Chapter 5 Sums of Additive Functions

v1 2018-19

The commonly seen **additive** arithmetic functions are $\omega(n)$, the number of distinct prime factors of n, and $\Omega(n)$, the number of prime divisors of n counted with multiplicity. So

$$\omega(n) = \sum_{p|n} 1 \quad \text{and} \quad \Omega(n) = \sum_{p^r|n} 1 = \sum_{p^a||n} a.$$

There are 'artificial' examples given by the logarithms of multiplicative functions, such as $\log d(n)$. We will concentrate on ω and Ω.

Little can be said of $\omega(n)$ for individual n, so we further concentrate on averages of $\omega(n)$

Theorem 1 *We have*

$$\sum_{n\le x} \omega(n) = x \log\log x + O(x),$$

and

$$\sum_{n\le x} \omega^2(n) \le x(\log\log x)^2 + O(x\log\log x). \tag{1}$$

The second result can be proved with equality, not just an upper bound, but we give the stated result for simplicity.

Proof Start from

$$\sum_{n\le x} \omega(n) = \sum_{n\le x}\sum_{p|n} 1 = \sum_{p\le x}\sum_{\substack{n\le x\\ p|n}} 1, \tag{2}$$

having interchanged the summations. Continuing,

$$\begin{aligned} &= \sum_{p\le x}\left[\frac{x}{p}\right] = \sum_{p\le x}\left(\frac{x}{p}+O(1)\right) = x\sum_{p\le x}\frac{1}{p} + O(\pi(x)) \\ &= x(\log\log x + O(1)) + O(x), \end{aligned}$$

using Merten's Theorem on the sum of reciprocals of primes, and the trivial $\pi(x) \le x$. Thus the required result follows.

Next

$$\sum_{n\le x}\omega^2(n)=\sum_{n\le x}\sum_{p|n}1\sum_{q|n}1=\sum_{p\le x}\sum_{q\le x}\sum_{\substack{n\le x\\ p|n,q|n}}1.$$

Then this double sum over pairs of primes (p,q) splits into either $p=q$ or $p\neq q$. That is,

$$\begin{aligned}\sum_{p\le x}\sum_{q\le x}\sum_{\substack{n\le x\\ p|n,q|n}}1 &= \sum_{p\le x}\sum_{\substack{n\le x\\ p|n}}1+\sum_{\substack{p\le x\\ p\neq q}}\sum_{q\le x}\sum_{\substack{n\le x\\ pq|n}}1\\ &= \sum_{p\le x}\left[\frac{x}{p}\right]+\sum_{\substack{p\le x\\ p\neq q}}\sum_{q\le x}\left[\frac{x}{pq}\right]\\ &\le \sum_{p\le x}\frac{x}{p}+\sum_{p\le x}\sum_{q\le x}\frac{x}{pq}. \qquad (3)\end{aligned}$$

Here we have used $[u]\le u$ for real u, and dropped the restriction $p\neq q$, increasing the sum. Continuing

$$\sum_{n\le x}\omega^2(n)\le x\sum_{p\le x}\frac{1}{p}+x\left(\sum_{p\le x}\frac{1}{p}\right)^2. \qquad (4)$$

By Merten's result again,

$$x\left(\sum_{p\le x}\frac{1}{p}\right)^2=x\left(\log\log x+O(1)\right)^2=x\left(\log\log x\right)^2+O(x\log\log x) \qquad (5)$$

The first sum on the right hand side of (4) is of the same magnitude as the error in (5). Combine to get stated result. ■

Note the inequality in (3) is due to two approximations: $[u]\le u$ for real u, and dropping the restriction $p\neq q$. The errors in these approximations can be estimated as $O(x\log\log x)$ in which case we can prove (1) with equality. See the Problem Sheet.

The two parts of Theorem 1 can be combined in

Theorem 2 *(Turán)*

$$\sum_{n\le x}(\omega(n)-\log\log x)^2 = O(x\log\log x)\,.$$

Proof

$$\begin{aligned}\sum_{n\le x}(\omega(n)-\log\log x)^2 &= \sum_{n\le x}\omega^2(n)-\log\log x\sum_{n\le x}\omega(n)+(\log\log x)^2\sum_{n\le x}1\\ &\le x\,(\log\log x)^2+O(x\log\log x)\\ &\qquad -(\log\log x)\,(x\log\log x+O(x))\\ &\qquad\quad +(\log\log x)^2\,(x+O(1))\\ &= O(x\log\log x)\end{aligned}$$

■

Note that with more work (in particular, with equality in (1)) this result can be improved to

$$\sum_{n\le x}(\omega(n)-\log\log x)^2 = x\log\log x+O(x)\,.$$

Return now to Theorem 2 and replace the $\log\log x$ term by $\log\log n$ in the terms of the series so they do not depend on x. For this we need a lemma.

Lemma 3

$$\sum_{3\le n\le x}(\log\log x-\log\log n)^2 = O(x)\,.$$

Proof The idea has been seen before, to split the sum into a long sum on which the summand changes little, with a remaining short sum on which the summand may change a lot.

In the long interval $\sqrt{x}<n\le x$ we have that

$$0=\log 1=\log\left(\frac{\log x}{\log x}\right)\le\log\left(\frac{\log x}{\log n}\right)<\log\left(\frac{\log x}{\log\sqrt{x}}\right)\le\log 2.$$

In the short interval $3\le n\le\sqrt{x}$ we have

$$\log 2<\log\left(\frac{\log x}{\log n}\right)\le\log\log x.$$

Thus, the contribution from the long interval is

$$\sum_{\sqrt{x}<n\le x} (\log\log x - \log\log n)^2 \le \sum_{n\le x} (\log 2)^2 \ll x. \tag{6}$$

And from the short interval

$$\sum_{3\le n\le\sqrt{x}} (\log\log x - \log\log n)^2 \le \sum_{3\le n\le\sqrt{x}} (\log\log x)^2 \le \sqrt{x}\,(\log\log x)^2 . \tag{7}$$

Combine (7) and (6) to get stated result. ■

In the following proof we make use of

$$(a+b)^2 \le 2\left(a^2+b^2\right).$$

This can be proved by starting with $(x-1)^2 \ge 0$. For then $x^2+1 \ge 2x$ and then, adding x^2+1 to both sides, $2x^2+2 \ge x^2+2x+1 = (x+1)^2$. Apply this with $x = a/b$ (if $b = 0$ the result is trivial).

Corollary 4

$$\sum_{3\le n\le x} (\omega(n) - \log\log n)^2 = O(x\log\log x)\,.$$

Proof Result follows from Theorem 2 and Lemma 3 used within

$$\begin{aligned}\sum_{3\le n\le x} (\omega(n) - \log\log n)^2 &= \sum_{3\le n\le x} (\omega(n) - \log\log x + \log\log x - \log\log n)^2 \\ &\le 2\sum_{n\le x} (\omega(n) - \log\log x)^2 \\ &\quad +2\sum_{3\le n\le x} (\log\log x - \log\log n)^2\end{aligned}$$

■

Corollary 5 *Let $\delta > 0$ be given. Then the number of $3 \le n \le x$ which do* ***not*** *satisfy*

$$|\omega(n) - \log\log n| < (\log\log n)^{1/2+\delta} \tag{8}$$

is $\ll x\,(\log\log x)^{-2\delta}$.

Proof The exceptional set is

$$\mathcal{E}(x) = \left\{3 \leq n \leq x : |\omega(n) - \log\log n| \geq (\log\log n)^{1/2+\delta}\right\}.$$

For the argument below let

$$\mathcal{E}_0(x) = \left\{n \in E\,(x) : n \geq \sqrt{x}\right\}.$$

Then $|\mathcal{E}\,(x)| = |\mathcal{E}_0\,(x)| + O(\sqrt{x})$. Consider first a sum over the integers in $\mathcal{E}_0\,(x)$,

$$\begin{aligned}
\sum_{n\in\mathcal{E}_0(x)} |\omega(n) - \log\log n|^2 &\geq \sum_{n\in v_0(x)} \left((\log\log n)^{1/2+\delta}\right)^2 \\
&\geq \left(\log\log\sqrt{x}\right)^{1+2\delta} \sum_{n\in E_0(x)} 1 \\
&\qquad \text{since } n \in \mathcal{E}_0\,(x) \implies n \geq \sqrt{x} \\
&\gg |\mathcal{E}_0\,(x)|\,(\log\log x)^{1+2\delta}.
\end{aligned}$$

Yet

$$\sum_{n\in\mathcal{E}_0(x)} |\omega(n) - \log\log n|^2 \leq \sum_{3\leq n\leq x} |\omega(n) - \log\log n|^2 \ll x\log\log x.$$

by Corollary 4. Combine the last two results as

$$|\mathcal{E}_0(x)| \ll \frac{x}{(\log\log x)^{2\delta}}.$$

Since $\sqrt{x}$ grows so much slower that this, the same bound holds for $|\mathcal{E}(x)|$. ■

Definition 6 *If a property $P\,(n)$ holds for all $n \leq x$ except for $n \in \mathcal{E}(x)$, ($\mathcal{E}$ for exceptional) and $|\mathcal{E}(x)| = o\,(x)$ we say that the property $P\,(n)$ holds* ***for almost all*** *n.*

In Corollary 5

$$\frac{|\mathcal{E}(x)|}{x} \ll \frac{1}{(\log\log x)^{2\delta}} \to 0$$

as $x \to \infty$, so $|\mathcal{E}(x)| = o\,(x)$. Thus, for any $\delta > 0$, $|\omega(n) - \log\log n| < (\log\log n)^{1/2+\delta}$ for almost all n.

Definition 7 *We say that a function $f(n)$ has **normal order** $F(n)$ if, for every $\varepsilon > 0$ the inequality*

$$(1-\varepsilon)\, F(n) < f(n) < (1+\varepsilon)\, F(n)$$

for almost all values of n.

This can be written as $|f(n) - F(n)| < \varepsilon F(n)$ for almost all n.

Corollary 8 *$\omega(n)$ has normal order $\log\log n$.*

Proof Choose $\delta = 1/4$ in Corollary 5 (only chosen so that $1/2 + \delta < 1$) so that $|\omega(n) - \log\log n| < (\log\log n)^{3/4}$ for almost all n.

Let $\varepsilon > 0$ be given. Then for $n > \exp\exp(1/\varepsilon^4)$ we have $(\log\log n)^{3/4} \leq \varepsilon \log\log n$. Thus $|\omega(n) - \log\log n| < \varepsilon \log\log n$ for almost all n.

Hence we have verified the definition that $\omega(n)$ has normal order $\log\log n$.

■

This was a result of Hardy and Ramanujan (1916). It says that almost all integers n have $\log\log n$ distinct prime divisors.

Turán's result holds with ω replaced by Ω and we have similarly that the number of $3 \leq n \leq x$ which do **not** satisfy

$$|\Omega(n) - \log\log n| < (\log\log n)^{1/2+\delta} \tag{9}$$

is $\ll x(\log\log x)^{-2\delta}$. Of course, the sets of $n \ll x(\log\log x)^{-2\delta}$. for which (8) fails and (9) fails may not be the same but the size of the union of the exceptions is still $\ll x(\log\log x)^{-2\delta}$, only the implies constant changes. Thus was can assume both (8) and (9) hold for almost all n. They can be combined in the one result for the divisor function:

Proposition 9 *For all $\varepsilon > 0$ we have*

$$(\log n)^{\log 2-\varepsilon} < d(n) < (\log n)^{\log 2+\varepsilon},$$

for almost all n.

The following result is fundamental to the proposition,

Lemma 10 *For $n \geq 1$,*

$$2^{\omega(n)} \leq d(n) \leq 2^{\Omega(n)}.$$

Proof of lemma The functions ω and Ω are additive so 2^ω and 2^Ω are multiplicative. The divisor function is also multiplicative. Since these functions are all positive we need only show the inequality on prime powers.

For the lower bound $d(p^a) = 1 + a \geq 2 = 2^{\omega(p^a)}$.

For the upper bound first note that it follows by induction that $1+a \leq 2^a$ for all integers $a \geq 1$. Hence

$$d(p^a) = 1 + a \leq 2^a = 2^{\Omega(p^a)}.$$

■

Proof of Proposition Let $\varepsilon > 0$ be given. Then, just as

$$(1 - \varepsilon/\log 2)\log\log n < \omega(n) < (1 + \varepsilon/\log 2)\log\log n$$

for almost all sufficiently large n followed from (8) then from (9) the same inequalities follow with ω replace by Ω.

Note next that

$$\begin{aligned} 2^{(1\pm\varepsilon/\log 2)\log\log n} &= \exp\left((\log 2 \pm \varepsilon)\log\log n\right) = \exp\left(\log\left(\log n\right)^{(\log 2\pm\varepsilon)}\right) \\ &= \left(\log n\right)^{(\log 2\pm\varepsilon)}. \end{aligned}$$

Hence, by the Lemma,

$$\left(\log n\right)^{(\log 2-\varepsilon)} < 2^{\omega(n)} \leq d(n) \leq 2^{\Omega(n)} < \left(\log n\right)^{(\log 2+\varepsilon)},$$

for almost all n. ■

The Proposition can be interpreted as saying that for almost all n the divisor function $d(n)$ is approximately

$$\begin{aligned} 2^{\log\log n} &= e^{\log 2\log\log n} = 2^{\log(\log n)^{\log 2}} = \left(\log n\right)^{\log 2} \\ &= \left(\log n\right)^{0.693\ldots}. \end{aligned}$$

Yet from the average result $\sum_{n\leq x} d(n) \sim x\log x$ we might have guessed the size of most $d(n)$ to be $\log n$.

The fact that $d(n)$ is almost always smaller than $\log n$ means that when it is larger it must be substantially larger. This can be seen in the $\sum_{n\leq x} d^2(n) \sim cx\log^3 x$ result for the squaring of $d(n)$ has amplified the substantially larger values of $d(n)$ and though they occur rarely they have contributed to the *cube* of the logarithm on the right hand side.