N. G. de BRUIJN

ASYMPTOTIC METHODS IN ANALYSIS

ASYMPTOTIC METHODS IN ANALYSIS

BY

N. G. DE BRUIJN

PROFESSOR OF MATHEMATICS AT THE UNIVERSITY OF AMSTERDAM, NETHERLANDS

1958

NORTH-HOLLAND PUBLISHING CO. – AMSTERDAM

P. NOORDHOFF LTD. – GRONINGEN

SOLE DISTRIBUTORS FOR U.S.A.:
INTERSCIENCE PUBLISHERS INC., NEW YORK

PRINTED IN THE NETHERLANDS

PREFACE

This book arose from a course of lectures given during the academic year 1954/'55 at the Mathematical Centre, Amsterdam, repeated in 1956/'57 in a course at Eindhoven, organized by the same institution. Its purpose is to teach asymptotic methods by explaining a number of examples in every detail, so as to suit beginners who seriously want to acquire some technique in attacking asymptotic problems.

Although asymptotics is by no means a new field, only in recent times have special courses and books been devoted to it. The reason may be that today university courses in analysis are condensed in favour of modern mathematics. The effect is that analytic techniques are not so widespread as they used to be. On the other hand there are so many questions of an asymptotic nature both in pure and applied mathematics, that we cannot afford to neglect the subject. Therefore it seems desirable to give a separate training in asymptotics to those who have only a general knowledge of analysis.

The reader will not find anything like a general theory in this book. Many asymptotic methods are very flexible, and in such cases it is not possible to formulate a single theorem covering all applications. Any attempt at generalization would actually result in a restriction.

Usually in mathematics one has to choose between saying more and more about less and less on the one hand, and saying less and less about more and more on the other. In a practical subject like this, however, the proper attitude should be to say as much as possible in a given situation. And as it is impossible to include all situations in one book, it seems to be imperative to teach by examples, stating generalities only when strictly necessary.

The choice of subjects in a book of this size is, of course, quite arbitrary. The reader will find quite an extensive treatment of the saddle point method (Ch. 4, 5, 6), and a great deal of attention has been paid to iteration (Ch. 8). On the other hand, with respect to Tauberian theorems (in Ch. 7), and the asymptotics of differential equations (Ch. 9), this book presents only a very small part of what the reader might expect. And what is worse, nothing of analytic number theory is included because this would require too much space. On the other hand, there are so many excellent books on analytic number theory that there is no need for another text-book. But no doubt anyone who wants to specialize in asymptotics, should study analytic number theory, with its great variety of beautiful asymptotic problems.

Many things in this book are not presented in the shortest possible form, as an attempt has been made to reveal, to a certain extent, the motives that lead to certain methods. Naturally one cannot go too far in this respect; a mathematician cannot possibly publish his waste-paper basket.

In some cases two or more different treatments of one and the same problem are given, so as to enable the reader to compare different methods. Several proofs for the Stirling theorem are included; the problem of sec. 4.7 is attacked a second time in Ch. 6; and the problem of the iterated sine is treated twice in Ch. 8.

On the whole the author has tried to discuss original problems and results, unless the inclusion of well-known things was strictly necessary. In a field like this it is, of course, very difficult to say whether something is new, especially when the ideas are certainly well-known. The contents of the following parts have probably never been published before: sec. 3.9, sec. 4.7, Ch. 6, Ch. 8 from sec. 8.4 onwards, and possibly even sec. 2.4, sec. 9.2 and sec. 9.3.

This book has not been written exclusively for mathematicians, but also for those physicists and engineers who have a certain maturity with respect to analysis, including some general knowledge of complex function theory. On the whole it will not be easy reading for any class of readers, asymptotics being a difficult subject that requires constant alertness and carefulness. For those who find the text occasionally too difficult it may be reassuring to know that the chapters of this book can be studied practically independently. The only exceptions are Chs. 5 and 6, which are based upon Ch. 4. And the introduction is, of course, fundamental for the whole book.

Most chapters start with simple things and end in quite hard examples. At the end of each chapter there are a few exercises. Even when these are quite difficult, they do not require methods beyond those treated in the text.

Hardly any references are given in this book, because the subject is so old that it is very difficult to give the proper ones. For a short bibliography of asymptotics we refer to A. ERDÉLYI, *Asymptotic Expansions*, Dover Publ., 1956, which also contains a much more extensive introduction to the asymptotics of differential equations than Ch. 9 of this book.

One warning should be given to all readers: this is not an encyclopaedia on asymptotic results. Not even the asymptotic behaviour of Bessel functions can be found in this book. Attention is focussed mainly on methods. The problems themselves are not of primary importance; their selection depends on instructiveness rather than on importance.

October, 1957. N. G. DE BRUIJN

CONTENTS

CHAPTER 1

INTRODUCTION

1.1. What is asymptotics?

This question is about as difficult to answer as the question "What is mathematics?" Nevertheless, we shall have to find some explanation for the word asymptotics.

It often happens that we want to evaluate a certain number, defined in a certain way, and that the evaluation involves a very large number of operations so that the direct method is almost prohibitive. In such cases we should be very happy to have an entirely different method for finding information about the number, giving at least some useful approximation to it. And usually this new method gives (as remarked by Laplace) the better results in proportion to its being more necessary: its accuracy improves when the number of operations involved in the definition increases. A situation like this is considered to belong to asymptotics.

This statement is very vague indeed. However, if we try to be more precise, a definition of the word asymptotics either excludes everything we are used to call asymptotics, or it includes almost the whole of mathematical analysis. It is hard to phrase the definition in such a way that Stirling's formula (1.1.1) belongs to asymptotics, and that a formula like $\int_0^\infty (1+x^2)^{-1}dx = \frac{1}{2}\pi$ does not. The obvious reason why the latter formula is not called an asymptotic formula is that it belongs to a part of analysis that already has got a name: the integral calculus. The safest and not the vaguest definition is the following one: Asymptotics is that part of analysis which considers problems of the type dealt with in this book.

A typical asymptotic result, and one of the oldest, is Stirling's formula just mentioned:

(1.1.1) $$\lim_{n\to\infty} n!/(e^{-n}n^n\sqrt{2\pi n}) = 1.$$

For each n, the number $n!$ can be evaluated without any theoretical difficulty, and the larger n is, the larger the number of necessary operations becomes. But Stirling's formula gives a decent approximation $e^{-n}n^n\sqrt{2\pi n}$, and the larger n is, the smaller its relative error becomes.

Several proofs of (1.1.1) and of its generalizations will be given in this book (see secs. 3.7, 3.10, 4.5, 6.9).

We quote another famous asymptotic formula, much deeper than the previous one and beyond the scope of this book. If x is a positive number, we denote by $\pi(x)$ the number of primes not exceeding x. Then the so-called Prime Number Theorem states that [1])

(1.1.2) $$\lim_{x\to\infty} \pi(x) \bigg/ \frac{x}{\log x} = 1.$$

The above formulas are limit formulas, and therefore they have, as they stand, little value for numerical purposes. For no single special value of n can we draw any conclusion from (1.1.1) about $n!$. It is a statement about infinitely many values of n, which, remarkably enough, does not state anything about any special value of n.

For the purpose of closer investigation of this feature, we abbreviate (1.1.1) to

(1.1.3) $$\lim_{n\to\infty} f(n) = 1, \quad \text{or} \quad f(n) \to 1 \quad (n \to \infty).$$

This formula expresses the mere existence of a function $N(\varepsilon)$ with the property that:

(1.1.4) for each $\varepsilon > 0$: $n > N(\varepsilon)$ implies $|f(n) - 1| < \varepsilon$.

When proving $f(n) \to 1$, one usually produces, hidden or not, information of the form (1.1.4) with explicit construction of a suitable function $N(\varepsilon)$. It is clear that the knowledge of $N(\varepsilon)$ actually means numerical information about f. However, when using the notation $f(n) \to 1$, this information is suppressed. So if we write (1.1.3), the knowledge of a function $N(\varepsilon)$ with the property (1.1.4) is replaced by the knowledge of the existence of such a function.

[1]) See A. E. Ingham, The Distribution of Primes, Cambridge 1932.

To a certain extent it is one of the reasons of the success of analysis that a notation has been found which suppresses that much information and still remains useful. Even with quite simple theorems, for instance $\lim a_n b_n = \lim a_n \cdot \lim b_n$, it is easy to see that the existence of the functions $N(\varepsilon)$ is easier to handle than the functions $N(\varepsilon)$ themselves.

1.2. The O-symbol

A weaker form of suppression of information is given by the Bachmann-Landau O-notation [1]). It does not suppress a function, but only a number. That is to say, it replaces the knowledge of a number with certain properties by the knowledge that such a number exists. The O-notation suppresses much less information than the limit notation, and yet it is easy enough to handle.

Assume that we have the following explicit information about the sequence $\{f(n)\}$:

(1.2.1) $$|f(n) - 1| \leq 3n^{-1} \qquad (n = 1, 2, 3, \ldots).$$

Then we clearly have a suitable function $N(\varepsilon)$ satisfying (1.1.4), viz. $N(\varepsilon) = 3\varepsilon^{-1}$. Therefore,

(1.2.2) $$f(n) \to 1 \qquad (n \to \infty).$$

It often happens that (1.2.2) is useless, and that (1.2.1) is satisfactory for some purpose on hand. And it often happens that (1.2.1) would remain as useful if the number 3 would be replaced by 10^5 or any other constant. In such cases, we could do with

(1.2.3) $$\begin{cases} \text{There exists a number } A \text{ (independent of } n\text{) such that} \\ \qquad |f(n) - 1| \leq An^{-1} \qquad (n = 1, 2, 3, \ldots). \end{cases}$$

The logical connections are given by

$$(1.2.1) \Rightarrow (1.2.3) \Rightarrow (1.2.2).$$

Now (1.2.3) is the statement expressed by the symbolism

(1.2.4) $$f(n) - 1 = O(n^{-1}) \qquad (n = 1, 2, 3, \ldots).$$

There are some minor differences between the various definitions

[1]) See E. Landau, Vorlesungen über Zahlentheorie, Leipzig 1927, vol. 2, p. 3 – 5.

of the O-symbol that occur in the literature, but these differences are unimportant. Usually the O-symbol is meant to represent the words "something that is in absolute value less than a constant number times". Instead, we shall use it in the sense of "something that is, in absolute value, at most a constant multiple of the absolute value of". So if S is any set, and if f and φ are real or complex functions defined on S, then the formula

$$f(s) = O(\varphi(s)) \qquad (s \in S), \tag{1.2.5}$$

means that there is a positive number A, not depending on s, such that

$$|f(s)| \leq A|\varphi(s)| \quad \text{for all } s \in S. \tag{1.2.6}$$

If, in particular, $\varphi(s) \neq 0$ for all $s \in S$, then (1.2.5) simply means that $f(s)/\varphi(s)$ is bounded throughout S.

We quote some obvious examples:

$$\begin{aligned} x^2 &= O(x) && (|x| < 2), \\ \sin x &= O(1) && (-\infty < x < \infty), \\ \sin x &= O(x) && (-\infty < x < \infty). \end{aligned}$$

Quite often we are interested in results of the type (1.2.6) only on part of the set S, especially on those parts of S where the information is non-trivial. For example, with the formula $\sin x = O(x)$ $(-\infty < x < \infty)$ the only interest lies in small values of $|x|$. But those uninteresting values of the variable sometimes give some extra difficulties, although these are not essential in any respect. An example is:

$$e^x - 1 = O(x) \qquad (-1 < x < 1).$$

We are obviously interested in small values of x here, but it is the fault of the large values of x that $e^x - 1 = O(x)$ $(-\infty < x < \infty)$ fails to be true. So a restriction to a finite interval is indicated, and it is of little concern what interval is taken.

On the other hand, there are cases where one has some trouble to find a suitable interval. Now in order to eliminate these non-essential minor inconveniences one uses a modified O-notation, which again suppresses some information. We shall explain it for the case where the interest lies in large positive values of x $(x \to \infty)$, but by obvious modifications we get similar notations for cases like

$x \to -\infty$, $|x| \to \infty$, $x \to c$, $x \uparrow c$ (i.e., x tends to c from the left).

The formula

$$(1.2.7) \qquad f(x) = O(\varphi(x)) \qquad (x \to \infty)$$

means that there exists a real number a such that

$$f(x) = O(\varphi(x)) \qquad (a < x < \infty).$$

In other words, (1.2.7) means that there exist numbers a and A such that

$$(1.2.8) \qquad |f(x)| \leq A|\varphi(x)| \text{ whenever } a < x < \infty.$$

Examples:

$$x^2 = O(x) \ (x \to 0); \qquad x = O(x^2) \ (x \to \infty);$$
$$e^{-x} = O(1) \ (x \to \infty); \qquad (\log x)^6 = O(x^{\frac{1}{2}}) \ (x \to \infty);$$
$$(\log x)^{-1} = O(1) \ (x \to \infty); \qquad (\sin x^{-1})^{-1} = O(x) \ (x \to \infty).$$

In many cases a formula of the type (1.2.7) can be replaced immediately by an O-formula of the type (1.2.5). For if (1.2.7) is given, and if f and φ are continuous in the interval $0 \leq x < \infty$, and if moreover $\varphi(x) \neq 0$ throughout this interval, then we have $f(x) = O(\varphi(x))$ $(0 \leq x < \infty)$. This follows from the fact that f/φ is continuous, and therefore bounded, over $0 \leq x \leq a$.

The reader should notice that so far we did not define what $O(\varphi(s))$ means; we only defined the meaning of some complete formulas. It is obvious that the isolated expression $O(\varphi(x))$ cannot be defined, at least not in such a way that (1.2.5) remains equivalent to (1.2.6). For, $f(s) = O(\varphi(s))$ obviously implies $2f(s) = O(\varphi(s))$. If $O(\varphi(s))$ in itself were to denote anything, we would infer $f(s) = O(\varphi(s)) = 2f(s)$, whence $f(s) = 2f(s)$.

The trouble is, of course, due to abusing the equality sign $=$. A similar situation arises if someone, because his typewriter lacks the sign $<$, starts to write $= L$ for the words "is less than", and so writes $3 = L(5)$. On being asked: "What does $L(5)$ stand for?", he has to reply "Something that is less than 5". Consequently, he rapidly gets the habit of reading L as "something that is less than", thus coming close to the actual words we used when introducing (1.2.5). After that, he writes $L(3) = L(5)$ (something that is less than 3 is something that is less than 5), but certainly not $L(5) = L(3)$.

He will not see any harm in $4 = 2 + L(3)$, $L(3) + L(2) = L(8)$.

The O-symbol is used in exactly the same manner as this person's L-symbol. We give a few examples:

$$O(x) + O(x^2) = O(x) \qquad (x \to 0).$$

This means: for any pair of functions f, g, such that

$$f(x) = O(x) \ (x \to 0), \qquad g(x) = O(x^2) \ (x \to 0),$$

we have

$$f(x) + g(x) = O(x) \qquad (x \to 0).$$

Analogously,

$$O(x) + O(x^3) = O(x^3) \qquad (x \to \infty),$$
$$e^{O(1)} = O(1) \qquad (-\infty < x < \infty),$$
$$e^{O(x)} = e^{O(x^2)} \qquad (x \to \infty).$$

We also write things like

$$e^x = 1 + x + O(x^2) \qquad (x \to 0),$$

which means that there is a function f such that $e^x = 1 + x + f(x)$, and $f(x) = O(x^2)$ $(x \to 0)$. And we write things like

$$x^{-1}O(1) = O(1) + O(x^{-2}) \qquad (0 < x < \infty).$$

This means that for any function $f(x)$ with $f(x) = O(1)$ $(0 < x < \infty)$ we can split $x^{-1}f(x)$ into two parts $g(x)$ and $h(x)$, such that $g(x) = O(1)$, $h(x) = O(x^{-2})$ $(0 < x < \infty)$. The proof is easy: take $g(x) = 0$ if $0 < x \leq 1$, $g(x) = x^{-1}f(x)$ if $x > 1$, $h(x) = x^{-1}f(x)$ if $0 < x \leq 1$, $h(x) = 0$ if $x > 1$.

The common interpretation of all these formulas can be expressed as follows. Any expression involving the O-symbol is to be considered as a class of functions. If the range $0 < x < \infty$ is considered, then $O(1) + O(x^2)$ denotes the class of all functions of the form $f(x) + g(x)$, with $f(x) = O(1)$ $(0 < x < \infty)$, $g(x) = O(x^{-2})$ $(0 < x < \infty)$. And $x^{-1}O(1) = O(1) + O(x^{-2})$ means that the class $x^{-1}O(1)$ is contained in the class $O(1) + O(x^{-2})$. Sometimes the left-hand-side of a relation is not a class, but a single function, as in (1.2.7). Then the relation means that the function on the left is a member of the class on the right.

It is obvious that the sign $=$ is really the wrong sign for such relations, because it suggests symmetry, and there is no such symmetry. For example, $O(x) = O(x^2)$ $(x \to \infty)$ is correct, but $O(x^2) = O(x)$ $(x \to \infty)$ is false. Once this warning has been given, there is, however, not much harm in using the sign $=$, and we shall maintain it, for no other reason than that it is customary.

Let φ and ψ be functions such that $\varphi(x) = O(\psi(x))$ $(x \to \infty)$ is true and $\psi(x) = O(\varphi(x))$ $(x \to \infty)$ is false. If a third function f satisfies

(1.2.9) $$f(x) = O(\varphi(x)) \qquad (x \to \infty),$$

then obviously it also satisfies

(1.2.10) $$f(x) = O(\psi(x)) \qquad (x \to \infty).$$

If (1.2.9) is true, it is called a refinement of (1.2.10). Formula (1.2.9) is called best possible if it cannot be refined, that is, if there are positive constants a and A such that $a|\varphi(x)| \leq |f(x)| \leq A|\varphi(x)|$ from a certain value of x onwards.

For example,

$$2x + x \sin x = O(x) \qquad (x \to \infty)$$

is best possible, since the left-hand side lies between x and $3x$. Also

$$\log(e^{2x \cos x} + e^x) = O(x) \qquad (x \to \infty)$$

is best possible. If $x > 0$, the logarithm is at most $\log(e^{2x} + e^x)$, and this is less than $\log(e^{2x} + e^{2x}) = \log 2 + 2x$. On the other hand we have $e^{2x \cos x} > 0$, whence the logarithm is certainly not less than $\log e^x = x$.

If m is a positive integer, then the estimate

(1.2.11) $$e^{-x} = O(x^{-m}) \qquad (x \to \infty)$$

holds ($x^m e^{-x}$ has its maximum at $x = m$, as far as positive values of x are concerned). But for no value of m (1.2.11) is best possible, since $e^{-x} = O(x^{-m-1})$ $(x \to \infty)$ is always a refinement.

We shall now discuss the matter of uniformity. We start with an example. Let S be a set of values of x, let k be a positive number, and let $f(x)$ and $g(x)$ be arbitrary functions. Then we have

(1.2.12) $$(f(x) + g(x))^k = O((f(x))^k) + O((g(x))^k) \qquad (x \in S).$$