac{|\mathcal{W}_{\hat{p}}|}{|\mathcal{W}|} = 1 - O(1/p)$.

Must make use of the weight $\varphi(v)/v$ to deal with them $\rightsquigarrow$ extra gain of factor $1 + 1/p$ in assymetric case

Thank you!