Integers with a divisor in a given interval

Kevin Ford   March 2006

Define $\tau(n, y, z) = |\{d|n : y < d \leq z\}|$.

By inclusion-exclusion,

$$H(x, y, z) := |\{n \leq x : \tau(n, y, z) \geq 1\}|$$

$$= \sum_{k \geq 1} (-1)^{k-1} \sum_{y < d_1 < \ldots < d_k \leq z} \left\lfloor \frac{x}{\text{lcm}[d_1, \ldots, d_k]} \right\rfloor.$$

The density of integers with a divisor in $(y, z]$ is

$$\varepsilon(y, z) = \lim_{x \to \infty} \frac{H(x, y, z)}{x},$$

$$= \sum_{k \geq 1} (-1)^{k-1} \sum_{y < d_1 < \ldots < d_k \leq z} \frac{1}{\text{lcm}[d_1, \ldots, d_k]}.$$

Besicovitch (1934): $\liminf_{y \to \infty} \varepsilon(y, 2y) = 0$

Erdős (1935): $\lim_{y \to \infty} \varepsilon(y, 2y) = 0$

Erdős (1960): $\varepsilon(y, 2y) = (\log y)^{-\delta + o(1)}$,

$$\delta = 1 - \frac{\log(e \log 2)}{\log 2} = 0.08607 \ldots$$
$H(x, y, z)$ when $z - y$ is small.

Let $z = e^\eta y$, $0 < \eta \leq 1$ and $y \leq \sqrt{x}$. Then

$$H(x, y, z) = \sum_{y < d \leq z} \left\lfloor \frac{x}{d} \right\rfloor + O \left( \sum_{y < d_1 < d_2 \leq z} \frac{x}{\text{lcm}(d_1, d_2)} \right) + O(z - y)$$

$$= x \sum_{y < d \leq z} \frac{1}{d} + O \left( x \sum_{m \leq \eta y} \frac{1}{m} \sum_{\frac{y}{m} < t_1 < t_2 \leq \frac{z}{m}} \frac{1}{t_1 t_2} \right) + O(\eta y)$$

$$= x \sum_{y < d \leq z} \frac{1}{d} + O(\eta^2 x \log y).$$

Therefore, if $z - y \to \infty$ and $z - y = o(y / \log y)$, then

$$H(x, y, z) \sim \eta x.$$

Tenenbaum, 1984: $H(x, y, z) \sim \eta x$ for $z - y \to \infty$ and $z \leq z_0(y)$, where

$$z_0(y) = y + \frac{y}{(\log y)^{\log 4 - 1}}.$$
$H(x, y, z)$ when $z$ is large.

Let $y^2 \leq z \leq \sqrt{x}$. By sieve methods, the number of $n \leq x$ that do not have a prime divisor in $(y, z]$ is

$$\ll x \prod_{y < p \leq z} \left(1 - \frac{1}{p}\right) \ll x \frac{\log y}{\log z}.$$ 

Thus, if $z \leq \sqrt{x}$ and $\frac{\log z}{\log y} \to \infty$, then

$$H(x, y, z) \sim x.$$ 

$H(x, y, z)$ for intermediate $z$.

Analysis much more difficult. Tenenbaum in 1984 gave reasonably sharp bounds when $z_0(y) \leq z \leq y^2$, e.g.

$$e^{-c\sqrt{\log \log y \log \log \log y}} \ll \frac{H(x, y, 2y)(\log y)^{\delta}}{x} \ll \frac{1}{\sqrt{\log \log y}},$$

where $\delta = 1 - \frac{1 + \log \log 2}{\log 2} = 0.08607 \ldots$. 
New results (2004)

\[ z = e^\eta y = y^{1+u}, \quad \eta = (\log y)^{-\beta}, \]

\[ \beta = \log 4 - 1 + \frac{\xi}{\sqrt{\log \log y}}, \]

\[ z_0(y) = y + \frac{y}{(\log y)^{\log 4 - 1}}, \]

\[ G(\beta) = \frac{1+\beta}{\log 2} \log \left( \frac{1+\beta e}{\log 2} \right) + 1 \quad (0 \leq \beta \leq \log 4 - 1) \]

**Theorem 1.** Uniformly in \( 100 \leq y \leq \sqrt{x} \), \( z \geq y + 1 \),

\[
H(x, y, z) \asymp \begin{cases} 
\eta = \log(z/y) & y + 1 \leq z \leq z_0(y) \\
\frac{\beta}{(1-\xi)(\log y)^{G(\beta)}} & z_0(y) \leq z \leq 2y \\
u^\delta(\log \frac{2}{u})^{-3/2} & 2y \leq z \leq y^2 \\
1 & z \geq y^2.
\end{cases}
\]

**Corollary.**

\[
H(x, y, 2y) \asymp \frac{x}{(\log y)^{\delta}(\log \log y)^{3/2}}.
\]
Short interval version

Theorem 2. For $y_0 \leq y \leq \sqrt{x}$, $z \geq y + 1$ and and
\[
\frac{x}{\log^{10} z} \leq \Delta \leq x,
\]
we have
\[
H(x, y, z) - H(x - \Delta, y, z) \asymp \frac{\Delta}{x} H(x, y, z).
\]

Here $y_0$ is a large constant.

Squarefree integers

Let $H^*(x, y, z)$ be the number of squarefree numbers $n \leq x$ with $\tau(n, y, z) \geq 1$.

Theorem 3. Suppose $y_0 \leq y \leq \sqrt{x}$, $y + 1 \leq z \leq x$ and
\[
\frac{x}{\log y} \leq \Delta \leq x.
\]
If $z \geq y + Ky^{1/5} \log y$, where $K$ is a large absolute constant, then
\[
H^*(x, y, z) - H^*(x - \Delta, y, z) \asymp \frac{\Delta}{x} H(x, y, z).
\]
Some applications (using $z \asymp y$ case)

1. (Erdős 1955/60, Linnik, A. I. Vinogradov).

Let $A(x) = |\{n = m_1 m_2 : m_i \leq \sqrt{x}\}|$.

**Corollary A.** $A(x) \asymp \frac{x}{(\log x)^\delta (\log \log x)^{3/2}}$.

2. Distribution of Farey gaps (Cobeli, Ford, Zaharescu, 2003). Farey fractions order $Q$: $\frac{0}{1}, \frac{1}{Q}, \frac{1}{Q-1}, \cdots, \frac{Q-1}{Q}, \frac{1}{1}$.

**Corollary B.** \# of distinct gaps in the sequence is $\sim \frac{Q^2}{(\log Q)^\delta (\log \log Q)^{3/2}}$.

3. Erdős function

$$\tau^+(n) = |\{k \in \mathbb{Z} : \tau(n, 2^k, 2^{k+1}) \geq 1\}|.$$

**Corollary C.**

$$\frac{1}{x} \sum_{n \leq x} \tau^+(n) \asymp \frac{(\log x)^{1-\delta}}{(\log \log x)^{3/2}}.$$
Divisors of shifted primes

\[ H(x, y, z; \mathcal{A}) = \left\{ n \leq x : n \in \mathcal{A}, \tau(n, y, z) \geq 1 \right\} \] for a set \( \mathcal{A} \subseteq \mathbb{N} \). Fix \( \lambda \neq 0 \), let \( P_\lambda = \{ p + \lambda : p \text{ prime} \} \).

**Theorem 4.** Let \( \lambda \) be a fixed non-zero integer. Let \( 1 \leq y \leq \sqrt{x} \) and \( y + 1 \leq z \leq x \). Then

\[
H(x, y, z; P_\lambda) \ll \lambda \begin{cases} 
\frac{H(x, y, z)}{\log x} & z \geq y + (\log y)^{2/3} \\
\frac{x}{\log x} \sum_{y < d \leq z} \frac{1}{\phi(d)} & y < z \leq y + (\log y)^{2/3}.
\end{cases}
\]

**Theorem 5.** For fixed \( \lambda, a, b \) with \( \lambda \neq 0 \) and \( 0 \leq a < b \leq 1 \), we have

\[
H(x, x^a, x^b; P_\lambda) \gg_{a, b, \lambda} \frac{x}{\log x}.
\]
Integers with exactly $r$ divisors in $(y, z]$

\[ H_r(x, y, z) = \left| \{ n \leq x : \tau(n, y, z) = r \} \right|, \]
\[ \varepsilon_r(y, z) = \lim_{x \to \infty} \frac{H_r(x, y, z)}{x}. \]

**Erdős conjecture, 1960:** $\lim_{y \to \infty} \frac{\varepsilon_1(y, 2y)}{\varepsilon(y, 2y)} = 0.$

**Tenenbaum conjectures, 1987:**

1. $\forall r \geq 1$, $\liminf_{y \to \infty} \frac{\varepsilon_r(y, 2y)}{\varepsilon(y, 2y)} > 0.$

2. $\forall r \geq 1$, if $z/y \to \infty$, then $\lim_{y \to \infty} \frac{\varepsilon_r(y, z)}{\varepsilon(y, z)} = 0.$

**Hall & Tenenbaum conjecture, 1988:**

\[ \lim_{y \to \infty} \frac{\varepsilon_r(y, 2y)}{\varepsilon(y, 2y)} = d_r > 0. \]
Tenenbaum’s 1987 paper

I. When $y + \frac{y}{(\log y)^{\log 4 - 1}} < z \leq 2y,$

$$\frac{1}{Z(y)} \ll_r \frac{\varepsilon_r(y, z)}{\varepsilon(y, z)} \leq 1.$$ 

II. When $2y \leq z \leq y^2,$

$$\frac{1}{Z(y) \log(z/y)} \ll_r \frac{\varepsilon_r(y, z)}{\varepsilon(y, z)} \ll_r \frac{Z(y)}{\log(z/y)^\delta},$$

Here $Z(y) = \exp\{c_r \sqrt{\log \log y \log \log \log y}\}.$
New bounds for $H_r(x, y, z)$ (2004)

**Theorem 6.** If $z \geq y + 1$, then
\[ \frac{\varepsilon_1(y, z)}{\varepsilon(y, z)} \asymp \frac{\log \log(z/y + 5)}{\log(z/y + 5)}. \]

**Theorem 7.** Suppose $r \geq 2$, $b > 0$ fixed and small, $C > 1$ fixed and large, and $y + \frac{y}{(\log y)^{\log 4 - 1 - b}} \leq z \leq y^C$, then
\[ \frac{\varepsilon_r(y, z)}{\varepsilon(y, z)} \asymp_{r, b, C} \frac{(\log \log(z/y + 5))^{\nu(r) + 1}}{\log(z/y + 5)}, \]
where $2^{\nu(r)} \parallel r$.

**Corollary D.** $\forall r \geq 1$, $\forall c > 1$, if $y \geq y_0(r)$, then
\[ \frac{\varepsilon_r(y, cy)}{\varepsilon(y, cy)} \gg_{r, c} 1. \]

**Corollary E.** If $z/y \to \infty$, then
\[ \frac{\varepsilon_r(y, z)}{\varepsilon(y, z)} \to 0. \]
Further bounds for $H_r(x, y, z)$

**Theorem 8** (Ford, Tenenbaum, 2006). If $y \leq \sqrt{x}$ and $y < z \leq y + \frac{y}{(\log y)^{\log 4 - 1 - o(1)}}$, then

$$
\sum_{r \geq 2} H_r(x, y, z) = o(H(x, y, z)).
$$

**Theorem 9** (Yong Hu, 2006). If $y^{20} \leq z \leq x^{1/4}$, then

$$
\frac{H_2(x, y, z)}{x} \leq \frac{\log \log y \log \log z}{\log z}
$$

and

$$
\frac{H_3(x, y, z)}{x} \leq \frac{\log \log z + (\log \log z - \log \log y)^2}{\log z}.
$$

**Conjecture:** Under the hypotheses of Theorem 9, for each $r$,

$$
\frac{H_r(x, y, z)}{x} \leq_r \frac{Q_r(\log \log y, \log \log z)}{\log z}
$$

for some polynomial $Q_r$. 
Proof ideas: $H(x, y, 2y)$

Suppose $n$ is squarefree, and $n_1 := \prod_{p | n, p \leq z^{10}} p \leq z^{10}$.

Assume $\{\log d : d | n_1\}$ is roughly uniformly distributed in $[0, \log n_1]$.

| 0 | $\log y$ | $\log z$ | $\log n_1$ |

If $\omega(n_1) = |\{p : p | n_1\}| = k$, then

$$\text{Prob}[\tau(n_1, y, z) = 1] \approx \frac{\log(z/y)^k}{\log n_1} \approx \frac{2^k}{\log z}.$$ 

With $k_0 = \left\lfloor \frac{\log \log z}{\log 2} \right\rfloor$, we predict that

$$H(x, y, 2y) \approx \sum_{k \geq k_0} |\{n \leq x : \omega(n_1) = k\}| \approx \sum_{k \geq k_0} \frac{x(\log \log z)^{k-1}}{(k - 1)! \log z} \approx \frac{x(\log \log y)^{k_0}}{k_0! \log y} \frac{1}{x} \approx \frac{1}{(\log y)^\delta (\log \log y)^{1/2}}.$$
Two principles

**Principle # 1.** The numbers \( \{ \log \log p : p|n_1 \} \) behave like uniformly distributed random numbers in \([0, \log \log z] \):

\[
\log \log p_j(n_1) \approx \frac{j \log \log z}{k} \quad (1 \leq j \leq k),
\]

where \( p_j(n_1) \) is the \( j \)-th smallest prime factor of \( n_1 \).

**Principle # 2.** The numbers \( \{ \log d : d|n_1 \} \) do not behave like uniformly distributed random numbers in \([0, \log n_1] \).

**Principle #1 \implies Principle # 2.** Reason: with high probability,

\[
\log \log p_j(n_1) \leq \frac{j \log \log z}{k} - c \sqrt{\log \log z}
\]

for some \( j \). This causes the numbers \( \{ \log d : d|n_1 \} \) to be grouped in isolated clumps. In fact, for most \( n_1 \),

\[
\text{Prob}[\tau(n_1, y, 2y) \geq 1] \approx \frac{2^k}{\log z} \exp\{-c \sqrt{\log \log z}\}.
\]
New ideas

Focus on **abnormal** integers, those satisfying

\[(1) \quad \log \log p_j(n_1) \geq \frac{j \log \log z}{k} - O(1) \quad (1 \leq j \leq k).\]

On average over \(n_1\) satisfying (1), \(\{\log d : d | n_1\}\) is distributed uniformly in \([0, \log n_1]\).

The probability that (1) occurs is \(\simeq \frac{1}{k} \simeq \frac{1}{\log \log y}\). This leads to a refined prediction

\[H(x, y, 2y) \asymp \frac{x}{(\log y)^\delta (\log \log y)^{3/2}},\]

which is the correct order. More generally, we need to estimate the likelihood that

\[(2) \quad \log \log p_j(n_1) \geq \frac{j \log \log z}{v} - u \quad (1 \leq j \leq k)\]

when \(v\) is close to \(k\) and \(1 \leq u \ll \sqrt{\log \log z}\).
Lower bound for $H_1(x, y, 2y)$

Say $d|n_1$ is $\eta$-isolated if there is no $f|n_1$ with $0 < |\log f/d| \leq \eta$. Let $I(m; \eta)$ be the number of $\eta$-isolated $d|m$. If $d$ is log 2-isolated and $y < d < 2y$, then $\tau(n, y, 2y = 1)$.

The probability that $\tau(n_1, y, z) = 1$ is about $\frac{I(n_1; \log 2)}{\log z}$.

If $\omega(n_1) = k_0 - c_1$ and

$$\log \log p_j(n_1) \geq \frac{j \log \log z}{k} - c_2 \quad (1 \leq j \leq k),$$

then (on average) $I(n_1)$ will be close to $2^{k_0} \approx \log z$, i.e. no clumps means lots of isolated divisors. This leads to the prediction that

$$H_1(x, y, 2y) \gg \frac{x}{(\log y)^\delta (\log \log y)^{3/2}} \gg H(x, y, 2y).$$
Bounding $H_r(x, y, z)$ for $10y \leq z \leq y^{1.1}$

Recall

$$\frac{\varepsilon_r(y, z)}{\varepsilon(y, z)} \approx_r \frac{(\log \log(z/y + 5))^{\nu(r)+1}}{\log(z/y + 5)}, \quad 2^{\nu(r)}\|r.$$  

For lower bound, set $\eta = \log(z/y)$ and consider $n = pqst,$

$$s \leq z/y, \quad \tau(s) \geq r + 1, \quad t < y^{1/4}, \quad P^-(t) > z/y,$$

$p$ is prime and $P^-(q) > z.$ Let $f$ be a $2\eta$-isolated divisor of $t,$ let $1 = d_1 < d_2 < \cdots < d_m = s$ be the divisors of $s,$ and $pfd_r \leq z < pfd_{r+1}.$ Then $\tau(n, y, z) = r.$

\[
\begin{array}{cccccccc}
 y & pfd_1 & \cdots & pfd_r & z & pfd_{r+1} \\
\end{array}
\]

Given $pst,$ number of $q$ is $\asymp \frac{x}{pst \log z}.$ Then

$$\sum_{z/fd_r < p \leq z/fd_{r+1}} \frac{1}{p} \asymp \frac{\log(d_{r+1}/d_r)}{\log z}.$$
Therefore,

\[ H(x, y, z) \gg \frac{x}{\log^2 z} \sum_{t<y^{1/4}} I(t; 2\eta) \sum_{s<z/y} \frac{\log(d_{t+1}/d_t)}{s}. \]

Write

\[ s = p_1 \cdots p_k, \quad p_1 < \cdots p_k. \]

If \( p_j > p_1 \cdots p_{j-1} \) for each \( j \), then

\[
\begin{align*}
    d_1 &= 1 \\
    d_2 &= p_1 \\
    d_3 &= p_2 \\
    d_4 &= p_2 p_1 \\
    d_5 &= p_3 \\
    d_6 &= p_3 p_1 \\
    d_7 &= p_3 p_2 \\
    d_8 &= p_3 p_2 p_1 \\
    d_9 &= p_4 \\
    d_{10} &= p_4 p_1 \\
    d_{11} &= p_4 p_2 \\
    d_{12} &= p_4 p_2 p_1, \text{ etc.}
\end{align*}
\]

i.e., the ordinary ordering of divisors of \( s \) coincides with the lexicographic ordering of the divisors. Then

\[
\frac{d_{t+1}}{d_t} = \frac{p_t}{p_1 \cdots p_{t-1}}, \quad t = \nu(r) + 1.
\]
Take $s$ so that

$$p_j > (p_1 \cdots p_{j-1})^2 \quad (1 \leq j \leq r),$$

then $\log(d_{r+1}/d_r) \gg \log p_t$, $t = \nu(r) + 1$. We obtain

$$\sum_s \frac{\log(d_{r+1}/d_r)}{s} \gg (\log z/y)(\log \log z/y)^t.$$

Upper bounds:

For simplicity assume $s$ is squarefree.

**Lemma 1.** Let $s = p_1 \cdots p_k$, $p_1 < \cdots < p_k$. For $1 \leq r \leq 2^k - 1$, \[ \frac{d_{r+1}(n)}{d_r(n)} \leq p_{\nu(r)+1}. \]

With lemma 1, we obtain

$$\sum_s \frac{\log(d_{r+1}/d_r)}{s} \ll (\log z/y)(\log \log z/y)^{\nu(r)+1}.$$
Open problems/future projects

(I) Asymptotic formulas for $H(x, y, 2y)$ and $A(x)$.

(II) Lower bounds for $H(x, y, z; P_\lambda)$ when $z = y^{1+o(1)}$.

(III) Triple and higher order factorizations, e.g.

$$H(x, y_1, z_1, y_2, z_2) = |\{n \leq x : d_1d_2|n, y_i < d_i \leq z_i\}|,$$

$$A_3(x) = |\{n = m_1m_2m_3 : \text{each } m_i \leq x^{1/3}\}|.$$

(IV) Analogs for algebraic integers, e.g. count the Gaussian integers with norm $\leq x$ and with a divisor in a given region of the plane (rectangle, section of an annulus, ...).

(V) Use methods/ideas to attack other types of divisor problems, such as the concentration function

$$\Delta(n) = \max_u \tau(n, u, eu).$$