The concentration of divisors

Kevin Ford
(joint work with Ben Green, Dimitrios Koukoulopoulos)

March 13, 2021

Random integers

How are the **prime factors** of a random integer $n \leqslant x$ distributed ?

(Erdős; Kubilius). Random integer $n \leqslant x$, $1 < a < b \leqslant \log\log x$. Then

$$\#\{p|n : e^{e^a} < p \leqslant e^{e^b}\} \approx \text{Poisson}(b-a)$$

and approx. independent for disjoint intervals $(a,b]$.

Idea: $\forall$ prime p, $\mathbb{P}(p|n) \approx 1/p$. By Mertens,

$$\sum_{e^{e^a} < p \leqslant e^{e^b}} \frac{1}{p} \approx b-a.$$

Divisors of a random integer $n \leqslant x$

Look at Divisors on a log-scale

- About $\log \log t$ prime factors below t;
- About $2^{\log \log t} = (\log t)^{\log 2}$ divisors below t;
- Prime factors $p_1, p_2, \ldots, p_k$. Then

$$\mathcal{D}_n := \{\log d : d|n\} = \{0, \log p_1\} + \{0, \log p_2\} + \cdots + \{0, \log p_k\}.$$

Much more complicated distribution.

Erdős' conjecture (1948)

Conjecture: Almost all n have two divisors d, d' with $d < d' \leqslant 2d$.

Heuristic: $\mathcal{D}_n - \mathcal{D}_n = \{\log(d'/d) : d|n, d'|n, (d, d') = 1\}$ has $3^k \approx (\log n)^{\log 3}$ elements, all in $[-\log n, \log n]$. There should be elements near 0 because $\log 3 > 1$.

Erdős' conjecture (1948)

Conjecture: Almost all n have two divisors d, d' with $d < d' \leqslant 2d$.

Heuristic: $\mathcal{D}_n - \mathcal{D}_n = \{\log(d'/d) : d|n, d'|n, (d,d') = 1\}$ has $3^k \approx (\log n)^{\log 3}$ elements, all in $[-\log n, \log n]$. There should be elements near 0.

The Erdős-Hooley Delta function

$\Delta(n) := \max_{u \geqslant 0} |\mathcal{D}_n \cap [u, u+1]| = \max_{u \geqslant 0} |\{d|n : e^u \leqslant d \leqslant e^{u+1}\}|.$

Erdős' conjecture roughly equivalent to $\Delta(n) \geqslant 2$ for most n.

Hooley (1979 paper): motivated by applications to Diophantine equations; further applications - Vaughan, Browning, la Bretèche, ...

$$\Delta(n) := \max_{u \geqslant 0} |\mathcal{D}_n \cap [u, u+1]| = \max_{u \geqslant 0} |\{d|n : e^u \leqslant d \leqslant e^{u+1}\}|.$$

Work of Maier-Tenenbaum

Maier-Tenenbaum, 1984. Erdős' conjecture is true. Moreover, for most n, $\Delta(n) \geqslant (\log\log n)^{H_1 - o(1)}$, where

$$H_1 = -\frac{\log 2}{\log(1 - 1/\log 3)} = 0.28754\ldots$$

Maier-Tenenbaum, 2009. For almost all n,

$$(\log\log n)^{H_2 + o(1)} \leqslant \Delta(n) \leqslant (\log\log n)^{\log 2 + o(1)},$$

where

$$H_2 = \frac{\log 2}{\log(\frac{1 - 1/\log 27}{1 - 1/\log 3})} = 0.33827\ldots$$

MT conjectured that $\Delta(n) = (\log\log n)^{H_2 + o(1)}$ for most n.

Maier-Tenenbaum main ideas

1. Focus of prime factors in $J = (y', y]$, where

$$y' = \exp\left\{(\log y)^{c-\varepsilon}\right\}, \quad c = 1 - \frac{1}{\log 3} = 0.089760\ldots$$

   Let $\mathcal{D}_n(J) := \{\log d : d|n, p|d \Rightarrow p \in J\}$. Show that $\mathcal{D}_n(J) - \mathcal{D}_n(J)$ nicely distributed in $[-\log n, \log n]$.

2. Use a small number of larger primes to fill any gap in $\mathcal{D}_n(J) - \mathcal{D}_n(J)$ near 0. Get $n_J = \prod_{p \in J, p|n} p$ has two close divisors.

3. Apply the above argument to many disjoint intervals $J_i = [y_i', y_i]$ to generate many close divisors of n. Get 1984 bounds.

4. 2009: Exploit the "unused primes" in J_j $(j < i)$ to augment the argument in (1), (2). Succeed with shorter intervals

$$(y'', y], \quad y'' = \exp\left\{(\log y)^{\theta-\varepsilon}\right\}, \quad \theta = \frac{1 - 1/\log 3}{1 - 1/\log 27} = 0.128857\ldots$$

New model (F, Green, Koukoulopoulos; 2019+)

$\#\{p|n, e^k < p \leqslant e^{k+1}\} \approx \text{Poisson}(1/k) \approx \text{Bernouilli}(1/k).$

Consider a random subset $\mathscr{A}$ of $\{1, 2, \ldots, N\}$, where

$$\mathbb{P}(k \in \mathscr{A}) = 1/k.$$

$$\mathscr{A} \;\leftrightarrow\; \{\log p : p|n\} \text{ for random } n \leqslant \mathrm{e}^N.$$

$$\Delta(n) \;\leftrightarrow\; F(\mathscr{A}) := \max_m \#\Big\{A \subset \mathscr{A} : \sum_{a \in A} a = m\Big\}$$

Example: $\mathscr{A} = \{1, 2, 4, 5, 7\}$. Then $F(\mathscr{A}) = 3$, corresponding to $k = 7$ or $k = 12$, e.g.

$$7 = 7 = 5 + 2 = 4 + 2 + 1.$$

(Setup) $\mathscr{A}$ is a random, harmonic weighted, subset of $\{1, \ldots, N\}$.

$$F(\mathscr{A}) := \max_m \#\Big\{ A \subset \mathscr{A} : \sum_{a \in A} a = m \Big\}$$

Corespondence $a \leftrightarrow \log p$, $\sum a \leftrightarrow \log d$.

Thm (FGK, 2019+). Let $\zeta = 0.3533227\ldots$ (a specific number). Then

$$F(\mathscr{A}) \geqslant (\log N)^{\zeta - o(1)} \quad \text{with prob.} \to 1 \text{ as } N \to \infty.$$

Corollary: For almost all n, $\Delta(n) \geqslant (\log\log n)^{\zeta - o(1)}$.

Compare with MT ('09) : $\Delta(n) \geqslant (\log\log n)^{0.33827\ldots - o(1)}$.

Conjecture (FGK). For most n, $\Delta(n) = (\log\log n)^{\zeta + o(1)}$.

Theorem (FGK, 2019+). Let

$$\beta_k := \sup\left\{c : F\big(\mathscr{A} \cap [N^c, N]\big) \geqslant k \text{ with prob.} \to 1 \text{ as } N \to \infty\right\}.$$

Then

$$\limsup_{k\to\infty} \frac{\log k}{\log(1/\beta_k)} \geqslant \zeta = 0.3533\dots$$

This corresponds to: Maximize c so that $n_J = \left\{d|n : p|d \Rightarrow p \in J\right\}$ has k close divisors for almost all n, with

$$J = \big(\exp\{(\log y)^c\}, y\big].$$

$$\beta_k := \sup\left\{c : F\big(\mathscr{A} \cap [N^c, N]\big) \geqslant k \text{ with prob.} \to 1 \text{ as } N \to \infty\right\}.$$

Theorem: FGK, 2019+ For each $k \geqslant 2$, let α_k be the supremum of numbers α so that almost all n have k divisors in

$$\left(y, y + \frac{y}{(\log y)^{\alpha}}\right]$$

for some y. Then, for all k, $\alpha_k \geqslant \dfrac{\beta_k}{1 - \beta_k}$.

Erdős-Hall,Maier-Tenenbaum: $\beta_2 = \frac{\log 3 - 1}{\log 3}$, $\alpha_2 = \log 3 - 1 \approx 0.0986$

M-T: For $2^{m-1} < k \leqslant 2^m$, $\quad \frac{\log 2}{k+1} \geqslant \alpha_k \geqslant \frac{(\log 3 - 1)^m 3^{m-1}}{(\log 27 - 1)^{m-1}} \approx k^{-1/0.33287}$

Conjecture (FGK): $\alpha_k = \frac{\beta_k}{1 - \beta_k}$.

Theorem (FGK, 2021+)

We have

$$\beta_3 = \frac{\log 3 - 1}{\log 3 + \frac{1}{\rho_1}} = 0.02616218\ldots \approx \frac{1}{38.223} \Rightarrow \alpha_3 \geqslant \frac{1}{37.223}$$

and

$$\beta_4 = \frac{\log 3 - 1}{\log 3 + \frac{1}{\rho_1} + \frac{1}{\rho_1\rho_2}} = 0.01295186\ldots \approx \frac{1}{77.209} \Rightarrow \alpha_4 \geqslant \frac{1}{76.209}$$

where

$$\rho_1 = \frac{\log\left(\frac{2}{\mathrm{e}-1}\right)}{\log(3/2)}, \quad \rho_2 = \frac{\log\left(\frac{2}{\mathrm{e}-1}\right)}{1 + \log\left(\frac{2}{\mathrm{e}-1}\right) - \log(1 + 2^{1-\rho_1})}.$$

Further,

$$\beta_k, \alpha_k \gtrsim k^{-1/0.3533\ldots}.$$

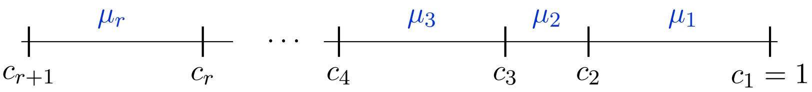
Problem: Find largest c so w.h.p., $\mathscr{A} \cap [N^c, N]$ has k equal subset sums.

Suppose $\sum_{a \in A_1} a = \cdots = \sum_{a \in A_k} a.$

Let B_ω, for $\omega \subseteq \{1, \ldots, k\}$, be the Venn diagram pieces.

Data to be optimized:

- Threshholds $c = c_{r+1} < c_r < \cdots < c_1 = 1.$
- Measures $\mu_1, \ldots, \mu_r$ where $\mu_j(\omega) = \dfrac{|B_\omega \cap (N^{c_{j+1}}, N^{c_j}]|}{|\mathscr{A} \cap (N^{c_{j+1}}, N^{c_j}]|}$

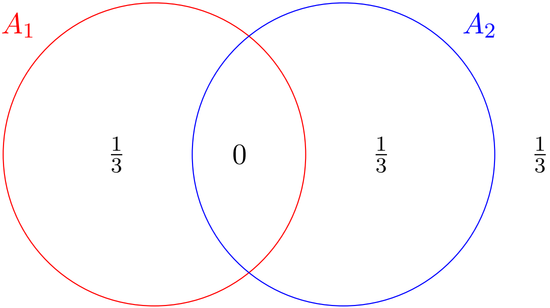


Maier-Tenenbaum: Optimal (uniform) measure for $k = 2$; sub-optimal choices for all $k > 2$.

Optimal choices for $k = 2$ (Maier-Tenenbaum, 1984)

$c_2 = 1 - \frac{1}{\log 3} \approx \frac{1}{11.14}$

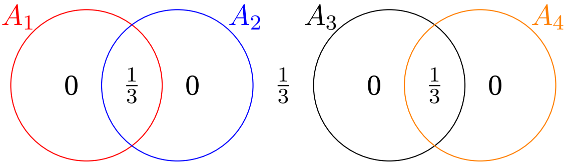
Optimal measure μ_1:



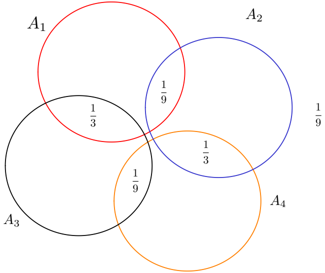
Choices for $k = 4$ (Maier-Tenenbaum, 2009)

$c_3 = c_2(1 - \frac{1}{\log 3}) \approx \frac{1}{86.457}$ $\quad$ $c_2 = \frac{1 - \frac{1}{\log 3}}{1 - \frac{1}{\log 27}} \approx \frac{1}{7.76}$

Optimal μ_1:

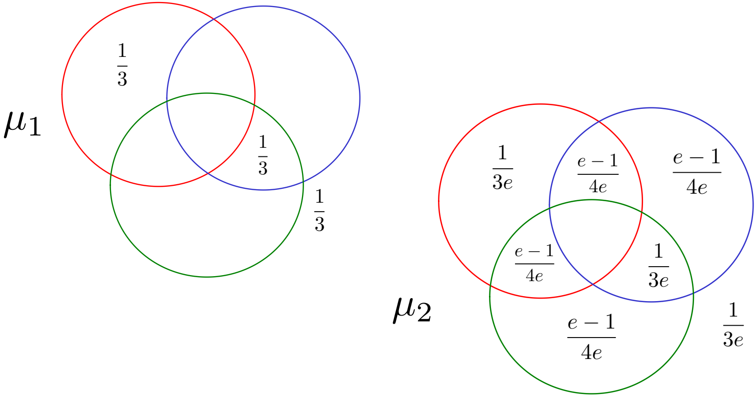


μ_2



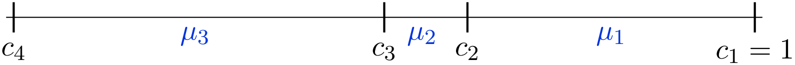
$k = 3$ case, new analysis - FGK

$c_3 = 0.026162\ldots \approx \frac{1}{38.223}$ $\qquad$ $c_2 = 0.2987701\ldots$



$k = 4$ case, new analysis - FGK

$c_4 = 0.012951\ldots \approx \frac{1}{77.2}$ $c_3 = 0.147909\ldots$ $c_2 = 0.152555\ldots$



- Identify parts of Venn diagram with $\omega \in \{0,1\}^4 =: \Omega$.
- μ_3 is nonzero on *every piece* of the Venn diagram except $A_1 \cap A_2 \cap A_3 \cap A_4$; that is, $\omega = 1111$.
- μ_2 is supported on $\mathrm{Span}_{\mathbb{Q}}(1111, 1010, 0001) \cap \Omega$;
- μ_1 is supported on $\mathrm{Span}_{\mathbb{Q}}(1111, 1010) \cap \Omega = (0000, 1111, 1010, 0101)$; mass 1/3 each on 1010, 0101 and 0000.

Linear algebra: some details

$$\sum_{a\in\mathcal{A}_1} a = \cdots = \sum_{a\in\mathcal{A}_k} a$$

1. Let $V_0 = \mathrm{Span}_{\mathbb{Q}}(11\cdots 1)$. Pieces $V_0 \cap \Omega$ don't matter.
2. Let $\omega_1 \in \Omega$ be the piece , not in V_0, containing the largest element, a_1, of $S = A_1 \cup \cdots A_k$. Let $V_1 = \mathrm{Span}_{\mathbb{Q}}(11\cdots 1, \omega_1)$. WHP $a_1 \approx N$.
3. Let ω_2 be the piece, not in V_1 with the largest element $a_2 = N^{c_2}$; Let $V_2 = \mathrm{Span}_{\mathbb{Q}}(11\cdots 1, \omega_1, \omega_2)$.
4. continue in this way, finishing with a sequence

$$c_{r+1} < c_r < \cdots < c_1 = 1$$

and a flag of vector spaces

$$\mathscr{V} : V_0 \leqslant V_1 \leqslant \cdots \leqslant V_r \leqslant \mathbb{Q}^k.$$

Definitions: For measure μ, subspace $W \leqslant \mathbb{Q}^k$, define the *entropy*

$$\mathbb{H}_\mu(W) := \sum_x \mu(x) \log \frac{1}{\mu(W+x)}$$

We say that $\mathscr{V}' : V_0' \leqslant V_1' \leqslant \cdots \leqslant V_r'$ is a *subflag* of $\mathscr{V}$ if $V_j' \leqslant V_j \quad \forall j$.

Entropy condition: Given a subflag $\mathscr{V}'$ of $\mathscr{V}$, let

$$e(\mathscr{V}') := \sum_{j=1}^{r} \Big[(c_j - c_{j+1}) \mathbb{H}_{\mu_j}(V_j') + c_j \dim(V_j'/V_{j-1}') \Big].$$

Def: γ_k is the supremum of $c > 0$ such that $\exists \; c_j, \mu_j, \mathscr{V}$ with

$$c_{r+1} = c, \qquad e(\mathscr{V}') \geqslant e(\mathscr{V}) \qquad (\forall \text{ subflags } \mathscr{V}').$$

Theorem (FGK, 2019+)

We (almost) have $\beta_k = \gamma_k$.

Where the constant $\zeta = 0.3533\ldots$ comes from

For $k = 2^r$ we consider a special *binary flag of order* r:
Identify $\mathbb{Q}^k$ with $\mathbb{Q}^{\mathcal{P}[r]}$, and for $i = 1, \ldots, r$ let V_i be the subspace of all $(x_S)_{S \subseteq [r]}$ for which $x_S = x_{S \cap [i]}$ for all $S \subseteq [r]$.

Theorem (FGK, 2019+)

We have $\beta_{2^r} \geqslant (\rho/2)^{r+o(1)}$, where $\rho = 0.28121134969637466015\ldots$ is the unique solution in $(0,1)$ of the equation

$$\frac{1}{1-\rho/2} = \log 2 + \sum_{j=1}^{\infty} \frac{1}{2^j} \log \Big(\frac{a_{j+1} + a_j^\rho}{a_{j+1} - a_j^\rho} \Big),$$

where the sequence a_j is defined by

$$a_1 = 2, \quad a_2 = 2 + 2^\rho, \quad a_j = a_{j-1}^2 + a_{j-1}^\rho - a_{j-2}^{2\rho} \qquad (j \geqslant 3).$$