The Erdős–Hooley Delta function

Dimitris Koukoulopoulos

Université de Montréal

32èmes Journées Arithmétiques
Nancy, 4 juillet 2023

The multiplicative structure of random integers

Question

Choose $n \in [1, x]$ uniformly at random. What is the distribution of its divisors?

Problem

Divisibility by different divisors could have dependencies due to common prime factors.

Easier question

What is the distribution of the set of prime factors $\{p|n\}$ of a randomly chosen n?

Warm-up: scale calibration

Prime factors

$$\mathbb{E}_{n\leqslant x}\left[\sum_{p|n,\ p\in[y,z]} 1\right] = \sum_{p\in[y,z]} \mathbb{P}_{n\leqslant x}(p|n) \sim \sum_{p\in[y,z]} \frac{1}{p} \sim \log\log z - \log\log y$$

Divisors

$$\mathbb{E}_{n\leqslant x}\left[\sum_{d|n,\ d\in[y,z]} 1\right] \sim \sum_{d\in[y,z]} \frac{1}{d} \sim \log z - \log y$$

The distribution of the prime factors

Prime factors form a Poisson Process

Let $I_1, \ldots, I_k$ be disjoint subintervals of $[1, x]$. Then

$$\mathbb{P}_{n \leqslant x}\Big(\#\{p|n,\ p \in I_j\} = m_j\ \forall j\Big) \approx \prod_{j=1}^{k} e^{-\lambda_j} \frac{\lambda_j^{m_j}}{m_j!}$$

with $\lambda_j = \sum_{p \in I_j} \frac{1}{p} \sim \log\log b_j - \log\log a_j \quad \text{if } I_j = [a_j, b_j]$.

Prime factors form a Brownian motion when normalized

$$\rho_N : [0,1] \to \mathbb{R}, \qquad \rho_N(t) := \frac{\#\big\{p|n : \log\log p \leqslant t \log\log x\big\} - t\log\log x}{\sqrt{t \log\log x}}$$

Then ρ_N converges in distribution to the *Brownian motion* in $[0, 1]$.

A cautionary tale about divisors

Theorem (Tenenbaum 1980)

If $\mathcal{N}$ is any **positive** density set of integers, then there is **no** weak limit for the distributions

$$F_n(u) := \frac{\#\{d|n : d \leqslant n^u\}}{\#\{d|n\}} \quad \text{as } n \to \infty \text{ over elements of } \mathcal{N}$$

Reason: integers with a divisor in a given short interval are sparse

$$\mathbb{P}_{n \leqslant x}\left(\exists d|n,\ n^u \leqslant d \leqslant n^{u+\varepsilon}\right) \asymp \varepsilon^c (\log(1/\varepsilon))^{-3/2} \qquad \text{(Ford 2008)}$$

with $c = 1 - \frac{1+\log\log 2}{\log 2} \approx 0.08607 > 0$.

Deeper reason

Small irregularities in distribution of large prime factors cause divisors to cluster (and thus to avoid a given target interval)

The DDT theorem

If $F_n(u) = \#\{d|n,\ d \leqslant n^u\}/\#\{d|n\}$, then

$$\mathbb{E}_{n\leqslant x}\Big[F_n(u)\Big] = \frac{2}{\pi}\arcsin\sqrt{u} + O\left(\frac{1}{\sqrt{\log x}}\right)$$

Various proofs:

- Deshouillers–Dress–Tenenbaum (1979)
- Sun Kai Leung (2022+)
- Hirth (1997)
- Haddad-K. (2023+) $\rightsquigarrow$ talk by Tony Haddad on Thursday

Measuring the concentration/clustering of divisors

$$\Delta(n) := \max_{u\in\mathbb{R}} \#\{d|n : e^u < d \leqslant e^{u+1}\}$$

Typical size of Δ? Average size? Other statistics?

Hooley's "new technique" (1979)

$$\sum_{n\leqslant x} \Delta(n) \ll x(\log x)^{4/\pi-1} \qquad (4/\pi - 1 \approx 0.273)$$

Compare to *trivial bound*

$$\sum_{n\leqslant x} \Delta(n) \leqslant \sum_{n\leqslant x} \#\{d|n\} \sim x\log x$$

- savings $\rightsquigarrow$ applications to counting solutions of Diophantine equations & inequalities (Hooley, Brüdern, La Bretèche–Tenenbaum, Sofos, Lartaux,...)

The typical size of Δ

Erdős's conjecture (1948): $\Delta(n) > 1$ for almost all n

Lower bounds: $\Delta(n) \geqslant (\log\log n)^{c+o(1)}$ a.s. with…

- Maier–Tenenbaum (1984): $c = -\log 2/\log(1-1/\log 3) \approx 0.28754$
- MT (2009): $c = \log 2/\log(\frac{1-1/\log 27}{1-1/\log 3}) \approx 0.33827$
- Ford–Green–K. (2023): $c = \eta \approx 0.35332$ (η has a precise theoretical dfn)

Upper bounds: $\Delta(n) \leqslant (\log\log n)^{C+o(1)}$ a.s. with…

- Maier–Tenenbaum (1985): $C = 1$
- Maier–Tenenbaum (2009): $C = \log 2$
- La Bretèche–Tenenbaum (2022+): $C = \frac{\log 2}{\log 2 + 1/\log 2 - 1} \approx 0.6102$

Conjecture (Ford–Green–K.): $\Delta(n) = (\log\log n)^{\eta+o(1)}$ for a.a. n

The average size of Δ

$$\delta(x) := \frac{1}{x} \sum_{n \leqslant x} \Delta(n)$$

Upper bounds

- Hooley (1979): $\delta(x) \ll (\log x)^{4/\pi - 1}$
- Hall–Tenenbaum (1986): $\delta(x) \ll_\varepsilon \exp\left((\sqrt{2} + \varepsilon)\sqrt{\log\log x \log\log\log x}\right)$
- La Bretèche–Tenenbaum (2022+): $\delta(x) \ll_\varepsilon \exp\left((\sqrt{2}\log 2 + \varepsilon)\sqrt{\log\log x}\right)$
- K.–Tao (2023+): $\delta(x) \ll (\log\log x)^{11/4}$

Lower bounds

- Hall–Tenenbaum (1984): $\delta(x) \gg \log\log x$

The Maier–Tenenbaum method with an FGK twist

Goal: show that $\Delta(n)$ is large for a typical n

Strategy

- Fix an integer $k \geqslant 2$ (MT take $k = 2$);
- Consider appropriate checkpoints $y_0 < y_1 < y_2 < \cdots < y_J$ (they depend on k);
- Let n_j be the part of n composed of all its primes factors in $(y_{j-1}, y_j]$;
- $\forall j = 1, \ldots, J$, find k **distinct** $d_{j,1}, \ldots, d_{j,k} | n_j$ s.t. $\log \max_i d_{j,i} < \log \min_i d_{j,i} + 1$.

Tensor trick...

- n has at least k^J distinct divisors $D_1, \ldots, D_{k^J}$ s.t. $\log \max D_m < \log \min D_m + J$;
- Thus $\Delta(n) \geqslant k^J / J$ by the pigeonhole principle.

Finding two divisors close together

Question: when does n_j have two divisors d, d' s.t. $0 < |\log(d'/d)| < 1$?

Heuristics

- The set $\mathcal{R}(n_j) := \{d'/d : (d, d') = 1,\ dd' | n_j\}$ has $\approx 3^{\omega(n_j)}$ elements;
- The set $\{\log(d'/d) : d'/d \in \mathcal{R}(n_j)\}$ is roughly well-distributed in $[-\log n_j, \log n_j]$;
- Then the needed pair (d, d') ought to exist if $3^{\omega(n_j)} > \log n_j$.
- Typically, $3^{\omega(n_j)} \approx (\log y_j / \log y_{j-1})^{\log 3}$ and $\log n_j \approx \log y_j$;
- So, the pair (d, d') ought to exist when $\log y_j > (\log y_{j-1})^{1/(1-1/\log 3)}$.

$$\rightsquigarrow \quad \Delta(n) > (\log \log n)^{-\frac{\log 2}{\log(1-1/\log 3)}} \quad \text{a.s.}$$

Finding k divisors close together

Strategy

- Consider $\mathcal{L}_k(n_j) := \bigcup\{(\log(d_1/d_k),\dots,\log(d_{k-1}/d_k)) + [0,1]^{k-1} : d_1,\dots,d_k|n\}$;
- Ensure $\mathcal{L}_k(n_j)$ is "as large as it can be"
- OK if $\log y_j > (\log y_{j-1})^{C_k}$ with C_k sufficiently large.

Main result (Ford–Green–K.)

We have $C_{2^r} \lessapprox (2/\rho)^r$, where $\rho = 0.28121\dots$ is the unique solution in $[0,1/3]$ of

$$\frac{1}{1-\rho/2} = \lim_{j\to\infty} \frac{\log a_j}{2^{j-2}},$$

where the sequence a_j is defined by

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

The average value of Δ

Theorem (K.–Tao)

$$\sum_{n\leqslant x}\Delta(n)\ll x(\log\log x)^{11/4}$$

Logarithmic weak $L^{1-o(1)}$ estimate

$$\mathbb{P}^{\log}_{n\leqslant x}\Big(\Delta(n)>\lambda\log\log x\Big):=\frac{1}{\log x}\sum_{\substack{n\leqslant x\\ \Delta(n)>\lambda\log\log x}}\frac{1}{n}\ll\frac{(\log\lambda)^{3/4}}{\lambda}.$$

Logarithmic weak $L^{1+\varepsilon}$ estimate

Fix $\varepsilon>0$ and assume that $\lambda>(\log x)^{\log 4-1+\varepsilon}$. There is a constant $c>0$ s.t.

$$\mathbb{P}^{\log}_{n\leqslant x}\Big(\Delta(n)>\lambda\log\log x\Big)\ll_\varepsilon\lambda^{-1-c\varepsilon^2}.$$

Ideas of proof: first approximation for Δ

- Let $\tau(a) := \#\{d|a\}$ and $\Delta(a;u) := \#\{d|a : e^u < d \leqslant e^{u+1}\}$. Then

$$\tau(a) = \int_{\mathbb{R}} \Delta(a;u)\mathrm{d}u = \int_{-1}^{\log a} \Delta(a;u)\mathrm{d}u \leqslant \Delta(a) \cdot (1 + \log a).$$

- Let $n_{\leqslant y}$ be the y-smooth/friable part of n. Then

$$\Delta(n) \geqslant \max_{2 \leqslant y \leqslant x} \Delta(n_{\leqslant y}) \geqslant \max_{2 \leqslant y \leqslant x} \frac{\tau(n_{\leqslant y})}{1 + \log n_{\leqslant y}} \approx \max_{2 \leqslant y \leqslant x} \frac{\tau(n_{\leqslant y})}{\log y}$$

- Fix A large and let $\mathcal{E} = \left\{ n \leqslant x : \max_{2 \leqslant y \leqslant x} \frac{\tau(n_{\leqslant y})}{\log y} > A \right\}$. Then

$$\mathbb{P}^{\log}_{\leqslant x}(\mathcal{E}) \asymp A^{-1}$$

Remark: $\mathcal{E}$ dominated by n's s.t. $\#\{p|n_{\leqslant y}\} \sim 2 \log\log y$ when $\log\log y \sim \frac{\log A}{\log 4 - 1}$.

Ideas of proof: bootstrapping to high moments

- $\mathcal{N}_1 := \{n \leqslant x\} \setminus \mathcal{E} \quad \Longrightarrow \quad \mathbb{P}^{\log}_{\leqslant x}\Big(\mathcal{N}_1\Big) = 1 - O\Big(\frac{1}{A}\Big).$

- $\Delta(n) \approx \mu_q(n)^{\frac{1}{q}} \qquad \text{with} \qquad q = \log\log x, \quad \mu_q(n) := \frac{1}{\tau(n)} \int_{\mathbb{R}} \Delta(n; u)^q \mathrm{d}u.$

- $\sum_{n \in \mathcal{N}_1} \frac{\mu_2(n)}{n} \lessapprox \log x \quad \overset{\text{Markov}}{\Longrightarrow} \quad \mathbb{P}^{\log}_{n \leqslant x}\Big(\underbrace{n \in \mathcal{N}_1, \ \mu_2(n) \leqslant A}_{=:\mathcal{N}_2}\Big) = 1 - O\Big(\frac{1}{A}\Big)$

- Iterating...

$$\sum_{n \in \mathcal{N}_{j-1}} \frac{\mu_j(n)}{n} \lessapprox (j-2)! A^{j-2} \log x \quad \overset{\text{Markov}}{\Longrightarrow} \quad \mathbb{P}^{\log}_{n \leqslant x}\Bigg(\underbrace{\begin{matrix} n \in \mathcal{N}_{j-1}, \\ \mu_j(n) \leqslant j! A^{j-1} \end{matrix}}_{=:\mathcal{N}_j}\Bigg) = 1 - O\Big(\frac{1}{j^2 A}\Big)$$

Ideas of proof: the inductive step

Let $p \nmid m$. Then $\Delta(pm; u) = \Delta(m; u) + \Delta(m; u - \log p)$

…and thus

$$\mu_j(pm) = \mu_j(m) + \sum_{\substack{a+b=j \\ 1 \leqslant a \leqslant j-1}} \binom{j}{a} \int_{\mathbb{R}} \frac{\Delta(m; u)^a \Delta(m; u - \log p)^b}{2\tau(m)} \mathrm{d}u$$

$$\leqslant \mu_j(m) + \sum_{\substack{a+b=j \\ 1 \leqslant a \leqslant j/2}} \binom{j}{a} \int_{\mathbb{R}} \frac{\Delta(m; u)^a \Delta(m; u - \log p)^b}{\tau(m)} \mathrm{d}u$$

$$\lesssim \mu_j(m) + O\bigg(\frac{\tau(m)}{\log p} \sum_{\substack{a+b=j \\ 1 \leqslant a \leqslant j/2}} \binom{j}{a} \mu_a(m) \mu_b(m) \bigg)$$

Thank you for your attention