INTRODUCTION TO ELEMENTARY ANALYTIC NUMBER THEORY

Masum Billal

Introduction to Elementary Analytic Number Theory

INTRODUCTION TO ELEMENTARY ANALYTIC NUMBER THEORY

Masum Billal

Introduction to Elementary Analytic Number Theory

Masum Billal

Monday 4^{th} October, 2021

Contents

Notations

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

$\operatorname{lcm}(a, b)$ 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. We will skip discussing the basic definitions since they are common in most introductory number theory texts. See Erdős, Möbius function

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

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

Note that the number of divisor function $\tau(n)$ is the summatory function of the unit function $u(n) = 1$. The sum of divisor function $\sigma(n)$ is the summatory function of the invariant function $f(n) = n$. An interesting property that we will repeatedly use 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. Recall that the number of divisor function $\tau(n) = \sum_{ab=n} 1$. We can generalize this as follows.

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

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

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

At this point, we will 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. 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

$$f(x) = o(g(x)) \tag{‡ 1.1}$$

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0 \tag{‡ 1.2}$$

Some trivial examples are $1/x = o(1)$, $x = o(x^2)$ and $2x^2 \neq o(x^2)$. 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 $x^2 \sim x^2 + x$. 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*. Now, we are interested

in the general number of divisors and general sum of divisors. Let us define the cumulative sum of these functions.

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

Notice the following.

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

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

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

We have that $T_k(x) = x\left(x^k + O(x^{k-1})\right) + O(x^{k+1}) = O(x^{k+1})$. Although this is a weak estimate, it shows a method of estimation. On this note, an interested reader can try and prove that

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

This is known as the *Pascal identity* (see Pascal,[1] for an English translation, see Knoebel et al.[2]). The reader may also be interested in MacMillan and Sondow.[3]

[1] Blaise Pascal: *Sommation des puissances numériques*. In: *Oeuvres complètes, Jean Mesnard, ed., Desclée-Brouwer, Paris* 3 (1964), pp. 341–367.

[2] Arthur Knoebel et al.: "Sums of numerical powers". In: *Mathematical Masterpieces: Further chronicles by the explorers*. Springer-Verlag, 2007, pp. 32–37.

[3] Kieren MacMillan and Jonathan Sondow: *Proofs of power sum and binomial coefficient congruences via Pascal's identity*. In: *The American Mathematical Monthly* 118.6 (2011), pp. 549–551. DOI: 10.4169/amer.math.monthly.118.06.549.

Using a similar approach, we will now estimate B. 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 $\{n/d : d \mid n\}$ actually coincide. 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

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

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

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

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*, 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 continuously 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 will demonstrate these ideas next. It is worth mentioning that Ramanujan also used such a technique. For example, in Aiyangar et al.,[4] we can definitely see what can only be described as the formula itself. It is unclear whether Ramanujan simply knew about this or he came up with the idea on his own while working on number of divisors. 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*.

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

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

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

Proof. □

[4]Ramanujan Srinivasa Aiyangar et al.: "Highly Composite Numbers". In: *Collected papers of Srinivasa ramanujan.* Cambridge University Press, 1927, pp. 78–128, Page 83, §4.

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

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

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

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

Clearly, this is just the number of pairs (a, b) such that $ab \le x$. We can divide the pairs in two classes. In the first class, $1 \le a \le \sqrt{x}$ 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

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

In the second case, since $a > \sqrt{x}$ and $b \le x/a$, we must have $b \le \sqrt{x}$. 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

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

Thus, the total number of such pairs is

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

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

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

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 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[5] used C and O in his original paper. Mascheroni[6] 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

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

[5] L. Euler: *E-43: De Progressionibus Harmonicis Observationes*. In: *Spectrum* (2020), pp. 133–141. DOI: 10.1090/spec/098/23.

[6] Lorenzo Mascheroni: *Adnotationes ad calculum Integralem Euleri*. Galeatii, 1790.

We can now use this to get

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

Thus, we get the following result.

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

Dirichlet[7] 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.

Average order of arithmetic functions. For an arithmetic function f, the average order of f is

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

Then Dirichlet's theorem on τ can be restated as *the average order of τ is $O(\log x)$*. Aiyangar et al.[8] 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[9] 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

$$T_3(x) = O(T(x)) + x\int_1^x \frac{T(t)}{t^2}dt$$
$$= O(x\log x) + x\int_1^x \frac{\log t}{t}dt$$

[7] G. Lejeune Dirichlet: "Über Die Bestimmung Der Mittleren Werthe". In: *G. Lejeune Dirichlet's Werke.* Ed. by L. Kronecker and L. Fuchs. Vol. 2. Druck Und Verlag Von Georg Reimer., 1897, pp. 49–66.

[8] Aiyangar et al. 1927.

[9] Edmund Landau: *Über die Anzahl der Gitterpunkte in geweissen Bereichen.* In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 19 (1912), pp. 687–772, Page 689.

Using the integration by parts formula,

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

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).

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

Bibliography

Aiyangar, Ramanujan Srinivasa et al.: "Highly Composite Numbers". In: *Collected papers of Srinivasa ramanujan*. Cambridge University Press, 1927, pp. 78–128.

Dirichlet, G. Lejeune: "Über Die Bestimmung Der Mittleren Werthe". In: *G. Lejeune Dirichlet's Werke*. Ed. by L. Kronecker and L. Fuchs. Vol. 2. Druck Und Verlag Von Georg Reimer., 1897, pp. 49–66.

Euler, L.: *E-43: De Progressionibus Harmonicis Observationes*. In: *Spectrum* (2020), pp. 133–141. DOI: 10.1090/spec/098/23.

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

Landau, Edmund: *Über die Anzahl der Gitterpunkte in geweissen Bereichen*. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 19 (1912), pp. 687–772.

MacMillan, Kieren and Jonathan Sondow: *Proofs of power sum and binomial coefficient congruences via Pascal's identity*. In: *The American Mathematical Monthly* 118.6 (2011), pp. 549–551. DOI: 10.4169/amer.math.monthly.118.06.549.

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

Pascal, Blaise: *Sommation des puissances numériques*. In: *Oeuvres complètes, Jean Mesnard, ed., Desclée-Brouwer, Paris* 3 (1964), pp. 341–367.

Chapter 2

Dirichlet Convolution and Generalization

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.

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

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

It is not so easy to see how the idea of Dirichlet convolution originates even though it is highly used in number theory. We can connect its origin with the zeta function.

$$\begin{aligned}\zeta(s) &= \frac{1}{1^s} + \frac{1}{2^s} + \dots \\ &= \sum_{i\geq 1} \frac{1}{i^s}\end{aligned}$$

Chapter 3

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 4

Dirichlet's Prime Number Theorem

Chapter 5

The Prime Number Theorem

Chapter 6

A Modest Introduction to Sieve Theory

Glossary

Erdős He is Erdős

Index