ELEMENTARY

ANALYTIC NUMBER THEORY

BY

MASUM BILLAL

THE PUBLISHER
2022

Preface

While analytic number theory is a very broad subject and there are a great many books on this topic, there are not many books that are truly introductory. There are some that are introductory enough such as Apostol[1] but they usually depend on abstract algebra and complex analysis heavily. In contrast to that, it can be argued that both topics are entirely excluded from this book. Thus, the primary objective of this book is to discuss analytic number theory in the most *elementary* way possible. Before I explain what I mean by elementary, I will mention a few more details.

(1) I neither discuss class number of quadratic forms nor do I follow Landau[2] to prove that $L(1,\chi) \neq 0$ for real non-principal Dirichlet character χ. Therefore, one of the two proofs I present for Dirichlet's prime number theorem is a mixture of the approach taken by Landau[3] and Apostol.[4] Thus, we avoid the unnecessary bulky calculation presented by Landau[5] while still keeping the proof of Dirichlet's theorem completely elementary. The other proof is due to Selberg.[6] Both will be included in Chapter 3.

(2) I will treat the reader to a topic I consider to be bittersweet–*Sieve Theory*. However, I will only discuss Brun's theorem on prime pairs and the idea behind Selberg's sieve; the reason being that I primarily intend to lay the groundwork for a solid foundation. One can consult Cojocaru and Murty[7] and Friedlander and Iwaniec[8] after reading this chapter. If I only discuss a lot of sieving techniques and prove a lot of theorems, that may cause a lack of sense in the mind of the reader as to why such methods are necessary and what leads one to think in such a way that allows us to prove such powerful results. Friedlander and Iwaniec[9] has an enormous discussion on the matter and it is mostly elementary but I think the text is a little difficult for a non-enthusiast.

(3) The reason I include sieve theory at all in this book is that this is the strongest and most interesting area in all of number theory. I will explain further. The development in recent number theory has been a little slow in comparison to the previous century. Brun's theorem on twin primes and Chen's theorem on almost primes related to Goldbach conjecture

---

[1]Tom Mike Apostol

1976 *Introduction to analytic number theory*, 1st ed., Undergraduate Texts in Mathematics, Springer New York, NY, DOI: 10.1007/978-1-4757-5579-4.

[2]Edmund Landau

1969 *Elementary number theory*, 1st ed., Chelsea, vol. 1, Part Two, Chapter III.

[3]Ibid., Part two, Chapter III, §3.

[4]Apostol, *Introduction to analytic number theory* cit., Chapter VI.

[5]Landau, *Elementary number theory* cit., Theorem 152.

[6]Atle Selberg

1949 "An elementary proof of Dirichlet's theorem about primes in an arithmetic progression", *The Annals of Mathematics*, 50, 2, pp. 297-304, DOI: 10.2307/1969454.

[7]Alina Carmen Cojocaru and Maruti Ram Pedaprolu Murty

2006 *An introduction to sieve methods and their applications*, Cambridge University Press.

[8]Jhn Friedlander and Henryk Iwaniec

2010 *Opera de cribro*, Colloquium Publications, American Mathematical Society, vol. 1.

[9]Ibid.

are still two of the most spectacular results in all of mathematics. Yet both of them are really old and I do not know of any improvements over these results that are of any real significance.

(4) Brun's theorem is of incredible historical importance; being the starting point of sieve methods today despite being over 100 years old. A story goes that Erdős was asked what he thought the strongest theorem was in elementary number theory. His response was Brun's theorem on twin prime pairs. Indeed, one can see that the result of Viggo Brun on twin primes is still the most spectacular result regarding twin primes which requires no analysis or deep results. Friedlander and Iwaniec[10] named their chapter on Brun's theorem *Brun's Sieve – The Big Bang*. It just goes to show how beautiful this result really is and I believe this theorem is the most underrated result in all of mathematics. Another story goes that Landau had been sitting on Brun's paper for 8 years, mostly due to the use of difficult notations by Brun. Later, Landau dedicated an entire chapter in his *Elementaren Zahlentheorie*, volume I.1 of Landau[11] to Brun's theorem. I personally believe this story to be true, and that Landau did this out of guilt that he had deprived the mathematics world from such influential results for so long. The other result, Chen's theorem is based on Selberg's sieve that states that an even number greater than 2 is the sum of a prime and an *almost prime* (product of two primes). This is still the best result available related to Goldbach's conjecture despite being 60 years old. As you can see, the most influential results related to the oldest problems in number theory are actually the fruits of sieve methods.

(5) I will discuss two elementary proofs of the prime number theorem. Again, this begs the question: *exactly what do we consider to be elementary?* which will be answered below. Both proofs will be included in Chapter 4.

(6) I initially wanted to discuss basic complex analysis with connection to the convergence of Dirichlet series but later decided not to include it at all. It simply does not go with the spirit of this book. Similarly, I did not follow Apostol[12] and discuss Dirichlet characters from a more general point of view using group theory.

(7) The reader may omit Chapter 6 entirely given that it is more of an opinion of mine than an actual mathematical discussion. The reason behind including this chapter is that I believe if this dispute had not occurred, we might have had a few more influential discoveries like the elementary proof of the prime number theorem from the collaboration of Erdős and Selberg. It was a crime that we were deprived of further collaboration between these two mathematical giants just because some third party that did not even witness the incidents first hand poked their noses where they did not belong; which consequently drove a wedge between Erdős and Selberg. It will be clear why I feel so strongly on the matter in the respective chapter. Even though I say this is purely my opinion, proper references will be provided for the history along with some relevant information such as letters between parties involved. I will attach the letters in their original form purely because of historical

[10] Ibid., Chapter VI.

[11] Edmund Landau

1927 *Vorlesungen über Zahlentheorie*, Hirzel, vol. 1.

[12] Apostol, *Introduction to analytic number theory* cit.

reasons and also add the corresponding textual versions since they can be difficult to read occasionally. Thus, despite this chapter being a personal opinion, the reader will have some (if not all) necessary relevant documents to draw their own conclusions.

I will now explain what I mean by elementary. My initial thought on elementary is that a result is elementary if it only involves what we learn in grades 1-12 (that is, before undergraduate study starts since the number of grades may differ depending on what country the reader is from). In that sense, basic calculus such as differentiation or integration is elementary. I believe this is an opinion mathematicians will share in general. For example, Landau,[13] Ingham[14] also consider basic integration along with basic properties of zeta functions to be elementary. I should warn the reader that elementary does not imply simplicity. In reality, it is often the exact opposite. Very frequently what we can prove by the use of deep/analytic methods can also be proven by elementary means, but with much more difficulty and an even greater amount of effort. The best example to demonstrate this is the elementary proof of prime number theorem by Selberg and Erdős. Selberg stated in his paper of the elementary proof of the prime number theorem that the proof used only the *simplest properties of logarithm*. And yet it took humanity over 150 years to produce an elementary proof of this theorem and a joint effort of two of the biggest mathematical giants of twentieth century. More context and specific details will be provided on this matter in Chapter 6, The Selberg and Erdős Dispute which will shed light on why this was such a difficult task.

Masum Billal
4:03 AM, 13 May, 2022

[13] Landau, *Elementary number theory* cit.
[14] Albert Edward Ingham
1932 *The distribution of prime numbers*, 1st ed., Cambridge University Press.

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 $1 \leqq a_1, \dots, a_k \leqq n$.

$\tau_k(n)$ Generalized number of divisors of n, $\tau_k(n) = \sum_{d_1 \cdots d_k = n} 1$. For $k = 1$, $\tau_1(n) = \tau(n)$, number of divisors of n.

$\sigma_k(n)$ Generalized sum of divisors of n, $\sigma_k(n) = \sum_{d \mid n} d^k$. For $k = 1$, $\sigma_1(n) = \sigma(n)$, sum of divisors of n.

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

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

$\lfloor x \rfloor$ Floor of x, greatest integer not exceeding x.

$I(n)$ Identity function, $I(n) = \lfloor \frac{1}{n} \rfloor$.

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

$\lambda(n)$ Liouville function of n, $\lambda(n) = (-1)^{\Omega(n)}$.

$\Lambda(n)$ 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.

$H(x)$ Harmonic sum for x, $H(x) = \sum_{n \leqq x} \frac{1}{x}$.

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

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

$\alpha \bullet \beta$ Generalized convolution of two arithmetic functions α and β.

γ Euler-Mascheroni constant.

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

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

For example, the prime counting function is

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

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

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

Associated with this is the average order of f.

Average Order. For an arithmetic function f,

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

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

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

holds for *almost* all $n \leqq x$ 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) 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.

§I: 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

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

[1]Godfrey Harold Hardy and Srinivasa Ramanujan Aiyangar
1917 “The normal number of prime factors of a number n”, *Quarterly Journal of Mathematics*, 48, pp. 76-92.

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 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

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

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 $0 \leqq \{x\} < 1$.

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

$$f(x) = o(g(x)) \tag{‡2}$$

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0 \tag{‡3}$$

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

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

[3]Paul Gustav Heinrich Bachmann
1894 *Analytische zahlentheorie*, vol. 2, Page 401.

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

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.

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) is $o(x)$. We can also say, the number of n satisfying (‡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*.

[5] Godfrey Harold Hardy and Marcel Riesz

1915 *The general theory of Dirichlet's series*, Cambridge University Press.

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.

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

Notice the following.

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

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

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

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 $\mathfrak{S}(x, k) = \sum_{n \leqq x} n^k$. This is known as the *Pascal identity* (see Pascal,[6] for an English

[6]Blaise Pascal
1964 "Sommation des puissances numériques", *Oeuvres complètes, Jean Mesnard, ed., Desclée-Brouwer, Paris*, 3, pp. 341-367.

translation, see Knoebel et al.[7]). Lehmer[8] proves that

$$\mathfrak{S}(x,k) = \frac{x^{k+1}}{k+1} + \Delta \tag{‡4}$$

where $|\Delta| \leqq x^k$. The reader may also be interested in MacMillan and Sondow.[9]

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

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

Beumer[10] also considers the generalization $\tau_k(n)$ in this exact form. Using this for T,

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

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

[8] Derrick Norman Lehmer
1900 "Asymptotic evaluation of certain Totient Sums", *American Journal of Mathematics*, 22, 4, pp. 293-335, DOI: `10.2307/2369728`, Chapter II, Theorem 1.

[9] Kieren MacMillan and Jonathan Sondow
2011 "Proofs of power sum and binomial coefficient congruences via Pascal's identity", *The American Mathematical Monthly*, 118, 6, pp. 549-551, DOI: `10.4169/amer.math.monthly.118.06.549`.

[10] Martin Beumer
1962 "The arithmetical function $\tau_k(n)$", *The American Mathematical Monthly*, 69, 8, pp. 777-781, DOI: `10.2307/2310778`, (§8).

Thus, we have the recursive result

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

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

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

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[11] states this formula which today is a cornerstone of analytic number theory.

THEOREM 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*

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

are known, then

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

In particular, if f is an arithmetic function,

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

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[12] also uses a method that can only be described as the partial summation formula. It is unclear if Ramanujan simply knew

[11] Niels Henrik Abel

1826 "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.*, 1, pp. 311-339, DOI: 10.1515/9783112347386-030.

[12] Srinivasa Ramanujan Aiyangar

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

about this. He essentially derives the partial summation formula while trying to express a sum of the form

$$\sum_{p \leqq x} \phi(p)$$

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

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

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 (‡4) taking $f(n) = n^k$ for $k \geqq 0$.

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

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

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

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

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

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

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

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

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

Thus, the total number of such pairs is

$$\sum_{a \leqq \sqrt{x}} \left\lfloor \frac{x}{a} \right\rfloor + \sum_{b \leqq \sqrt{x}} \left(\left\lfloor \frac{x}{b} \right\rfloor - \lfloor \sqrt{x} \rfloor \right) = 2 \sum_{n \leqq \sqrt{x}} \left\lfloor \frac{x}{n} \right\rfloor - \lfloor x \rfloor^2 \tag{‡5}$$

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

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

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 3 (Divergence of Harmonic Sum). *For $x \geqq 1$,*

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

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 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[13] (republished in Euler[14]) used C and O in his original paper. Mascheroni[15] 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 Divergence of Harmonic Sum. Applying this, we have

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

We can now use this to get

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

Thus, we get the following result.

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

[13]Leonhard Euler

1740 "De Progressionibus Harmonicis Observationes", *Commentarii academiae scientiarum Petropolitanae*, 7, pp. 150-161.

[14]Leonhard Euler

2020b "E-43: De Progressionibus Harmonicis Observationes", *Spectrum*, pp. 133-141, DOI: `10.1090/spec/098/23`.

[15]Lorenzo Mascheroni

1790 *Adnotationes ad calculum Integralem Euleri*, Galeatii.

Dirichlet[16] actually proves the more precise result given below.

Theorem 5 (Dirichlet's average order of τ theorem).

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

where γ is the Euler-Mascheroni constant.

Then Dirichlet's theorem on τ can be restated as *the average order of τ is $O(\log x)$*. Aiyangar[17] points out in his paper that the error term $O(1/\sqrt{x})$ in Dirichlet's theorem can be improved to $O\left(x^{-2/3+\epsilon}\right)$ or $O\left(x^{-2/3}\log x\right)$ as Landau[18] 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,

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

[16] Johann Peter Gustav Lejeune Dirichlet

1897 "Über Die Bestimmung Der Mittleren Werthe", in, *G. Lejeune Dirichlet's Werke*, ed. by Leopold Kronecker and László Fuchs, Druck Und Verlag Von Georg Reimer., vol. 2, pp. 49-66.

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

[18] Edmund Landau

1912a "Über die Anzahl der Gitterpunkte in geweissen Bereichen", *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 19, pp. 687-772, Page 689.

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).

THEOREM 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[19] 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 $S_k(x)$.

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

[19]Edmund Landau

1912b "Über eine idealtheoretische funktion", *Transactions of the American Mathematical Society*, 13, 1, pp. 1-21, DOI: 10.1090/s0002-9947-1912-1500901-6, Page 2.

Here, we can see that the function

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

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

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

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

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

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

THEOREM 7. *Let s be a positive real number other than 1. Then*

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

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

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

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

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

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

From this, we finally get the following.

THEOREM 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 $\varphi(n)$.

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

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.

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

Lehmer[20] used the notation $\varphi_k(n)$ but today $J_k(n)$ is used more often. Jordan[21] first discussed this function and Lehmer[22] 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

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

$$J_k(n^m) = n^{k(m-1)}J_k(n)$$

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

Lehmer[23] proves the following which he calls the *fundamental theorem*.

$$J_k(mn) = J_k(n) \prod_{p^e \| m} \left(p^{ke} - p^{k(e-1)}\lambda(n,p)\right) \tag{‡6}$$

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.

$$\sum_{d|n} J_k(d) = n^k \tag{‡7}$$

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.

§II: 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[24] is a very good read on this subject. Euler's first contribution in this matter is Euler[25] 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.[26] Euler[27] (republished in Euler[28]) proves the following fundamental result which essentially gives a new proof of infinitude of primes.

Theorem 9 (Euler Product Identity). *Let s be a positive integer. Then*

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

where p extends over all primes.

[23]Ibid., Theorem VI.

[24]Raymond Ayoub

1974 "Euler and the zeta function", *The American Mathematical Monthly*, 81, 10, pp. 1067-1086, DOI: `10.2307/2319041`.

[25]Leonhard Euler

1738 "De summatione innumerabilium progressionum", *Commentarii academiae scientiarum Petropolitanae*, 5, pp. 91-105.

[26]Leonhard Euler

2020a "E-20: De summatione innumerabilium Progressionum", *Spectrum*, pp. 52-64, DOI: `10.1090/spec/098/10`.

[27]Leonhard Euler

1744 "Variae Observationes circa series infinitas", *Commentarii academiae scientiarum Petropolitanae*, 9, pp. 160-188.

[28]Leonhard Euler

2020c "E-72: Variae Observationes circa series Infinitas", *Spectrum*, pp. 249-260, DOI: `10.1090/spec/098/41`.

One of the results in Euler[29] is the following which we shall prove later.

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

Here, $\sim$ is the asymptotic equivalence we have already defined. Even though Euler is the main architect behind the development of ζ, Riemann[30] is the first one to consider ζ for complex s instead of real s only. By tradition, we write $s = \sigma + it$ where $\sigma = \mathfrak{R}(s)$ is the real part of s and $t = \mathfrak{I}(s)$ is the imaginary part of s.

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

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

So, ζ is a special case of $\mathfrak{D}$ when $a(n) = 1$ for all n. Hardy and Riesz[31] considers the following as *general Dirichlet series*

$$\sum_{n \geqq 1} a_n e^{-\lambda_n s} \tag{‡8}$$

where (λ_n) is a strictly increasing sequence of real numbers that tend to infinity. Following this, Hardy and Riesz[32] calls $\mathfrak{D}$ the *ordinary Dirichlet series* when $\lambda_n = \log n$. Dirichlet[33] considers real values of s and proves a number of important theorems. As Hardy states, Jensen[34] discusses the first theorems where s is complex involving the nature of convergence of ‡8. Cahen[35] 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.*

EULER PRODUCT. Let s be a complex number and f be a bounded multiplicative function. Then *Euler product* is a special case of Dirichlet series that can be written as

$$\prod_p \sum_{i \geqq 1} \frac{f(p^i)}{p^{is}}$$

---

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

[30] Bernhard Riemann

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

[31] Hardy and Riesz, *The general theory of Dirichlet's series* cit., § 1, Page 1.

[32] Ibid.

[33] Johann Peter Gustav Lejeune Dirichlet

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

[34] Johan Ludwig William Valdemar Jensen

1884 "OM RÆKKERS KONVERGENS", *Tidsskrift for mathematik*, 5th ser., 2, pp. 63-72, ISSN: 09092528, 24460737, http://www.jstor.org/stable/24540057; Johan Ludwig William Valdemar Jensen

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

[35] Eugène Cahen

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

where p extends over all primes. From the fundamental theorem of arithmetic,

$$\sum_{n\geqq 1}\frac{f(n)}{n^s}=\prod_{p}\sum_{i\geqq 1}\frac{f(p^i)}{p^{is}}$$

If f is completely multiplicative, then the sum inside the product becomes a geometric series and we have

$$\sum_{n\geqq 1}\frac{f(n)}{n^s}=\prod_{p}\frac{1}{1-\dfrac{f(p)}{p^s}}$$

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

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

Then we have

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

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

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

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

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

THEOREM 10. *Let f and g be multiplicative arithmetic functions. Then $f*g$ is also multiplicative.*

Proof. □

THEOREM 11 (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 $\Omega(12) = 3$.

THEOREM 12 (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.

[36] August Ferdinand Möbius

1832 "Über eine besondere art von Umkehrung der Reihen." *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 9, pp. 105-123, DOI: `10.1515/crll.1832.9.105`.

[37] Cojocaru and Murty, *An introduction to sieve methods and their applications* cit., Page 4, Theorem 1.2.3.

THEOREM 13 (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. This is not as difficult as it looks. We mainly need to look at $\mathbb{D}_n$, the set of divisors of n for $n \in \mathbb{D}$. If $m \mid n$, then $\mathbb{D}_m \in \mathbb{D}_n$. Let $M(n) = \max\{m \in \mathbb{D}: n \mid m\}$ for $n \in \mathbb{D}$, $N(n) = M(n)/n$ and $P(n) = \prod_{p|N(n)} p$. Since

$$\begin{aligned} F(n) &= \sum_{\substack{n|d \\ d\in\mathbb{D}}} f(d) \\ &= \sum_{k|N(n)} f(nk) \end{aligned}$$

For a prime p, if $\nu_p(k) > 1$, then $\mathbb{D}_{nk} \in \mathbb{D}_{np}$, so we don't need to consider any of $F(nk)$ separately for $nk \in \mathbb{D}$. We only need to consider the set of sets

$$\{\mathbb{D}_{nq}: q \mid P(n)\}$$

Note that for distinct $q, r \in \mathbb{D}_{P(n)}$, $\mathbb{D}_{nq} \cap \mathbb{D}_{nr} = \mathbb{D}_{nqr}$. Thus, we can easily see that

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

From this, it is pretty obvious that unless $q = 1$, all the other terms cancel out. Indeed, for $q \mid P(n), f(nq)$ appears $\binom{\omega(P(n))}{\omega(q)}$ times in the sum $\sum_{q|P(n)} f(nq)\mu(q)$ and

$$\binom{m}{0} + \binom{m}{2} + \dots = \binom{m}{1} + \binom{m}{3} + \dots$$

So, running q through all divisors of $P(n)$, the conclusion follows. □

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*

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

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 14. *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. $\sum_{d|n} \mu(d) = I(n)$.

§ III: 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.

We proved before that

$$\sum_{n \leqq x} \tau(n) = 2 \sum_{n \leqq \sqrt{x}} \left\lfloor \frac{x}{n} \right\rfloor - \left\lfloor \sqrt{x} \right\rfloor^2$$

[38] Pentti Haukkanen

2000 "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*, 6, 4, pp. 118-124, DOI: `https://nntdm.net/volume-06-2000/number-4/118-124/`, Theorem 2.2.

In a similar manner, we can also prove the following.

$$\sum_{n \leqq x} \sigma(n) = \frac{1}{2}\left(\sum_{n \leqq \sqrt{x}} \left\lfloor \frac{x}{n} \right\rfloor + \sum_{n \leqq \sqrt{x}} (2n+1)\left\lfloor \frac{x}{n} \right\rfloor - \left\lfloor \sqrt{x} \right\rfloor^2 - \left\lfloor \sqrt{x} \right\rfloor^3\right)$$

Note that, in both cases, we are able to express the partial sum of a multiplicative function up to x in terms of a combination of some partial sums of some other functions up to $\sqrt{x}$. The generalization of this method is known as the *Dirichlet hyperbola method*.

THEOREM 15 (Dirichlet Hyperbola Method). *Let f and g be arithmetic functions. If h is the Dirichlet convolution of f and g, then*

$$\sum_{n \leqq x} h(n) = \sum_{n \leqq a} f(n) G\left(\frac{x}{n}\right) + \sum_{n \leqq b} g(n) F\left(\frac{x}{n}\right) - F(a)G(b)$$

where F and G are the partial sums of f and g respectively.

$$F(x) = \sum_{n \leqq x} f(n)$$
$$G(x) = \sum_{n \leqq x} g(n)$$

Specially when $a = b$,

$$\sum_{de \leqq x} f(d)g(e) = \sum_{n \leqq \sqrt{x}} \left(f(n) G\left(\frac{x}{n}\right) + g(n) F\left(\frac{x}{n}\right)\right) - F(\sqrt{x})G(\sqrt{x})$$

Next, we will discuss generalizations of Dirichlet convolution. Let f and g be arithmetic functions such that $g(x) = 0$ if $0 < x < 1$. Then the *general convolution* of f and g is

$$f \circ g(x) = \sum_{n \leqq x} f(n) g\left(\frac{x}{n}\right)$$

We can easily prove the following.

THEOREM 16 (General convolution theorem). *Let f, g and h be arithmetic functions. Then*

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

From this, we can also get the general Möbius inversion formula.

THEOREM 17. *Let f, g be an arithmetic functions and f^{-1} be the Dirichlet inverse of f. If*

$$G(x) = \sum_{n \leqq x} f(n) g\left(\frac{x}{n}\right)$$

then

$$g(x) = \sum_{n \leqq x} f^{-1}(n) G\left(\frac{x}{n}\right)$$

§IV: GENERALIZATION OF GENERAL CONVOLUTION

Let $\mathbf{x} = (x_1, \dots, x_k)$ and $\mathbf{a} = (a_1, \dots, a_n)$ be vectors of positive real numbers. $\{\sqrt[\mathbf{a}]{\mathbf{x}}\}$ denotes the largest positive integer n for which $n^{a_i} \leqq x_i$ for some $1 \leqq i \leqq k$. That is,

$$\max\{\sqrt[\mathbf{a}]{\mathbf{x}}\} = \max\{\lfloor \sqrt[a_1]{x_1} \rfloor, \dots, \lfloor \sqrt[a_k]{x_k} \rfloor\}$$

For a positive integer n, let $n^{\mathbf{a}} \leqq \mathbf{x}$ denote that $n \leqq \max\{\sqrt[\mathbf{a}]{\mathbf{x}}\}$.

Let f be a real or complex valued function defined in k variables. For a vector of positive real numbers $\mathbf{a}$, let $\mathbf{x}/\mathbf{a}$ denote the vector $(x_1/a_1, \dots, x_k/a_1)$, $\lfloor \mathbf{x}/\mathbf{a} \rfloor$ denote the vector $(\lfloor x_1/a_1 \rfloor, \dots, \lfloor x_k/a_k \rfloor)$ and

$$\begin{aligned} f(\mathbf{x}) &= f(x_1, \dots, x_k) \\ f\left(\frac{\mathbf{x}}{n^{\mathbf{a}}}\right) &= f\left(\frac{x_1}{n^{a_1}}, \dots, \frac{x_k}{n^{a_k}}\right) \\ f\left(\left\lfloor \frac{\mathbf{x}}{n^{\mathbf{a}}} \right\rfloor\right) &= f\left(\left\lfloor \frac{x_1}{n^{a_1}} \right\rfloor, \dots, \left\lfloor \frac{x_k}{n^{a_k}} \right\rfloor\right) \end{aligned}$$

GENERALIZED CONVOLUTION. Let the generalized convolution of an arithmetic function α and a function f defined for k real numbers and a positive integer a be

$$(\alpha \bullet f)(\mathbf{x}, \mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \alpha(n) f\left(\frac{\mathbf{x}}{n^{\mathbf{a}}}\right) \tag{‡9}$$

We have the next theorem about the associativity of $\bullet$ convolution.

THEOREM 18 (Associativity of Generalized Convolution). *Let $\mathbf{x}$ be a vector of k positive real numbers, α, β be arithmetic functions, a be a fixed positive integer and $f(x_1, \dots, x_k)$ be a real or complex valued multivariate function. Then*

$$(\alpha \bullet (\beta \bullet f))(\mathbf{x}, \mathbf{a}) = ((\alpha * \beta) \bullet f)(\mathbf{x}, \mathbf{a})$$

where $f * g$ is the usual Dirichlet convolution of arithmetic functions f and g.

Proof. From the definition,

$$\begin{aligned} (\beta \bullet f)(\mathbf{x}, \mathbf{a}) &= \sum_{m^{\mathbf{a}} \leqq \mathbf{x}} \beta(m) \left(\frac{x_1}{m^{a_1}}, \dots, \frac{x_k}{m^{a_k}}\right) \\ (\alpha \bullet (\beta \bullet f))(\mathbf{x}, \mathbf{a}) &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \alpha(n) \left((\beta \bullet f)\left(\frac{\mathbf{x}}{n^{\mathbf{a}}}, \mathbf{a}\right)\right) \\ &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \alpha(n) \left((\beta \bullet f)\left(\frac{x_1}{n^{a_1}}, \dots, \frac{x_k}{n^{a_k}}\right)\right) \\ &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \alpha(n) \sum_{m^{a} \leqq \mathbf{x}/n^{\mathbf{a}}} \beta(m) f\left(\frac{x_1}{m^a n^a}, \dots, \frac{x_k}{m^a n^a}\right) \end{aligned}$$

We can collect the m and n together and write

$$\begin{aligned}
(\alpha \bullet (\beta \bullet f))(\mathbf{x}, \mathbf{a}) &= \sum_{(mn)^{\mathbf{a}} \leqq \mathbf{x}} \alpha(n)\beta(m) f\left(\frac{x_1}{m^a n^a}, \dots, \frac{x_k}{m^{a_k} n^{a_k}}\right) \\
&= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \left(\sum_{d|n} \alpha(d)\beta\left(\frac{n}{d}\right)\right) f\left(\frac{x_1}{n^{a_1}}, \dots, \frac{x_k}{n^{a_k}}\right) \\
&= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} (\alpha * \beta) f\left(\frac{\mathbf{x}}{n^{\mathbf{a}}}\right) \\
&= (\alpha * \beta) \bullet f(\mathbf{x}, \mathbf{a})
\end{aligned}$$

□

Theorem 19 (Inversion of Generalized Convolution). *Let α be an arithmetic function and f be a real or complex valued multivariate function. If*

$$g(\mathbf{x}, \mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \alpha(n) f(\mathbf{x}, \mathbf{a})$$

and α^{-1} is the Dirichlet inverse of α, then

$$f(\mathbf{x}, \mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} \alpha^{-1}(n) g\left(\frac{\mathbf{x}}{n^{\mathbf{a}}}\right)$$

Proof. First, we see that

$$\begin{aligned}
(I \bullet f)(\mathbf{x}, \mathbf{a}) &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} I(n) f(\mathbf{x}, \mathbf{a}) \\
&= I(1) f(\mathbf{x}, \mathbf{a}) + \sum_{\substack{n^{\mathbf{a}} \leqq \mathbf{x} \\ n > 1}} I(n) f(\mathbf{x}, \mathbf{a}) \\
&= f(\mathbf{x}, \mathbf{a})
\end{aligned}$$

Since $g = \alpha \bullet f$, we will use Associativity of Generalized Convolution on α^{-1} and g. We have

$$(\alpha^{-1} \bullet (\alpha \bullet f))(\mathbf{x}, \mathbf{a}) = ((\alpha^{-1} * \alpha) \bullet f)(\mathbf{x}, \mathbf{a})$$

From the definition of Dirichlet inverse, $\alpha^{-1} * \alpha = I$. So, we have

$$\begin{aligned}
(\alpha^{-1} \bullet g)(\mathbf{x}, \mathbf{a}) &= (\alpha^{-1} \bullet (\alpha \bullet f))(\mathbf{x}, \mathbf{a}) \\
&= ((\alpha^{-1} * \alpha) \bullet f)(\mathbf{x}, \mathbf{a}) \\
&= (I \bullet f)(\mathbf{x}, \mathbf{a}) \\
&= f(\mathbf{x}, \mathbf{a})
\end{aligned}$$

Thus, we have the theorem. □

If we set $\mathbf{a} = (1), k = 1$ and $\mathbf{x} = (x)$ for a real number x in Associativity of Generalized Convolution and Inversion of Generalized Convolution, we have the usual General convolution theorem and Inversion of Generalized Convolution.

THEOREM 20. *Let f and g be arithmetic functions and $h = f * g$. If*

$$F(\mathbf{x},\mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} f(n)$$

$$G(\mathbf{x},\mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} g(n)$$

$$H(\mathbf{x},\mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} h(n)$$

then we have

$$\begin{aligned} H(\mathbf{x},\mathbf{a}) &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} f(n)\left(\frac{\mathbf{x}}{n^{\mathbf{a}}},\mathbf{a}\right) \\ &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} g(n) F\left(\frac{\mathbf{x}}{n^{\mathbf{a}}},\mathbf{a}\right) \end{aligned}$$

Proof. We can write H as follows.

$$\begin{aligned} H(\mathbf{x},\mathbf{a}) &= \sum_{n^{\mathbf{a}} \leqq (\mathbf{x})} h(n) \\ &= \sum_{(de)^{\mathbf{a}} \leqq \mathbf{x}} f(d)g(e) \\ &= \sum_{d^{\mathbf{a}} \leqq \mathbf{x}} f(d) \sum_{e^{\mathbf{a}} \leqq \mathbf{x}/d^{\mathbf{a}}} g(e) \\ &= \sum_{d^{\mathbf{a}} \leqq \mathbf{x}} f(d) G\left(\frac{\mathbf{x}}{d^{\mathbf{a}}},\mathbf{a}\right) \end{aligned}$$

We can prove the other part similarly by fixing e and letting d run through for f instead of g. □

As a corollary, we have the following theorem.

THEOREM 21. *Let f be an arithmetic function. If*

$$F(\mathbf{x},\mathbf{a}) = \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} f(n)$$

then we have

$$\begin{aligned} \sum_{n \leqq \sqrt[\mathbf{a}]{\mathbf{x}}} \sum_{d \mid n} f(d) &= \sum_{n \leqq \sqrt[\mathbf{a}]{\mathbf{x}}} \left\lfloor \frac{\sqrt[\mathbf{a}]{\mathbf{x}}}{n} \right\rfloor f(n) \\ &= \sum_{n^{\mathbf{a}} \leqq \mathbf{x}} F\left(\frac{\sqrt[\mathbf{a}]{\mathbf{x}}}{n}\right) \end{aligned}$$

We will now see some applications of generalized convolution $\bullet$. Let s be a fixed positive integer and f be defined as

$$f(\mathbf{x}) = \prod_{i=1}^{k} \lfloor x_i \rfloor$$

For Jordan function $J_s(n)$, we will omit the index s because the context is clear. Moreover, define the function u as

$$u(n) = n^s$$

From (‡7),

$$\prod_{d|n} J(d) = n^s$$

Using Möbius Inversion, we also get that

$$\mu * u = J \tag{‡10}$$

Let $F(\mathbf{x})$ be the number of vectors of positive integers $(a_1, \dots, a_k)$ such that $1 \leqq a_i \leqq x_i$ and $\gcd(a_1, \dots, a_k) = 1$. Then we have

$$F(\mathbf{x}) = (\mu \bullet f)(\mathbf{x}, \mathbf{1})$$

The total number of vectors such that $1 \leqq a_i \leqq x_i$ is $x_1 \cdots x_k$. Consider an arbitrary vector $(a_1, \dots, a_k)$. If $g = \gcd(a_1, \dots, a_k) > 1$, then every a_i has to be divisible by g. Then the number of such vectors is

$$\begin{aligned} t(g) &= \left(\frac{\mathbf{a}}{g}\right) \\ &= \left\lfloor \frac{a_1}{g} \right\rfloor \cdots \left\lfloor \frac{a_k}{g} \right\rfloor \end{aligned}$$

We can see that the $t(p)$ vectors which has all elements divisible by p also has all vectors which are divisible by a multiple of p. So, if g is composite, and has r prime factors, every vector of the $t(g)$ vectors is also divisible by any of those r prime factors. Using a simple principle of inclusion and exclusion, we see that the number of vectors divisible by g has the sign $\mu(g)$. So, the total number of vectors where they have a common factor other than 1 is

$$\sum_{2 \leqq g \leqq \min(\mathbf{x})} \mu(g) \left\lfloor \frac{x_1}{g} \right\rfloor \cdots \left\lfloor \frac{x_k}{g} \right\rfloor$$

Then the number of vectors where $\gcd(a_1, \dots, a_k) = 1$ is

$$x_1 \cdots x_k + \left(\sum_{2 \leqq g \leqq \min(\mathbf{x})} \mu(g) \left\lfloor \frac{x_1}{g} \right\rfloor \cdots \left\lfloor \frac{x_k}{g} \right\rfloor \right) = \sum_{n \leqq \min(\mathbf{x})} \mu(n) f(\mathbf{x}, \mathbf{1})$$

Thus, we have the result. As a consequence of this result, we can prove the next result using the fact that the number of non-decreasing sequences $(a_1, \dots, a_k)$ such that $1 \leqq a_i \leqq a_{i+1} \leqq n$ is $\binom{n+k-1}{k}$.

Let $B(n,k)$ be the number of vectors of non-decreasing sequences $(a_1, \dots, a_k)$ such that $1 \leqq a_1 \leqq \dots \leqq a_k \leqq n$ and $\gcd(a_1, \dots, a_k) = 1$. If for a positive integer m, $\mathbf{m} = \underbrace{(m, \dots, m)}_{k \text{ times}}$ and

$$f(\mathbf{m}) = \binom{m+k-1}{k}$$

then we have

$$B(n,k) = (\mu \bullet f)(\mathbf{n}, \mathbf{1})$$

Next, let S be the sum

$$S(\mathbf{x}) = \sum_{1 \leqq a_i \leqq x_i} g(\mathbf{a})^s$$

where $g(\mathbf{a}) = \gcd(a_1, \dots, a_k)$ for the vector of positive integers $\mathbf{a} = (a_1, \dots, a_k)$. Then we have

$$S(\mathbf{x}) = \sum_{n \leqq \mathbf{x}} J_s(n) \prod_{i=1}^{k} \left\lfloor \frac{x_i}{n} \right\rfloor \tag{‡11}$$

$$= \sum_{n \leqq \mathbf{x}} \mu(n) \left(\sum_{i \leqq \mathbf{x}/n} i^s \prod_{j=1}^{k} \left\lfloor \frac{x_j}{ni} \right\rfloor \right) \tag{‡12}$$

(‡11) follows from Associativity of Generalized Convolution and (‡10).

$$(J_s \bullet f)(\mathbf{x}, \mathbf{1}) = (\mu \bullet (u \bullet f))(\mathbf{x}, \mathbf{1})$$

So, we will only prove (‡12). Consider the vector $(a_1, \dots, a_k)$ and $g = \gcd(a_1, \dots, a_k)$. Letting $a_i = gb_i$, we have that $\gcd(b_1, \dots, b_k) = 1$. The number of such vectors is $(\mu \bullet f)(\mathbf{x}, \mathbf{1})$. Each of these vectors contribute g^s to the sum, so for a particular g, the contribution of g in the sum is

$$g^s (\mu \bullet f)(\mathbf{x}, \mathbf{1})$$

Then by the principle of inclusion and exclusion, we have that

$$\begin{aligned} S(\mathbf{x}) &= \sum_{n \leqq \mathbf{x}} n^s (\mu \bullet f)(\mathbf{x}, \mathbf{1}) \\ &= (u \bullet (\mu \bullet f))(\mathbf{x}, \mathbf{1}) \end{aligned}$$

Remark. We could prove this result without using $\bullet$ convolution as well. For example, in the case $s = 1$, if $d \mid g$ and $d < g$, then g has already appeared in the vectors of d. Thus, we cannot consider any d that shares a common factor with g. $n \leqq g$ will contribute a new sum to the vectors only if $\gcd(n, g) = 1$. So, the total sum of $g(\mathbf{a})$ with $\gcd(a_1, \dots, a_k) = g$ is $\varphi(g)$. Generalizing this for arbitrary s, we can easily see that the contributed sum for g is

$$J_s(g) \sum_{n \leqq \min(\mathbf{x})/g} f(\mathbf{x}/n)$$

Bertrand to Tschebischeff

We said before that *almost* all natural numbers are composite. A major objective of this book is to discuss how often the primes occur. The same question has bugged mathematicians for a centuries. It was Gauss who first observed that the change in the distribution of primes in every interval $[x, x+1000]$ was around $1/\log x$. Thus, the rough estimate

$$\pi(x) = \int_2^x \frac{1}{\log t} dt$$

was made which is now known as *logarithmic integral*. Gauss conjectured (see Landau[1]) around 1792 or 1793 that

$$\lim_{x \to \infty} \frac{\frac{\pi(x)}{x}}{\log x} = 1$$

Tchebycheff[2] is the first one to make any substantial progress on the matter. But instead of going straight into discussing the findings of Tschebischeff, we will first try to understand the reasoning behind his approach. It is very difficult to do so and there is no real concrete intuition behind it. But we will make an attempt anyway. Imagine you are given the numbers from 1 to n. If one of these numbers was missing, we could find the missing number by subtracting the sum of the remaining numbers from the sum $1 + \dots + n$. Now what happens if two numbers go missing? Does the same technique apply? Let us give it a try. If a and b are the missing numbers, then we only know that the subtracted value is going to be $a + b$. However, there are $a + b - 1$ possible pairs of positive integers for which $a + b$ can be achieved. So, how do we know exactly which pair is the answer? Definitely we need another clue for this.

[1] Edmund Landau

1911 "Handbuch der Lehre von der Verteilung der Primzahlen", *Monatshefte für Mathematik und Physik*, 22, 1 (Dec. 1911), DOI: `10.1007/bf01742852`, Page 37.

[2] Tchebycheff

1852 "Mémoire sur les nombres premiers." fre, *Journal de Mathématiques Pures et Appliquées*, pp. 366-390, `http://eudml.org/doc/234762`.

Now, if you were allowed to ask questions for getting clues (directly asking for the numbers is not allowed obviously), what kind of questions would you want to ask? Assume that you are allowed only one question in this scenario regarding only one of the numbers (again, except asking what the actual number is). One idea would be to ask what the multiple of that number (say, multiple of 3) is. But this is again asking for the numbers directly. So assume that this type of questioning is also not allowed either. But now we can ask questions such as what is the sum when the first number multiplied by 3 is added to the second number? We can form equations from such questions and then easily find the answers. Let us now make this a little more challenging. Imagine the same problem, except now you do not have to find the numbers explicitly. Rather you just have to find the number of missing numbers. However, you can ask fewer questions this time. What should be our approach now? If we just keep randomly asking questions to form linear equations like before, we will run out of questions sooner.

§I: SOME ELEMENTARY THEOREMS ON PRIME NUMBERS

We will start with the sum of reciprocals of primes. Euler[3] proved that the sum of reciprocals of all primes diverges although his method was a bit questionable.

THEOREM 22 (Divergence of sum of reciprocals of primes). *The sum*

$$\sum_{p} \frac{1}{p} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \dots$$

where the sum is taken over all primes diverges.

This proof is due to Landau[4](this is a translation of Landau.[5] It contains the first half of the first volume.).

Proof. From Divergence of Harmonic Sum, we already know that

$$\sum_{n \geqq 1} \frac{1}{n} = \prod_{p} \frac{1}{1 - \frac{1}{p}}$$

and that this sum diverges. Now,

$$\log\left(\sum_{n \geqq 1} \frac{1}{n}\right) = \sum_{p} \left(-\log\left(1 - \frac{1}{p}\right)\right)$$

[3]Leonhard Euler
1737 "Variae Observationes circa series Infinitas", *Commentarii Academiae Scientiarum Petropolitanae*, 9, pp. 160-188, DOI: 10.1090/spec/098/41.
[4]Landau, *Elementary number theory* cit., Part Two. Chapter I, Theorem 114.
[5]Landau, *Vorlesungen über Zahlentheorie* cit.

Setting $\eta := 1/p$ and using

$$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \ldots$$

we have

$$\begin{aligned}\log\left(\sum_{n\geqq 1}\frac{1}{n}\right) &= \sum_p\left(\eta + \frac{\eta^2}{2} + \frac{\eta^3}{3} + \ldots\right)\\ &< \sum_p(\eta + \eta^2 + \ldots)\\ &= \sum_p \frac{\eta}{1-\eta}\\ &< 2\sum_p \eta\\ &= 2\sum_p \frac{1}{p}\end{aligned}$$

If $\sum_p \frac{1}{p}$ does not diverge, then

$$\log\left(\sum_{n\geqq 1}\frac{1}{n}\right) < C$$

for a constant C. Thus,

$$\sum_{n\geqq 1}\frac{1}{n} < e^C$$

and does not converge either. This is impossible. So, the original sum must diverge. □

Mertens[6] actually proved that

$$\sum_{p\leqq x}\frac{1}{p} = \log\log x$$

This is the first formal proof since Euler's method wasn't exactly clean. Euler[7] uses

$$\log\left(\frac{1}{1-x}\right) = \sum_{n\geqq 1}\frac{x^n}{n}$$

[6]Franz Mertens
1874 "Ein Beitrag Zur analytischen zahlentheorie", *Journal für die reine und angewandte Mathematik Band 78*, 78, pp. 46-62, DOI: `10.1515/9783112389843-002`.
[7]Leonhard Euler
1748 *Introductio in Analysin Infinitorum*, vol. 1, Page 228.

and sets $x := 1$ to conclude

$$\sum_{n \geqq 1} \frac{1}{n} = \infty$$

This is indeed true, however, Euler's statement is vague. Today it is known that

$$\sum_{p \leqq x} \frac{1}{p} \geqq \log\log(x+1) - \frac{\pi^2}{6}$$

Two Elementary Proofs of Legendre-Dirichlet Prime Number Theorem

§I: First Proof

§II: Proof by Selberg

Two Elementary Proofs of the Prime Number Theorem

A Modest Introduction to Sieve Theory

A composite positive integer n has at least one prime factor not exceeding $\sqrt{x}$. Thus, the number of primes in the interval $[\sqrt{x}, x]$ is

$$\begin{aligned}\pi(x) - \pi(\sqrt{x}) + 1 &= \lfloor x \rfloor - \sum_{p \leqq x} \left\lfloor \frac{x}{p} \right\rfloor + \sum_{p_1 < p_2 \leqq x} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \sum_{p_1 < p_2 < p_3 \leqq x} \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor + \dots \\ &= \sum_{n \leqq x} \mu(n) \left\lfloor \frac{x}{n} \right\rfloor\end{aligned}$$

Now, $\lfloor x \rfloor = x + O(1)$, so

$$\begin{aligned}\pi(x) - \pi(\sqrt{x}) + 1 &= x \sum_{\substack{n \leqq x \\ \varrho(n) \leqq \sqrt{x}}} \frac{\mu(n)}{n} + O\left(\sum_{\substack{n \leqq x \\ \varrho(n) \leqq \sqrt{x}}} \mu(n) \right) \\ &= x \prod_{p \leqq \sqrt{x}} \left(1 - \frac{1}{p}\right) + O(2^{\pi(\sqrt{x})})\end{aligned}$$

The last line is true since there are $\pi(\sqrt{x})$ primes not exceeding $\sqrt{x}$ and $|\mu(n)| = 1$ for all square-free $n \leqq x$ such that $\varrho(n) \leqq \sqrt{x}$. However, this is not particularly useful.

The Selberg and Erdős Dispute

Glossary

Erdős He is Erdős

Articles

Niels Henrik Abel

1826 "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.*, 1, pp. 311-339, DOI: 10.1515/9783112347386-030.

Raymond Ayoub

1974 "Euler and the zeta function", *The American Mathematical Monthly*, 81, 10, pp. 1067-1086, DOI: 10.2307/2319041.

Martin Beumer

1962 "The arithmetical function $\tau_k(n)$", *The American Mathematical Monthly*, 69, 8, pp. 777-781, DOI: 10.2307/2310778.

Eugène Cahen

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

Leonhard Euler

1737 "Variae Observationes circa series Infinitas", *Commentarii Academiae Scientiarum Petropolitanae*, 9, pp. 160-188, DOI: 10.1090/spec/098/41.

1738 "De summatione innumerabilium progressionum", *Commentarii academiae scientiarum Petropolitanae*, 5, pp. 91-105.

1740 "De Progressionibus Harmonicis Observationes", *Commentarii academiae scientiarum Petropolitanae*, 7, pp. 150-161.

1744 "Variae Observationes circa series infinitas", *Commentarii academiae scientiarum Petropolitanae*, 9, pp. 160-188.

2020a "E-20: De summatione innumerabilium Progressionum", *Spectrum*, pp. 52-64, DOI: 10.1090/spec/098/10.

2020b "E-43: De Progressionibus Harmonicis Observationes", *Spectrum*, pp. 133-141, DOI: 10.1090/spec/098/23.

2020c "E-72: Variae Observationes circa series Infinitas", *Spectrum*, pp. 249-260, DOI: 10.1090/spec/098/41.

Godfrey Harold Hardy and Srinivasa Ramanujan Aiyangar

1917 "The normal number of prime factors of a number n", *Quarterly Journal of Mathematics*, 48, pp. 76-92.

Pentti Haukkanen

2000 "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*, 6, 4, pp. 118-124, doi: `https://nntdm.net/volume-06-2000/number-4/118-124/`.

Johan Ludwig William Valdemar Jensen

1884 "OM RÆKKERS KONVERGENS", *Tidsskrift for mathematik*, 5th ser., 2, pp. 63-72, issn: 09092528, 24460737, `http://www.jstor.org/stable/24540057`.

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

Edmund Landau

1911 "Handbuch der Lehre von der Verteilung der Primzahlen", *Monatshefte für Mathematik und Physik*, 22, 1 (Dec. 1911), doi: `10.1007/bf01742852`.

1912a "Über die Anzahl der Gitterpunkte in geweissen Bereichen", *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 19, pp. 687-772.

1912b "Über eine idealtheoretische funktion", *Transactions of the American Mathematical Society*, 13, 1, pp. 1-21, doi: `10.1090/s0002-9947-1912-1500901-6`.

Derrick Norman Lehmer

1900 "Asymptotic evaluation of certain Totient Sums", *American Journal of Mathematics*, 22, 4, pp. 293-335, doi: `10.2307/2369728`.

Kieren MacMillan and Jonathan Sondow

2011 "Proofs of power sum and binomial coefficient congruences via Pascal's identity", *The American Mathematical Monthly*, 118, 6, pp. 549-551, doi: `10.4169/amer.math.monthly.118.06.549`.

Franz Mertens

1874 "Ein Beitrag Zur analytischen zahlentheorie", *Journal für die reine und angewandte Mathematik Band 78*, 78, pp. 46-62, doi: `10.1515/9783112389843-002`.

August Ferdinand Möbius

1832 "Über eine besondere art von Umkehrung der Reihen." *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 9, pp. 105-123, doi: `10.1515/crll.1832.9.105`.

Blaise Pascal

1964 "Sommation des puissances numériques", *Oeuvres complètes, Jean Mesnard, ed., Desclée-Brouwer, Paris*, 3, pp. 341-367.

Bernhard Riemann

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

Atle Selberg

1949 "An elementary proof of Dirichlet's theorem about primes in an arithmetic progression", *The Annals of Mathematics*, 50, 2, pp. 297-304, doi: 10.2307/1969454.

Tchebycheff

1852 "Mémoire sur les nombres premiers." fre, *Journal de Mathématiques Pures et Appliquées*, pp. 366-390, http://eudml.org/doc/234762.

Books

Srinivasa Ramanujan Aiyangar

1927 "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, pp. 78-128.

Tom Mike Apostol

1976 *Introduction to analytic number theory*, 1st ed., Undergraduate Texts in Mathematics, Springer New York, NY, DOI: `10.1007/978-1-4757-5579-4`.

Paul Gustav Heinrich Bachmann

1894 *Analytische zahlentheorie*, vol. 2.

Alina Carmen Cojocaru and Maruti Ram Pedaprolu Murty

2006 *An introduction to sieve methods and their applications*, Cambridge University Press.

Johann Peter Gustav Lejeune Dirichlet

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

1897 "Über Die Bestimmung Der Mittleren Werthe", in, *G. Lejeune Dirichlet's Werke*, ed. by Leopold Kronecker and László Fuchs, Druck Und Verlag Von Georg Reimer., vol. 2, pp. 49-66.

Leonhard Euler

1748 *Introductio in Analysin Infinitorum*, vol. 1.

Jhn Friedlander and Henryk Iwaniec

2010 *Opera de cribro*, Colloquium Publications, American Mathematical Society, vol. 1.

Godfrey Harold Hardy

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

Godfrey Harold Hardy and Marcel Riesz

1915 *The general theory of Dirichlet's series*, Cambridge University Press.

Albert Edward Ingham

1932 *The distribution of prime numbers*, 1st ed., Cambridge University Press.

Camille Jordan

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

Arthur Knoebel, Reinhard Laubenbacher, Jerry Lodder, and David Pengelley

2007 "Sums of numerical powers", in, *Mathematical Masterpieces: Further chronicles by the explorers*, Springer-Verlag, pp. 32-37.

Edmund Landau

1909 *Handbuch der Lehre von der Verteilung der Primzahlen*, vol. 2.

1927 *Vorlesungen über Zahlentheorie*, Hirzel, vol. 1.

1969 *Elementary number theory*, 1st ed., Chelsea, vol. 1.

Lorenzo Mascheroni

1790 *Adnotationes ad calculum Integralem Euleri*, Galeatii.

Index