Towards a high-dimensional theory of divisors of integers

Dimitris Koukoulopoulos

Université de Montréal

Number Theory Web Seminar
28 October 2021

Anatomical investigations

- Choose an integer $n \in [1, x]$ uniformly at random. What can we say about its “multiplicative structure”?
- Distribution of its prime factors $p_1 < p_2 < \cdots$?
- Distribution of its divisors $d_1 < d_2 < \cdots$?

$p_1(n) < \cdots < p_k(n)$ prime factors of n, with $k = \omega(n) = \#\{p|n\}$

- Hardy–Ramanujan: $\omega(n) \sim \log\log x$ for a "typical" $n \leqslant x$.
- Landau–Selberg–Delange: $\omega(n)$ has a *perturbed* Poisson distribution with parameter $\lambda = \log\log x$.
- If $I_1, \ldots, I_m$ are disjoint (+technical conditions), then the RVs $\#\{p|n : p \in I_j\}$ are approximately independent and Poisson with $\lambda_j = \sum_{p \in I_j} 1/p$.
- If $\ell \to \infty$, the vector $(\frac{\log\log p_i(n)}{\log\log x})_{i=\ell}^{k-\ell}$ is approximately distributed like $(\xi_i)_{i=\ell}^{k-\ell}$, where $0 \leqslant \xi_1 \leqslant \cdots \leqslant \xi_k \leqslant 1$ are uniform order statistics. In particular, $\log\log p_j(n) \sim j$ typically.
- Billingsley: The vector $(\frac{\log p_{k-i}(n)}{\log n})_{i=0}^{\ell-1}$ is approximately distributed like $(V_i)_{i=1}^{\ell}$, where $(V_1, V_2, \ldots)$ has Poisson–Dirichlet distribution.

Two problems by Erdős about divisors

Problem A (1948)

Is it true that almost all integers have two divisors $d < d' \leqslant 2d$?

- **Motivation:** understand local properties of the sequence of divisors $d_1(n) < d_2(n) < \cdots$ of a "random" integer n.
- We typically have $\log\log d_j(n) \sim \frac{\log j}{\log 2}$.
- Naively, one might then guess $\log \frac{d_{j+1}(n)}{d_j(n)} \approx j^{1/\log 2 - 1} \to \infty$.

Problem B (1955)

How many integers are there in the $N \times N$ multiplication table?

- **Generalization:** how many integers $\leqslant x$ have a divisor in $[y, z]$?
- When $x = N^2$, $y = N/2$, $z = N$, this is $\approx$ Problem B.

Two generalizations

Problem A*

How big can $k = k(x)$ be so that almost all integers $n \leqslant x$ have k divisors $d_1 < d_2 < \cdots < d_k \leqslant 2d_1$?

- **Reformulation:** $\Delta(n) := \max_y \#\{d|n : y \leqslant d \leqslant 2y\}$.
- How big is $\Delta(n)$ for a typical n?
- Hooley's original motivation: study $\sum_{n\leqslant x} \Delta(n)$ and apply results to various Diophantine problems.

Problem B*

How many integers are there in the $N_1 \times \cdots \times N_k$ multiplication table?

- **Reformulation:** how many integers $\leqslant x$ have a factorization $n = d_1 \cdots d_k$ with $d_j \in [y_j, z_j]$ for $j = 1, \ldots, k$?

Integers with two divisors close together

$$\mathcal{R}(n) = \bigcup_{d,d'|n,\ d\neq d'} [\log \tfrac{d'}{d} - \log 2, \log \tfrac{d'}{d} + \log 2]$$

- $\exists d, d'|n$ s.t. $d < d' \leqslant 2d \quad \Longleftrightarrow \quad 0 \in \mathcal{R}(n)$
- We have two competing constraints on measure of $\mathcal{R}(n)$:
  - Geometric: $\mathcal{R}(n) \subset [-\log(2n), \log(2n)]$;
  - Combinatorial: $\#\{d'/d : d, d'|n\} \approx 3^{\omega(n)} \approx (\log n)^{\log 3}$ (typically).
- Thus, unless there is too much overlap between different intervals, $\mathrm{meas}(\mathcal{R}(n)) \gg \log n$.
- For most n with $\mathrm{meas}(\mathcal{R}(n)) \gg \log n$, we may locate a ratio d'/d close to 1 w.h.p. (uses that $\mathcal{R}(mpp') \supset \mathcal{R}(m) + \log(p'/p)$)

Theorem (Maier–Tenenbaum (1984))

For almost all n, there are divisors $d < d' < d \cdot (1 + (\log n)^{1-\log 3+o(1)})$

Integers with many divisors close together

$$\mathcal{R}(n) = \bigcup_{d,d'|n,\ d\neq d'} [\log \tfrac{d'}{2d}, \log \tfrac{2d'}{d}], \quad n_{y,z} = \prod_{p|n,\ y<p\leqslant z} p.$$

- For typical n, we have the following competing constraints:
  - Geometric: $\mathcal{R}(n_{y,z}) \subset [-C\log z, C\log z]$;
  - Combinatorial: $\#\{d'/d : d, d'|n_{y,z}\} = 3^{\omega(n_{y,z})} \approx 3^{\log \frac{\log z}{\log y}}$.
- If $\log z > (\log y)^{\frac{\log 3}{\log 3-1}}$, we have more than $\log z$ ratios d'/d with $d, d'|n_{y,z}$, so we can find a ratio ≈ 1 w.h.p.
- Use J disjoint intervals $[y_j, z_j]$ to get 2^J divisors of n close together.

Theorem (Maier–Tenenbaum (1984))

$\Delta(n) \geqslant (\log\log n)^{H_1 - o(1)}$ *for a.a. n, where* $H_1 = \frac{\log 2}{\log \frac{\log 3}{\log 3-1}} = 0.28754\ldots$

Integers with a divisor in a given interval

- $\exists d|n,\ d \in [x^\theta, 2x^\theta] \iff \theta \log x \in \mathcal{L}(n) := \bigcup_{d|n} [\log \frac{d}{2}, \log d]$
- $\mathbb{P}_{n\leqslant x}\Big(\theta \log x \in \mathcal{L}(n)\Big) \overset{?}{\approx} \mathbb{E}_{n\leqslant x}\Big[\frac{\text{meas}(\mathcal{L}(n))}{\log x}\Big] \overset{??}{\approx} \mathbb{E}_{n\leqslant x}\Big[\min\{1, \frac{\tau(n)}{\log x}\}\Big]$
- Erdős–Tenenbaum: need $\omega(n) = \frac{\log\log x}{\log 2} + O(1)$, i.e. $\tau(n) \asymp \log x$.

$$\rightsquigarrow \quad \mathbb{P}_{n\leqslant x}\Big(\exists d|n,\ d \in [x^\theta, 2x^\theta]\Big) \approx (\log x)^{-\delta}(\log\log x)^{-1/2}$$

$$\text{with} \quad \delta = 1 - \frac{1+\log\log 2}{\log 2} = 0.08607\ldots$$

- Ford: we also need $\#\{p|n,\ p \leqslant t\} \leqslant \frac{\log\log t}{\log 2} + O(1)$ for all $t \leqslant x$.

Theorem (Ford (2004))

$$\mathbb{P}_{n\leqslant x}\Big(\exists d|n,\ d \in [y, 2y]\Big) \asymp (\log y)^{-\delta}(\log\log y)^{-3/2} \quad (3 \leqslant y \leqslant \sqrt{x}).$$

The k-dimensional multiplication table

$$h(x,\vec{y}) := \mathbb{P}_{n\leqslant x}\Big(\exists d_1\cdots d_{k-1}|n,\ d_i\in[y_i/2,y_i]\ \forall i\Big) =? \quad (y_1\leqslant y_2\leqslant\cdots)$$

- $\mathcal{L}_k(n;\vec{y}) := \bigcup_{\substack{d_1\cdots d_{k-1}|n \\ p|d_i\ \Rightarrow\ p\leqslant y_i}} [\log\frac{d_1}{2},\log d_1]\times\cdots\times[\log\frac{d_{k-1}}{2},\log d_{k-1}]$
- If $\omega_i(n)=\#\{p|n : y_{i-1}<p\leqslant y_i\}$ for $i=1,\ldots,k-1$, then

$$\mathrm{meas}(\mathcal{L}_k(n;\vec{y})) \overset{?}{\approx} \min\Big\{\prod_{i=1}^{k-1}(k-i+1)^{\omega_i(n)},\prod_{i=1}^{k-1}\log y_i\Big\}.$$

- If $\lambda_i=\log\log y_i-\log\log y_{i-1}$, the above suggests that

$$h(x,\vec{y}) \approx \max_{\substack{\vec{m}\in\mathbb{N}^{k-1} \\ \prod_{i=1}^{k-1}(k-i+1)^{m_i}\asymp\prod_{i=1}^{k-1}\log y_i}} \prod_{i=1}^{k-1}\frac{e^{-\lambda_i}\lambda_i^{m_i}}{m_i!}.$$

Theorem (K. (2014))

1. *If* $k \leqslant 6$, *then* $h(x,\vec{y}) \asymp \max\limits_{\substack{\vec{m}\in\mathbb{N}^{k-1} \\ \prod_{i=1}^{k-1}(k-i+1)^{m_i} \asymp \prod_{i=1}^{k-1}\log y_i}} \prod_{i=1}^{k-1} \frac{e^{-\lambda_i}\lambda_i^{m_i}}{m_i!} \quad (*)$.
2. $(*)$ *is true if* $k \geqslant 7$ *and* $\log y_{k-1} \leqslant (\log y_1)^{1+\varepsilon}$ *with* $\varepsilon = \varepsilon_k$ *small.*
3. *Let* $k = 7$, $y_1 = \cdots = y_5 \leqslant y_6$ *and* $c = 55.8247430495071898 6\ldots$
   If $\log y_6 \leqslant (\log y_1)^{c-\varepsilon}$, *then* $(*)$ *holds; but if* $\log y_6 \geqslant (\log y_1)^{c+\varepsilon}$, *then* $h(x,\vec{y})$ *is smaller than the RHS of* $(*)$.

Key observation: there are low-dimensional constraints of geometro-combinatorial nature:

- $$\mathcal{L}_7(n;\vec{y}) = \bigcup_{\substack{d_1\cdots d_6 \mid n \\ p\mid d_i \Rightarrow p\leqslant y_i\ \forall i\leqslant 6}} \prod_{i\leqslant 5}[\log \tfrac{d_i}{2}, \log d_i] \times [\log \tfrac{d_6}{2}, \log d_6]$$
- $$\mathrm{meas}(\mathcal{L}_7(n;\vec{y})) \leqslant \sum_{\substack{d_1\cdots d_5 \mid n \\ p\mid d_i \Rightarrow p\leqslant y_i\ \forall i\leqslant 5}} (\log 2)^5 \cdot \mathrm{meas}(\mathcal{L}_2(\tfrac{n}{d_1\cdots d_5}; y_6)).$$

Improving the 1984 Maier–Tenenbaum result

- MT 1984: suppose we can find J disjoint intervals $[y_j, z_j]$ for which $\exists d_j, d_j' | n$ distinct s.t.:
  - d_j, d_j' only consist of primes in $[y_j, z_j]$;
  - $d_j \approx d_j'$.

  Then n has 2^J prime factors close together.

- MT 2009: use primes in $[y_1, z_1]$ to find d_1, d_1'. Then, use primes in $[y_2, z_2]$ and remaining primes in $[y_1, z_1]$ to find d_2, d_2', etc.

$\Delta(n) \geqslant (\log\log n)^{H_2 - o(1)}$ a.s. with $H_2 = \frac{\log 2}{\log(\frac{1-1/\log 27}{1-1/\log 3})} = 0.33827\ldots$

- Ford–Green–K. (2019 $\rightarrow$ 2022?): locate J disjoint intervals $[y_j, z_j]$ s.t. the primes from each $[y_j, z_j]$ yield k distinct divisors $d_{j,1} \approx d_{j,2} \approx \cdots \approx d_{j,k}$.

$\Delta(n) \geqslant (\log\log n)^{H_3 - o(1)}$ a.s. with $H_3 = 0.3533227727\ldots$

The linear algebra of k divisors close together

- We want to understand when there are distinct $d_1, \ldots, d_k$ that all divide $n_{y,z} = \prod_{p|n,\ y<p\leqslant z} p$ and satisfy the linear system

$$\Big(\sum_{p|d_1} \log p, \sum_{p|d_2} \log p, \ldots, \sum_{p|d_k} \log p\Big) = O(1) \pmod{\vec{1}}.$$

- For each $\vec{\omega} = (\omega_1, \ldots, \omega_k) \in \{0,1\}^k$, let

$$D_{\vec{\omega}} := \prod_{p \in \mathcal{P}_{\vec{\omega}}} p \quad \text{with} \quad \mathcal{P}_{\vec{\omega}} = \{p | n_{y,z} \quad \text{s.t} \quad p|d_i \iff \omega_i = 1\},$$

so that $(\log d_1, \ldots, \log d_k) = \sum_{\vec{\omega}} \vec{\omega} \log D_{\vec{\omega}}$.

- Conditionally on $\Omega = \{\vec{\omega} \in \{0,1\}^k \setminus \{\vec{0}, \vec{1}\} : D_{\vec{\omega}} > 1\}$, the distribution of $(\log d_1, \ldots, \log d_k) \bmod \vec{1}$ is controlled by:
  - $V = \mathrm{Span}(\Omega)/\langle \vec{1} \rangle$ (geometric)
  - the distribution of $\{p | D_{\vec{\omega}}\}$ with $\vec{\omega} \in \Omega$ (combinatorial)

Geometric constraints

- The "longest" dim of $\sum_{\vec{\omega}\in\Omega}\vec{\omega}\log D_{\vec{\omega}}$ is onto $\vec{\omega}^1 \in \Omega$ s.t.
$$P^+(D_{\vec{\omega}^1}) \geqslant P^+(D_{\vec{\omega}}) \quad \forall\vec{\omega}\in\Omega$$
- The second longest dim is onto $\vec{\omega}^2 \in \Omega \setminus \mathrm{Span}(\vec{\omega}^1)$ s.t.
$$P^+(D_{\vec{\omega}^2}) \geqslant P^+(D_{\vec{\omega}}) \quad \forall\vec{\omega}\in\Omega \setminus \mathrm{Span}(\vec{\omega}^1).$$
- Having defined $\vec{\omega}^1,\dots,\vec{\omega}^j$ with span $V_j \leqslant V$, the $(j+1)$-th longest dimension is the projection onto $\vec{\omega}^{j+1} \in \Omega \setminus V_j$ s.t.
$$P^+(D_{\vec{\omega}^{j+1}}) \geqslant P^+(D_{\vec{\omega}}) \quad \forall\vec{\omega}\in\Omega \setminus V_j.$$
- This terminates after $r = \dim(V)$ steps. We then find that $\sum_{\vec{\omega}\in\Omega}\vec{\omega}\log D_{\vec{\omega}}$ are contained in an r-dim rectangle of r-volume
$$\lessapprox \prod_{1\leqslant j\leqslant r} \log P^+(D_{\vec{\omega}^j}).$$

- Given $c \in (0,1)$ and $k \in \mathbb{Z}_{\geqslant 2}$, consider the following data:
  - $\langle \vec{1} \rangle = V_0 \leqslant V_1 \leqslant \cdots \leqslant V_r \leqslant \mathbb{Q}^k$
  - $1 \geqslant c_1 \geqslant \cdots c_r \geqslant c$, where $\log y = (\log z)^c$
  - μ_i probability measure supported on $V_i \cap \{0,1\}^k$
- To construct $d_1, \dots, d_k$ close together, we consider configurations s.t.
  - $D_{\vec{\omega}} > 1$ if-f $\vec{\omega} \in V_r \setminus \{\vec{0}, \vec{1}\}$;
  - $P^+(D_{\vec{\omega}}) \leqslant \exp\{(\log z)^{c_j}\}$ for all $\omega \in V_j \cap \{0,1\}^k$;
  - $\#\{p | D_{\vec{\omega}} : c_{j+1} < \frac{\log\log p}{\log\log z} \leqslant c_j\} \sim \mu_j(\vec{\omega})(c_j - c_{j+1}) \log\log z$.
- To avoid all geometro-combinatorial constraints, we need $\forall V_j' \leqslant V_j$

$$\sum_{1 \leqslant j \leqslant r} (c_j - c_{j+1}) \mathbb{H}_{\mu_j}(V_j') + \sum_{1 \leqslant j \leqslant r} c_j \dim(V_j'/V_{j-1}') \geqslant \sum_{1 \leqslant j \leqslant r} c_j \dim(V_j/V_{j-1}),$$

where $\mathbb{H}_\mu(W)$ is the μ-entropy of the partition of $\mathbb{Q}^k$ into W-cosets.

Theorem (Ford–Green–K. (2019 $\to$ 2022?))

A random integer has k divisors close together composed of primes in $[\exp\{(\log z)^c, z]$ "iff" there are V_j, c_j, μ_j as above.

The binary system

Binary flag of order r

Let $k = 2^r$, 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 (Ford–Green–K. (2019 $\to$ 2022?))

A random integer has 2^r divisors close together composed of primes in $[\exp\{(\log z)^c, z]$ w.h.p. if $c \geqslant (\rho/2)^{r+o(1)}$, where $\rho = 0.2812\ldots$ is s.t.

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

where $a_1 = 2,\ a_2 = 2 + 2^\rho,\ a_j = a_{j-1}^2 + a_{j-1}^\rho - a_{j-2}^{2\rho}$ for $j \geqslant 3$.

Corollary

$\Delta(n) \geqslant (\log\log n)^{\frac{\log 2}{\log(2/\rho)}+o(1)}$ a.s. (conjectured to be optimal)

A unified point of view

Given a set of integers $\mathcal{N}$ and a linear map $\Psi : \mathbb{R}^\ell \to W$, where W is a real vector space, determine

$$\mathbb{P}_{n\in\mathcal{N}}\left(\begin{array}{l}\exists d_1,\dots,d_\ell | n \quad \text{s.t.} \quad \Psi(\log d_1,\dots,\log d_\ell) = \vec{p} + O(1) \\ \text{and the sets } \{p|d_i\},\ i=1,\dots,\ell, \text{ satisfy certain conditions}\end{array}\right).$$

- Problem A: $\mathcal{N} = \{\text{typical integers in } [1,x]\}$, $\ell = 2$, $W = \mathbb{R}$, $\Psi(t,s) = t - s$, $\vec{p} = \vec{0}$, $(d_1, d_2) = 1$.
- Problem B: $\mathcal{N} = \mathbb{Z} \cap [1,x]$, $\ell = 1$, $W = \mathbb{R}$, $\Psi(t) = t$, $\vec{p} = \log y$, d is y-smooth
- Problem B*: $\mathcal{N} = \mathbb{Z} \cap [1,x]$, $\ell = k-1$, $W = \mathbb{R}^{k-1}$, $\Psi(\vec{x}) = \vec{x}$, $\vec{p} = (\log y_i)_{i=1}^{k-1}$, $(d_i, d_j) = 1$ and d_i is y_i-smooth $\forall i \neq j$.
- Problem A*: $\mathcal{N} = \{\text{typical integers in } [1,x]\}$, $\ell = k$, $W = \mathbb{R}^k/\langle\vec{1}\rangle$, $\Psi(\vec{t}) = \vec{t}$, complicated constraints.

Some future directions

1. For each $k \geqslant 2$, calculate optimal α_k s.t. w.h.p. here are k divisors $d_1, \ldots, d_k$ of n with $|\log(d_i/d_j)| \leqslant (\log n)^{-\alpha_k+o(1)}$ $\forall i, j$

MT: $\alpha_2 = \log 3 - 1 \approx 0.09861$;
FGK (work in progress): $\alpha_3 \approx 0.026865$, $\alpha_4 \approx 0.0131218$.

2. Determine order of magnitude of k-dim multiplication table for $k \geqslant 7$ for all possible side-lengths.

3. Upper bounds for $\Delta(n)$.

Thank you for your attention