

A Basic Introduction of Bachmann-Landau Asymptotic Notation

Riddhiman Bhattacharya

Contents

1	Introduction	2
2	Theory	3
2.1	Big- O	3
2.2	Little- o	6
2.3	Bachmann-Landau equations	8
3	Application: Average Orders of Divisor Functions	12
4	Conclusion	18

Chapter 1

Introduction

Paul Bachmann introduced the O notation in his book *Die Analytische Zahlentheorie*¹ [2] while working on asymptotic estimates and handling error terms in analytic number theory. His ideas served as inspiration for Edmund Landau, who further refined the concept and introduced a stronger variant known as the little- o notation in his book *Handbuch der Lehre von der Verteilung der Primzahlen*² [3]. These notations are now extensively used in the field of analytic number theory.

Although mathematicians generally grasp the meaning of big- O and little- o notation intuitively, effectively communicating their precise definitions can sometimes pose challenges. Therefore, this paper aims to thoroughly comprehend the essence of big- O and little- o notation and elucidate how to appropriately apply them in various mathematical contexts.

¹In English *Analytic Number Theory*

²In English *Handbook of the theory of distribution of prime numbers*

Chapter 2

Theory

Both big- O notation and little- o notation are mathematical tools used to compare the behavior of one function concerning another function as their inputs approach certain limits. Typically, these limits are chosen as $+\infty$ (referred to as infinite asymptotics) or 0 (referred to as infinitesimal asymptotics). However, the definition provided here expands this concept to encompass any real limit point $a \in \mathbb{R}_{\text{ext}}$. Despite this generalization, the primary and most commonly encountered scenarios are still when $a = 0$ (infinitesimal asymptotics) and $a = \infty$ (infinite asymptotics). Here, \mathbb{R}_{ext} denotes the set of extended real numbers, which is defined as $\mathbb{R}_{\text{ext}} = \mathbb{R} \cup \{\pm\infty\}$ according to the work by Rudin [4].

2.1 Big- O

In the family of Bachmann-Landau notations, we can define each notation using either a direct limit-based approach or a quantifier-based approach.

Definition 1. Let $a \in \overline{\mathbb{R}}$ be a real number or $\pm\infty$, and consider real functions f and g defined on a neighborhood of a , with g being nonzero on this neighborhood. We say that $f(x) = O_a(g(x))$ or simply $f(x) = O(g(x))$ as x approaches a if one of the following equivalent conditions holds:

1. There exists $\delta \in (0, \infty)$ such that

$$\sup_{x \in N_\delta(a)} \left| \frac{f(x)}{g(x)} \right| < \infty;$$

2. There exist positive constants k and δ such that for all $x \in N_\delta(a)$, we have $|f(x)| \leq k|g(x)|$.

Here, $N_\delta(a)$ represents the neighborhood of a with radius δ . For finite a , this neighborhood is $(a - \delta, a + \delta)$.

For infinite values, we handle it differently by defining $N_\delta(\infty) = (\delta, \infty)$ and $N_\delta(-\infty) = (-\infty, -\delta)$.

The set implicit in the first supremum is denoted as

$$S_\delta = \left\{ \left| \frac{f(x)}{g(x)} \right| : x \in (a - \delta, a + \delta) \right\}.$$

Let $s_\delta = \sup S_\delta$. If $s_\delta < \infty$ for some δ , then we observe that s_δ acts as an upper bound for the set S_δ . Consequently, for all x in the neighborhood N_δ , we have $|f(x)| \leq s_\delta |g(x)|$. Hence, the first condition (the existence of a finite supremum) implies the second condition (the existence of k, δ such that $|f(x)| \leq k |g(x)|$ for $x \in N_\delta$).

Conversely, let's assume that there exist positive constants k and δ such that for all x in the neighborhood $N_\delta(a)$, we've $|f(x)| \leq k |g(x)|$. This means that k serves as an upper bound for S_δ . By utilizing, the least-upper-bound property of the real numbers, we can deduce that S_δ has a finite supremum. Therefore, the second condition implies the existence of a finite supremum $s_\delta < \infty$ for S_δ .

Thus, we've shown that the two conditions, namely the existence of a finite supremum for S_δ and the existence of k, δ such that $|f(x)| \leq k |g(x)|$ for $x \in N_\delta$, are equivalent. (As evident in this argument, the first condition essentially indicates that $|f(x)/g(x)|$ should be bounded above in some neighborhood of a .)

Note, g must be nonzero is essential for making sense of the fraction $f(x)/g(x)$. However, it suffices for g to be nonzero over some neighborhood of a to make the definition of O_a applicable. This can be achieved by finding a δ that is smaller than the radius of the neighborhood in which g is nonzero.

The original definition, found in [3], focuses on the case $a = \infty$ since analytic number theory primarily deals with infinite limits. Conversely, the case when $a = 0$ becomes useful in analysis, particularly when approximating functions by truncating higher-order terms in their Taylor series.

While O_a might seem of limited use for finite nonzero values of a , it is preferable to defining O_0 and

O_∞ separately because it emphasizes the commonality and unity between both concepts.

Examples:

- For $a = \infty$, O_∞ considers the behavior of functions as they approach infinity, which is crucial in analytic number theory.
- On the other hand, when $a = 0$, O_0 becomes significant in analysis, particularly for approximating functions using their Taylor series and discarding higher-order terms.

Although O_a may not find extensive use for finite nonzero values of a , having a unified definition that encompasses both O_0 and O_∞ highlights the connection between these concepts and their applicability in different mathematical domains.

When it comes to Bachmann-Landau's notations, understanding the true essence is more important than a definition. A helpful way to grasp big- O notation is to view it as analogous to the \leq symbol used with numbers:

$f(x) = O_a(g(x))$ means that $f(x)$ is less than or equal to $g(x)$ in the limit, up to a constant factor.

(If f does not dominate g , then a is not greater than b)

Examples:

Let's examine some examples to illustrate the concepts of Bachmann-Landau notation:

- Verification of $100x^2 = O_\infty(x^3)$ using both conditions:
 - Using the first condition, for any $\delta \in (0, \infty)$, we find that $\sup_{x>\delta} \left| \frac{100x^2}{x^3} \right| = \frac{100}{\delta} < \infty$. Thus, $100x^2$ is bounded by x^3 as x approaches ∞ .
 - Using the second condition, we observe that for $x > 1$, $100x^2 \leq 100x^3$. In the analogy, $100x^2$ is "less than or equal to" x^3 in the limit to infinity since x^2 grows slower than x^3 as x approaches ∞ .
- Relationship between $f = O_\infty(g)$ and $f = O_{-\infty}(g)$ for functions that are either even or odd: - If both f and g are either even or odd functions, then $f = O_\infty(g)$ if and only if $f = O_{-\infty}(g)$. Consequently, $100x^2 = O_{-\infty}(x^3)$.
- Disproving $100x^2 = O_0(x^3)$ but confirming $x^3 = O_0(100x^2)$: - It is false that $100x^2 = O_0(x^3)$, as

assuming so would lead to the absurdity that there exist $\delta, k \in (0, \infty)$ such that $|x| < \delta$ implies $x \geq k/100$. - However, we find that $x^3 = O_0(100x^2)$, since for $|x| < 100$, $x^3 \leq 100x^2$. Visually, the graph of $100x^2$ lies below the graph of x^3 for $x > 100$ (implying $100x^2 \leq x^3$ as $x \rightarrow \infty$), and the graph of x^3 lies below the graph of $100x^2$ for $x < 100$ (indicating $x^3 \leq 100x^2$ as $x \rightarrow 0$).

- An exaggerated example: $x^2 = O_\infty(x^{100})$ but $x^{100} = O_0(x^2)$.
- Condition for $f = O_a(g)$ at a finite a : - If f does not have any asymptote at a and g is nonzero in some neighborhood of a , then we have $f = O_a(g)$. This is because $|f/g|$ achieves its supremum over a sufficiently small neighborhood.

And thus the notation O_a is not commonly used because it is most relevant when $a = 0$ or $a = \infty$. As a result, the literature often presents $f(x) = O(g(x))$ as $x \rightarrow a$, or sometimes omits the "as $x \rightarrow a$ " altogether since the value of a is usually clear from the context.

From this point forward, we'll drop the a subscripts, and in some cases, like in the following theorem, we'll also omit the variable x .

Theorem 1 (Transitivity). *If $f = O(g)$ and $g = O(h)$, then $f = O(h)$. (If a is not specified, neither explicitly nor from context, as in this theorem, then assume any arbitrary $a \in \overline{\mathbb{R}}$.)*

Proof. By definition, there exist $k, l, \delta, \epsilon \in (0, \infty)$ such that $|f(x)| \leq k|g(x)|$ for all $x \in N_\delta(a)$ and $|g(x)| \leq l|h(x)|$ for all $x \in N_\epsilon(a)$. Therefore, we've:

$$|f(x)| \leq k|g(x)| \leq kl|h(x)|$$

for all $x \in N_{\min(\delta, \epsilon)}(a)$.

2.2 Little-o

Once, we've got a way to compare functions using the \leq notation, we wonder if a similar notion exists for the $<$ comparison. This leads us to the concept of little- o .

Definition 2. Given $a \in \overline{\mathbb{R}}$ and real functions f and g defined on a neighborhood of a , where g is nonzero on this neighborhood, we say that $f(x) = o_a(g(x))$ or $f(x) = o(g(x))$ as x approaches a if one of the following equivalent conditions holds:

1. The limit of $f(x)/g(x)$ as x approaches a is 0.
2. For all $k > 0$, there exists $\delta \in (0, \infty)$ such that $|f(x)| \leq k|g(x)|$ for all $x \in N_\delta(a)$.¹

The equivalence in this case can be established using the definition of a limit. If $\lim_{x \rightarrow a} f(x)/g(x) = 0$, then for all $k > 0$, there exists a $\delta > 0$ such that $|f(x)/g(x)| < k$ for all $x \in N_\delta(a)$, and vice versa.

Again, it's worth noting that the definition given by Landau in [3] assumes $a = \infty$.

To gain an intuitive understanding of little- o , consider the following:

$$f(x) = o_a(g(x)) \text{ means that } kf(x) < g(x) \text{ in the limit for any } k.$$

$$(g \text{ dominates } f \iff b \text{ is greater than } a)$$

Thus, little- o is stronger than big- O , as $f(x) = o_a(g(x))$ implies $f(x) = O_a(g(x))$, but the reverse is not necessarily true.

Example:

- As x approaches infinity, $\sin x$ grows slower than x , denoted as $\sin x = o_\infty(x)$, since $\sin x/x$ tends to 0 as x goes to infinity. However, as x approaches 0, $\sin x$ does not satisfy $\sin x = o_0(x)$ because the limit of $\sin x/x$ is 1 and not 0. Similarly, x is not $o_0(\sin x)$ either, as $x/\sin x$ also has a limit of 1. The fact that $\sin x = o_\infty(x)$ can also be shown using the second condition. For any $k > 0$, we can find a δ such that $\sin x \leq kx$ for all $x > \delta$. For example, $\delta = 1/k$ works.
- If the limit of a function $f(x)$ as x approaches 0 is 0, then we can say $f = o_0(1)$. Consequently, we have $\sin x = o_0(1)$ and $\tan x = o_0(1)$ since their limits as x approaches 0 are 0.
- In a previous section 2.1, it was shown that $100x^2 = O_\infty(x^3)$. Moreover, since $100x^2/x^3 = 100/x$ tends to 0 as x approaches infinity, we have $100x^2 = o_\infty(x^3)$. However, it is important to note that while $bx^3 = O_\infty(x^3)$ for any real constant b , the same is not true when big- O is replaced by little- o .
- The relationship $f(x) = o_a(g(x))$ holds if and only if $f(x+a) = o_0(g(x+a))$ for real functions f and g defined all over \mathbb{R} .

¹The k is placed on the right side here to align with the definition of big- O , but it actually belongs on the left side because it is g that is always greater no matter which multiple of f we take.

As with big- O notation, we often omit the subscript and sometimes the variable when using little- o notation.

Theorem 2 (Transitivity). *If $f = o(g)$ and $g = o(h)$, then $f = o(h)$.*

Proof. Let m be a positive real number. Since $f = o(g)$, there exist positive real numbers δ and ϵ such that $|f(x)| \leq \sqrt{m}|g(x)|$ for all x in the neighborhood $N_\delta(a)$. Similarly, since $g = o(h)$, there exist positive real numbers δ' and ϵ' such that $|g(x)| \leq \sqrt{m}|h(x)|$ for all x in the neighborhood $N_{\epsilon'}(a)$.

Now, consider the neighborhood $N_{\min(\delta, \epsilon')}$ of a . For all x in this neighborhood, we've:

$$|f(x)| \leq \sqrt{m}|g(x)| \leq m|h(x)|.$$

Therefore, I've shown that for any positive real number m , there exists a neighborhood around a such that $|f(x)| \leq m|h(x)|$ for all x in this neighborhood. This precisely means that $f = o(h)$.

2.3 Bachmann-Landau equations

We introduce a new type of equation that allows the presence of big- O 's and little- o 's on either side, leading to equations like $x^2 + O(x^3)o(1) = o(x^4)$. These equations are referred to as *Bachmann-Landau equations*. However, there is some confusion as mathematicians usually don't differentiate between Bachmann-Landau equations and normal equations.

To clarify this, it's better to consider normal equations as a specific case within the broader class of Bachmann-Landau equations. The Bachmann-Landau equations make claims about the existence of certain functions satisfying particular properties. The following definition describes this general situation, without getting too formal.

Definition 3. Given a real number $a \in \overline{\mathbb{R}}$, we represent X^i and Y^j as placeholders for either O or o with $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. Let f_i and g_j be real functions defined and nonzero on a neighborhood of a for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. A Bachmann-Landau equation in x has the form:

$$F(X^1(f_1(x)), X^2(f_2(x)), \dots, X^m(f_m(x))) = G(Y^1(g_1(x)), Y^2(g_2(x)), \dots, Y^n(g_n(x)))$$

where F and G are expressions involving the placeholders X^i and Y^j . This equation implies the following: for any real functions f_1^*, \dots, f_m^* defined on a neighborhood of a such that $f_i^*(x) = X^i(f_i(x))$, there exist real functions g_1^*, \dots, g_n^* defined on an equal or larger neighborhood of a such that $g_j^*(x) = Y^j(g_j(x))$, and the equation holds:

$$F(f_1^*(x), f_2^*(x), \dots, f_m^*(x)) = G(g_1^*(x), g_2^*(x), \dots, g_n^*(x))$$

for all x for which the left side is defined.

Note that Bachmann-Landau equations are *not symmetric*, as the following examples show. They are thus to be read left-to-right.

Examples:

- We will verify the Bachmann-Landau equation $x^2 + O(x^3) = o(1)$ as $x \rightarrow \infty$. Let f and g be real functions defined everywhere such that $|f(x)| \leq k|x|^3$ for all $x \in (-\delta, \delta)$ for some finite $\delta > 0$, and $\lim_{x \rightarrow 0} g(x) = 0$. We need to show that $\frac{x^2 + f(x)g(x)}{x^4} \rightarrow 0$ as $x \rightarrow \infty$. By analyzing the limits, we find that $f(x)/x^4 \rightarrow 0$ as $x \rightarrow \infty$, which is trivial to prove.
- We will verify $n^{O(1)} = e^{O(n)}$ as $n \rightarrow \infty$. Let f be a real function defined on $\mathbb{R}_{>0}$ such that $f = O(1)$. We need to show that $f(n) \log n = O(n)$. By defining N and k based on f , we find that $|f(n)| \log n \leq kn$ for $n > \max(N, 1)$.
- However, $e^{O(n)} \neq n^{O(1)}$. This means there exists some $f(n) = O(n)$ for which there is no $g(n) = O(1)$ satisfying $e^{f(n)} = n^{g(n)}$. Taking $f(n) = n$ as an example, the only possible function g would be $g(n) = n/\log n$, but $n/\log n$ is not $O(1)$.
- The Riemann hypothesis implies $\pi(x) = \text{Li}(x) + O(\sqrt{x} \log x)$, where $\pi(x)$ is the number of primes from 1 to x and $\text{Li}(x) = \int_2^x dt/\log t$. This equation means $|\pi(x) - \text{Li}(x)| \leq k\sqrt{x} \log x$ for some real $k > 0$ and all sufficiently large x .
- If $F = G$ and $G = H$ are valid Bachmann-Landau equations, then $F = H$ is also valid according to the definition. This allows us to write chains of equalities.

Fortunately, we don't have to go through such long calculations when we are dealing with Bachmann-Landau equations. There are some recurring ways in which Bachmann-Landau equations are manip-

ulated that are explained in the following two theorems. Assume a fixed value of a .

Theorem 3. *Let $\lambda \in \mathbb{R}$, and let f, g be real functions defined and nonzero on some neighbourhood of a . Then we have $O(f)O(g) = O(fg)$, $fO(g) = O(fg)$, $\lambda O(f) = O(f)$, and $O(\lambda f) = O(f)$. If f and g are nonnegative, then $O(f) + O(g) = O(\max(f, g))$ and in particular $O(f) + O(f) = O(f)$.*

Each of the equations in the statement of the above theorem is a Bachmann-Landau equation, as the proof makes clear.

Proof. Suppose $f_1 = O(f)$ and $g_1 = O(g)$, so there exist $\delta, \epsilon, k, l \in (0, \infty)$ such that $|f_1(x)| \leq k|f(x)|$ for all $x \in N_\delta(a)$ and $|g_1(x)| \leq l|g(x)|$ for all $x \in N_\epsilon(a)$. We know that $N_\delta(a) \cap N_\epsilon(a) = N_\gamma(a)$ for some $\gamma \in (0, \infty)$.² Let $m = \max(k, l)$. Then $|f_1(x)/f(x)| \leq m$ and $|g_1(x)/g(x)| \leq m$ for all $x \in N_\gamma(a)$, therefore

$$\left| \frac{f_1(x)g_1(x)}{f(x)g(x)} \right| \leq m^2$$

for all $x \in N_\gamma(a)$; this proves the first equation. The second and third equations follow from

$$\left| \frac{f(x)g_1(x)}{f(x)g(x)} \right| = \left| \frac{g_1(x)}{g(x)} \right| \leq l \quad \forall x \in N_\epsilon(a) \quad \text{and} \quad \left| \frac{\lambda f_1(x)}{f(x)} \right| = |\lambda| \frac{f_1(x)}{f(x)} \leq k|\lambda| \quad \forall x \in N_\delta(a).$$

The fourth equation follows from the fact that

$$\sup_{N_\kappa(a)} \left| \frac{f_2(x)}{\lambda f(x)} \right| < \infty \implies \sup_{N_\kappa(a)} \left| \frac{f_2(x)}{f(x)} \right| < \infty$$

for any real function f_2 defined on a neighbourhood of a . Finally, if f and g are both nonnegative, then for all $x \in N_\gamma(a)$ we have

$$\begin{aligned} |f_1(x) + g_1(x)| &\leq |f_1(x)| + |g_1(x)| \\ &\leq m(|f(x)| + |g(x)|) \\ &= m(f(x) + g(x)) \\ &\leq 2m \max(f(x), g(x)) = 2m |\max(f(x), g(x))|. \end{aligned}$$

²We have $\gamma = \min(\delta, \epsilon)$ if a is finite, and $\gamma = \pm \max(\delta, \epsilon)$ if $a = \pm\infty$.

The exact same theorem holds with big- O replaced by little- o .

Theorem 4. *Let $\lambda \in \mathbb{R}$, and let f, g be real functions defined and nonzero on some neighbourhood of a . Then we have $o(f)o(g) = o(fg)$, $fo(g) = o(fg)$, $\lambda o(f) = o(f)$, and $o(\lambda f) = o(f)$. If f and g are nonnegative, then $o(f) + o(g) = o(\max(f, g))$ and in particular $o(f) + o(f) = o(f)$.*

Proof. Let f_1 and g_1 be real functions defined on a neighborhood of a such that $f_1 = o(f)$ and $g_1 = o(g)$. The fact that $\lim_{x \rightarrow a} f_1(x)/f(x) = \lim_{x \rightarrow a} g_1(x)/g(x) = 0$ implies that $\lim_{x \rightarrow a} \frac{f(x)g_1(x)}{f(x)g(x)} = \lim_{x \rightarrow a} \frac{g_1(x)}{g(x)} = 0$ and $\lim_{x \rightarrow a} \frac{f_1(x)g_1(x)}{f(x)g(x)} = \left(\lim_{x \rightarrow a} \frac{f_1(x)}{f(x)} \right) \left(\lim_{x \rightarrow a} \frac{g_1(x)}{g(x)} \right) = 0$. This shows that $o(f)o(g) = o(fg)$ and $fo(g) = o(fg)$. Moreover, if $\lim_{x \rightarrow a} \lambda f_1(x)/f(x)$ holds, then $\lambda o(f) = o(f)$. Similarly, if $f_2 = o(\lambda f)$, then $\lim_{x \rightarrow a} f_2(x)/(\lambda f(x)) = 0$, and consequently, $\lim_{x \rightarrow a} f_2(x)/f(x) = 0$, which implies $o(\lambda f) = o(f)$. Now, suppose f and g are non-negative, which means f_1 and g_1 are also nonnegative. For any $m \in (0, \infty)$, there exist $\delta, \epsilon \in (0, \infty)$ such that $f_1(x) \leq mf(x)$ for all x in a neighborhood $N_\delta(a)$ and $g_1(x) \leq mg(x)$ for all x in a neighborhood $N_\epsilon(a)$. Consequently, $f_1(x) + g_1(x) \leq m(f(x) + g(x))$ for all x in a neighborhood $N_\zeta(a)$, where $\zeta = \min(\delta, \epsilon)$.

Chapter 3

Application: Average Orders of Divisor Functions

We explore some very basic estimates in analytic number theory that require big- O to explain the use of Bachmann-Landau notation in day to day mathematics. I took them from Chapter 3 of [1] and involved the partial summations of standard arithmetic functions.

Let's define the notations first which we'll require

Definition 4.

1. Given a natural number n , let $\tau(n)$ denote the number of positive divisors of n , and given $s \in \mathbb{R}$ let $\sigma_s(n)$ denote the sum of the s -th powers of these divisors.
2. For $s \in \mathbb{R}_{>0}$, define

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

This is called the *Riemman Zeta-Function*.

3. Define $C = \lim_{n \rightarrow \infty} (1 + 1/2 + 1/3 + \dots + 1/n - \log n)$. This is called *Euler's constant*.

There is a special way to write such functions that run over the divisors of a variable number. We

write

$$\tau(n) = \sum_{d|n} 1 \quad \text{and} \quad \sigma_s(n) = \sum_{d|n} d^s.$$

In particular, this clarifies the fact that σ_0 and τ are the same function, called the *divisor counting function*. Also σ_1 is often written just as σ and called the *divisor sum function*. The function ζ is the heart of number theory. Note, the infinite sum $\sum n^{-s}$ converges if and only if $s > 1$, and thus the case $0 < s < 1$ has to be considered separately; we'll not go into details about how this second expression is derived. As for the constant C , it features in the theorems below and is not to be confused with the constant $e = \lim_{n \rightarrow \infty} (1 + 1/n)^n$. Although it is not strictly necessary, we'll use the fact that $\zeta(2) = \pi^2/6$.

One of the goals of number theoretic research is to understand deeply the behaviour of arithmetic functions like τ and σ_s , and analytic number theory focuses on their behaviour for large n . However, functions in number theory are notorious for fluctuating; τ , for example, takes the value 2 for every prime n , of which there are infinitely many. So, instead of focusing on $\tau(n)$ specifically, we can instead question the *average* value of $\tau(n)$ over a large sample of numbers, leading to the following.

Definition 5. Given a function $f : \mathbb{N} \rightarrow \mathbb{R}$, the *average order* of f is the function $\tilde{f}(x) = \frac{1}{x} \sum_{n \leq x} f(n)$.

It turns out that we can compute big- O estimates for $\tilde{\tau}$ and $\tilde{\sigma}_s$, but before that we need some basic results.

Theorem 5 (Euler's Summation Formula). *Given a real function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is continuously differentiable on the interval $[x, y]$, we have*

$$\sum_{x < n \leq y} f(n) = \int_x^y f(t) dt + \int_x^y (t - \lfloor t \rfloor) f'(t) dt + f(x)(x - \lfloor x \rfloor) - f(y)(y - \lfloor y \rfloor).$$

Proof. Observe

$$\begin{aligned}
\int_{\lfloor x \rfloor}^{\lfloor y \rfloor} \lfloor t \rfloor f'(t) dt &= \sum_{n=\lfloor x \rfloor+1}^{\lfloor y \rfloor} \int_{n-1}^n \lfloor t \rfloor f'(t) dt \\
&= \sum_{n=\lfloor x \rfloor+1}^{\lfloor y \rfloor} \int_{n-1}^n (n-1) f'(t) dt \\
&= \sum_{n=\lfloor x \rfloor+1}^{\lfloor y \rfloor} (n-1)(f(n) - f(n-1)) \\
&= \lfloor y \rfloor f(\lfloor y \rfloor) - \lfloor x \rfloor f(\lfloor x \rfloor) - \sum_{x < n \leq y} f(n),
\end{aligned}$$

and so

$$\begin{aligned}
\sum_{x < n \leq y} f(n) &= \lfloor y \rfloor f(\lfloor y \rfloor) - \lfloor x \rfloor f(\lfloor x \rfloor) - \int_{\lfloor x \rfloor}^{\lfloor y \rfloor} \lfloor t \rfloor f'(t) dt \\
&= \lfloor y \rfloor f(\lfloor y \rfloor) - \lfloor x \rfloor f(\lfloor x \rfloor) - \left(\int_x^y \lfloor t \rfloor f'(t) dt + \int_{\lfloor x \rfloor}^x \lfloor t \rfloor f'(t) dt - \int_{\lfloor y \rfloor}^y \lfloor t \rfloor f'(t) dt \right) \\
&= \lfloor y \rfloor f(\lfloor y \rfloor) - \lfloor x \rfloor f(\lfloor x \rfloor) - \left(\int_x^y \lfloor t \rfloor f'(t) dt + \lfloor x \rfloor (f(x) - f(\lfloor x \rfloor)) - \lfloor y \rfloor (f(y) - f(\lfloor y \rfloor)) \right) \\
&= \lfloor y \rfloor f(y) - \lfloor x \rfloor f(x) - \left(\int_x^y \lfloor t \rfloor f'(t) dt \right).
\end{aligned}$$

Now doing integration by parts gives

$$\int_x^y t f'(t) dt = \int_x^y f(t) dt + y f(y) - x f(x),$$

and adding this to the last expression finishes the proof.

Theorem 6 (Exponential Sums). *For $x \geq 1$ we've*

1. $\sum_{n \leq x} 1/n = \log x + C + O(1/x);$
2. $\sum_{n \leq x} 1/n^s = x^{1-s}/(1-s) + \zeta(s) + O(x^{-s})$ if $s \in \mathbb{R}_{>0} - \{1\};$
3. $\sum_{n > x} 1/n^s = O(x^{1-s})$ if $s > 1;$
4. $\sum_{n \leq x} n^u = x^{u+1}/(u+1) + O(x^u)$ if $u \geq 0.$

Let's understand what the significance of these equations. The first of the above equations says that if we define the function $F : \mathbb{R}_{>1} \rightarrow \mathbb{R}$ by $F(x) = \sum_{n \leq x} 1/n$, then the function $F(x) - \log x - C$ grows slower than $1/x$ as $x \rightarrow \infty$ which means there exists a constant k such that for all sufficiently large x we've

$$C + \log x - \frac{k}{x} \leq F(x) \leq C + \log x + \frac{k}{x}.$$

The other three equations can similarly be converted into such inequalities.

Proof.

1. Taking $f(t) = 1/t$ in Theorem 5 gives

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

Here, we used the fact that $-(x - \lfloor x \rfloor)/x = O(1/x)$, which is true because $x - \lfloor x \rfloor \leq 1$ for all x .

Now observe

$$\int_1^x \frac{t - \lfloor t \rfloor}{t^2} dt = \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt - \int_x^\infty \frac{t - \lfloor t \rfloor}{t^2} dt,$$

where the first improper integral converges by the comparison test with $\int_1^\infty t^{-2} dt$. Denote the second integral as a function $G(x)$ for $x \geq 1$. We have $0 \leq G(x) \leq \int_x^\infty t^{-2} dt = 1/x$, so actually

$$\int_1^x \frac{t - \lfloor t \rfloor}{t^2} dt = \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt + O\left(\frac{1}{x}\right)$$

and putting this in the previous equation we get

$$\sum_{n < x} \frac{1}{n} = \log x + \left(1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt \right) + O\left(\frac{1}{x}\right).$$

Note that the two $O(1/x)$ have been combined into one. Moving $\log x$ to the left side and taking the limit as $x \rightarrow \infty$, the $O(1/x)$ term disappears and we are left with

$$\left(1 - \int_1^\infty \frac{t - \lfloor t \rfloor}{t^2} dt \right) = \lim_{x \rightarrow \infty} \left(\sum_{n < x} \frac{1}{n} - \log x \right) = C,$$

finishing the proof.

2. Let $f(t) = 1/t^s$ in Theorem 5 and observe as in part (a) that

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

where

$$C(s) = 1 - \frac{1}{1-s} - s \int_1^\infty \frac{t - \lfloor t \rfloor}{t^{s+1}} dt.$$

If $s > 1$, then in the limit the left side converges to $\zeta(s)$ while the non-constant terms on the right vanish, leaving $C(s) = \zeta(s)$. If $s < 1$, then moving the x^{1-s} term to the left and taking the limit again gives $C(s) = \zeta(s)$, finishing the proof.

3. We 've

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

from part (b).

4. Simply put $f(t) = t^u$ in Theorem 5 and reduce as in parts (a) and (b).

Using these results, it's now possible to prove the two theorems below that determine the average order of τ and σ_s . Unfortunately, the proof will not be given here due to the lack of space, but it can be found, as mentioned above, in Chapter 3 of [1].

Theorem 7 (Dirichlet's Formula). *For $x \geq 1$ we have*

$$\sum_{n \leq x} \tau(n) = x \log x + (2C - 1)x + O(\sqrt{x})$$

Although the error term $O(\sqrt{x})$ has been improved to smaller powers of x , the infimum of all u for which the error term is $O(x^u)$ is an open problem called *Dirichlet's divisor problem*.

Theorem 8 (Divisor Sum Estimates). *For $x \geq 1$ we have*

$$\sum_{n \leq x} \sigma_s(n) = \begin{cases} \pi^2 x^2 / 12 + O(x \log x) & \text{if } s = 1 \\ \pi^2 x / 6 + O(\log x) & \text{if } s = -1 \\ \zeta(s+1) x^{s+1} / (s+1) + O(x^{\max(1,s)}) & \text{if } s > 0 \text{ and } s \neq 1 \\ \zeta(1-s) x + O(x^{\max(0,1+s)}) & \text{if } s < 0 \text{ and } s \neq -1. \end{cases}$$

As a direct corollary, we get the average orders; simply divide by a factor of x in each of the above-cited formulas. We ignore the error term when mentioning the average order, so the average order of τ is $\log x$ and that of σ_1 is $\pi^2 x / 12$.

Theorem 9. *For $x \geq 1$ we've $\tilde{\tau}(x) = \log x + (2C - 1) + O(x^{-1/2})$ and*

$$\tilde{\sigma}_s(x) = \begin{cases} \pi^2 x / 12 + O(\log x) & \text{if } s = 1 \\ \pi^2 / 6 + O(\log x / x) & \text{if } s = -1 \\ \zeta(s+1) x^s / (s+1) + O(x^{\max(0,s-1)}) & \text{if } s > 0 \text{ and } s \neq 1 \\ \zeta(1-s) + O(x^{\max(-1,s)}) & \text{if } s < 0 \text{ and } s \neq -1. \end{cases}$$

The above result can be interpreted as saying that as $x \rightarrow \infty$, the function τ is “on average” $\log x + (2C - 1)$, since the error term goes to 0 in the limit; similarly, σ_{-1} is “on average” $\pi^2 / 6$ for large x .

Chapter 4

Conclusion

In this investigation, I've introduced and defined two important notations, O and o , from their fundamental concepts. The theoretical aspects of their usage have been thoroughly discussed, primarily focusing on their application in dealing with infinite asymptotic estimates. However, it's important to note that both O and o have broader applications, including handling error terms in infinitesimal limits.

The formalization of equations involving Bachmann-Landau notation proved to be challenging due to the inherent vagueness in its interpretation. I've presented a sustained example to illustrate the practical use of O in real mathematics. The example exclusively focused on employing O for infinite asymptotic estimates, but it is crucial to recognize that these notations can be utilized in various mathematical contexts.

Bibliography

- [1] Tom M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag, 1976.
- [2] Paul Gustav Heinrich Bachmann. *Die Analytische Zahlentheorie*. Leipzig B.G. Teubner, 1894. URL: <https://archive.org/details/dieanalytischeza00bachuoft/>.
- [3] Edmund Landau. *Handbuch der Lehre von der Verteilung der Primzahlen*. Leipzig B.G. Teubner, 1909. URL: <https://archive.org/details/handbuchderlehre01landuoft/>.
- [4] Walter Rudin. *Principles of Mathematical Analysis*. McGraw Hill, 1953.