Bachmann-Landau Notation

1 Introduction

Paul Bachmann (1837-1920) and Edmund Landau (1877-1930) were German mathematicians who are both remembered for their numerous contributions to number theory. It was in Bachmann's book *Die Analytische Zahlentheorie*[1] [Bac94] on analytic number theory, a branch of mathematics best characterised by the application of analytic methods to number theoretic problems, that big-O notation was first introduced to mathematics, which he used for handling the error terms in his asymptotic estimates. Inspired by Bachmann's idea, Landau introduced a stronger variant of big-O notation, which we now call little-o, in his two-volume treatise titled *Handbuch der Lehre von der Verteilung der Primzahlen*[2] [Lan09]. Today, these two notations, along with other similarly defined notions (like Θ-, ω-, or Ω-notation), are used in a wide variety of areas ranging, of course, from analytic fields like analytic number theory to even computer science.

I myself have often come across Bachmann-Landau notation in several contexts, but in each such instance, I have had to look up and struggle with their definitions since their meaning used to escape me whenever they appeared in equations. The usage of big-O and little-o seems to be intuitively clear to mathematicians yet communicating it properly often poses a difficulty. Thus this investigation is directed towards understanding big-O and little-o notation for what they are and especially how they are to be applied in appropriate mathematical contexts.

2 Theory

Both big-O and little-o notation describe the size of a function relative to another function "in the limit". Usually, this limit point is taken to be ∞ (*infinite* asymptotics), or 0 (*infinitesimal* asymptotics). The definition given here generalises this notion to any limit point $a \in \overline{\mathbb{R}}$, but the main use cases will remain $a = 0$ and $a = \infty$. Here $\overline{\mathbb{R}}$ represents the *extended real numbers*, namely $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. A good reference for the basic properties of the extended reals would be the beginning of [Rud53].

[1] Translates literally to *Analytic Number Theory.*

[2] Translates to *Handbook of the theory of distribution of prime numbers.*

2.1 Big-O

There are generally two ways to define each notation in the Bachmann-Landau family of notations: a direct limit-based approach, and a quantifier-based approach.

Definition 1. Given $a \in \overline{\mathbb{R}}$ and real functions f, g that are defined on a neighbourhood of a such that g is nonzero on this neighbourhood, we say that $f(x) = O_a(g(x))$ or $f(x) = O(g(x))$ as $x \to a$ if one of the following equivalent conditions hold:

1. there exists $\delta \in (0, \infty)$ such that
$$\sup_{N_\delta(a)} \left| \frac{f(x)}{g(x)} \right| < \infty;$$
2. there exist $k, \delta \in (0, \infty)$ such that $x \in N_\delta(a)$ implies $|f(x)| \le k\,|g(x)|$.

Here $N_\delta(d)$ denotes the neighbourhood of a of radius δ, which is $(a - \delta, a + \delta)$ for finite a. For infinite values, we can define it differently: set $N_\delta(\infty) = (\delta, \infty)$ and $N_\delta(-\infty) = (-\infty, -\delta)$.

The set implicit in the first supremum is

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

Let $s_\delta = \sup S_\delta$. If $s_\delta < \infty$ for some δ, then as s_δ is an upper bound of S_δ, we have $|f(x)| \le s_\delta\,|g(x)|$ for all x such that $x \in N_\delta$; thus the first condition implies the second. Conversely, suppose that there exist $k, \delta \in (0, \infty)$ such that $x \in N_\delta(a)$ implies $|f(x)| \le k\,|g(x)|$. Then k is an upper bound of S_δ, and the least-upper-bound property of $\mathbb{R}$ then shows that S_δ has a finite supremum. Thus the two conditions are equivalent.[3]

Note that g is required to be nonzero in order for the fraction $f(x)/g(x)$ to make sense, but that it is sufficient that it be nonzero over some neighbourhood of a for the definition to work: we can rewrite the definition to ask for such a δ that is less than the radius of the neighbourhood in which g is nonzero.

The definition found in [Lan09] is for $a = \infty$ only, since analytic number theory is concerned primarily with infinite limits; the case $a = 0$, on the other hand, is useful in analysis, for example when truncating the higher order terms in a Taylor series. As the below examples show, O_a is rather useless for finite nonzero values of a. But the above definition is still preferred to defining O_0 and O_∞ separately because it highlights the "sameness" of both notions.

When it comes to Bachmann-Landau notations, however, much more important than the rigorous definition is the understanding. The right way to understand big-O is that big-O is to functions what $\le$ is to numbers:

$$f(x) = O_a(g(x)) \text{ means } f(x) \le g(x) \text{ in the limit, up to a constant factor.}$$
$$(f \text{ does not dominate } g \iff a \text{ is not greater than } b)$$

[3] As is apparent in this argument, the first condition basically says that $|f(x)/g(x)|$ should be bounded above in some neighbourhood of a.

Examples:

- Let us verify that $100x^2 = O_\infty(x^3)$ using both conditions. Using the first condition, for any $\delta \in (0, \infty)$ we have

$$\sup_{x>\delta} \left| \frac{100x^2}{x^3} \right| = \sup_{x>\delta} \frac{100}{x} = \frac{100}{\delta} < \infty.$$

  Using the second condition, we know that for $x > 1$ we have $100x^2 \leq 100x^3$. In terms of the above analogy, we can observe that $100x^2$ is "less than or equal to" x^3 in the limit to infinity, as x^2 grows slower than x^3 as $x \to \infty$.

- If f and g are functions that are each either even or odd, then $f = O_\infty(g)$ if and only if $f = O_{-\infty}(g)$. Thus $100x^2 = O_{-\infty}(x^3)$.

- But $100x^2 \neq O_0(x^3)$; otherwise, there would exist $\delta, k \in (0, \infty)$ such that $|x| < \delta$ implies $x \geq k/100$, which is absurd. We do have, however, that $x^3 = O_0(100x^2)$, as for $|x| < 100$ we have $x^3 \leq 100x^2$. There is a graphical way to understand the relations $100x^2 = O_\infty(x^3)$ and $x^3 = O_0(100x^2)$. The graph of $100x^2$ is below the graph of x^3 after $x = 100$, meaning that $100x^2 \leq x^3$ as $x \to \infty$; similarly, as the graph of x^3 is below the graph of $100x^2$ for all $x < 100$, we see that $x^3 \leq 100x^2$ as $x \to 0$.

- To give a more exaggerated example, we have $x^2 = O_\infty(x^{100})$ but $x^{100} = O_0(x^2)$.

- Given a finite a, if f does not have any asymptote at a and g is nonzero in some neighbourhood of a, then we have $f = O_a(g)$. This is because $|f/g|$ achieves the supremum over a sufficiently small neighbourhood.

It is thus evident that O_a is not very interesting except when $a = 0$ or $a = \infty$, so it makes sense that the symbol O_a is not standard; the literature generally exhibits $f(x) = O(g(x))$ as $x \to a$, or even hides the "as $x \to a$" since the value of a is usually clear from the context (either 0 or ∞). We will omit the a subscripts from now on, and sometimes, as in the following theorem, the variable x too.

Theorem 1 (Transitivity). *If $f = O(g)$ and $g = O(h)$, then $f = O(h)$.*[4]

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 $x \in N_\epsilon(a)$. It follows that

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

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

2.2 Little-o

Once we have an analogue of $\leq$ for functions, we are motivated to consider if such a notion exists for $<$. This gives rise to little-o.

[4] If a is not specified, neither explicitly nor from context, as in this theorem, then assume any arbitrary $a \in \overline{\mathbb{R}}$.

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

1. we have $\lim_{x\to a} f(x)/g(x) = 0$;
2. for all $k > 0$, there exists $\delta \in (0, \infty)$ such that $|f(x)| \leq k\,|g(x)|$ for all $x \in N_\delta(a)$.[5]

The equivalence in this case follows by the definition of a limit: if $\lim_{x\to a} f(x)/g(x) = 0$, 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, the definition given by Landau in [Lan09] assumes $a = \infty$.

As with big-O, there is a simple way to understand little-o:

$$f(x) = o_a(g(x)) \text{ means that } kf(x) < g(x) \text{ in the limit for } \textit{any } k.$$
$$(g \text{ dominates } f \iff b \text{ is greater than } a)$$

Thus little-o is stronger than big-O, in that $f(x) = o_a(g(x))$ implies $f(x) = O_a(g(x))$ but not vice versa.

Examples:

- We have $\sin x = o_\infty(x)$ as $\sin x/x$ clearly tends to 0 as $x \to \infty$. However, from elementary calculus we know that $\lim_{x\to 0} \sin x/x = 1 \neq 0$, so $\sin x \neq o_0(x)$. Similarly $x \neq o_0(\sin x)$ either, as the limit of $x/\sin x$ is also 1. We can see that $\sin x = o_\infty(x)$ using the second condition too. Pick $k > 0$, so we must find δ such that $\sin x \leq kx$ for all $x > \delta$. But this is easy: $\sin x \leq 1 = k \cdot 1/k \leq kx$ for all $x > 1/k$, so $\delta = 1/k$ works.
- If $\lim_{x\to 0} f(x) = 0$, then $f = o_0(1)$. In particular we have $\sin x = o_0(1)$ and $\tan x = o_0(1)$.
- Remember from last section that $100x^2 = O_\infty(x^3)$. Since $100x^2/x^3 = 100/x \to 0$ as $x \to \infty$, we have $100x^2 = o_\infty(x^3)$. However, note that while $bx^3 = O_\infty(x^3)$ for any real b, the same does not hold with big-O replaced by little-o.
- We have $f(x) = o_a(g(x))$ if and only if $f(x + a) = o_0(g(x + a))$ for real functions f and g defined all over $\mathbb{R}$.

As with big-O, we will almost always omit the subscript and sometimes the variable when using little-o.

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

Proof. Given $m \in (0, \infty)$, there exist $\delta, \epsilon \in (0, \infty)$ such that $|f(x)| \leq \sqrt{m}\,|g(x)|$ for all $x \in N_\delta(a)$ and $|g(x)| \leq \sqrt{m}\,|h(x)|$ for all $x \in N_\epsilon(a)$. So we have

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

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

[5] The k here actually belongs on the left side because it is g that is always greater no matter which multiple of f we take; it is placed on the right side just to make it appear similar to the definition of big-O.

2.3 Bachmann-Landau equations

We now define a new type of equation in which big-O's and little-o's are allowed to appear on either side any number of times, so equations like $x^2 + O(x^3)o(1) = o(x^4)$ will make sense. Mathematicians usually do not distinguish between these so-called *Bachmann-Landau* equations and normal equations, which creates some confusion. A better way is to think of the class of normal equations as a subclass of the class of Bachmann-Landau equations; the latter make a claim about the existence of some functions satisfying certain properties. The definition below describes a very general situtation without too much focus on formalisation, as that would distract from the topic at hand.

Definition 3. Given $a \in \overline{\mathbb{R}}$, let X^i, Y^j represent any of O and o for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. Let f_i, g_j be real functions defined and nonzero on a neighbourhood of a, for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. An equation of the form

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

for some expressions F and G is called a *Bachmann-Landau equation in x* and means the following: for any real functions $f_1^*, \ldots, f_m^*$ defined on a neighbourhood of a such that $f_i^*(x) = X^i(f_i(x))$, there exist real functions $g_1^*, \ldots, g_n^*$ defined on an equal or larger neighbourhood of a such that $g_j^*(x) = Y^j(g_j(x))$ and

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

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:

- Consider the Bachmann-Landau equation given above: $x^2 + O(x^3)o(1) = o(x^4)$, as $x \to \infty$. Let us check using the above definition that this equation holds. Take arbitrary real functions f and g defined everywhere such that $|f(x)| \le k\,|x|^3$ for all $x \in (-\delta, \delta)$ for some finite $\delta > 0$ and $\lim_{x\to 0} g(x) = 0$; we must verify that $x^2 + f(x)g(x) = o(x^4)$, or in other words that $\frac{x^2+f(x)g(x)}{x^4} \to 0$ as $x \to \infty$. Since

$$\lim_{x\to 0} \frac{x^2 + f(x)g(x)}{x^4} = \lim_{x\to 0} \frac{x^2}{x^4} + \left(\lim_{x\to 0} g(x)\right)\left(\lim_{x\to 0} \frac{f(x)}{x^4}\right)$$

  if the very last limit converges, it suffices to show that $f(x)/x^4 \to 0$ as $x \to \infty$. But this is trivial: given $\epsilon > 0$, we have $\left|f(x)/x^4\right| \le k/\,|x| < \epsilon$ for all $|x| > k/\epsilon$.

- Let us verify that $n^{O(1)} = e^{O(n)}$ as $n \to \infty$; note that here the variable is n instead of x, as is common in number theory, and note also that since left side might not make sense for $n < 0$ all domains will be $\mathbb{R}_{>0}$. Take any real function f defined on $\mathbb{R}_{>0}$ such that $f = O(1)$; it suffices to show that $f(n)\log n = O(n)$. By definition, there exist finite $N, k > 0$ such that $n > N$ implies $|f(n)| \le k$, and we know that $\log n < n$ for all $n > 1$.[6] Thus for all $n > \max(N, 1)$ we have $|f(n)| \log n \le kn$, as desired.

- However, $e^{O(n)} \ne n^{O(1)}$. This is to say, there exists some $f(n) = O(n)$ for which there is no $g(n) = O(1)$ satisfying $e^{f(n)} = n^{g(n)}$. But this is not hard: just take $f(n) = n$. Then the only possible function g would be $g(n) = n \log_n(e) = n/\log n$, but it is easy to check that $n/\log n \ne O(1)$. Of course, here $a = \infty$.

- The Swedish mathematician Helge von Koch showed in [Koc01] that the Riemann hypothesis implies

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x} \log x);$$

here $\pi(x)$ is the number of primes from 1 to x and $\mathrm{Li}(x) = \int_2^x dt/\log t$. This equation means that $|\pi(x) - \mathrm{Li}(x)| \le k\sqrt{x} \log x$ for some real $k > 0$ and all sufficiently large x, or explicitly that

$$\mathrm{Li}(x) - k\sqrt{x} \log x \le \pi(x) \le \mathrm{Li}(x) + \sqrt{x} \log x.$$

Fortunately, we do not 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 manipulated 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 $|f_1(x)| \le k\,|f(x)|$ for all $x \in N_\delta(a)$ and $|g_1(x)| \le l\,|g(x)|$ for all $x \in N_\epsilon(a)$. We know that $N_\delta(a) \cap N_\epsilon(a) = N_\gamma(a)$ for some $\gamma \in (0, \infty)$.[7] Let $m = \max(k, l)$. Then $|f_1(x)/f(x)| \le m$ and $|g_1(x)/g(x)| \le m$ for all $x \in N_\gamma(a)$, therefore

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

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

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

[6] That $\log n < n$ for $n \ge 1$ follows from the *racetrack principle*; see [Spi11] for details.
[7] We have $\gamma = \min(\delta, \epsilon)$ if a is finite, and $\gamma = \pm \max(\delta, \epsilon)$ if $a = \pm\infty$.

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

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

□

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, g_1 be real functions defined on a neighbourhood of a such that $f_1 = o(f)$ and $g_1 = o(g)$. Then $\lim_{x\to a} f_1(x)/f(x) = \lim_{x\to a} g_1(x)/g(x) = 0$ implies

$$\lim_{x\to a} \frac{f(x)g_1(x)}{f(x)g(x)} = \lim_{x\to a} \frac{g_1(x)}{g(x)} = 0 \qquad \text{and} \qquad \lim_{x\to a} \frac{f_1(x)g_1(x)}{f(x)g(x)} = \left(\lim_{x\to a} \frac{f_1(x)}{f(x)} \right) \left(\lim_{x\to a} \frac{g_1(x)}{g(x)} \right) = 0;$$

thus $o(f)o(g) = o(fg)$ and $fo(g) = o(fg)$. Also $\lim_{x\to a} \lambda f_1(x)/f(x)$ implies $\lambda o(f) = o(f)$, and if $f_2 = o(\lambda f)$, then $\lim_{x\to a} f_2(x)/(\lambda f(x)) = 0$ and therefore $\lim_{x\to a} f_2(x)/f(x) = 0$, showing $o(\lambda f) = o(f)$. Now suppose f and g are nonnegative, implying that are f_1 and g_1, and take any $m \in (0, \infty)$. There exist $\delta, \epsilon \in (0, \infty)$ such that we have $f_1(x) \leq mf(x)$ for all $x \in N_\delta(a)$ and $g_1(x) \leq mg(x)$ for all $x \in N_\epsilon(a)$. Then $f_1(x) + g_1(x) \leq m(f(x) + g(x))$ for all $x \in N_\zeta(a)$, where $\zeta = \min(\delta, \epsilon)$. □

Examples:

- abc.

3 Applications

To demonstrate the use Bachmann-Landau notation in everyday mathematics, two different example areas are covered. In the first place, some very basic estimates in analytic number theory that require big-O are presented. These are taken from Chapter 3 of [Apo76]. Secondly, to show the use of little-o in the context of infinitesimal asymptotics, we define and develop the theory of differentiation using little-o.

3.1 Analytic Number Theory

3.2 Differentiation

4 Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent mattis, nibh ut venenatis maximus, nibh ex finibus libero, sed fermentum lorem ex ut erat. Proin quis risus vel velit scelerisque laoreet. Maecenas ut eros mattis turpis tincidunt scelerisque quis ut augue. Sed viverra sem eu enim viverra facilisis. Etiam laoreet felis ac dolor malesuada tincidunt. Morbi id ante mauris. Sed laoreet aliquet ante. Interdum et malesuada fames ac ante ipsum primis in faucibus. Fusce leo diam, varius posuere dignissim sit amet, volutpat eu est. Aliquam erat volutpat. Maecenas tincidunt leo elit, mollis dignissim turpis malesuada ut. Ut congue sollicitudin scelerisque. Maecenas sed lectus eu nibh accumsan auctor in dictum tellus.

References

[Apo76] Tom M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag, 1976.

[Bac94] Paul Gustav Heinrich Bachmann. *Die Analytische Zahlentheorie*. Leipzig B.G. Teubner, 1894. URL: `https://archive.org/details/dieanalytischeza00bachuoft/`.

[Koc01] Helge von Koch. "Sur la distribution des nombres premiers". In: (1901). DOI: `10.1007/BF02403071`.

[Lan09] Edmund Landau. *Handbuch der Lehre von der Verteilung der Primzahlen*. Leipzig B.G. Teubner, 1909. URL: `https://archive.org/details/handbuchderlehre01landuoft/`.

[Rud53] Walter Rudin. *Principles of Mathematical Analysis*. McGraw Hill, 1953.

[Spi11] Mike Spivey. *How to prove* $\log n < n$*?* 2011. URL: `https://math.stackexchange.com/q/65203`.