On the Duffin-Schaeffer Conjecture: 2

James Maynard

University of Oxford
Joint work with D. Koukoulopoulos (Montreal)

SFB Online Talk Series
September 2020

Talk Plan

This talk will give an introduction to the key new technical ideas in the proof.

1. Reduce to addive combinatorial problem about sets with large GCDs
2. Reformulate as a combinatorial problem about graphs
3. Overview of the iterative argument for graphs

Main Aim

Recall that we are given a function $\Delta : \mathbb{Z}_{>0} \to \mathbb{R}_{\geq 0}$.

The aim is to show 'quasi independence' on average

Main Aim

$$\sum_{q,r} \text{meas}(E_q^* \cap E_r^*) \leq 10^6 \sum_{q,r} \text{meas}(E_q^*)\,\text{meas}(E_r^*)$$

(for some suitable range of summation for q, r)

Here

$$E_q := \bigcup_{\substack{1 \leq a \leq q \\ \gcd(a,q)=1}} \left[\frac{a}{q} - \Delta(q), \frac{a}{q} + \Delta(q)\right].$$

Special case

Let's focus on a special case to emphasize the main ideas.

- $\Delta(q) \in \{q^{-1-c}, 0\}$ for some constant $c \in [0,1]$.
- $\Delta(q) = 0$ if q has a repeated prime factor or a factor $< 10^{10^{10}}$.
- $\Delta(q)$ is supported on a union of dyadic intervals $[x_1, 2x_1] \cup [x_2, 2x_2] \cup \dots$ with $x_{i+1} > x_i^2$ and

$$\sum_{q \in [x_i, 2x_i]} \phi(q)\Delta(q) \in [1,2].$$

Special case

Let's focus on a special case to emphasize the main ideas.

- $\Delta(q) \in \{q^{-1-c}, 0\}$ for some constant $c \in [0,1]$.
- $\Delta(q) = 0$ if q has a repeated prime factor or a factor $< 10^{10^{10}}$.
- $\Delta(q)$ is supported on a union of dyadic intervals $[x_1, 2x_1] \cup [x_2, 2x_2] \cup \dots$ with $x_{i+1} > x_i^2$ and

$$\sum_{q\in[x_i,2x_i]} \phi(q)\Delta(q) \in [1,2].$$

In this setup, we have that

$$\mathrm{meas}(E_q^*) \ll \frac{\phi(q)}{q} q^{-c}$$

($A \ll B$ means $A = O(B)$ means $|A| \leq cB$ for a constant $c > 0$)

Aim

Let $\mathcal{S} = \{q \in [x, 2x] : \Delta(q) \neq 0\}$. Then

$$\sum_{q,r\in\mathcal{S}} \mathrm{meas}(E_q^* \cap E_r^*) = O(1).$$

Simplification II

Lemma (Pollington-Vaughan bound)

$$\sum_{\substack{p|q \text{ or } r\\ p\geq t}} \frac{1}{p} \leq 11 \quad \Rightarrow \quad \mathrm{meas}(E_q^*\cap E_r^*) \leq 100(\log t)\,\mathrm{meas}(E_q^*)\,\mathrm{meas}(E_r^*).$$

Moreover, if $\gcd(q,r) \leq x^{1-c}/t$ *then*

$$\sum_{\substack{p|q \text{ or } r\\ p\geq t}} \frac{1}{p} \leq 11 \quad \Rightarrow \quad \mathrm{meas}(E_q^* \cap E_r^*) \leq 100\,\mathrm{meas}(E_q^*)\,\mathrm{meas}(E_r^*).$$

Simplification II

Lemma (Pollington-Vaughan bound)

$$\sum_{\substack{p|q \text{ or } r\\ p\geq t}} \frac{1}{p} \leq 11 \quad \Rightarrow \quad \operatorname{meas}(E_q^* \cap E_r^*) \leq 100(\log t) \operatorname{meas}(E_q^*) \operatorname{meas}(E_r^*).$$

Moreover, if $\gcd(q,r) \leq x^{1-c}/t$ *then*

$$\sum_{\substack{p|q \text{ or } r\\ p\geq t}} \frac{1}{p} \leq 11 \quad \Rightarrow \quad \operatorname{meas}(E_q^* \cap E_r^*) \leq 100 \operatorname{meas}(E_q^*) \operatorname{meas}(E_r^*).$$

This reduces to showing for every set $\mathcal{S} \subset [x, 2x]$ and every $t \geq 1$

$$\sum_{\substack{q,r\in\mathcal{E}_t\\ \gcd(q,r)\geq x^{1-c}/t}} \frac{\phi(q)}{q}\frac{\phi(r)}{r} = O\Big(\frac{x^{2c}}{t}\Big),$$

where

$$\sum_{q\in\mathcal{S}} \frac{\phi(q)}{q} = O(x^c), \qquad \mathcal{E}_t := \Big\{(q,r)\in\mathcal{S}^2 : \sum_{\substack{p|qr/\gcd(q,r)^2\\ p\geq t}} \frac{1}{p} \geq 10\Big\}.$$

Simplification II

This reduces to showing for every set $\mathcal{S} \subset [x, 2x]$ and every $t \geq 1$

$$\sum_{\substack{q,r\in\mathcal{E}_t \\ \gcd(q,r)\geq x^{1-c}/t}} \underbrace{\frac{\phi(q)}{q}\frac{\phi(r)}{r}}_{\text{weights} \approx 1} = O\Big(\frac{x^{2c}}{t}\Big),$$

where

$$\sum_{q\in\mathcal{S}} \underbrace{\frac{\phi(q)}{q}}_{\text{weights} \approx 1} = O(x^c), \qquad \mathcal{E}_t := \Big\{(q,r)\in\mathcal{S}^2 : \sum_{\substack{p|qr/\gcd(q,r)^2 \\ p\geq t}} \frac{1}{p} \geq 10\Big\}.$$

Simplification II

This reduces to showing for every set $\mathcal{S}\subset[x,2x]$ and every $t\geq 1$

$$\sum_{\substack{q,r\in\mathcal{E}_t\\ \gcd(q,r)\geq x^{1-c}/t}}\underbrace{\frac{\phi(q)}{q}\frac{\phi(r)}{r}}_{\text{weights}\approx 1}=O\Big(\frac{x^{2c}}{t}\Big),$$

where

$$\sum_{q\in\mathcal{S}}\underbrace{\frac{\phi(q)}{q}}_{\text{weights}\approx 1}=O(x^c),\qquad \mathcal{E}_t:=\Big\{(q,r)\in\mathcal{S}^2:\underbrace{\sum_{\substack{p|qr/\gcd(q,r)^2\\ p\geq t}}\frac{1}{p}\geq 10}_{\text{many prime factors condition}}\Big\}.$$

Simplification II

This reduces to showing for every set $\mathcal{S} \subset [x, 2x]$ and every $t \geq 1$

$$\sum_{\substack{q,r \in \mathcal{E}_t \\ \underbrace{\gcd(q,r) \geq x^{1-c}/t}_{\text{GCD condition}}}} \underbrace{\frac{\phi(q)}{q}\frac{\phi(r)}{r}}_{\text{weights} \approx 1} = O\Big(\frac{x^{2c}}{t}\Big),$$

where

$$\sum_{q \in \mathcal{S}} \underbrace{\frac{\phi(q)}{q}}_{\text{weights} \approx 1} = O(x^c), \qquad \mathcal{E}_t := \Big\{(q,r) \in \mathcal{S}^2 : \underbrace{\sum_{\substack{p \mid qr/\gcd(q,r)^2 \\ p \geq t}} \frac{1}{p} \geq 10}_{\text{many prime factors condition}} \Big\}.$$

Model problem

If we ignore the $\phi(q)/q$ weights, this simplifies to: If $\#\mathcal{S} = x^c$ and $\mathcal{S} \subseteq [x, 2x]$, is it the case that

$$\#\Big\{(q,r) \in \mathcal{E}_t : \gcd(q,r) \geq x^{1-c}/t\Big\} = O(x^{2c}/t)?$$

This is interesting even for $t = 1$ (i.e. ignoring the small prime factors condition).

Model problem

If we ignore the $\phi(q)/q$ weights, this simplifies to: If $\#\mathcal{S} = x^c$ and $\mathcal{S} \subseteq [x, 2x]$, is it the case that

$$\#\Big\{(q,r) \in \mathcal{E}_t : \gcd(q,r) \geq x^{1-c}/t\Big\} = O(x^{2c}/t)?$$

This is interesting even for $t = 1$ (i.e. ignoring the small prime factors condition).

Problem (Model problem)

What sets $\mathcal{S} \subseteq [x, 2x]$ have $\#\mathcal{S} \approx x^c$ and 1% of pairs $(s_1, s_2) \in \mathcal{S}^2$ with $\gcd(s_1, s_2) \geq x^{1-c}$?

We hope to say that any such set must be very structured, and so can then be analysed explicitly.

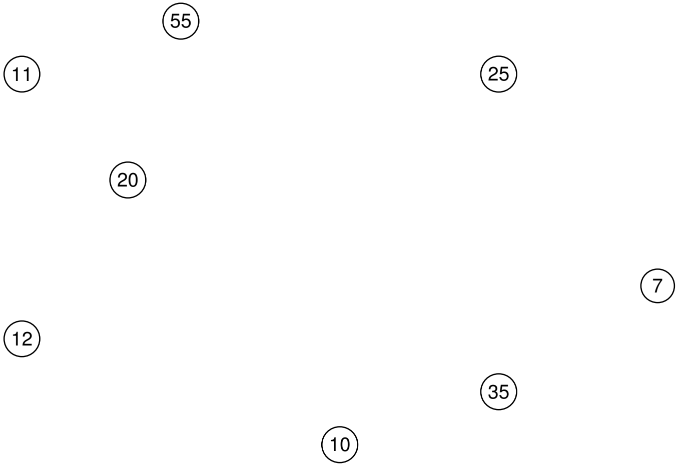
GCD Graphs

We visualise these sets as a graph

GCD Graphs

We visualise these sets as a graph

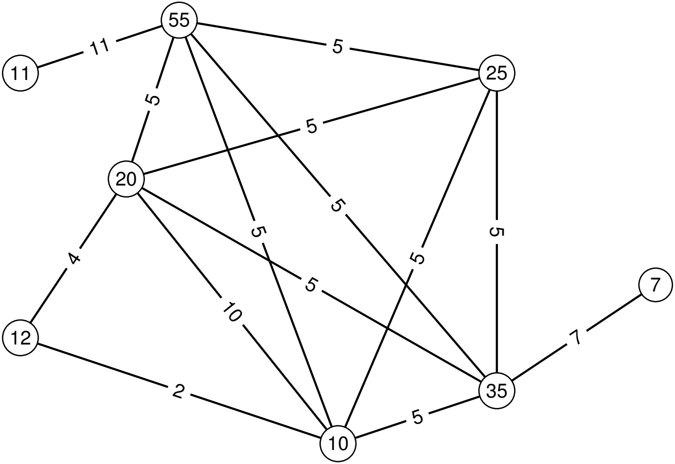
Vertices: Elements of the set.



GCD Graphs

We visualise these sets as a graph

Vertices: Elements of the set. **Edges:** Pairs with large GCD.



A positive proportion of pairs have a large GCD, so a **dense graph**.

Technical point: actually we consider bipartite version

GCD Graphs

Technical point: actually we consider bipartite version

Vertices: Elements of the set (two copies).

GCD Graphs

Technical point: actually we consider bipartite version

Vertices: Elements of the set (two copies).

GCD Graphs

Technical point: actually we consider bipartite version

Vertices: Elements of the set (two copies).

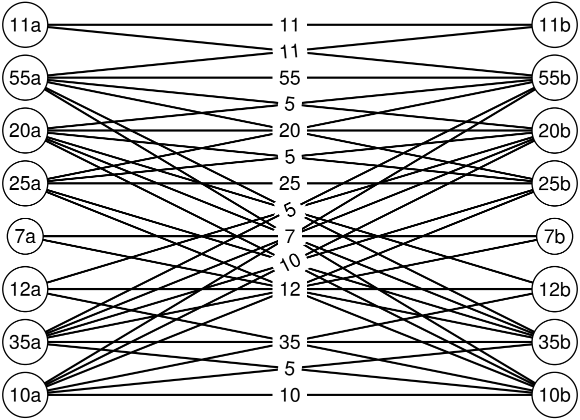
Edges: Pairs with large GCD.

GCD Graphs

Technical point: actually we consider bipartite version

Vertices: Elements of the set (two copies).

Edges: Pairs with large GCD.



Setup

We're given a dense graph with edges corresponding to big GCDs.

We want to either:

- Find a highly structured subset coming from the same divisor.
- Show that the graph must have few vertices.

Setup

We're given a dense graph with edges corresponding to big GCDs.

We want to either:

- Find a highly structured subset coming from the same divisor.
- Show that the graph must have few vertices.

Key Idea: We will repeatedly throw away a few 'unstructured' vertices/edges, to form a sequence of subgraphs

$$G_1 \supseteq G_2 \supseteq G_3 \supseteq \dots$$

where we increase 'structure' at each step.

Setup

We're given a dense graph with edges corresponding to big GCDs.

We want to either:

- Find a highly structured subset coming from the same divisor.
- Show that the graph must have few vertices.

Key Idea: We will repeatedly throw away a few 'unstructured' vertices/edges, to form a sequence of subgraphs

$$G_1 \supseteq G_2 \supseteq G_3 \supseteq \dots$$

where we increase 'structure' at each step.

- We don't try to control the (finite) number of iterations.
- Since the final graph cannot be iterated further, we hope that it will be simple to explicitly analyse.
  (e.g. all vertices are a multiple of d, so everything connected.)
- If we increase structure at each step, we hope that we can understand our original graph from the final one.

The basic manouvre

The key step to pass from G_i to G_{i+1} is to restrict each vertex class to elements which are a multiple of p or coprime to p.

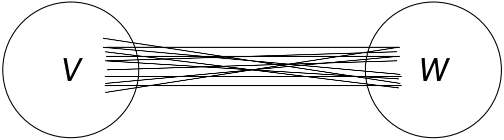
The basic manouvre

The key step to pass from G_i to G_{i+1} is to restrict each vertex class to elements which are a multiple of p or coprime to p.

Let G_i have vertex sets V and W. Choose a prime p arbitrarily. Define

$$V_p := \{v \in V : p|v\}, \qquad V_{\widehat{p}} := \{v \in V : p \nmid v\},$$
$$W_p := \{w \in W : p|w\}, \qquad W_{\widehat{p}} := \{w \in V : p \nmid w\}.$$

G_i :



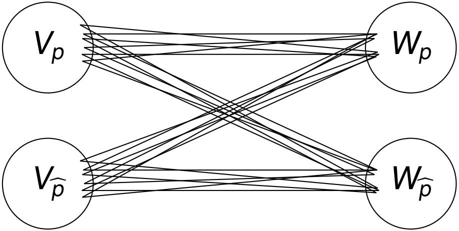
The basic manouvre

The key step to pass from G_i to G_{i+1} is to restrict each vertex class to elements which are a multiple of p or coprime to p.

Let G_i have vertex sets V and W. Choose a prime p arbitrarily. Define

$$V_p := \{v \in V : p|v\}, \qquad V_{\widehat{p}} := \{v \in V : p \nmid v\},$$
$$W_p := \{w \in W : p|w\}, \qquad W_{\widehat{p}} := \{w \in V : p \nmid w\}.$$

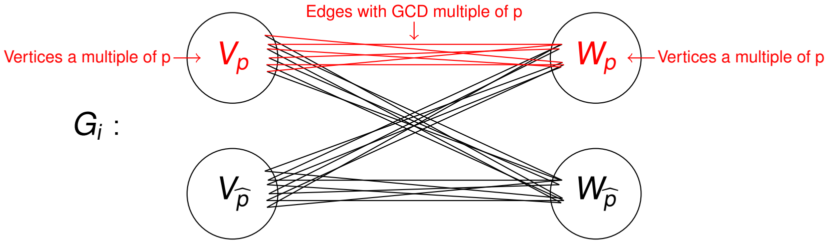
G_i :



The key step to pass from G_i to G_{i+1} is to restrict each vertex class to elements which are a multiple of p or coprime to p.

Let G_i have vertex sets V and W. Choose a prime p arbitrarily. Define

$$V_p := \{v \in V : p|v\}, \qquad V_{\widehat{p}} := \{v \in V : p \nmid v\},$$
$$W_p := \{w \in W : p|w\}, \qquad W_{\widehat{p}} := \{w \in V : p \nmid w\}.$$

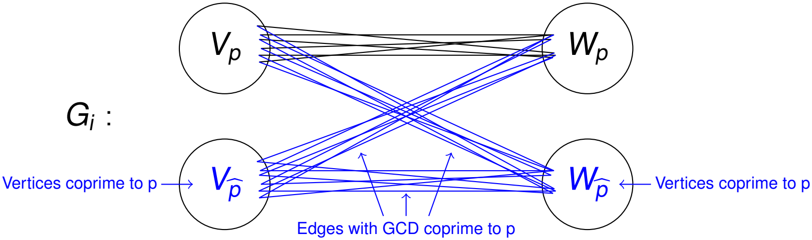


The basic manouvre

The key step to pass from G_i to G_{i+1} is to restrict each vertex class to elements which are a multiple of p or coprime to p.

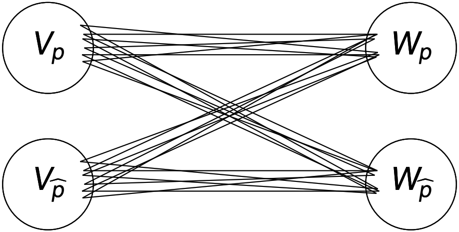
Let G_i have vertex sets V and W. Choose a prime p arbitrarily. Define

$$V_p := \{v \in V : p|v\}, \qquad V_{\widehat{p}} := \{v \in V : p \nmid v\},$$
$$W_p := \{w \in W : p|w\}, \qquad W_{\widehat{p}} := \{w \in V : p \nmid w\}.$$



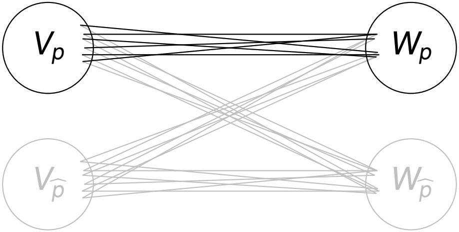
The basic manouvre II

We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



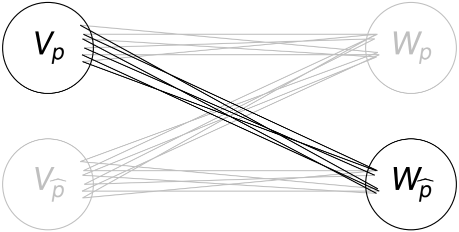
The basic manouvre II

We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



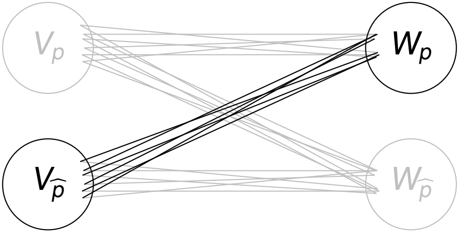
The basic manouvre II

We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



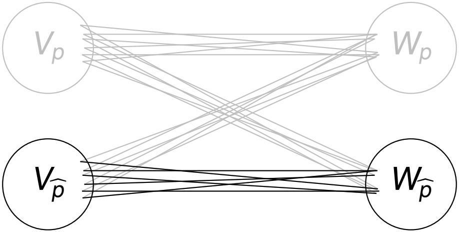
The basic manouvre II

We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



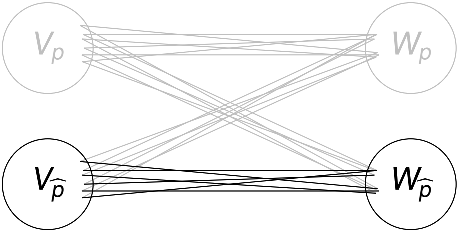
The basic manouvre II

We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



The basic manouvre II

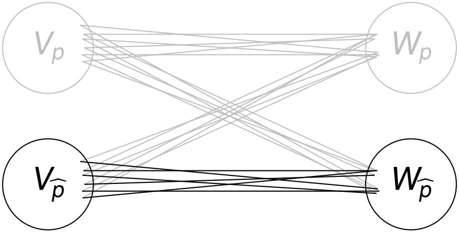
We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



If we do this repeatedly for every prime that occurs in any GCD, eventually all GCDs will come from the same fixed divisor.

The basic manouvre II

We will throw away either V_p or $V_{\widehat{p}}$ and either W_p or $W_{\widehat{p}}$ and let G_{i+1} be the resulting graph.



If we do this repeatedly for every prime that occurs in any GCD, eventually all GCDs will come from the same fixed divisor.

Question

How do we maintain control over this procedure?

'Quality'

Question

How do we maintain control over this procedure?

We define a statistic which increases at each iteration, and controls our original graph.

'Quality'

Question

How do we maintain control over this procedure?

We define a statistic which increases at each iteration, and controls our original graph.

We need to keep track of previous iterations, so consider a bipartite graph G with a set P of primes.

Definition (Quality of a graph)

*We define the **quality** of G with vertex sets V, W and set of primes P as*

$$q(G) = \delta^{10} \cdot \#V \cdot \#W \cdot \prod_{p \in P} p$$

where δ is the edge density.

P is the set of all primes where we have chosen $(V_p, W_{\widehat{p}})$ or $(V_{\widehat{p}}, W_p)$ in earlier iterations.

Quality II

Why is this definition of any use?

Quality II

Why is this definition of any use? Recall our setup:

- Begin with G_{start}, a dense bipartite graph with vertex sets in $[x, 2x]$ and edges formed by joining v, w if $\gcd(v, w) \geq x^{1-c}$.
- Apply the iteration procedure until all primes dividing any GCD have been accounted for, leaving a subgraph G_{end}.

Quality II

Why is this definition of any use? Recall our setup:

- Begin with G_{start}, a dense bipartite graph with vertex sets in $[x, 2x]$ and edges formed by joining v, w if $\gcd(v, w) \geq x^{1-c}$.
- Apply the iteration procedure until all primes dividing any GCD have been accounted for, leaving a subgraph G_{end}.

Lemma (Quality controls our original graph)

Let G_{start} have edge set E_{start}. Then

$$\#E_{start} \ll \frac{q(G_{start})}{q(G_{end})} x^{2c}$$

Quality II

Why is this definition of any use? Recall our setup:

- Begin with G_{start}, a dense bipartite graph with vertex sets in $[x, 2x]$ and edges formed by joining v, w if $\gcd(v, w) \geq x^{1-c}$.
- Apply the iteration procedure until all primes dividing any GCD have been accounted for, leaving a subgraph G_{end}.

Lemma (Quality controls our original graph)

Let G_{start} have edge set E_{start}. Then

$$\#E_{start} \ll \frac{q(G_{start})}{q(G_{end})} x^{2c}$$

Lemma (Quality increment)

If $\max(\#V_{\widehat{p}}/\#V, \#W_{\widehat{p}}/\#W) \geq 10^{40}/p$ then we can choose a subgraph G' from $V_p, V_{\widehat{p}}, W_p, W_{\widehat{p}}$ with

$$q(G') \geq q(G).$$

Quality III

Lemma (Quality controls our original graph)

Let G_{start} have edge set E_{start}. Then

$$\#E_{start} \ll \frac{q(G_{start})}{q(G_{end})} x^{2c}$$

Lemma (Quality increment)

If (conditions) then there is a subgraph G' with

$$q(G') \geq q(G).$$

Quality III

Lemma (Quality controls our original graph)

Let G_{start} have edge set E_{start}. Then

$$\#E_{start} \ll \frac{q(G_{start})}{q(G_{end})} x^{2c}$$

Lemma (Quality increment)

If (conditions) then there is a subgraph G' with

$$q(G') \geq q(G).$$

If we pretended that we can always find G', then we see that we can ensure $q(G_{end}) \geq q(G_{start})$, so

$$\#E_{start} \ll x^{2c}$$

This is almost what we want! If $\#E_{start} \approx x^{2c}$ then iterations must lose nothing, which is very close to giving our structure theorem.

Application to DS

For the DS problem we only need to save in the upper bound based on the threshold t for small prime factors. In this case

$$E_{start} = \mathcal{E}_t := \Bigg\{(q,r) \in \mathcal{S}^2 : \sum_{\substack{p \mid qr/\gcd(q,r)^2 \\ p \geq t}} \frac{1}{p} \geq 10\Bigg\}.$$

Application to DS

For the DS problem we only need to save in the upper bound based on the threshold t for small prime factors. In this case

$$E_{start} = \mathcal{E}_t := \Big\{(q,r) \in \mathcal{S}^2 : \sum_{\substack{p|qr/\gcd(q,r)^2 \\ p \geq t}} \frac{1}{p} \geq 10\Big\}.$$

Using the fact that the density of numbers with $\sum_{p|n, p\geq t} 1/p \geq 10$ is $O(e^{-t})$ we can refine the lemma.

For the DS problem we only need to save in the upper bound based on the threshold t for small prime factors. In this case

$$E_{start} = \mathcal{E}_t := \Big\{(q,r) \in \mathcal{S}^2 : \sum_{\substack{p|qr/\gcd(q,r)^2 \\ p \geq t}} \frac{1}{p} \geq 10\Big\}.$$

Using the fact that the density of numbers with $\sum_{p|n, p\geq t} 1/p \geq 10$ is $O(e^{-t})$ we can refine the lemma.

Lemma (Quality controls our DS graph)

Let G_{start} have edge set $\mathcal{E}_t$. Then either

$$\#\mathcal{E}_t \ll \frac{q(G_{start})}{q(G_{end})} x^{2c} e^{-t}$$

or $q(G_{end}) \geq e^t q(G_{start})$.

Thus in either case $\#\mathcal{E}_t \ll x^{2c} e^{-t}$, which is what we need.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq ... \supseteq$.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq$.
- To go from $G = G_i$ to G_{i+1}, we choose a prime p which divides $gcd(v, w)$ for some $v \in V_i$, $w \in W_i$ but that doesn't divide all $v \in V_i, w \in W_i$.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq$.
- To go from $G = G_i$ to G_{i+1}, we choose a prime p which divides $gcd(v, w)$ for some $v \in V_i$, $w \in W_i$ but that doesn't divide all $v \in V_i, w \in W_i$.
- By Lemma, there is a subgraph $G' \in \{G_{p,p}, G_{\widehat{p},p}, G_{p,\widehat{p}}, G_{\widehat{p},\widehat{p}}\}$ with $q(G') \geq q(G)$. If $G' = G_{p,\widehat{p}}$ or $G_{\widehat{p},p}$ then $P_{i+1} := P_i \cup \{p\}$ and otherwise $P_{i+1} := P_i$. We set $G_{i+1} := G'$.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq$.
- To go from $G = G_i$ to G_{i+1}, we choose a prime p which divides $gcd(v, w)$ for some $v \in V_i$, $w \in W_i$ but that doesn't divide all $v \in V_i, w \in W_i$.
- By Lemma, there is a subgraph $G' \in \{G_{p,p}, G_{\widehat{p},p}, G_{p,\widehat{p}}, G_{\widehat{p},\widehat{p}}\}$ with $q(G') \geq q(G)$. If $G' = G_{p,\widehat{p}}$ or $G_{\widehat{p},p}$ then $P_{i+1} := P_i \cup \{p\}$ and otherwise $P_{i+1} := P_i$. We set $G_{i+1} := G'$.
- Since each prime can only be chosen once, this procedure terminates when there are no primes left.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq$.
- To go from $G = G_i$ to G_{i+1}, we choose a prime p which divides $gcd(v, w)$ for some $v \in V_i$, $w \in W_i$ but that doesn't divide all $v \in V_i, w \in W_i$.
- By Lemma, there is a subgraph $G' \in \{G_{p,p}, G_{\widehat{p},p}, G_{p,\widehat{p}}, G_{\widehat{p},\widehat{p}}\}$ with $q(G') \geq q(G)$. If $G' = G_{p,\widehat{p}}$ or $G_{\widehat{p},p}$ then $P_{i+1} := P_i \cup \{p\}$ and otherwise $P_{i+1} := P_i$. We set $G_{i+1} := G'$.
- Since each prime can only be chosen once, this procedure terminates when there are no primes left.
- Since we can't iterate any more, if $p | \gcd(v, w)$ for some $v \in V_{end}, w \in W_{end}$ then $p|v$ for all v and $p|w$ for all w.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq$.
- To go from $G = G_i$ to G_{i+1}, we choose a prime p which divides $gcd(v, w)$ for some $v \in V_i$, $w \in W_i$ but that doesn't divide all $v \in V_i, w \in W_i$.
- By Lemma, there is a subgraph $G' \in \{G_{p,p}, G_{\widehat{p},p}, G_{p,\widehat{p}}, G_{\widehat{p},\widehat{p}}\}$ with $q(G') \geq q(G)$. If $G' = G_{p,\widehat{p}}$ or $G_{\widehat{p},p}$ then $P_{i+1} := P_i \cup \{p\}$ and otherwise $P_{i+1} := P_i$. We set $G_{i+1} := G'$.
- Since each prime can only be chosen once, this procedure terminates when there are no primes left.
- Since we can't iterate any more, if $p | \gcd(v, w)$ for some $v \in V_{end}, w \in W_{end}$ then $p|v$ for all v and $p|w$ for all w.
- If $p \in P_{end}$ then $p|v$ for all v and $p \nmid w$ for all w, or vice versa.

Idealised iteration procedure

- We start with $G_{start} = (V_{start}, W_{start}, E_{start})$ with $P_{start} = \emptyset$.
- We then form a finite sequence of subgraphs $G_{start} = G_0 \supseteq G_1 \supseteq G_2 \supseteq \dots \supseteq$.
- To go from $G = G_i$ to G_{i+1}, we choose a prime p which divides $gcd(v, w)$ for some $v \in V_i$, $w \in W_i$ but that doesn't divide all $v \in V_i, w \in W_i$.
- By Lemma, there is a subgraph $G' \in \{G_{p,p}, G_{\widehat{p},p}, G_{p,\widehat{p}}, G_{\widehat{p},\widehat{p}}\}$ with $q(G') \geq q(G)$. If $G' = G_{p,\widehat{p}}$ or $G_{\widehat{p},p}$ then $P_{i+1} := P_i \cup \{p\}$ and otherwise $P_{i+1} := P_i$. We set $G_{i+1} := G'$.
- Since each prime can only be chosen once, this procedure terminates when there are no primes left.
- Since we can't iterate any more, if $p | \gcd(v, w)$ for some $v \in V_{end}, w \in W_{end}$ then $p|v$ for all v and $p|w$ for all w.
- If $p \in P_{end}$ then $p|v$ for all v and $p \nmid w$ for all w, or vice versa.
- Thus there are a, b such that $a|v$ for all $v \in V_{end}$, $b|w$ for all $w \in W_{end}$, and $(v, w) \in E_{end} \Rightarrow \gcd(v, w) = \gcd(a, b) \geq x^{1-c}$.

Quality increment

Thus we're in a good position if we can always get a quality increment.

Quality increment

Thus we're in a good position if we can always get a quality increment.

Recall:

Lemma (Quality increment)

If $\max(\#V_{\widehat{p}}/\#V, \#W_{\widehat{p}}/\#W) \geq 10^{40}/p$ *then we can choose a subgraph G' from $V_p, V_{\widehat{p}}, W_p, W_{\widehat{p}}$ with*

$$q(G') \geq q(G).$$

This adequately gives quality increments **provided** α_p, β_p **are not both close to 1**.

(The argument generalizes easily to weighted graphs - DS problem has $\phi(q)/q$ weights on vertices.)

Difficult case

Question

What if $\alpha_p, \beta_p \approx 1$*?*

This is a more technical situation, and we **cannot** obtain a quality increment in general.

Difficult case

Question

What if $\alpha_p, \beta_p \approx 1$*?*

This is a more technical situation, and we **cannot** obtain a quality increment in general.

This is because the model problem has a counterexample!

Example (S. Chow)

If $\mathcal{S} = \{n!/j : j \in [n/2, n]\} \subseteq [(n-1)!, 2(n-1)!]$ *then:*

- *All pairs* $s_1, s_2 \in \mathcal{S}$ *have* $\gcd(s_1, s_2) \geq n!/n^2$.
- *There is no* $d \in \mathbb{Z}$ *of size* $\gg (n-1)!/\#\mathcal{S}$ *which divides a positive proportion of elements.*

Difficult case II

Question

So what do we do?

Difficult case II

Question

So what do we do?

The $\phi(q)/q$ weights save us!

These dampen down the contribution of numbers with many small prime factors, which is what happens when $\alpha_p, \beta_p \approx 1$ for many primes p.

Difficult case II

Question

So what do we do?

The $\phi(q)/q$ weights save us!

These dampen down the contribution of numbers with many small prime factors, which is what happens when $\alpha_p, \beta_p \approx 1$ for many primes p.

$$\sum_{q,r\in E_{end}} \frac{\phi(q)}{q}\frac{\phi(r)}{r} \leq \frac{\phi(a)}{a}\frac{\phi(b)}{b}\frac{x}{a}\frac{x}{b}e^{-t}$$

We gain a factor $\phi(a)\phi(b)/(ab)$. This allows us to lose a factor of $(1-1/p)^2$ in quality if we restrict to V_p and W_p.

Difficult case II

Question

So what do we do?

The $\phi(q)/q$ weights save us!

These dampen down the contribution of numbers with many small prime factors, which is what happens when $\alpha_p, \beta_p \approx 1$ for many primes p.

$$\sum_{q,r\in E_{end}} \frac{\phi(q)}{q}\frac{\phi(r)}{r} \le \frac{\phi(a)}{a}\frac{\phi(b)}{b}\frac{x}{a}\frac{x}{b}e^{-t}$$

We gain a factor $\phi(a)\phi(b)/(ab)$. This allows us to lose a factor of $(1-1/p)^2$ in quality if we restrict to V_p and W_p.

- To account for this, we use a different (more technical) definition of quality.
- This change in argument means the iteration needs to be done it two stages.
- This difficult situation is precisely that of the counterexample.

Summary

The key argument is with ‘GCD graphs’.

- We define $q(G)$ which captures a useful notion of ‘structure’.

Summary

The key argument is with ‘GCD graphs’.

- We define $q(G)$ which captures a useful notion of ‘structure’.
- We repeatedly pass to subsets of the vertices which are all a multiple of some prime p, or all coprime to p. A ‘compression argument’.

Summary

The key argument is with 'GCD graphs'.

- We define $q(G)$ which captures a useful notion of 'structure'.
- We repeatedly pass to subsets of the vertices which are all a multiple of some prime p, or all coprime to p. A 'compression argument'.
- We do this in a way that increases $q(G)$ at each iteration.

Summary

The key argument is with 'GCD graphs'.

- We define $q(G)$ which captures a useful notion of 'structure'.
- We repeatedly pass to subsets of the vertices which are all a multiple of some prime p, or all coprime to p. A 'compression argument'.
- We do this in a way that increases $q(G)$ at each iteration.
- When this process finaly terminates the resulting graph is very simple to analyse since all vertices on the left are a multiple of a and all on the right a multiple of b and GCDs are $\gcd(a, b)$.

Summary

The key argument is with 'GCD graphs'.

- We define $q(G)$ which captures a useful notion of 'structure'.
- We repeatedly pass to subsets of the vertices which are all a multiple of some prime p, or all coprime to p. A 'compression argument'.
- We do this in a way that increases $q(G)$ at each iteration.
- When this process finaly terminates the resulting graph is very simple to analyse since all vertices on the left are a multiple of a and all on the right a multiple of b and GCDs are $\gcd(a, b)$.
- This essentially puts us in the Erdős-Vaaler situation and so compatible with bounds from the 'anatomy of integers'.

Summary

The key argument is with ‘GCD graphs’.

- We define $q(G)$ which captures a useful notion of ‘structure’.
- We repeatedly pass to subsets of the vertices which are all a multiple of some prime p, or all coprime to p. A ‘compression argument’.
- We do this in a way that increases $q(G)$ at each iteration.
- When this process finaly terminates the resulting graph is very simple to analyse since all vertices on the left are a multiple of a and all on the right a multiple of b and GCDs are $\gcd(a, b)$.
- This essentially puts us in the Erdős-Vaaler situation and so compatible with bounds from the ‘anatomy of integers’.
- Extra care is needed when almost all elements are a multiple of many small primes p. In this case the $\phi(q)/q$ weights which are natural in the DS problem save us!

Questions

Thank you for listening.