Towards a high-dimensional theory of divisors of integers

Dimitris Koukoulopoulos

Université de Montréal

Number Theory Web Seminar
28 October 2021
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” \( n \leq x \).

- **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 \( \left( \frac{\log \log p_{i}(n)}{\log \log x} \right)^{k-\ell}_{i=\ell} \) is approximately distributed like \( (\xi_i)^{k-\ell}_{i=\ell} \), where \( 0 \leq \xi_1 \leq \cdots \leq \xi_k \leq 1 \) are uniform order statistics. In particular, \( \log \log p_j(n) \sim j \) typically.

- **Billingsley**: The vector \( \left( \frac{\log p_{k-i}(n)}{\log n} \right)^{\ell-1}_{i=0} \) is approximately distributed like \( (V_i)^{\ell}_{i=1} \), 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 \( d \leq d' \leq 2d ? \)

- **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 \( \log \frac{d_{j+1}(n)}{d_j(n)} \approx j^{1/(\log 2 - 1)} \rightarrow \infty \).

Problem B (1955)
How many integers are there in the \( N \times N \) multiplication table?

- **Generalization:** how many integers \( \leq x \) 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 $n \leq x$ have $k$ divisors $d_1 < d_2 < \cdots < d_k \leq 2d_1$?

- **Reformulation:** $\Delta(n) := \max_y \#\{d|n : y \leq d \leq 2y\}$.
- How big is $\Delta(n)$ for a typical $n$?
- Hooley’s original motivation: study $\sum_{n \leq x} \Delta(n)$ 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 $\leq x$ 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

\[ R(n) = \bigcup_{d,d' \mid n, d \neq d'} [\log \frac{d'}{d} - \log 2, \log \frac{d'}{d} + \log 2] \]

- \[ \exists d, d' \mid n \text{ s.t. } d < d' \leq 2d \iff 0 \in R(n) \]

- We have two competing constraints on measure of \( R(n) \):
  - Geometric: \( R(n) \subset [-\log(2n), \log(2n)] \);
  - Combinatorial: \( \#\{d'/d : d, d' \mid n\} \approx 3^{\omega(n)} \approx (\log n)^{\log 3} \) (typically).

- Thus, unless there is too much overlap between different intervals, \( \text{meas}(R(n)) \gg \log n \).

- For most \( n \) with \( \text{meas}(R(n)) \gg \log n \), we may locate a ratio \( d'/d \) close to 1 w.h.p. (uses that \( R(mpp') \supset R(m) + \log(p'/p)) \)

\textbf{Theorem (Maier–Tenenbaum (1984))}

\textit{For almost all } \( n \), \textit{there are divisors } \( d < d' < d \cdot (1 + (\log n)^{1-\log 3 + o(1)}) \)
Integers with many divisors close together

\[ \mathcal{R}(n) = \bigcup_{d, d' \mid n, \ d \neq d'} [\log \frac{d'}{2d}, \log \frac{2d'}{d}], \quad n_{y, z} = \prod_{p \mid n, \ y < p \leq z} p. \]

- For typical \( n \), we have the following competing constraints:
  - Geometric: \( \mathcal{R}(n_{y, z}) \subset [-C \log z, C \log z] \);
  - Combinatorial: \( \# \{d'/d : d, d' \mid n_{y, z}\} = 3^{\omega(n_{y, z})} \approx 3^{\log \log \frac{z}{\log y}} \).

- If \( \log z > (\log y)^{\frac{\log 3}{\log 3 - 1}} \), we have more than \( \log z \) ratios \( d'/d \) with \( d, d' \mid n_{y, z} \), so we can find a ratio \( \approx 1 \) w.h.p.

- Use \( J \) disjoint intervals \([y_j, z_j]\) to get \( 2^J \) divisors of \( n \) close together.

**Theorem (Maier–Tenenbaum (1984))**

\[ \Delta(n) \geq (\log \log n)^{H_1 - o(1)} \text{ for a.a. } n, \text{ where } H_1 = \frac{\log 2}{\log \frac{\log 3}{\log 3 - 1}} = 0.28754 \ldots \]
Integers with a divisor in a given interval

\[ \exists d \mid n, \quad d \in [x^\theta, 2x^\theta] \iff \theta \log x \in \mathcal{L}(n) := \bigcup_{d \mid n} [\log \frac{d}{2}, \log d] \]

\[ \mathbb{P}_{n \leq x} \left( \theta \log x \in \mathcal{L}(n) \right) \approx \mathbb{E}_{n \leq x} \left[ \frac{\text{meas}(\mathcal{L}(n))}{\log x} \right] \approx \mathbb{E}_{n \leq x} \left[ \min\{1, \frac{\tau(n)}{\log x}\} \right] \]

**Erdős–Tenenbaum:** need \( \omega(n) = \frac{\log \log x}{\log 2} + O(1) \), i.e. \( \tau(n) \asymp \log x \).

\[ \mathbb{P}_{n \leq x} \left( \exists d \mid n, \quad d \in [x^\theta, 2x^\theta] \right) \asymp (\log x)^{-\delta} (\log \log x)^{-1/2} \]

with \( \delta = 1 - \frac{1 + \log \log 2}{\log 2} = 0.08607 \ldots \)

**Ford:** we also need \( \#\{p \mid n, \quad p \leq t\} \leq \frac{\log \log t}{\log 2} + O(1) \) for all \( t \leq x \).

**Theorem (Ford (2004))**

\[ \mathbb{P}_{n \leq x} \left( \exists d \mid n, \quad d \in [y, 2y] \right) \asymp (\log y)^{-\delta} (\log \log y)^{-3/2} \quad (3 \leq y \leq \sqrt{x}) \]
The $k$-dimensional multiplication table

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

- $\mathcal{L}_k(n; \vec{y}) := \bigcup_{\substack{d_1 \cdots d_{k-1} \mid n \\ p \mid d_i \Rightarrow p \leq y_i}} \left[ \log \frac{d_1}{2}, \log d_1 \right] \times \cdots \times \left[ \log \frac{d_{k-1}}{2}, \log d_{k-1} \right]$

- If $\omega_i(n) = \# \{ p \mid n : y_{i-1} < p \leq y_i \}$ for $i = 1, \ldots, k-1$, then

  \[
  \text{meas}(\mathcal{L}_k(n; \vec{y})) \approx \min \left\{ \prod_{i=1}^{k-1} (k-i+1)^{\omega_i(n)}, \prod_{i=1}^{k-1} \log y_i \right\}.
  \]

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

  \[
  h(x, \vec{y}) \approx \max_{\vec{m} \in \mathbb{N}^{k-1}} \prod_{i=1}^{k-1} e^{-\lambda_i} \frac{\lambda_i^{m_i}}{m_i!}.
  \]
Theorem (K. (2014))

1. If $k \leq 6$, then 
   $$h(x, \vec{y}) \approx \max_{\vec{m} \in \mathbb{N}^{k-1}} \prod_{i=1}^{k-1} \frac{e^{-\lambda_i} \chi_i^{m_i}}{m_i!} \prod_{i=1}^{k-1} (k-i+1)^{m_i} \prod_{i=1}^{k-1} \log y_i$$

2. $(\ast)$ is true if $k \geq 7$ and $\log y_{k-1} \leq (\log y_1)^{1+\varepsilon}$ with $\varepsilon = \varepsilon_k$ small.

3. Let $k = 7$, $y_1 = \cdots = y_5 \leq y_6$ and $c = 55.82474304950718986\ldots$
   If $\log y_6 \leq (\log y_1)^{c-\varepsilon}$, then $(\ast)$ holds; but if $\log y_6 \geq (\log y_1)^{c+\varepsilon}$,
   then $h(x, \vec{y})$ is smaller than the RHS of $(\ast)$.

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

- $L_7(n; \vec{y}) = \bigcup_{d_1 \cdots d_6 | n} \prod_{i=1}^{5} [\log \frac{d_i}{2}, \log d_i] \times [\log \frac{d_6}{2}, \log d_6]$
- $p | d_i \Rightarrow p \leq y_i \ \forall i \leq 6$
- $\text{meas}(L_7(n; \vec{y})) \leq \sum_{d_1 \cdots d_5 | n} (\log 2)^5 \cdot \text{meas}(L_2(\frac{n}{d_1 \cdots d_5}; y_6))$. 

\[10/18\]
Improving the 1984 Maier–Tenenbaum result

MT 1984: suppose we can find $J$ disjoint intervals $[y_j, z_j]$ for which
\[ \exists d_j, d'_j \mid n \text{ distinct s.t.:} \]
- $d_j, d'_j$ 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.

\[ \Delta(n) \geq (\log \log n)^{H_2 - o(1)} \text{ a.s. with } H_2 = \frac{\log 2}{\log \left(\frac{1 - 1/\log 27}{1 - 1/\log 3}\right)} = 0.33827 \ldots \]

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

\[ \Delta(n) \geq (\log \log n)^{H_3 - o(1)} \text{ a.s. with } H_3 = 0.3533227727 \ldots \]
The linear algebra of $k$ divisors close together

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

\[
\left( \sum_{p|d_1} \log p, \sum_{p|d_2} \log p, \ldots, \sum_{p|d_k} \log p \right) = O(1) \pmod{\vec{1}}.
\]

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

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

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 $(\log d_1, \ldots, \log d_k) \pmod{\vec{1}}$ is controlled by:

▶ $V = \text{Span}(\Omega)/\langle \vec{1} \rangle$ (geometric)
▶ the distribution of $\{p|D_{\vec{\omega}}\}$ with $\vec{\omega} \in \Omega$ (combinatorial)
Geometric constraints

- The “longest” dim of $\sum_{\tilde{\omega} \in \Omega} \tilde{\omega} \log D_{\tilde{\omega}}$ is onto $\tilde{\omega}^1 \in \Omega$ s.t.
  \[ P^+(D_{\tilde{\omega}^1}) \geq P^+(D_{\tilde{\omega}}) \quad \forall \tilde{\omega} \in \Omega \]

- The second longest dim is onto $\tilde{\omega}^2 \in \Omega \setminus \text{Span}(\tilde{\omega}^1)$ s.t.
  \[ P^+(D_{\tilde{\omega}^2}) \geq P^+(D_{\tilde{\omega}}) \quad \forall \tilde{\omega} \in \Omega \setminus \text{Span}(\tilde{\omega}^1). \]

- Having defined $\tilde{\omega}^1, \ldots, \tilde{\omega}^j$ with span $V_j \leq V$, the $(j + 1)$-th longest dimension is the projection onto $\tilde{\omega}^{j+1} \in \Omega \setminus V_j$ s.t.
  \[ P^+(D_{\tilde{\omega}^{j+1}}) \geq P^+(D_{\tilde{\omega}}) \quad \forall \tilde{\omega} \in \Omega \setminus V_j. \]

- This terminates after $r = \dim(V)$ steps. We then find that $\sum_{\tilde{\omega} \in \Omega} \tilde{\omega} \log D_{\tilde{\omega}}$ are contained in an $r$-dim rectangle of $r$-volume
  \[ \ll \prod_{1 \leq j \leq r} \log P^+(D_{\tilde{\omega}^j}). \]
Given $c \in (0, 1)$ and $k \in \mathbb{Z}_{\geq 2}$, consider the following data:

- $\langle \vec{1} \rangle = V_0 \leq V_1 \leq \cdots \leq V_r \leq \mathbb{Q}^k$
- $1 \geq c_1 \geq \cdots \geq c_r \geq c$, where $\log y = (\log z)^c$
- $\mu_j$ probability measure supported on $V_i \cap \{0, 1\}^k$

To construct $d_1, \ldots, d_k$ close together, we consider configurations s.t.

- $D_{\vec{\omega}} > 1$ if-$f$ $\vec{\omega} \in V_r \setminus \{\vec{0}, \vec{1}\}$;
- $P^+(D_{\vec{\omega}}) \leq \exp\{(\log z)^{c_j}\}$ for all $\vec{\omega} \in V_j \cap \{0, 1\}^k$;
- $\#\{p | D_{\vec{\omega}} : c_{j+1} < \frac{\log \log p}{\log \log z} \leq c_j\} \sim \mu_j(\vec{\omega})(c_j - c_{j+1}) \log \log z.$

To avoid all geometro-combinatorial constraints, we need $\forall V_j' \leq V_j$

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

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

**Theorem (Ford–Green–K. (2019 → 2022?))**

A random integer has $k$ divisors close together composed of primes in $[\exp\{(\log z)^c, z\}$ “iff” there are $V_j, c_j, \mu_j$ as above.
The binary system

Binary flag of order $r$

Let $k = 2^r$, identify $\mathbb{Q}^k$ with $\mathbb{Q}^{p[r]}$, 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 → 2022?))

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

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

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 $j \geq 3$.

Corollary

$\Delta(n) \geq (\log \log n)^{\log 2 \over \log(2/\rho) + o(1)}$ 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

\[
\mathbb{P}_{n \in \mathcal{N}} \left( \exists d_1, \ldots, d_\ell \mid n \hspace{1em} \text{s.t.} \hspace{1em} \Psi(\log d_1, \ldots, \log d_\ell) = \vec{p} + O(1) \right. \\
\left. \text{and the sets } \{p \mid d_i\}, i = 1, \ldots, \ell, \text{ satisfy certain conditions} \right).
\]

- **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:** \( \mathcal{N} = \mathbb{Z} \cap [1, x], \ell = 1, W = \mathbb{R}, \Psi(t) = t, \vec{p} = \log y, d \text{ is } y\text{-smooth} \)

- **Problem B***: \( \mathcal{N} = \mathbb{Z} \cap [1, x], \ell = k - 1, W = \mathbb{R}^{k-1}, \Psi(\vec{x}) = \vec{x}, \vec{p} = (\log y_i)^{k-1}_{i=1}, (d_i, d_j) = 1 \text{ and } d_i \text{ is } y_i\text{-smooth } \forall i \neq j. \)

- **Problem A***: \( \mathcal{N} = \{\text{typical integers in } [1, x]\}, \ell = k, W = \mathbb{R}^k / \langle \vec{1} \rangle, \Psi(\vec{t}) = \vec{t}, \text{ complicated constraints.} \)
Some future directions

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

   **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 $k \geq 7$ for all possible side-lengths.

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

Thank you for your attention