Introduction to Elementary Analytic Number Theory

Masum Billal

Monday 1st November, 2021

Contents

Notations

gcd(a, b) Greatest common divisor of a and b

lcm(a, b) Least common multiple of a and b

$\varphi(n)$ Euler's totient function of n

$\tau(n)$ Number of divisors of n

$\sigma(n)$ Sum of divisors of n

$\omega(n)$ Number of distinct prime divisors of n

$\Omega(n)$ Number of total prime divisors of n

$\lambda(n)$ Liouville function of n

$\mu(n)$ Möbius function of n

$\vartheta(x)$ Tchebycheff function of the first kind

$\psi(x)$ Tchebycheff function of the second kind

$\zeta(s)$ Zeta function of the complex number s

Chapter 1

Arithmetic Functions

In this chapter, we will discuss some generalized arithmetic functions and their asymptotic behavior. By asymptotic behavior, we mean that we want to understand how a function $f(x)$ grows as x tends to infinity. A common way of analyzing growth of an arithmetic function f is to consider the order of an arithmetic function.

Order of Arithmetic Function. The order of an arithmetic function f is defined by the asymptotic $\lim_{x\to\infty} f(x)$. To understand the growth of f, we often analyze the asymptotic of partial summation

$$\lim_{x\to\infty}\sum_{n\leq x} f(n)$$

For example, the prime counting function is

$$\pi(x)=\sum_{n\leq x} C(n)$$

where $C(n)$ is the characteristic function of, that is, $C(n)=1$ if n is a prime otherwise $C(n)=0$. One of the biggest questions we will try to answer is how $\lim_{x\to\infty}\pi(x)$ behaves.

Summatory Function. For an arithmetic function f, the *summatory function* of f is defined as

$$F(n)=\sum_{d\in\mathbb{S}} f(d)$$

where $\mathbb{S}$ is some set possibly dependent on n. When $\mathbb{S}$ is the set of divisors of n, the number of divisor function $\tau(n)$ is the summatory function of the unit function $u(n)=1$ and the sum of divisor function $\sigma(n)$ is the summatory function of the invariant function $f(n)=n$. Another summatory function is the partial summation

$$\sum_{n\leq x} f(n)$$

Associated with this is the average order of f.

Average Order. For an arithmetic function f,

$$\lim_{x\to\infty} \frac{\sum_{n\leq x} f(x)}{x}$$

is the *average order*. In this context, a very interesting way of analyzing growth is the *normal order* of f. The concept of normal numbers arises from Hardy and Aiyangar.*

Normal Order. Let f and F be arithmetic functions such that

$$(1-\epsilon)F(n) < f(n) < (1+\epsilon)F(n) \tag{‡ 1.1}$$

holds for *almost* all $n \leq x$ as $x \to \infty$. Then we say that F is the *normal order* of f. A trivial(?) example of normal order is that almost all positive integers not exceeding x are composite if x is sufficiently large. We should probably elaborate on what we mean by *almost* here. One interpretation is that the number of positive integers not exceeding x which are prime is very small compared to x. Similarly, f is of order F means that the number of positive integers n not exceeding x which do not satisfy ‡ 1.1 is very small compared to x.

An interesting property in summatory functions is that

$$\begin{aligned}\sum_{i=1}^{n} F(i) &= \sum_{i=1}^{n} \sum_{d|i} f(d)\\ &= \sum_{i=1}^{n} \left\lfloor \frac{n}{i} \right\rfloor f(i)\end{aligned}$$

Here, the last equation is true because there are $\lfloor n/i \rfloor$ multiples of i not exceeding n.

1.1 Order of Arithmetic Functions

Recall that the number of divisor function $\tau(n) = \sum_{ab=n} 1$. We can generalize this as follows.

Generalized Number of Divisors. The *generalized number of divisor* function is defined as

$$\tau_k(n) = \sum_{d_1 \cdots d_k = n} 1$$

So $\tau_k(n)$ is the number of ways to write n as a product of k positive integers. Similarly,

*Godfrey Harold Hardy and Srinivasa Ramanujan Aiyangar, "The normal number of prime factors of a number n", *Quarterly Journal of Mathematics*, vol. xlviii (1917), pp. 76-92.

we can take the sum of divisor function and generalize it.

Generalized Sum of Divisors. The *generalized sum of divisor* function can be defined as

$$\sigma_k(n) = \sum_{d|n} d^k$$

At this point, we should discuss some asymptotic notions.

Big O. Let f and g be two real or complex valued functions. We say that

$$f(x) = O(g(x))$$

if there is a positive real constant C such that

$$|f(x)| \leq Cg(x)$$

for all sufficiently large x. It is also written as $f(x) \ll g(x)$ or $g(x) \gg f(x)$. When we say g is an asymptotic estimate of f, we mean that

$$f(x) = g(x) + O(h(x))$$

for two functions g and h as $x \to \infty$. Here, h is the *error term* which obviously should be of lower magnitude than g. In particular, $f(x) = O(1)$ means that f is bounded above by some positive constant. Some trivial examples are $x^2 = O(x^3)$, $x + 1 = O(x)$ and $x^2 + 2x = O(x^2)$. We usually want $g(x)$ to be as small as possible to avoid triviality. A useful example is

$$\lfloor x \rfloor = x + O(1)$$

since $x = \lfloor x \rfloor + \{x\}$ and $0 \leq \{x\} < 1$.

Small O. Let f and g be two real or complex valued functions. Then the following two statements are equivalent

$$f(x) = o(g(x)) \tag{‡ 1.2}$$

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0 \tag{‡ 1.3}$$

Some trivial examples are $1/x = o(1)$, $x = o(x^2)$ and $2x^2 \neq o(x^2)$. Landau[†] states that the symbol O had been first used by Bachmann.[‡] Hardy[§] uses the notations $\prec$ and $\succ$

[†] Edmund Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, 1909, vol. ii, Page 883 (second volume is paged consecutively after first volume).

[‡] Paul Gustav Heinrich Bachmann, *Analytische zahlentheorie*, 1894, vol. ii, Page 401.

[§] Godfrey Harold Hardy, *Orders of Infinity: The 'Infinitärcalcül' of Paul Du Bois-Reymond, Cambridge Tracts in Mathematics*, Cambridge University Press, 1910.

respectively but they are no longer in practice. Hardy and Riesz[¶] adopted the notations small o and big O and today these are the primary notations for this purpose.

It should be evident that having an estimate with respect to O asymptotic formulas is more desirable than o formulas. By nature, O formulas give us a better understanding and a specific estimate whereas o does not always say as much. Moreover, working with O is a lot easier than working with o. For example,

$$\sum O(f(x)) = O\left(\sum f(x)\right)$$
$$\int O(f(x))dx = O\left(\int f(x)dx\right)$$

Or consider the possibility that we can very easily deal with constants that would otherwise pop up here and there unnecessarily. With the help of O,

$$O(1) + c = O(1)$$
$$O(cf(x)) = O(f(x))$$

and so on.

Equivalence. Let f and g be two real or complex valued functions. We say that they are *asymptotically equivalent* if

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = 1$$

and we denote it by $f \sim g$. So, we can say that g is an asymptotic formula for f. An example is $x^2 \sim x^2 + x$. Another example in connection with normal order is that f has normal order F if the number of n not satisfying ‡ 1.1 is $o(x)$. We can also say, the number of n satisfying ‡ 1.1 is $\sim x$. Note the following.

$$f \sim g \iff |f(x) - g(x)| = o(g(x))$$

We will use these symbols extensively throughout the book. It is of utmost importance that the reader gets well familiarized with these notions since they will be crucial in understanding much of this book. The primary motivation behind these asymptotic notions is to get an as precise as possible idea about the *order of magnitude* of a certain function. This is why we will be leaning more towards $x^2 + 2x = O(x^2)$ than $x^2 + 2x = O(x^3)$ even though both are mathematically correct. The reason is, even though $x^2 + 2x = O(x^3)$ is true, it is taking away a great portion of the accuracy to which we suppose $x^2 + 2x$ should be measured with. On the other hand, we easily see that we cannot have $x^2 + 2x = O(x^\epsilon)$ for $\epsilon < 2$. Under the same philosophy, we define the order of magnitude equivalence.

Definition. If f and g are functions such that both $f(x) \ll g(x)$ and $g(x) \ll f(x)$ hold, then we write $f \asymp g$ and say that f and g have the *same order of magnitude*.

[¶] Godfrey Harold Hardy and Marcel Riesz, *The general theory of Dirichlet's series*, Cambridge University Press, 1915.

Now, we are interested in the order of general number of divisors and general sum of divisors. Let us define the cumulative sum of these functions.

$$S_k(x) = \sum_{n\leq x} \sigma_k(n)$$
$$T_k(x) = \sum_{n\leq x} \tau_k(n)$$

Notice the following.

$$\begin{aligned} S_k(x) &= \sum_{n\leq x}\sum_{d\mid n} d^k \\ &= \sum_{n\leq x} \left\lfloor \frac{x}{n} \right\rfloor n^k \\ &= \sum_{n\leq x} \left(\frac{x}{n} + O(1)\right) n^k \\ &= x\sum_{n\leq x} n^{k-1} + O\left(\sum_{n\leq x} n^k\right) \end{aligned}$$

We can use this to establish an asymptotic for $T_k(x)$ if we can establish the asymptotic of $A_2(x)$. We will get to that in a moment. First, let us take care of the summation within the big O bracket. We have the trivial inequality that

$$\begin{aligned} \sum_{n\leq x} n^k &\leq \sum_{n\leq x} x^k \\ &= x^k \sum_{n\leq x} 1 \\ &= \lfloor x \rfloor x^k \\ &= (x + O(1))x^k \\ &= x^{k+1} + O(x^k) \end{aligned}$$

We have that $S_k(x) = x\left(x^k + O(x^{k-1})\right) + O(x^{k+1}) = O(x^{k+1})$. Although weak, we get an estimate this way. On this note, an interested reader can try and prove that

$$(n+1)^{k+1} - 1 = \sum_{i=0}^{k} \binom{k+1}{i} \mathfrak{S}(n, i)$$

where $\mathfrak{S}(x, k) = \sum_{n\leq x} n^k$. This is known as the *Pascal identity* (see Pascal,[‖] for an

[‖] Blaise Pascal, "Sommation des puissances numériques", *Oeuvres complètes, Jean Mesnard, ed., Desclée-Brouwer, Paris*, vol. iii (1964), pp. 341-367.

English translation, see Knoebel et al.**). Lehmer†† proves that

$$\mathfrak{S}(x,k) = \frac{x^{k+1}}{k+1} + \Delta \tag{‡ 1.4}$$

where $|\Delta| \le x^k$. The reader may also be interested in MacMillan and Sondow.‡‡

We shall try to estimate T in a similar fashion. First, see that

$$\begin{aligned}\tau_k(n) &= \sum_{d_1 \cdots d_k = n} 1\\ &= \sum_{d_k \mid n} \sum_{d_1 \cdots d_{k-1} = n/d_k} 1\\ &= \sum_{d \mid n} \tau_{k-1}\left(\frac{n}{d}\right)\end{aligned}$$

Note that the two sets $\{d : d \mid n\}$ and $\{n/d : d \mid n\}$ are actually the same. So, we get

$$\tau_k(n) = \sum_{d \mid n} \tau_{k-1}(d)$$

Using this for T,

$$\begin{aligned}T_k(x) &= \sum_{n \le x} \tau_k(n)\\ &= \sum_{n \le x} \sum_{d \mid n} \tau_{k-1}(d)\\ &= \sum_{n \le x} \left\lfloor \frac{x}{n} \right\rfloor \tau_{k-1}(n)\\ &= \sum_{n \le x} \left(\frac{x}{n} + O(1)\right) \tau_{k-1}(n)\\ &= x \sum_{n \le x} \frac{\tau_{k-1}(n)}{n} + O\left(\sum_{n \le x} \tau_{k-1}(n)\right)\end{aligned}$$

Thus, we have the recursive result

$$T_k(x) = x \sum_{n \le x} \frac{\tau_{k-1}(n)}{n} + O(T_{k-1}(x))$$

**Arthur Knoebel et al., “Sums of numerical powers”, in, *Mathematical Masterpieces: Further chronicles by the explorers*, Springer-Verlag, 2007, pp. 32-37.

††Derrick Norman Lehmer, “Asymptotic evaluation of certain Totient Sums”, *American Journal of Mathematics*, vol. xxii, no. 4 (1900), pp. 293-335, doi: `10.2307/2369728`, Chapter II, Theorem 1.

‡‡Kieren MacMillan and Jonathan Sondow, “Proofs of power sum and binomial coefficient congruences via Pascal’s identity”, *The American Mathematical Monthly*, vol. cxviii, no. 6 (2011), pp. 549-551, doi: `10.4169/amer.math.monthly.118.06.549`.

It gets nontrivial how to proceed from here. Consider the *harmonic sum*

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

It does not seem easy to calculate H accurately, however, we can make a decent attempt to estimate H. The tool that is best suited for carrying out such an estimation is the *Abel partial summation formula*. Abel[§§] states this formula which today is a cornerstone of analytic number theory.

Theorem 1.1 (Abel partial summation formula). *Let $\{a_n\}$ be a sequence of real numbers and f be a continuous differentiable function in $[y, x]$. If the partial sums of $\{a_n\}$ is*

$$A(x) = \sum_{n \leq x} a_n$$

are known, then

$$\sum_{y < n \leq x} a_n f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt$$

In particular, if f is an arithmetic function,

$$\sum_{n \leq x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt$$

Proof. □

It is not straightforward to realize how such a formula can be as influential as we are describing it to be. Notice that, the formula essentially converts a discreet sum into an integral, which occasionally may be calculable. If the integral is not calculable, we may be able to estimate its value sometimes. We should mention that Aiyangar[¶¶] also uses a method that can only be described as the partial summation formula. It is unclear if Ramanujan simply knew about this. Considering he shows the calculation instead of just mentioning the formula, it is certainly possible he came up with the idea on his own, possibly before he had been working on that particular paper. He essentially derives the partial summation formula while trying to express a sum of the form $\sum_{p \leq x} \phi(p)$ with respect to $\pi(x), \phi(x)$ and an integral where $\pi(x)$ is the number of primes not exceeding x. A consequence of Abel partial summation formula is the celebrated *Euler's summation formula*.

[§§] Niels Henrik Abel, "Untersuchungen über die Reihe: $1 + (m/1)x + m \cdot (m-1)/(1 \cdot 2) \cdots x^2 + m \cdot (m-1) \cdot (m-2)/(1 \cdot 2 \cdot 3) \cdots x^3 + \ldots$", *Journal für Math.*, vol. i (1826), pp. 311-339, doi: 10.1515/9783112347386-030.

[¶¶] Srinivasa Ramanujan Aiyangar, "Highly Composite Numbers", in, *Collected papers of Srinivasa Ramanujan*, ed. by Godfrey Harold Hardy et al., Cambridge University Press, 1927, pp. 78-128, Page 83, §4.

Theorem 1.2 (Euler's summation formula). *Let f be a continuous differentiable function in $[y, x]$. Then*

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)dt + \int_y^x \{t\}f(t)dt + \{y\}f(y) - \{x\}f(x)$$

where $\{t\} = t - \lfloor t \rfloor$ is the fractional part of t.

Proof. □

As an application of Euler's summation formula, we can derive a result similar to ‡ 1.4 taking $f(n) = n^k$ for $k \geq 0$.

$$\begin{aligned}
\mathfrak{S}_k(x) &= \sum_{n\leq x} n^k \\
&= \int_1^x t^k dt + k\int_1^x t^{k-1}(t - \lfloor t \rfloor)dt + 1 - (x - \lfloor x \rfloor)x^k \\
&= \frac{x^{k+1}}{k+1} - \frac{1}{k+1} + O\left(k\int_1^x t^{k-1}dt\right) + O(x^k) \\
&= \frac{x^{k+1}}{k+1} + O(x^k)
\end{aligned}$$

Setting $a_n = \tau_{k-1}(n)$ and $f(n) = 1/n$ in Abel partial summation formula, we get

$$\sum_{n\leq x} \frac{\tau_{k-1}(n)}{n} = \frac{T_{k-1}(x)}{x} - \int_1^x -\frac{T_{k-1}(t)}{t^2}dt$$

Thus, we have a result where we can inductively get to the final expression. First, let us see the case $k = 2$.

$$\sum_{n\leq x} \tau(n) = \sum_{n\leq x} \left\lfloor \frac{x}{n} \right\rfloor$$

Clearly, this is just the number of pairs (a, b) such that $ab \leq x$. We can divide the pairs in two classes. In the first class, $1 \leq a \leq \sqrt{x}$ and in the second one, $a > \sqrt{x}$. In the first case, for a fixed a, there are $\lfloor x/a \rfloor$ possible choices for a valid value of b. So, the number of pairs in the first case is

$$\sum_{a\leq\sqrt{x}} \left\lfloor \frac{x}{a} \right\rfloor$$

In the second case, since $a > \sqrt{x}$ and $b \leq x/a$, we must have $b \leq \sqrt{x}$. For a fixed b, there are $\lfloor x/b \rfloor - \sqrt{x}$ choices for a valid value of a, the choices namely are

$$\lfloor x \rfloor + 1, \dots, \left\lfloor \frac{x}{b} \right\rfloor$$

Then the number of pairs in this case is

$$\sum_{b \leq \sqrt{x}} \left\lfloor \frac{x}{b} \right\rfloor - \sqrt{x}$$

Thus, the total number of such pairs is

$$\sum_{a \leq \sqrt{x}} \left\lfloor \frac{x}{a} \right\rfloor + \sum_{b \leq \sqrt{x}} \left(\left\lfloor \frac{x}{b} \right\rfloor - \sqrt{x} \right) = 2 \sum_{n \leq \sqrt{x}} \left\lfloor \frac{x}{n} \right\rfloor - \lfloor x \rfloor^2 \tag{‡ 1.5}$$

For getting past this sum, we have to deal with the sum

$$\begin{aligned} \sum_{n \leq \sqrt{x}} \lfloor x/n \rfloor &= \sum_{n \leq \sqrt{x}} \left(\frac{x}{n} + O(1) \right) \\ &= x \sum_{n \leq \sqrt{x}} \frac{1}{n} + O(\sqrt{x}) \\ &= xH(\sqrt{x}) + O(\sqrt{x}) \end{aligned}$$

Setting $a_n = 1$ and $f(n) = 1/n$ in Abel partial summation formula, we get

$$H(x) = \frac{A(x)}{x} - \int_1^x -\frac{A(t)}{t^2} dt$$

Here, $A(x) = \lfloor x \rfloor = x + O(1)$. Using this,

$$\begin{aligned} H(x) &= 1 + O\left(\frac{1}{x}\right) + \int_1^x \left(\frac{1}{t} + \frac{O(1)}{t^2} \right) dt \\ &= 1 + O\left(\frac{1}{x}\right) + \int_1^x \frac{1}{t} dt + O\left(\int_1^x \frac{1}{t^2} dt \right) \\ &= 1 + O\left(\frac{1}{x}\right) + \log x + O\left(1 - \frac{1}{x}\right) \end{aligned}$$

Thus, we have the following result.

Theorem 1.3.

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

where C is a constant.

We get a more precise formulation of $H(x)$ by considering the limit $x \to \infty$ which removes $O(1/x)$ from the expression since this limit would be 0.

Theorem 1.4. *There is a constant γ such that*

$$\gamma = \lim_{x\to\infty} (H(x) - \log x)$$

This constant γ is now known as *Euler's constant* or *Euler-Mascheroni's constant*, although, neither Euler nor Mascheroni used the notation γ for this constant. Euler[***] (republished in Euler[†††]) used C and O in his original paper. Mascheroni[‡‡‡] used A and a. Today it is not known whether γ is even irrational. For now, we will not require the use of γ, so we will use Theorem 1.3. Applying this, we have

$$\begin{aligned}\sum_{n\le\sqrt{x}} \left\lfloor \frac{x}{n} \right\rfloor &= xH(\sqrt{x}) + O(\sqrt{x})\\ &= x\left(C + \log\sqrt{x} + O\left(\frac{1}{\sqrt{x}}\right)\right) + O(\sqrt{x})\\ &= \frac{1}{2}x\log x + Cx + O\left(\frac{x}{\sqrt{x}}\right) + O(\sqrt{x})\\ &= \frac{1}{2}x\log x + O(x)\end{aligned}$$

We can now use this to get

$$\begin{aligned}\sum_{n\le x}\tau(n) &= 2\sum_{n\le\sqrt{x}} \lfloor x/n \rfloor - \lfloor\sqrt{x}\rfloor^2\\ &= x\log x + O(x)\end{aligned}$$

Thus, we get the following result.

$$\frac{\sum_{n\le x}\tau(n)}{x} = \log x + O(1)$$

Lejeune[§§§] actually proves the more precise result given below.

Theorem 1.5 (Dirichlet's average order of τ theorem)**.**

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

where γ is the Euler-Mascheroni constant.

***Leonhard Euler, "De Progressionibus Harmonicis Observationes", *Commentarii academiae scientiarum Petropolitanae*, vol. vii (1740), pp. 150-161.

†††Leonhard Euler, "E-43: De Progressionibus Harmonicis Observationes", *Spectrum* (2020), pp. 133-141, doi: 10.1090/spec/098/23.

‡‡‡Lorenzo Mascheroni, *Adnotationes ad calculum Integralem Euleri*, Galeatii, 1790.

§§§Dirichlet Peter Gustav Lejeune, "Über Die Bestimmung Der Mittleren Werthe", in, *G. Lejeune Dirichlet's Werke*, ed. by L. Kronecker and L. Fuchs, Druck Und Verlag Von Georg Reimer., 1897, vol. ii, pp. 49-66.

Then Dirichlet's theorem on τ can be restated as *the average order of τ is $O(\log x)$.* Aiyangar[¶¶¶] points out in his paper that the error term $O(1/\sqrt{x})$ in Dirichlet's theorem can be improved to $O\left(x^{-\frac{2}{3}+\epsilon}\right)$ or $O\left(x^{-2/3}\log x\right)$ as Landau[17] shows.

We can now get back to estimating T. Using Abel partial summation formula, we were able to deduce

$$T_k(x) = O\left(T_{k-1}(x)\right) + x\int_1^x \frac{T_{k-1}(t)}{t^2}dt$$

Using Dirichlet's average order of τ theorem, $T(x) = x\log x + O(x)$, so

$$\begin{aligned}
T_3(x) &= O(T(x)) + x\int_1^x \frac{T(t)}{t^2}dt \\
&= O(x\log x) + x\int_1^x \frac{\log t + O(1)}{t}dt \\
&= O(x\log x) + x\int_1^x \frac{\log t}{t}dt + xO\left(\int_1^x \frac{1}{t}dt\right) \\
&= O(x\log x) + x\int_1^x \frac{\log t}{t}dt
\end{aligned}$$

Using integration by parts,

$$\begin{aligned}
\int \frac{\log t}{t}dt &= \log t\int \frac{1}{t} - \int\left(\frac{1}{t}\int \frac{1}{t}dt\right)dt \\
&= \log^2 t - \int \frac{\log t}{t}dt
\end{aligned}$$

Thus, we get

$$\int_1^x \frac{\log t}{t}dt = \frac{1}{2}\log^2 x$$

which in turn gives

$$T_3(x) = \frac{1}{2}x\log^2 x + O(x\log x)$$

We leave it as an exercise for the reader to prove the following (from what we have already developed, induction is one way to go about it).

[¶¶¶] Aiyangar, "Highly Composite Numbers" cit.

[17] Edmund Landau, "Über die Anzahl der Gitterpunkte in geweissen Bereichen", *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. xix (1912), pp. 687-772, Page 689.

Theorem 1.6. *Let k be a positive integer. Then*

$$T_k(x) = \frac{1}{(k-1)!} x \log^{k-1} x + O\left(x \log^{k-2} x\right)$$

The reason we do not write $T_k(x)$ as $O\left(x \log^{k-1} x\right)$ directly is because in this case, we already know the constant multiplier of $x \log^{k-1} x$ which is not ugly. Usually, we write $O(f(x))$ when we do not know what the constant multiplier of $f(x)$ is or when it gets too big to keep track of. Landau[18] states a sharper result.

$$T_k(x) = x\left(\sum_{m=0}^{k-1} b_m \log^m x\right) + O\left(x^{1-\frac{1}{k}}\right) + O\left(x^{1-\frac{1}{k}} \log^{k-2} x\right)$$

Let us now turn our attention to improving the asymptotic of $S_k(x)$.

$$\begin{aligned} S_k(x) &= \sum_{n\le x}\sum_{d\mid n} d^k \\ &= \sum_{n\le x}\sum_{m\le x/n} m^k \\ &= \sum_{n\le x} \mathfrak{S}_k\left(\frac{x}{n}\right) \\ &= \sum_{n\le x} \frac{x^{k+1}}{(k+1)n^{k+1}} + O\left(\frac{x^k}{n^k}\right) \\ &= \frac{x^{k+1}}{k+1}\sum_{n\le x}\frac{1}{n^{k+1}} + O\left(x^k \sum_{n\le x}\frac{1}{n^k}\right) \end{aligned}$$

Here, we can see that the function

$$\sum_{n\le x}\frac{1}{n^k}$$

occurs repeatedly. It is in fact, the partial sum of the famous Euler's *zeta function*.

Zeta Function. For a complex number s, the zeta function $\zeta(s)$ is defined as

$$\zeta(s) = \sum_{n\ge 1}\frac{1}{n^s}$$

We will discuss zeta function in details in Section 1.2. For now, let us establish a result

[18] Edmund Landau, "Über eine idealtheoretische funktion", *Transactions of the American Mathematical Society*, vol. xiii, no. 1 (1912), pp. 1-21, doi: 10.1090/s0002-9947-1912-1500901-6, Page 2.

similar to Dirichlet's average order of τ theorem for partial sums of ζ. Setting $f(n) := n^{-s}$ and $a_n = 1$ in Abel partial summation formula, $A(x) = \lfloor x \rfloor = x + O(1)$ and

$$\begin{aligned}
\sum_{n\leq x}\frac{1}{n^s} &= \lfloor x\rfloor x^{-s} - \int_1^x (t+O(1))f'(t)dt\\
&= x^{1-s} + O\left(x^{-s}\right) + s\int_1^x t^{-s}dt + O\left(s\int_1^x t^{-s-1}dt\right)\\
&= x^{1-s} + \frac{s}{1-s}\left(x^{1-s}-1\right) + O\left(\int_1^x t^{-s-1}dt\right)\\
&= \frac{x^{1-s}}{1-s} + C + O(x^{-s})
\end{aligned}$$

Similar to γ, we can take $x \to \infty$ and get the following result.

Theorem 1.7. *Let s be a positive real number other than 1. Then*

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

where C is a constant similar to Euler-Mascheroni constant dependent on s and

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

Furthermore, if $0 < s < 1$, then $C = \zeta(s)$ since $x^{1-s} \to 0$.

We can now get back to estimating $S_k(x)$.

$$\begin{aligned}
S_k(x) &= \frac{x^{k+1}}{k+1}\sum_{n\leq x}\frac{1}{n^{k+1}} + O\left(x^k\sum_{n\leq x}\frac{1}{n^k}\right)\\
&= \frac{x^{k+1}}{k+1}\left(\frac{x^{-k}}{-k} + \zeta(k+1) + O(x^{-k-1})\right) + O\left(x^k\left(\frac{x^{1-k}}{1-k} + \zeta(k) + O(x^{-k})\right)\right)\\
&= \frac{x}{-k(k+1)} + \frac{x^{k+1}}{k+1}\zeta(k+1) + O(x^{k+1-k-1}) + O\left(\frac{x}{1-k} + x^k\zeta(k) + O(1)\right)\\
&= \frac{x^{k+1}}{k+1}\zeta(k+1) + O(x) + O(1) + O(x+x^k)
\end{aligned}$$

From this, we finally get the following.

Theorem 1.8. *Let k be a positive integer. Then*

$$S_k(x) = \frac{x^{k+1}}{k+1}\zeta(k+1) + O(x^{\max(1,k)})$$

We leave the case when k is a negative integer as an exercise. Next, we consider a generalization of the Euler's totient function $\varphi(n)$.

$$\varphi(x,a)=\sum_{\substack{n\le x\\ \gcd(n,a)=1}} 1$$

For a positive integer n, $\varphi(n)=\varphi(n,n)$ and *Jordan function* is a generalization of φ.

Jordan Function. Let n and k be positive integers. Then the Jordan function $J_k(n)$ is the number of k tuples of positive integers not exceeding n that are relatively prime to n.

$$J_k(n)=\sum_{\substack{1\le a_1,\ldots,a_k\le n\\ \gcd(a_1,\ldots,a_k,n)=1}} 1$$

Lehmer[19] used the notation $\varphi_k(n)$ but today $J_k(n)$ is used more often. Jordan[20] first discussed this function and Lehmer[21] developed some asymptotic results. Jordan totient function is interesting not only because it is a generalization of Euler's totient function but also because it has many interesting properties. For example, similar to φ, we can show that

$$J_k(n)=\prod_{p^e\|n} p^{k(e-1)}(p-1)$$
$$J_k(n^m)=n^{k(m-1)}J_k(n)$$

Lehmer[22] proves the following which he calls the *fundamental theorem*.

$$J_k(mn)=J_k(n)\prod_{p^e\|m}\left(p^{ke}-p^{k(e-1)}\lambda(n,p)\right) \tag{‡ 1.6}$$

where $\lambda(n,p)=0$ if $p\mid n$ otherwise $\lambda(n,p)=1$. We leave the proof of this result and the following to the reader.

$$\sum_{d\mid n} J_k(d)=n^k$$

Like $\sigma_k(n)$, $J_k(n)$ is also related to the sum $\mathfrak{S}(x,k)$. But we do not derive the order of $J_k(n)$ yet.

[19] Lehmer, "Asymptotic evaluation of certain Totient Sums" cit.
[20] Camille Jordan, *Traiteé des substitutions et des équations algébriques*, Gauthier-Villars, Paris, 1870, Page 95 – 97.
[21] Lehmer, "Asymptotic evaluation of certain Totient Sums" cit.
[22] Ibid., Theorem VI.

1.2 Dirichlet Series and Dirichlet Convolution

We encountered ζ when we tried to develop an asymptotic for $S_k(x)$. The function ζ has quite a rich history. Today ζ is mostly called Riemann's zeta function, however, Euler is the first one to investigate this function. Euler started working on ζ around 1730. During that period, the value of $\zeta(2)$ was unknown and of high interest among prominent mathematicians. Ayoub[23] is a very good read on this subject. Euler's first contribution in this matter is Euler[24] where he proves that $\zeta(2) \approx 1.644934$. The paper was first presented to the St. Petersburg Academy on March 5, 1731 and republished in Euler.[25] Euler[26] (republished in Euler[27]) proves the following fundamental result which essentially gives a new proof of infinitude of primes.

Theorem 1.9 (Euler's identity). *Let s be a positive integer. Then*

$$\zeta(s) = \prod_p \frac{p^s}{p^s - 1}$$

where p extends over all primes.

One of the results in Euler[28] is the following which we shall prove later.

$$\sum_{n \le x} \frac{1}{p} \sim \log \sum_{n \le x} \frac{1}{x}$$

Here, $\sim$ is the asymptotic equivalence we have already defined. Even though Euler is the main architect behind the development of ζ, Riemann[29] is the first one to consider ζ for complex s instead of real s only. By tradition, we write $s = \sigma + it$ where $\sigma = \Re(s)$ is the real part of s and $t = \Im(s)$ is the imaginary part of s.

Dirichlet Series. For a complex number s, a *Dirichlet series* is a series of the form

$$\mathfrak{D}_a(s) = \sum_{n \ge 1} \frac{a(n)}{n^s}$$

[23] Raymond Ayoub, "Euler and the zeta function", *The American Mathematical Monthly*, vol. lxxxi, no. 10 (1974), pp. 1067-1086, doi: 10.2307/2319041.

[24] Leonhard Euler, "De summatione innumerabilium progressionum", *Commentarii academiae scientiarum Petropolitanae*, vol. v (1738), pp. 91-105.

[25] Leonhard Euler, "E-20: De summatione innumerabilium Progressionum", *Spectrum* (2020), pp. 52-64, doi: 10.1090/spec/098/10.

[26] Leonhard Euler, "Variae Observationes circa series infinitas", *Commentarii academiae scientiarum Petropolitanae*, vol. ix (1744), pp. 160-188.

[27] Leonhard Euler, "E-72: Variae Observationes circa series Infinitas", *Spectrum* (2020), pp. 249-260, doi: 10.1090/spec/098/41.

[28] Euler, "Variae Observationes circa series infinitas" cit.

[29] Bernhard Riemann, "Ueber die anzahl der primzahlen unter einer gegebenen grösse", *Monatsberichte der Berliner Akademie* (Nov. 1859), pp. 136-144, doi: 10.1017/cbo9781139568050.008.

So, ζ is a special case of $\mathfrak{D}$ when $a(n) = 1$ for all n. HARDY and RIESZ[30] considers the following as *general Dirichlet series*

$$\sum_{n\geq 1} a_n e^{-\lambda_n s} \tag{‡ 1.7}$$

where (λ_n) is a strictly increasing sequence of real numbers that tend to infinity. Following this, HARDY and RIESZ[31] calls $\mathfrak{D}$ the *ordinary Dirichlet series* when $\lambda_n = \log n$. LEJEUNE[32] considers real values of s and proves a number of important theorems. As Hardy states, JENSEN[33] discusses the first theorems where s is complex involving the nature of convergence of ‡ 1.7. CAHEN[34] makes the *first attempt to construct a systematic theory of the function $\mathfrak{D}_f(s)$ although much of the analysis which it contains is open to serious criticism, has served—and possibly just for that reason—as the starting point of most of the later researches in the subject.*

CONVERGENCE.

JENSEN[35] proves the following theorem of fundamental importance.

THEOREM 1.10. *Suppose $\mathfrak{D}_f(s)$ is convergent for the complex number $s = \omega + it$, then it is convergent for any value of s for which $\mathfrak{R}(s) > \omega$.*

Consider the Dirichlet series for two arithmetic functions f and g.

$$\mathfrak{D}_f(s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$$
$$\mathfrak{D}_g(s) = \sum_{n\geq 1} \frac{g(n)}{n^s}$$

Then we have

$$\mathfrak{D}_f(s)\mathfrak{D}_g(s) = \sum_{n\geq 1} \frac{f(n)}{n^s} \sum_{n\geq 1} \frac{g(n)}{n^s}$$

[30] HARDY and RIESZ, *The general theory of Dirichlet's series* cit., §1, Page 1.

[31] Ibid.

[32] DIRICHLET PETER GUSTAV LEJEUNE, *Vorlesungen Über Zahlentheorie*, ed. by R. Dedekind, Cambridge University Press, 1879.

[33] JOHAN LUDWIG WILLIAM VALDEMAR JENSEN, "OM RÆKKERS KONVERGENS", *Tidsskrift for mathematik*, 5th ser., vol. II (1884), pp. 63-72, ISSN: 09092528, 24460737, http://www.jstor.org/stable/24540057; JOHAN LUDWIG WILLIAM VALDEMAR JENSEN, "Sur une généralisation d'un théorème de Cauchy", *Comptes Rendus* (Mar. 1888).

[34] E. CAHEN, "Sur la fonction $\zeta(s)$ de Riemann et sur des fonctions analogues", fr, *Annales scientifiques de l'École Normale Supérieure*, vol. 3e série, 11 (1894), pp. 75-164, DOI: 10.24033/asens.401.

[35] JENSEN, "OM RÆKKERS KONVERGENS" cit.

Now, imagine we want to write this product as another Dirichlet series. Then it would be of the form

$$\mathfrak{D}_h(s) = \sum_{n\geq 1} \frac{h(n)}{n^s}$$

The coefficients $h(n)$ of $\mathfrak{D}_h(s)$ is determined as follows.

$$h(n) = \sum_{de=n} f(d)g(e)$$

After a little observation, it seems quite obvious that this is indeed correct. In fact, this is what we call Dirichlet convolution today.

DIRICHLET CONVOLUTION. For two arithmetic functions f and g, the *Dirichlet product* or *Dirichlet convolution* of f and g is defined as

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

THEOREM 1.11. *Let f and g be multiplicative arithmetic functions. Then $f * g$ is also multiplicative.*

Proof. □

THEOREM 1.12 (Associativity of Dirichlet Convolution). *Dirichlet convolution is associative. That is, if f, g and h are arithmetic functions, then*

$$(f * g) * h = f * (g * h)$$

Proof. □

An interesting function associated with Dirichlet convolution and summatory functions is the *Möbius function* μ, defined in MÖBIUS.[36]

$$\mu(n) = \begin{cases} 0 & \text{if } p^2 \mid n \text{ for some prime } p \\ (-1)^{\omega(n)} & \text{otherwise} \end{cases}$$

where $\omega(n)$ is the number of distinct prime divisors of n. On the other hand, $\Omega(n)$ is the total number of prime divisors of n. So, $\omega(12) = 2$ whereas $\Omega(12) = 3$.

[36] AUGUST FERDINAND MÖBIUS, "Über eine besondere art von Umkehrung der Reihen." *Journal für die reine und angewandte Mathematik (Crelles Journal)*, vol. IX (1832), pp. 105-123, DOI: 10.1515/crll.1832.9.105.

Theorem 1.13 (Möbius inversion). *Let f be an arithmetic function and F be the summatory function*

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

Then

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

Proof. □

Following Cojocaru and Murty,[37] let us define *dual convolution*.

Divisor Closed Set. A set of positive integers $\mathbb{S}$ is a *divisor closed set* if $d \mid n$, then $d \in \mathbb{S}$ holds for all $n \in \mathbb{S}$.

Dual Convolution. Let f and g be arithmetic functions. Then the dual convolution of f and g is the arithmetic function h defined as

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

where $\mathbb{D}$ is a divisor closed set.

Theorem 1.14 (Dual Möbius Inversion). *Let f be an arithmetic function and F be the summatory function*

$$F(n) = \sum_{\substack{n|d \\ d\in\mathbb{D}}} f(d)$$

where $\mathbb{D}$ is a divisor closed set. Then

$$f(n) = \sum_{\substack{n|d \\ d\in\mathbb{D}}} \mu\left(\frac{d}{n}\right) f(d)$$

Proof. □

While discussing inversion, we should also mention Dirichlet inverse.

Dirichlet Inverse. Let f be an arithmetic function such that $f(1) \neq 0$. Then the *Dirichlet inverse* of f is a function g such that $f * g = I$ where I is the *identity function*

$$I(n) = \left\lfloor \frac{1}{n} \right\rfloor$$

[37] Alina Carmen Cojocaru and Maruti Ram Pedaprolu Murty, *An introduction to sieve methods and their applications*, Cambridge University Press, 2006, Page 4, Theorem 1.2.3.

This inverse g can be expressed recursively.

$$g(1) = \frac{1}{f(1)}$$
$$g(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d<n}} f\left(\frac{n}{d}\right) g(d)$$

Haukkanen[38] proves the following closed formula to find the Dirichlet inverse of an arithmetic function f which we do not prove here.

Theorem 1.15. *Let f be an arithmetic function such that $f(1) = 1$. Then the Dirichlet inverse of f is*

$$f^{-1}(n) = \sum_{k=1}^{\Omega(n)} (-1)^k \sum_{\substack{d_1 \cdots d_k = n \\ d_1, \ldots, d_{k>1}}} f(d_1) \cdots f(d_k)$$

We leave the following as exercise.

1. If f is a multiplicative arithmetic function, then the Dirichlet inverse f^{-1} is also multiplicative.
2. If f and $f * g$ are multiplicative functions, then g is also multiplicative.
3. $\sum_{d|n} \mu(d) = I(n)$.

1.3 General Convolution and Dirichlet Hyperbola Method

In this chapter, we will discuss Dirichlet convolution and its generalization, use Dirichlet derivative to prove the Selberg identity, establish some results using generalized convolution and finally, prove the fundamental identity of Selberg.

[38] Pentti Haukkanen, "Expressions for the Dirichlet Inverse of an Arithmetical Function", *Notes on Number Theory and Discrete Mathematics, ISSN 1310-5132 Volume 6, 2000, Number 4, Pages 118—124*, vol. vi, no. 4 (2000), pp. 118-124, doi: `https://nntdm.net/volume-06-2000/number-4/118-124/`, Theorem 2.2.

Chapter 2

Bertrand to Tchebycheff

Definition. Tchebycheff function of the first kind or *Tchebycheff's theta function* is defined as

$$\vartheta(x) = \sum_{p \leq x} \log p$$

Definition. Tchebycheff function of the second kind or *Tchebycheff's psi function* is defined as

$$\begin{aligned}
\psi(x) &= \sum_{n \leq x} \Lambda(n) \\
&= \sum_{p^e \leq x} \log p \\
&= \sum_{p \leq x} \left\lfloor \frac{\log x}{\log p} \right\rfloor \log p
\end{aligned}$$

Chapter 3

A Modest Introduction to Sieve Theory

Chapter 4

Proof of Prime Number Theorems

Articles

Abel, Niels Henrik, "Untersuchungen über die Reihe: $1 + (m/1)x + m \cdot (m-1)/(1 \cdot 2) \cdots x^2 + m \cdot (m-1) \cdot (m-2)/(1 \cdot 2 \cdot 3) \cdots x^3 + \ldots$", *Journal für Math.*, vol. i (1826), pp. 311-339, doi: 10.1515/9783112347386-030.

Ayoub, Raymond, "Euler and the zeta function", *The American Mathematical Monthly*, vol. lxxxi, no. 10 (1974), pp. 1067-1086, doi: 10.2307/2319041.

Cahen, E., "Sur la fonction $\zeta(s)$ de Riemann et sur des fonctions analogues", fr, *Annales scientifiques de l'École Normale Supérieure*, vol. 3e série, 11 (1894), pp. 75-164, doi: 10.24033/asens.401.

Euler, Leonhard, "De Progressionibus Harmonicis Observationes", *Commentarii academiae scientiarum Petropolitanae*, vol. vii (1740), pp. 150-161.

— "De summatione innumerabilium progressionum", *Commentarii academiae scientiarum Petropolitanae*, vol. v (1738), pp. 91-105.

— "E-20: De summatione innumerabilium Progressionum", *Spectrum* (2020), pp. 52-64, doi: 10.1090/spec/098/10.

— "E-43: De Progressionibus Harmonicis Observationes", *Spectrum* (2020), pp. 133-141, doi: 10.1090/spec/098/23.

— "E-72: Variae Observationes circa series Infinitas", *Spectrum* (2020), pp. 249-260, doi: 10.1090/spec/098/41.

— "Variae Observationes circa series infinitas", *Commentarii academiae scientiarum Petropolitanae*, vol. ix (1744), pp. 160-188.

Hardy, Godfrey Harold and Srinivasa Ramanujan Aiyangar, "The normal number of prime factors of a number n", *Quarterly Journal of Mathematics*, vol. xlviii (1917), pp. 76-92.

Haukkanen, Pentti, "Expressions for the Dirichlet Inverse of an Arithmetical Function", *Notes on Number Theory and Discrete Mathematics, ISSN 1310-5132 Volume 6, 2000, Number 4, Pages 118—124*, vol. vi, no. 4 (2000), pp. 118-124, doi: https://nntdm.net/volume-06-2000/number-4/118-124/.

Jensen, Johan Ludwig William Valdemar, "OM RÆKKERS KONVERGENS", *Tidsskrift for mathematik*, 5th ser., vol. ii (1884), pp. 63-72, issn: 09092528, 24460737, http://www.jstor.org/stable/24540057.

— "Sur une généralisation d'un théorème de Cauchy", *Comptes Rendus* (Mar. 1888).

Landau, Edmund, "Über die Anzahl der Gitterpunkte in geweissen Bereichen", *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. xix (1912), pp. 687-772.

— "Über eine idealtheoretische funktion", *Transactions of the American Mathematical Society*, vol. xiii, no. 1 (1912), pp. 1-21, doi: 10.1090/s0002-9947-1912-1500901-6.

Lehmer, Derrick Norman, "Asymptotic evaluation of certain Totient Sums", *American Journal of Mathematics*, vol. xxii, no. 4 (1900), pp. 293-335, doi: 10.2307/2369728.

MacMillan, Kieren and Jonathan Sondow, "Proofs of power sum and binomial coefficient congruences via Pascal's identity", *The American Mathematical Monthly*, vol. cxviii, no. 6 (2011), pp. 549-551, doi: 10.4169/amer.math.monthly.118.06.549.

Möbius, August Ferdinand, "Über eine besondere art von Umkehrung der Reihen." *Journal für die reine und angewandte Mathematik (Crelles Journal)*, vol. ix (1832), pp. 105-123, doi: 10.1515/crll.1832.9.105.

Pascal, Blaise, "Sommation des puissances numériques", *Oeuvres complètes, Jean Mesnard, ed., Desclée-Brouwer, Paris*, vol. iii (1964), pp. 341-367.

Riemann, Bernhard, "Ueber die anzahl der primzahlen unter einer gegebenen grösse", *Monatsberichte der Berliner Akademie* (Nov. 1859), pp. 136-144, doi: 10.1017/cbo9781139568050.008.

Books

Aiyangar, Srinivasa Ramanujan, "Highly Composite Numbers", in, *Collected papers of Srinivasa Ramanujan*, ed. by Godfrey Harold Hardy, Peruvemba Venkatesvara Seshu Aiyar, and Bertram Martin Wilson, Cambridge University Press, 1927, pp. 78-128.

Bachmann, Paul Gustav Heinrich, *Analytische zahlentheorie*, 1894, vol. ii.

Cojocaru, Alina Carmen and Maruti Ram Pedaprolu Murty, *An introduction to sieve methods and their applications*, Cambridge University Press, 2006.

Hardy, Godfrey Harold, *Orders of Infinity: The 'Infinitärcalcül' of Paul Du Bois-Reymond, Cambridge Tracts in Mathematics*, Cambridge University Press, 1910.

Hardy, Godfrey Harold and Marcel Riesz, *The general theory of Dirichlet's series*, Cambridge University Press, 1915.

Jordan, Camille, *Traiteé des substitutions et des équations algébriques*, Gauthier-Villars, Paris, 1870.

Knoebel, Arthur, Reinhard Laubenbacher, Jerry Lodder, and David Pengelley, "Sums of numerical powers", in, *Mathematical Masterpieces: Further chronicles by the explorers*, Springer-Verlag, 2007, pp. 32-37.

Landau, Edmund, *Handbuch der Lehre von der Verteilung der Primzahlen*, 1909, vol. ii.

Lejeune, Dirichlet Peter Gustav, "Über Die Bestimmung Der Mittleren Werthe", in, *G. Lejeune Dirichlet's Werke*, ed. by L. Kronecker and L. Fuchs, Druck Und Verlag Von Georg Reimer., 1897, vol. ii, pp. 49-66.

— *Vorlesungen Über Zahlentheorie*, ed. by R. Dedekind, Cambridge University Press, 1879.

Mascheroni, Lorenzo, *Adnotationes ad calculum Integralem Euleri*, Galeatii, 1790.

Index