Introduction to

ELEMENTARY ANALYTIC NUMBER THEORY

MASUM BILLAL

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, $\varphi(n)$ is the number of positive integers not exceeding n which are relatively prime to n.

$J_k(n)$ Jordan function of n, the number of tuples $(a_1, \dots, a_k)$ such that $\gcd(a_1, \dots, a_k, n) = 1$ and $1 \leq a_1, \dots, a_k \leq n$.

$\tau(n)$ 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

$\mu(n)$ Möbius function of n, $\mu(n) = (-1)^{\omega(n)}$ if n is square-free, otherwise $\mu(n) = 0$.

$\lambda(n)$ Liouville function of n, $\lambda(n) = (-1)^{\Omega(n)}$ if n is square-free, otherwise $\lambda(n) = 0$.

$\Lambda(n)$ Von Mangoldt Function of n. $\Lambda(n) = \log p$ if $n = p^e$ for some positive integer e, otherwise $\Lambda(n) = 0$.

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

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

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

$\alpha * \beta$ Dirichlet convolution of two arithmetic functions α and β.

$\alpha \circ \beta$ General convolution of two arithmetic functions α and β.

γ Euler-Mascheroni constant.

§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. Erdős

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 n, 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.[1]

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

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

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

§§1.1 ORDER OF SOME 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, we can take

[1] 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.

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[2] states that the symbol O had been first used by Bachmann.[3] Hardy[4] uses the notations $\prec$ and $\succ$ respectively but they are no longer in practice. Hardy and Riesz[5] adopted the notations small o and big O and today these are the primary notations for this purpose.

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

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

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

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

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*.

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.

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