Rational approximations of irrational numbers

Dimitris Koukoulopoulos

Université de Montréal

(joint work with James Maynard (Oxford))

Number Theory Down Under
28 September 2021

Rational approximations I

Fundamental Question

Let $x \in \mathbb{R} \setminus \mathbb{Q}$. Find fractions a/q that approximate it "well".

- q must be small (fractions of "low complexity")
- the error $|x - a/q|$ must be small
- possible additional constraint: q must lie in a restricted set of denominators (e.g. primes, squares etc.)

Rational approximations II

- Decimal expansion: $x \approx a/10^n$ with error $\approx 1/10^n$ (typically).
- Dirichlet: for every irrational x, the inequality $|x - a/q| < 1/q^2$ has infinitely many solutions $(a, q) \in \mathbb{Z} \times \mathbb{N}$.
- Continue fractions: algorithm for constructing best possible rational approximations.

Improving Dirichlet's theorem I

Definition (Irrationality measure)

$\mu(x) := \sup\{\nu \geqslant 0 : |x - a/q| \leqslant q^{-\nu} \text{ infinitely often}\}$

- If $x \in \mathbb{Q}$, then $\mu(x) = 1$.
- Dirichlet: if $x \in \mathbb{R} \setminus \mathbb{Q}$, then $\mu(x) \geqslant 2$.
- Roth: if x is an algebraic irrational, then $\mu(x) = 2$.
- Zeilberger–Zudilin (2020): $\mu(\pi) \leqslant 7.10320533\ldots$

Improving Dirichlet's theorem II

Theorem (Zaharescu (1995))

Fix $\varepsilon > 0$. For every irrational x, there are infinitely many pairs $(a, q) \in \mathbb{Z} \times \mathbb{N}$ such that $\left|x - \frac{a}{q^2}\right| \leqslant \frac{1}{q^{8/3-\varepsilon}}$.

Theorem (Matomäki (2009))

Fix $\varepsilon > 0$. For every irrational x, there are infinitely many primes p and integers a such that $\left|x - \frac{a}{p}\right| \leqslant \frac{1}{p^{4/3-\varepsilon}}$.

Remark

Hard open problem: improve the exponents to 3 and 2, respectively.

These would be best possible (we'll see a justification shortly).

Metric Diophantine approximation

Basic principles

- Approximating specific numbers leads to hard open problems.
- Focus on proving results about almost all numbers.
- Exclusion of small pathological sets $\rightsquigarrow$ simple, general results

Khinchin's theorem

Definition

Given "admissible margins of error" $\Delta_1, \Delta_2, \ldots \geqslant 0$, let

$$\mathcal{A} := \left\{x \in [0,1] : \left|x - \tfrac{a}{q}\right| < \Delta_q \text{ for } \infty\text{-many } (a,q) \in \mathbb{Z} \times \mathbb{N}\right\}$$

Remark

We may focus on $[0,1]$ WLOG by periodicity.

Theorem (Khinchin (1924))

1. *If* $\sum_{q \geqslant 1} q\Delta_q < \infty$, *then* $\text{meas}(\mathcal{A}) = 0$.
2. *If* $\sum_{q \geqslant 1} q\Delta_q = \infty$ *and* $q^2\Delta_q \searrow$, *then* $\text{meas}(\mathcal{A}) = 1$.

Corollary

Let $\varepsilon > 0$. *For a.a.* $x \in \mathbb{R}$, *there are* ∞*-many* (a,q) *such that* $\left|x - \frac{a}{q}\right| < \frac{1}{q^2 \log^{1-\varepsilon} q}$, *but only finitely many s.t.* $\left|x - \frac{a}{q}\right| < \frac{1}{q^2 \log^{1+\varepsilon} q}$.

Theorem (Borel–Cantelli)

Let $E_1, E_2, \ldots$ be events in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $E = \limsup_{j\to\infty} E_j$ be the event that ∞-many occur.

1. *If $\sum \mathbb{P}(E_j) < \infty$, then $\mathbb{P}(E) = 0$.*
2. *If $\sum \mathbb{P}(E_j) = \infty$ and the E_j's are independent, then $\mathbb{P}(E) = 1$.*

- We may write $\mathcal{A} = \limsup_{q\to\infty} \mathcal{A}_q$, where

$$\mathcal{A}_q := [0,1] \cap \bigcup_{0\leqslant a\leqslant q} \left(\frac{a}{q} - \Delta_q, \frac{a}{q} + \Delta_q\right).$$

- $\mathrm{meas}(\mathcal{A}_q) = 2q\Delta_q$
- Khinchin: the $\mathcal{A}_q$'s are sufficiently "quasi-independent" when $q^2\Delta_q \searrow$

Generalizing Khinchin

Remark

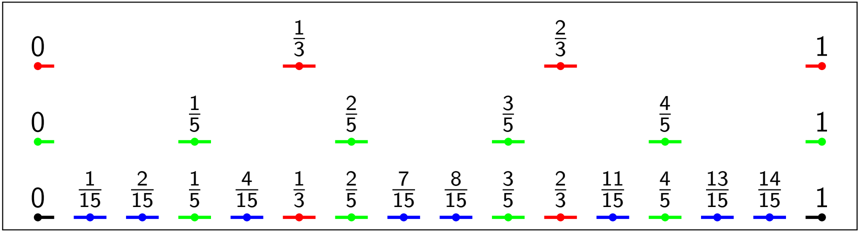
If $q^2\Delta_q \searrow$, then either $\Delta_q = 0$ for all large q, or $\operatorname{supp}(\Delta) = \mathbb{N}$.

Conclusion: must remove this condition to restrict denominators.

Proposition (Duffin–Schaeffer (1941))

Khinchin's theorem fails in full generality, i.e. we may find $\Delta_1, \Delta_2, \ldots$ *s.t.* $\sum q\Delta_q = \infty$ *and yet* $\operatorname{meas}(\mathcal{A}) = 0$.

(Recall: $\mathcal{A}$=set of "approximable numbers" with errors $< \Delta_q$)



Example with $\mathcal{A}_3, \mathcal{A}_5 \subset \mathcal{A}_{15}$

The Duffin–Schaeffer conjecture

- $\mathcal{A}^* := \left\{x \in [0,1] : \left|x - \frac{a}{q}\right| < \Delta_q \text{ for } \infty\text{-many reduced } \frac{a}{q}\right\}$
- This is the limsup of the sets

$$\mathcal{A}_q^* := \left\{x \in [0,1] : \left|x - \frac{a}{q}\right| < \Delta_q \text{ for some } a \in \mathbb{Z} \text{ co-prime to } q\right\}$$

that have measure $2\varphi(q)\Delta_q$, where $\varphi(q) = \#(\mathbb{Z}/q\mathbb{Z})^\times$.

Conjecture (Duffin-Schaeffer (1941))

1. *If* $\sum \varphi(q)\Delta_q < \infty$, *then* $\text{meas}(\mathcal{A}^*) = 0$.
2. *If* $\sum \varphi(q)\Delta_q = \infty$, *then* $\text{meas}(\mathcal{A}^*) = 1$.

Theorem (K.–Maynard (2019 $\to$ 2020))

The Duffin–Schaeffer Conjecture (DSC) is true.

The history of the conjecture

- Duffin–Schaeffer (1941): DSC is true if the errors Δ_q are supported on "not-too-abnormal integers", i.e. if

$$\limsup_{Q\to\infty} \frac{\sum_{q\leqslant Q} w_q \cdot \frac{\varphi(q)}{q}}{\sum_{q\leqslant Q} w_q} > 0 \qquad \text{with weights } w_q = q\Delta_q$$

Corollary. $|x - \frac{a}{p}| < \frac{1}{p^2}$ and $|x - \frac{a}{q^2}| < \frac{1}{q^3}$ i.o., for a.a. x.

- Gallagher (1961): there is a 0–1 law, i.e. $\text{meas}(\mathcal{A}^*) \in \{0, 1\}$.
- Erdős (1970) – Vaaler (1978): DSC is true if $\Delta_q = O(1/q^2)$. (Note: $\sum \varphi(q)\Delta_q = \infty$ implies then $\sum_{q\in\text{supp}(\Delta)} 1/q = \infty$.)
- Pollington–Vaughan (1990): DSC true in all dimensions > 1.
- Many authors: DSC is true when there is "extra divergence".
- Beresnevich–Velani (2006): DSC implies a generalized DSC for Hausdorff measures (via their general *Mass Transference Principle*).

Some corollaries of DSC

1. - Recall: $\mathcal{A}$ is the set of approximable numbers without constraints on GCDs. What is the correct 0–1 law for $\mathcal{A}$?
   - When $x \in \mathbb{R} \setminus \mathbb{Q}$ and $\Delta_q \to 0$, it is easy to check that

     $$|x - \tfrac{a}{q}| < \Delta_q \text{ i.o.} \quad \Longleftrightarrow \quad |x - \tfrac{a}{q}| < \Delta'_q \text{ with } \gcd(a,q) = 1 \text{ i.o.},$$

     where $\Delta'_q := \sup_{m \geqslant 1} \Delta_{mq}$.
   - This observation led Catlin to conjecture:

     $$\mathrm{meas}(\mathcal{A}) = 1 \quad \Longleftrightarrow \quad \sum \varphi(q)\Delta'_q = \infty$$

   - DSC readily implies Catlin's conjecture (which is the correct generalization to Khinchin)
2. Beresnevich–Velani: if $\sum_q \varphi(q)\Delta_q < \infty$, then

   $$\dim(\mathcal{A}^*) = \inf\left\{ s > 0 : \sum \varphi(q)\Delta_q^s < \infty \right\}.$$

Reduction to a pair-correlation estimate

Theorem (Borel–Cantelli)

Let $E_1, E_2, \ldots$ be events in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $E = \limsup_{j\to\infty} E_j$ be the event that ∞-many occur.

1. *If $\sum \mathbb{P}(E_j) < \infty$, then $\mathbb{P}(E) = 0$.*
2. *If $\sum \mathbb{P}(E_j) = \infty$ and the E_j's are independent, then $\mathbb{P}(E) = 1$.*

- Turan used Cauchy–Schwarz to prove Borel–Cantelli when

$$\mathbb{P}(E_i \cap E_j) \leqslant (1+\varepsilon)\mathbb{P}(E_i)\mathbb{P}(E_j) \quad \text{on average over } i \neq j.$$

- For DSC, we know the limsup satisfies a 0–1 law, so we only need to show that

$$\mathbb{P}(E_i \cap E_j) \leqslant 10^{10^{10}}\mathbb{P}(E_i)\mathbb{P}(E_j) \quad \text{on average over } i \neq j.$$

A special but crucial case

- $\mathcal{S} \subset [Q, 2Q]$ set of denominators
- $\sum_{q \in \mathcal{S}} \frac{\varphi(q)}{q} =: Q/D$, so that $D \gg 1$. (Think $\#\mathcal{S} \approx Q/D$.)
- $\Delta_q := \frac{D}{Q} \cdot \frac{1_{q \in \mathcal{S}}}{q}$, so that $\sum_{q \in \mathcal{S}} \mathrm{meas}(\mathcal{A}_q^*) = 1$.
- Can we prove $\mathrm{meas}(\bigcup_{q \in \mathcal{S}} \mathcal{A}_q^*) \gg 1$?

A special but crucial case, ctd.

Recall: $\mathcal{S} \subset [Q, 2Q]$, $\sum_{q \in \mathcal{S}} \frac{\varphi(q)}{q} = Q/D$, $\Delta_q = \frac{D}{Q} \cdot \frac{1_{q \in \mathcal{S}}}{q}$

- Pollington–Vaughan: $\mathrm{meas}(\bigcup_{q \in \mathcal{S}} \mathcal{A}_q^*) \gg 1$, unless $\exists t \geqslant 10^{10^{10}}$ s.t. there are $\geqslant t^{-1} \# \mathcal{S}^2$ pairs $(q, r) \in \mathcal{S} \times \mathcal{S}$ with:
  1. the number $qr/\gcd(q, r)^2$ has too many prime factors $> t$;
  2. $\gcd(q, r) > D/t$.
- Condition 1 occurs for $\ll e^{-t} Q^2$ pairs $(q, r) \in [Q, 2Q]^2$. This is sufficient if $D \asymp 1$ (Erdős–Vaaler argument).
- If D is large, we must exploit Condition 2. We show it induces "structure" on $\mathcal{S}$.
- If $d > D/t$ and $\mathcal{S} \subset \{q \in [Q, 2Q] : d | q\}$, then Condition 2 is satisfied for **all** pairs (q, r). Is some converse statement also true?

A combinatorial problem

The guiding model problem

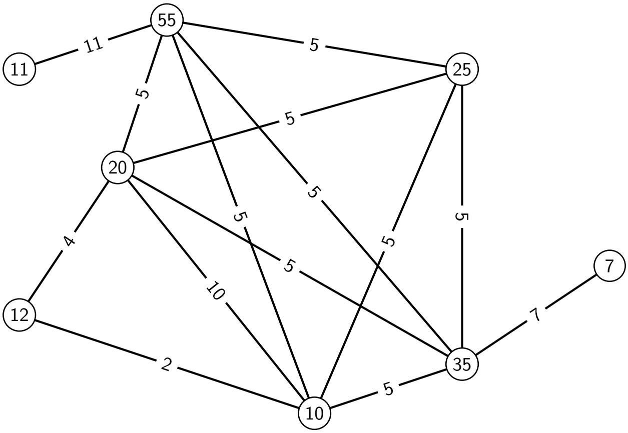
Let $\mathcal{S} \subset [Q, 2Q]$ a set of Q/D integers. Suppose there are $\geqslant \#\mathcal{S}^2/t$ pairs $(q, r) \in \mathcal{S} \times \mathcal{S}$ such that $\gcd(q, r) > D/t$. Must there exist some integers $d > D/t$ that divides $\gg t^{-100}\#\mathcal{S}$ elements of $\mathcal{S}$?

- If such a d exists, replace $\mathcal{S}$ with $\mathcal{S}' = \{m : dm \in \mathcal{S}\}$.
- $\#\mathcal{S}' \gg (Q/D)t^{-100}$ and $\mathcal{S} \subset [1, 2tQ/D]$, so $\mathcal{S}'$ is "dense".
- In addition, $qr/\gcd(q, r)^2 = mn/\gcd(m, n)^2$ when $(q, r) = (dm, dn)$
- Hence, Condition 2 carries through to the "dense" set $\mathcal{S}'$, and the Erdős–Vaaler argument completes the proof.

The graph of dependencies

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

- $\mathcal{S} \subset [Q, 2Q]$ is a set of $\approx Q/D$ integers;
- $\mathcal{E} = \{(q, r) \in \mathcal{S} \times \mathcal{S} : \gcd(q, r) > D\}$.



(graph by J. Maynard)

An iterative compression algorithm

- *Simplify*: $\mathcal{S}$ contains only square-free integers.
- *Technical manoeuvre:* necessary to view G as a bipartite graph $(\mathcal{S}, \mathcal{S}, \mathcal{E})$.
- Divise an algorithm that produces a nested sequence bipartite graphs $G^{(j)} = (\mathcal{V}^{(j)}, \mathcal{W}^{(j)}, \mathcal{E}^{(j)})$

$$G = G^{\text{start}} =: G^{(0)} \supset G^{(1)} \supset \cdots \supset G^{(J)} =: G^{\text{end}}$$

  and distinct primes $p_1, p_2, \ldots, p_J$ such that:
  1. For each $i \leqslant j$, the prime p_i either divides all elements of $V^{(j)}$ or none of them (and similarly with $W^{(j)}$).
  2. G^{end} has all large GCDs due to a universal divisor. Hence, we can analyze it using the Erdős–Vaaler argument.
  3. the graph $G^{(j)}$ has better “quality” than $G^{(j-1)}$ (i.e. more edges than naively expected). This ensures that the EV argument on G^{end} gives us non-trivial bounds on G^{start}.

The quality increment argument I

First attempt: ensure $\delta(G^{(j)}) \geqslant \delta(G^{(j-1)})$ at each step, mimicking Roth's density increment strategy.

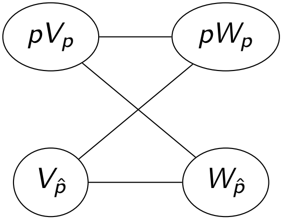
This fails because we lose control on edge set of G^{end};

The quality increment argument II

Second attempt: consider the following notation:

- $(\mathcal{V}, \mathcal{W}) = (\mathcal{V}^{(j-1)}, \mathcal{W}^{(j-1)})$ and $p = p_j$;
- $\mathcal{V}_p = \{v \in \mathcal{V} : p|v\}$, $\mathcal{V}_{\hat{p}} = \mathcal{V} \setminus \mathcal{V}_p$

We then have four options for $(\mathcal{V}^{(j)}, \mathcal{W}^{(j)})$:



1. $(\mathcal{V}_p, \mathcal{W}_p)$: gain factor of p left and right; but lose a factor of p in the GCDs (that affects both sides).
2. $(\mathcal{V}_{\hat{p}}, \mathcal{W}_{\hat{p}})$: no gained factors of p, so balanced situation.
3. $(\mathcal{V}_p, \mathcal{W}_{\hat{p}})$: we gain a factor of p on the left, and nothing on the right. BUT the GCDs are not affected, so we gain a factor of p overall. Hence, we can afford a large loss of edges.
4. $(\mathcal{V}_{\hat{p}}, \mathcal{W}_p)$: as in case 3, we gain a factor of p.

The quality increment argument III

Second attempt continued: ensure that $p_j^{\sigma_j} \# \mathcal{E}^{(j)} \geqslant \# \mathcal{E}^{(j-1)}$ at each step, where $\sigma_j = 0$ in the symmetric Cases 1,2 and $\sigma_j = 1$ in the asymmetric Cases 3,4.

This would allow control of $\mathcal{E}^{\text{start}}$ in terms of $\mathcal{E}^{\text{end}}$, but we cannot show it can be made to increase.

Third attempt: ensure that
$\delta(G^{(j)})^{10} p_j^{\sigma_j} \# \mathcal{E}^{(j)} \geqslant \delta(G^{(j-1)})^{10} \# \mathcal{E}^{(j-1)}$.

This almost works. Stumbling block: the Model Problem as stated is false! We must take account the weights $\varphi(q)/q$.

Fourth attempt: ensure that
$\delta(G^{(j)})^{10} p_j^{\sigma_j} (1 - 1/p_j)^{-\tau_j} \# \mathcal{E}^{(j)} \geqslant \delta(G^{(j-1)})^{10} \# \mathcal{E}^{(j-1)}$ at each step, where $\tau_j = 1$ in Case 1 where everything is divisible by p_j, and $\tau_j = 0$ otherwise.

Sam Chow's counterexample

$$\mathcal{S} = \{P/j : j|P,\ x/2 \leqslant j \leqslant x\} \quad \text{with} \quad P = \prod_{p \leqslant x} p.$$

- all pairwise GCDs here are $\geqslant P/x^2$
- no fixed integer of size $\gg P/x^2$ dividing a positive proportion of elements of $\mathcal{S}$
- notice that if $p \leqslant x/\log x$, then the proportion of $\mathcal{S}$ divisible by p is $\sim 1 - 1/p$.
- The case when $\#\mathcal{V}_p \sim (1 - 1/p)\#\mathcal{V}$ turns out to be the critical case in our "quality increment argument", and where we need to make use of the weights $\varphi(q)/q$.

Thank you for your attention