N. G. de Bruijn begins his book *Asymptotic Methods in Analysis* by considering a Big \( \mathcal{O} \) notation that helps us understand Big Oh. If we write \( L(5) \) for a number whose absolute value is less than 5 (but we don’t say what the number is), then we can perform certain calculations without knowing the full truth. For example, we can deduce formulas such as \( 1 + L(5) = L(6) \); \( L(2) + L(3) = L(5) \); \( L(2)L(3) = L(6) \); \( e^{L(5)} = L(e^5) \); and so on. But we cannot conclude that \( L(5) = L(3) = L(2) \), since the left side might be 4 \( \neq 0 \). In fact, the most we can say is \( L(5) = L(3) = L(8) \).

Bachmann’s \( \mathcal{O} \)-notation is similar to \( L \)-notation but it’s even less precise: \( \mathcal{O}(\alpha) \) stands for a number whose absolute value is at most a constant times \( |\alpha| \). We don’t say what the number is and we don’t even say what the constant is. Of course the notion of a “constant” is nonsense if there is nothing variable in the picture, so we use \( \mathcal{O} \)-notation only in contexts when there’s at least one quantity (say \( n \)) whose value is varying. The formula

\[
f(n) = \mathcal{O}(g(n)) \quad \text{for all } n
g(n)
\]

meets in this context that there is a constant \( C \) such that

\[
|f(n)| \leq C|g(n)| \quad \text{for all } n;
\]

and when \( \mathcal{O}(g(n)) \) stands in the middle of a formula it represents a function \( f(n) \) that satisfies \( (9.12) \). The values of \( f(n) \) are unknown, but we do know that they aren’t too large. Similarly, de Bruijn’s ‘\( L(n) \)’ represents an unspecified function \( f(n) \) whose values satisfy \( f(n) \ll |n| \). The main difference between \( L \) and \( \mathcal{O} \) is that \( \mathcal{O} \)-notation involves an unspecified constant \( C \); each appearance of \( \mathcal{O} \) might involve a different \( C \), but each \( C \) is independent of \( n \).

For example, we know that the sum of the first \( n \) squares is

\[
\square_n = \frac{1}{2}n(n+1)(n+1) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n.
\]

We can write

\[
\square_n = \mathcal{O}(n^3)
\]

because \( |\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n| \leq \frac{1}{3}|n^3| + \frac{1}{2}|n^2| + \frac{1}{6}|n| \leq |n|^3 \) for all integers \( n \). Similarly, we have the more specific formula

\[
\square_n = \frac{1}{3}n^3 + \mathcal{O}(n^2);
\]

we can also be sloppy and throw away information, saying that

\[
\square_n = \mathcal{O}(n^{10}).
\]

Nothing in the definition of \( \mathcal{O} \) requires us to give a best possible bound.