Student Seminar in Elementary Number Theory

Average orders of arithmetical functions

Patrick Edera, Nina Goldhirsch

November 11, 2021

1 Introduction

The topic of this lecture is the average orders of arithmetical functions, where an arithmetical function is a complex-valued function defined on the set of positive integers. In particular we are interested of the behavior of these types of functions for large values of n.

Many arithmetical functions behave irregularly and fluctuate as n increases and that's why it is often more interesting to study the arithmetical mean which usually behave more regularly. We will work with sums of the form

$$\sum_{k\leq x} f(k) \qquad\qquad x \in \mathbb{R},\ 1 \leq k \leq [x]$$

If $f(n)$ is an arithmetical function and $g(n)$ is any simple function of n such that

$$f(1) + f(2) + \ldots + f(n) \sim g(1) + \ldots + g(n)$$

where $f \sim g$ as $x \to \infty$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$. In this case we say that $f(n)$ is of the *average order* of $g(n)$.

Our goal is to find the average order of:

- $d(n)$, the number of divisor of n: $\sum_{d|n} 1$
- $\sigma_\alpha(n)$, the sum of the αth powers of the positive divisors of n, including 1 and n: $\sum_{d|n} d^\alpha$
- $\varphi(n)$, the number of positive integers not greater than n that are coprime with n

To do this our principal reference is the book by *T.M. Apostol*

2 Some elementary asymptotic formulas

Let's start this section with recalling the big oh notation:

Definition 1. *Let f,g : $\mathbb{R} \to \mathbb{R}$ and $a \in \mathbb{R}\cup\{\pm\infty\}$. Then is: $f(x) = \mathcal{O}(g(x))$ for $x \to a$, when $\limsup_{x\to a} \left|\frac{f(x)}{g(x)}\right| < \infty$. Also if $g(x) > 0$ for all $x \leq a$ we write $f(x) = \mathcal{O}(g(x))$ to mean that the quotient $\frac{f(x)}{g(x)}$ is bounded for $x \leq a$: that is, there exists a constant $M > 0$ such that $|f(x)| \leq Mg(x)$ for all $x \leq a$.*

With this definition we can now give some asymptotic formulas:

Theorem 2.1. *if $x \geq 1$ we have:*

(a) $\sum\limits_{n \leq x} \frac{1}{n} = \log(x) + C + \mathcal{O}(\frac{1}{x})$
where $C = \lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n} - \log(n)\right)$ is called Euler's constant

(b) $\sum\limits_{n \leq x} \frac{1}{n^s} = \frac{x^{1-s}}{1-s} + \zeta(s) + \mathcal{O}(x^{-s}) \qquad \text{if} s > 0, s \neq 1$
where

$$\zeta(s) = \begin{cases} \sum\limits_{n=1}^{\infty} \frac{1}{n^s} & \text{if } s > 1, \\ \lim_{x\to\infty} \left(\sum_{n\leq x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s}\right) & \text{if } 0 < s < 1 \end{cases}$$

(c) $\sum\limits_{n > x} \frac{1}{n^s} = \mathcal{O}(x^{1-s}) \qquad \text{if } s > 1$

(d) $\sum\limits_{n \leq x} n^{\alpha} = \frac{x^{\alpha+1}}{\alpha+1} + \mathcal{O}(x^{\alpha}) \qquad \text{if } \alpha \geq 0$

For the proof of this theorem we require:

Theorem 2.2 (Euler's summation formula)**.** *If f has a continuous derivative f' on the interval $[y, x]$, where $0 < y < x$, then*

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)\,dt + \int_y^x (t - [t])f'(t)\,dt + f(x)([x] - x) - f(y)([y] - y)$$

Proof. Let m be the smallest integer on which we want to sum also let $m = [y + 1]$, and let k be the greatest, also $k = [x]$.
For integers n and $n - 1$ in $[y, x]$ we have

$$\int_{n-1}^{n} [t]f'(t)\,dt = \int_{n-1}^{n} (n-1)f'(t)\,dt = (n-1)(f(n) - f(n-1)) =$$
$$= nf(n) - (n-1)f(n-1) - f(n)$$

Summing from $n = m + 1$ to $n = k$ we find:

$$\int_m^k [t]f'(t)\,dt = \sum_{n=m+1}^{k} (nf(n) - (n-1)f(n-1)) - \sum_{n=m+1}^{k} f(n) =$$

$$= \sum_{n=m+1}^{k} (nf(n) - (n-1)f(n-1)) - \sum_{n=m+1}^{k} f(n) - f(m) + f(m) =$$

$$= kf(k) - mf(m) - \sum_{y<n\le x} f(n) + f(m)$$

Therefore

$$\int_y^x [t]f'(t)\,dt = \int_y^m [t]f'(t)\,dt + \int_m^k [t]f'(t)\,dt + \int_k^x [t]f'(t)\,dt =$$

$$= \int_y^m (m-1)f'(t)\,dt + \int_m^k [t]f'(t)\,dt + \int_k^x kf'(t)\,dt =$$

$$= mf(m) - f(m) - mf(y) + f(y) + kf(k) - mf(m) - \sum_{y<n\le x} f(n) + f(m) + kf(x) - kf(k) =$$

$$= kf(x) - (m-1)f(y) - \sum_{y<n\le x} f(n)$$

Also we get:

$$\sum_{y<n\le x} f(n) = -\int_y^x [t]f'(t)\,dt + [x]f(x) - [y]f(y) \tag{1}$$

Finally if we calculate a last integral by integrating by parts

$$\int_y^x tf'(t)\,dt = xf(x) - yf(y) - \int_y^x f(t)\,dt$$

and combine it with (1) we obtain the desired formula. □

Now we are ready to prove Theorem 2.1:

Proof. (a) We apply the Euler's summation formula with $f(t) = 1/t$:

$$\sum_{n\le x} \frac{1}{n} = \int_1^x \frac{1}{t}\,dt - \int_1^x \frac{t-[t]}{t^2}\,dt + 1 - \frac{x-[x]}{x} =$$

$$= \log(x) - \int_1^x \frac{t-[t]}{t^2}\,dt + 1 + \mathcal{O}\left(\frac{1}{x}\right) =$$

$$= \log(x) + 1 - \int_1^\infty \frac{t-[t]}{t^2}\,dt + \int_x^\infty \frac{t-[t]}{t^2}\,dt + \mathcal{O}\left(\frac{1}{x}\right)$$

Since $\int_1^\infty \frac{t-[t]}{t^2}\,dt$ is dominated by $\int_1^\infty \frac{1}{t^2}\,dt$ and

$$0 \le \int_x^\infty \frac{t-[t]}{t^2}\,dt \le \int_x^\infty \frac{1}{t^2}\,dt = \frac{1}{x}$$

we have that

$$\sum_{n\leq x}\frac{1}{n} = \log(x) + 1 - \int_1^{\infty}\frac{t-[t]}{t^2}\,dt + \mathcal{O}\left(\frac{1}{x}\right)$$

that is equal to

$$= \log(\mathrm{x}) + C + \mathcal{O}\left(\frac{1}{x}\right)$$

with $C = 1 - \int_1^{\infty}\frac{t-[t]}{t^2}\,dt$. It's easy to see that C is exactly the Euler's constant, in fact if we let $x \to \infty$ in (a) we find that

$$\lim_{x\to\infty}\left(\sum_{n\leq x}\frac{1}{n} - \log(x)\right) = 1 - \int_1^{\infty}\frac{t-[t]}{t^2}\,dt$$

(b) We apply the Euler's summation formula with $f(x) = x^{-s}$:

$$\sum_{n\leq x}\frac{1}{n^s} = \int_1^{x}\frac{1}{t^s}\,dt - s\int_1^{x}\frac{t-[t]}{t^{s+1}}\,dt + 1 - \frac{x-[x]}{x^s} =$$

$$= \frac{x^{1-s}}{1-s} - \frac{1}{1-s} + 1 - s\int_1^{\infty}\frac{t-[t]}{t^{s+1}}\,dt + \mathcal{O}(x^{-s})$$

Therefore

$$\sum_{n\leq x}\frac{1}{n^s} = \frac{x^{1-s}}{1-s} + C(s) + \mathcal{O}(x^{-s}) \tag{2}$$

Where

$$C(s) = 1 - \frac{1}{1-s} - s\int_1^{\infty}\frac{t-[t]}{t^{s+1}}\,dt$$

Now we have to divide into two cases:
if $s > 1$, $\sum_{n\leq x}\frac{1}{n^s}$ approaches $\zeta(s)$ as $x \to \infty$ and the terms x^{1-s} and x^{-s} both approach 0. From the definition of $\zeta(s)$ and the fact that $C(s)$ doesn't depend on x, by making x tend to infinity in (2), we obtain that $C(s) = \zeta(s)$ if $s > 1$.
If instead $0 < s < 1$ and as above taking x tend to infinity in (2), we have that $x^{-s} \to 0$. By the fact that $C(s)$ doesn't depend on x we can see that

$$\lim_{x\to\infty}\left(\sum_{n\leq x}\frac{1}{n^s} - \frac{x^{1-s}}{1-s}\right) = C(s)$$

Therefore, by definition, $C(s)$ is also equal to $\zeta(s)$ if $0 < s < 1$.

(c) To prove (c) we use (b) with $s > 1$ to obtain:

$$\sum_{n>x}\frac{1}{n^s} = \zeta(s) - \sum_{n\leq x}\frac{1}{n^s} = \frac{x^{1-s}}{1-s} + \mathcal{O}(x^{-s}) = \mathcal{O}(x^{1-s})$$

since $x^{-s} \leq x^{1-s}$

(d) We apply the Euler's summation formula with $f(t) = t^\alpha$:

$$\begin{aligned}\sum_{n\leq x} n^\alpha &= \int_1^x t^\alpha\,dt + \alpha\int_1^x t^{\alpha-1}(t-[t])\,dt + 1 - (x-[x])x^\alpha \\ &\overset{\alpha\geq 0}{=} \frac{x^{\alpha+1}}{\alpha+1} - \frac{1}{\alpha+1} + \mathcal{O}\left(\alpha\int_1^x t^{\alpha-1}\,dt\right) + \mathcal{O}(x^\alpha) \\ &= \frac{x^{\alpha+1}}{\alpha+1} + \mathcal{O}(x^\alpha)\end{aligned}$$

□

3 The average order of $d(n)$

Now we are ready to derive the Dirichlet's asymptotic formula for the partial sums of the divisor function $d(n)$.

Theorem 3.1. *For all $x \geq 1$ we have*

$$\sum_{n\leq x} d(n) = x\log(x) + (2C-1)x + \mathcal{O}(\sqrt{x}) \tag{3}$$

where C is Euler's constant

Proof. Since $d(n)$ is the number of divisor of n, we can write $d(n) = \sum_{d|n} 1$, and so we obtain

$$\sum_{n\leq x} d(n) = \sum_{n\leq x}\sum_{d|n} 1$$

This is a double sum extended over n and d. Since $d|n$ we can write $n = qd$ and extend the sum over all pairs of positive integers q, d with $qd \leq x$. Thus

$$\sum_{n\leq x} d(n) = \sum_{\substack{q,d \\ qd\leq x}} 1 \tag{4}$$

This can be interpreted as a sum extended over certain lattice points in the qd-plane (see figure 1). The lattice points with $qd = n$ lie on a hyperbola, so the sum in (4) counts the number of lattice points which lie on the hyperbolas corresponding to $n = 1, 2, ..., [x]$.

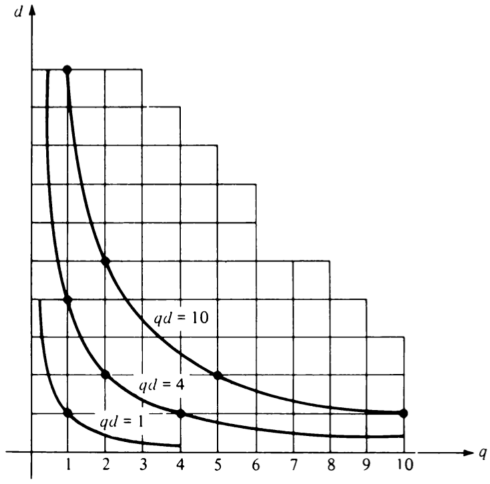

Figure 1: Lattice Points

Now we fix $d \leq x$ and we can count first those lattice points on the horizontal line segment $1 \leq q \leq \frac{x}{d}$, and then sum over all $d \leq x$. Thus (4) becomes

$$\sum_{n \leq x} d(n) = \sum_{d \leq x} \sum_{q \leq \frac{x}{d}} 1$$

Now we can use part (d) of Theorem 2.1 with $\alpha = 0$ to obtain

$$\sum_{q \leq \frac{x}{d}} 1 = \frac{x}{d} + \mathcal{O}(1)$$

Using this along with Theorem 2.1 (a) we find

$$\sum_{n \leq x} d(n) = \sum_{d \leq x} \left(\frac{x}{d} + \mathcal{O}(1) \right) = x \sum_{d \leq x} \frac{1}{d} + \mathcal{O}(x) =$$

$$= x \left(\log(x) + C + \mathcal{O}(\frac{1}{x}) \right) + \mathcal{O}(x)$$

$$= x\log(x) + \mathcal{O}(x)$$

This is a weak version of (3) which implies

$$\sum_{n \leq x} d(n) \sim x\log(x) \qquad \text{as} \quad x \to \infty$$

and gives $\log(n)$ as the average order of $d(n)$. In this case we say that $\sum_{n \leq x} d(n)$ is asymptotic to $x\log(x)$ as $x \to \infty$, which implies that

$$\lim_{x \to \infty} \frac{\sum_{n \leq x} d(n)}{x\log(x)} = 1$$

To prove the more precise formula (3) we return to the sum (4) which counts the number of lattice points in a hyperbolic region and take advantage of the fact that the hyperbola is symmetric to the line $q = d$. So we obtain that the total number of lattice points in the region is equal to twice the number below the line $q = d$ plus the number on the bisecting line segment.

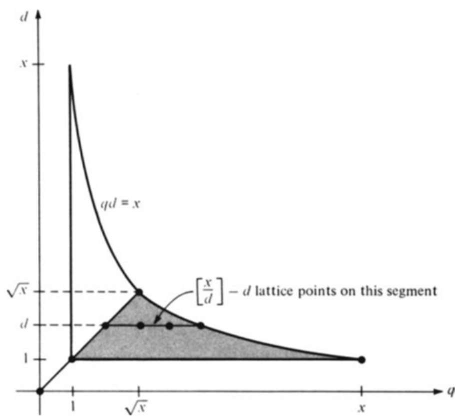

Figure 2: Total number of lattice points

From figure 2 we see that:

$$\sum_{n \le x} d(n) = 2 \sum_{d \le \sqrt{x}} \left(\left[\frac{x}{d} \right] - d \right) + [\sqrt{x}]$$

Now we use the relation $[y] = y + \mathcal{O}(1)$ and parts (a) and (d) of theorem 2.1 to obtain

$$\sum_{n \le x} d(n) = 2 \sum_{d \le \sqrt{x}} \left(\frac{x}{d} - d + \mathcal{O}(1) \right) + \mathcal{O}(\sqrt{x}) =$$

$$= 2x \sum_{d \le \sqrt{x}} \frac{1}{d} - 2 \sum_{d \le \sqrt{x}} d + \mathcal{O}(\sqrt{x}) =$$

$$= 2x \left\{ \log \sqrt{x} + C + \mathcal{O}\left(\frac{1}{\sqrt{x}} \right) \right\} - 2 \left\{ \frac{x}{2} + \mathcal{O}(\sqrt{x}) \right\} + \mathcal{O}(\sqrt{x}) =$$

$$= x \log(x) + (2C - 1)x + \mathcal{O}(\sqrt{x})$$

□

4 The average order of $\sigma_\alpha(n)$

We are going to calculate separately the cases $\alpha = 1$, $\alpha > 0$, and $\alpha < 0$.

Theorem 4.1. *For all $x \geq 1$ we have*

$$\sum_{n \leq x} \sigma_1(n) = \frac{1}{2}\zeta(2)x^2 + \mathcal{O}\left(x\log(x)\right)$$

Proof.

$$\sum_{n \leq x} \sigma_1(n) = \sum_{n \leq x}\sum_{q|n} q =$$

$$= \sum_{\substack{q,d \\ qd \leq x}} q = \sum_{d \leq x}\sum_{q \leq \frac{x}{d}} q \overset{Thm2.1(d)}{=} \sum_{d \leq x}\left(\frac{1}{2}\left(\frac{x}{d}\right)^2 + \mathcal{O}\left(\frac{x}{d}\right)\right) =$$

$$= \frac{x^2}{2}\sum_{d \leq x}\frac{1}{d^2} + \mathcal{O}\left(x\sum_{d \leq x}\frac{1}{d}\right) =$$

$$\overset{Thm2.1(a),(b)}{=} \frac{x^2}{2}\left(-\frac{1}{x} + \zeta(2) + \mathcal{O}\left(\frac{1}{x^2}\right)\right) + \mathcal{O}\left(x\log(x)\right) =$$

$$= \frac{1}{2}\zeta(2)x^2 + \mathcal{O}\left(x\log(x)\right)$$

□

Theorem 4.2. *If $x \geq 1$ and $\alpha > 0$, $\alpha \neq 1$, we have*

$$\sum_{n \leq x} \sigma_\alpha(n) = \frac{\zeta(\alpha+1)}{\alpha+1}x^{\alpha+1} + \mathcal{O}\left(x^\beta\right)$$

where $\beta = \max\{1, \alpha\}$

Proof.

$$\sum_{n \leq x} \sigma_\alpha(n) = \sum_{n \leq x}\sum_{q|n} q^\alpha = \sum_{d \leq x}\sum_{q \leq \frac{x}{d}} q^\alpha =$$

$$\overset{Thm2.1(d)}{=} \sum_{d \leq x}\left\{\frac{1}{\alpha+1}\left(\frac{x}{d}\right)^{\alpha+1} + \mathcal{O}\left(\frac{x^\alpha}{d^\alpha}\right)\right\} =$$

$$= \frac{x^{\alpha+1}}{\alpha+1}\sum_{d \leq x}\frac{1}{d^{\alpha+1}} + \mathcal{O}\left(x^\alpha\sum_{d \leq x}\frac{1}{d^\alpha}\right) =$$

$$\overset{Thm2.1(b)}{=} \frac{x^{\alpha+1}}{\alpha+1}\left\{\frac{x^{-\alpha}}{-\alpha}+\zeta(\alpha+1)+\mathcal{O}\left(x^{-\alpha-1}\right)\right\}+\mathcal{O}\left(x^{\alpha}\left\{\frac{x^{1-\alpha}}{1-\alpha}+\zeta(\alpha)+\mathcal{O}\left(x^{-\alpha}\right)\right\}\right)=$$

$$=\frac{x}{-\alpha(\alpha+1)}+\frac{x^{\alpha+1}}{\alpha+1}\zeta(\alpha+1)+\mathcal{O}(1)+\mathcal{O}\left(\frac{x}{1-\alpha}\right)+\mathcal{O}(x^{\alpha}\zeta(\alpha))+\mathcal{O}(1)=$$

$$=\mathcal{O}(x)+\frac{x^{\alpha+1}}{\alpha+1}\zeta(\alpha+1)+\mathcal{O}(1)+\mathcal{O}(x)+\mathcal{O}(x^{\alpha})=$$

$$=\frac{\zeta(\alpha+1)}{\alpha+1}x^{\alpha+1}+\mathcal{O}(x)+O(1)+\mathcal{O}\left(x^{\alpha}\right)=$$

$$=\frac{\zeta(\alpha+1)}{\alpha+1}x^{\alpha+1}+\mathcal{O}\left(x^{\beta}\right)$$

where $\beta=\max\{1,\alpha\}$ □

Finally, to find the average order of $\sigma_{\alpha}(n)$ for negative α we write $\alpha=-\beta$, where $\beta>0$

Theorem 4.3.

$$\sum_{n\leq x}\sigma_{-\beta}(n)=\begin{cases}\zeta(\beta+1)x+\mathcal{O}(x^{\delta}) & \textit{if } \beta\neq 1\\ \zeta(2)x+\mathcal{O}(\log\text{ (x)}) & \textit{if } \beta=1\end{cases}$$

Proof.

$$\sum_{n\leq x}\sigma_{-\beta}(n)=\sum_{n\leq x}\sum_{d|n}\frac{1}{d^{\beta}}=\sum_{d\leq x}\frac{1}{d^{\beta}}\sum_{q\leq\frac{x}{d}}1=$$

$$\overset{Thm2.1(d)}{=}\sum_{d\leq x}\frac{1}{d^{\beta}}\left\{\frac{x}{d}+\mathcal{O}(1)\right\}=$$

$$=x\sum_{d\leq x}\frac{1}{d^{\beta+1}}+\mathcal{O}\left(\sum_{d\leq x}\frac{1}{d^{\beta}}\right)$$

The last term is $\mathcal{O}(\log(\text{x}))$ if $\beta=1$ and $\mathcal{O}\left(x^{\delta}\right)$ if $\beta\neq 1$. Since

$$x\sum_{d\leq x}\frac{1}{d^{\beta+1}}\overset{Thm2.1(b)}{=}\frac{x^{1-\beta}}{-\beta}+\zeta(\beta+1)x+\mathcal{O}\left(x^{-\beta}\right)=\zeta(\beta+1)x+\mathcal{O}\left(x^{1-\beta}\right)$$

this completes the proof.

□

5 The average order of $\varphi(n)$

Before we calculate the average order for the partial sums of Euler's totient we have to state some definitions and theorems that we're going to use later. First of all we recall that for any integer n, we define the Möbius function $\mu(n)$ as the sum of the primitive nth roots of unity. This function is defined by the following three properties:

- $\mu(1) = 1$;
- $\mu(n) = (-1)^k$, if n is a product of k different primes;
- $\mu(n) = 0$, otherwise; that is, if n is divisible by a square different from 1.

Definition 2. *If f and g are two arithmetical functions we define their Dirichlet product (or Dirichlet convolution) to be the arithmetical function h defined by the equation*

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right)$$

Now we can state the following theorem without proof:

Theorem 5.1. *Given two functions $F(s)$ and $G(s)$ represented by Dirichlet series,*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \text{ for } \sigma > a$$

and

$$G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s} \text{ for } \sigma > b$$

Then in the half-plane where both series converge absolutely we have

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s} \tag{5}$$

where $h = f * g$, the Dirichlet convolution of f and g :

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right)$$

With this theorem we can evaluate the series $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2}$ as follows. Since both series $\sum n^{-s}$ and $\sum \mu(n) n^{-s}$ converge absolutely for $\sigma > 1$, we can multiply them out and by taking $f(n) = 1$ and $g(n) = \mu(n)$ in (5) we get $\sum_{n=1}^{\infty} \frac{1}{n^s} \cdot \sum_{m=1}^{\infty} \frac{\mu(m)}{m^s} = \sum_{\nu=1}^{\infty} \frac{c_\nu}{\nu^s}$, where $c_\nu = \sum_{k|v} \mu(k)$. From the previous lesson we know that $\sum_{d|n} \mu(d) = \begin{cases} 1, & \text{if } n = 1 \\ 0, & \text{if } n > 1 \end{cases}$. So we get $c_1 = 1$, and $c_n = 0$ for $n > 1$. This finally implies that $\zeta(s) \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = 1 \quad \text{if } \sigma > 1$ and also $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$. So we obtain the relation $\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}$, which we are going to use for the following theorem.

Theorem 5.2.

$$\sum_{n \le x} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2} + \mathcal{O}\left(\frac{1}{x}\right)$$

Proof. By part (c) of Theorem 2.1 and by the result above we obtain:

$$\begin{aligned} \sum_{n \le x} \frac{\mu(n)}{n^2} &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} - \sum_{n > x} \frac{\mu(n)}{n^2} \\ &= \frac{6}{\pi^2} + \mathcal{O}\left(\sum_{n > x} \frac{1}{n^2}\right) = \frac{6}{\pi^2} + \mathcal{O}\left(\frac{1}{x}\right) \end{aligned}$$

□

We know from the previous presentation that:

$$\sum_{d|n} \varphi(d) = n \tag{6}$$

and that we can relate the Möbius function and the Euler's totien function with the first Möbius inversion formula (which has been demonstrated in previous presentations):

Theorem 5.3. *If f is an arithmetical function, and*

$$g(n) = \sum_{d|n} f(d)$$

then

$$f(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right)$$

From this theorem and (6) it follows that:

$$\varphi(n) = \sum_{d|n} \mu(d)\frac{n}{d} = n\sum_{d|n} \frac{\mu(d)}{d}$$

We are now ready to calculate the average order of $\varphi(n)$:

Theorem 5.4. *For $x > 1$ we have*

$$\sum_{n\leq x} \varphi(n) = \frac{3}{\pi^2}x^2 + \mathcal{O}(x\log x)$$

so the average order of $\varphi(n)$ is $3n/\pi^2$.

Proof.

$$\begin{aligned}
\sum_{n\leq x} \varphi(n) &= \sum_{n\leq x}\sum_{d|n} \mu(d)\frac{n}{d} = \sum_{\substack{q,d \\ qd\leq x}} \mu(d)q = \sum_{d\leq x} \mu(d) \sum_{q\leq x/d} q \\
&= \sum_{d\leq x} \mu(d) \left\{\frac{1}{2}\left(\frac{x}{d}\right)^2 + \mathcal{O}\left(\frac{x}{d}\right)\right\} \\
&= \frac{1}{2}x^2 \sum_{d\leq x} \frac{\mu(d)}{d^2} + \mathcal{O}\left(x \sum_{d\leq x} \frac{1}{d}\right) \\
&= \frac{1}{2}x^2 \left\{\frac{6}{\pi^2} + \mathcal{O}\left(\frac{1}{x}\right)\right\} + \mathcal{O}(x\log(x)) = \frac{3}{\pi^2}x^2 + \mathcal{O}(x\log(x))
\end{aligned}$$

□

References

- *Introduction to Analytic Number Theory; Chapter 3*
  T.M. Apostol
  Springer 1976

- *Introduction to Analytic Number Theory; Chapter VI*
  K. Chandrasekharan
  Springer 1968

- *An introduction to the theory of numbers (5th edition); Chapter XVIII*
  G.H. Hardy and E.M. Wright
  Oxford University Press 1980