BIBLIOTHECA MATHEMATICA

A Series of Monographs on Pure and Applied Mathematics

Volume IV

★

Edited with the cooperation of

THE „MATHEMATISCH CENTRUM”

and

THE „WISKUNDIG GENOOTSCHAP”
at Amsterdam

★

Editors

N. G. DE BRUIJN
D. VAN DANTZIG
J. DE GROOT

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

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:

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

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])

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

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

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

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

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

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 $\{f(n)\}$:

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

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

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

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

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

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

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

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

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

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

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

Examples:

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

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

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

then obviously it also satisfies

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

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

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

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

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

For, we have

$$|(f+g)^k| \le (|f|+|g|)^k \le \{2\max(|f|,|g|)\}^k \le$$
$$\le 2^k \max(|f|^k, |g|^k) \le 2^k(|f|^k + |g|^k).$$

Formula (1.2.12) means that A and B can be found such that

$$|(f(x)+g(x))^k| \le A|f(x)|^k + B|g(x)|^k \qquad (x \in S),$$

and it should be noted that A and B depend on k, or rather, that we have not shown the existence of suitable A and B not depending on k.

On the other hand, in

$$\left(\frac{k}{x^2+k^2}\right)^k = O\left(\frac{1}{x^k}\right) \qquad (1 < x < \infty) \tag{1.2.13}$$

the constant involved in the O-symbol can be chosen independently of k ($0 < k < \infty$). For, we have $x^2 + k^2 = (x-k)^2 + 2kx \ge 2kx$, whence

$$\left(\frac{k}{x^2+k^2}\right)^k \le \frac{1}{(2x)^k}.$$

We have $2^{-k} < 1$ for all $k > 0$. It follows that there is a number A, not depending on k (viz. $A = 1$), such that

$$\left(\frac{k}{x^2+k^2}\right)^k \le \frac{A}{x^k} \qquad (1 < x < \infty,\ k > 0).$$

This fact is expressed by saying that (1.2.13) holds uniformly with respect to k.

We can also look upon (1.2.13) from a different point of view. The function $k^k(x^2+k^2)^{-k}$ is a function of the two variables x and k, and therefore it can be considered as a function of a variable point in the x-k-plane. Now the uniformity of (1.2.13) expresses the same thing as

$$\left(\frac{k}{x^2+k^2}\right)^k = O(x^{-k}) \qquad (1 < x < \infty,\ 0 < k < \infty).$$

The set S referred to in (1.2.6) specializes to the subset of the x-k-plane described by $1 < x < \infty$, $0 < k < \infty$.

Careful uniform estimates are often required in situations of the following type. We want to have an O-formula for a function f. We have some expression for $f(x)$, which we split into two parts; the way in which the splitting is made, depends on some parameter t. Estimating both parts uniformly in x and t we get, for example,

$$f(x) = O(x^{2t}) + O(x^{4t-2}) \qquad (x > 1,\ t > 1).$$

We now want to choose t such that the right-hand-side is as small as possible. As the formula holds uniformly, we may take t equal to some function of x. So the question is to minimize $x^{2t} + x^{4t-2}$, if x is given. The minimum is attained at $t = (2x^2)^{1/3}$, and then both terms have the same order, viz. $x^{8/3}$. So $f(x) = O(x^{8/3})$ $(x > 1)$.

In order to obtain this result it is not necessary to determine the exact minimum. We can argue as follows. Let t_0 be such that x^{2t} and x^{4t-2} have the same order, $t_0 = x^{2/3}$, say. This gives the optimal result, for by taking a larger value of t we increase the first term, and a smaller value of t increases the second term. In both cases the final result is worse (or at least not better) than the result with t_0.

In O-formulas involving conditions like $x \to \infty$, there are two constants involved (A and a in (1.2.8)). We shall speak of uniformity with respect to a parameter k only if both A and a can be chosen independent of k.

Example: For each individual $k > 0$ we have

$$k^2(1+kx^2)^{-1} = O(x^{-1}) \qquad (x \to \infty),$$

but this does not hold uniformly. If it did, there would be positive numbers A and a, both independent of k, such that

$$k^2(1+kx^2)^{-1} < Ax^{-1} \qquad (x > a,\ k > 0).$$

If we put $k = x^2$, we get $A(1+x^4) > x^5$ whenever $x > a$, which is impossible.

On the other hand, *one* of the two constants A and a can be chosen independent of k. We can take $a = k$, $A = 1$, as

$$k^2(1+kx^2)^{-1} < kx^{-2} < x^{-1} \qquad (x > k,\ k > 0).$$

We can also take $a = 1$, $A = k$, since

$$k^2(1+kx^2)^{-1} < kx^{-2} < kx^{-1} \qquad (x > 1,\ k > 0).$$

1.3. The o-symbol

The expression

(1.3.1) $$f(x) = o(\varphi(x)) \qquad (x \to \infty)$$

means that $f(x)/\varphi(x)$ tends to 0 when $x \to \infty$. This is a stronger assertion than the corresponding O-formula: (1.3.1) implies (1.2.7), as convergence implies boundedness from a certain point onwards.

Furthermore we adopt the same conventions we introduced for the O-symbol: = is to be read as "is", and "o" is to be read as "something that tends to zero, multiplied by". Some examples are

$$\begin{aligned}
\cos x &= 1 + o(x) && (x \to 0),\\
e^{o(x)} &= 1 + o(x) && (x \to 0),\\
n! &= e^{-n}n^n\sqrt{2\pi n}(1 + o(1)) && (n \to \infty),\\
n! &= e^{-n+o(1)}n^n\sqrt{2\pi n} && (n \to \infty),\\
o(f(x)\,g(x)) &= o(f(x))\,O(g(x)) && (x \to 0),\\
o(f(x)\,g(x)) &= f(x)\,o(g(x)) && (x \to 0).
\end{aligned}$$

In asymptotics, o's are less important than O's, because they hide so much information. If something tends to zero, we usually wish to know how rapid the convergence is.

1.4. Asymptotic equivalence

We say that $f(x)$ and $g(x)$ are asymptotically equivalent as $x \to \infty$, if the quotient $f(x)/g(x)$ tends to unity. Our notation is

$$f(x) \sim g(x) \qquad (x \to \infty).$$

The notation will also be used for all other ways of passing to a limit (e.g., $x \to -\infty$, $x \to 0$, $x \downarrow 0$, $|z| \to 0$).

Properly speaking, the symbol $\sim$ is superfluous, as $f(x) \sim g(x)$ can be conveniently written as $f(x) = g(x)\,(1 + o(1))$, or as $f(x) = e^{o(1)}g(x)$.

Examples:

$$\begin{aligned}
x + 1 &\sim x && (x \to \infty),\\
\sinh x &\sim \tfrac{1}{2}e^x && (x \to \infty),\\
n! &\sim e^{-n}n^n\sqrt{2\pi n} && (n \to \infty) && (\text{cf.}(1.1.1)),\\
\pi(x) &\sim x/(\log x) && (x \to \infty) && (\text{cf.}(1.1.2)).
\end{aligned}$$

When asking for the "asymptotic behaviour" of a given function $f(x)$ as $x \to \infty$, say, one means to ask for asymptotic information of any kind. But usually it means asking for a simple function $g(x)$ which is asymptotically equivalent to $f(x)$. Here "simple" means that its explicit evaluation does *not* become extremely hard if x is very large. From a certain point of view $n!$ is much simpler than $e^{-n}n^n(2\pi n)^{\frac{1}{2}}$, but from the asymptotic point of view the latter expression is the simpler.

The words "asymptotic formula for $f(x)$" are, accordingly, usually taken in the same restricted sense, viz. of an equivalence formula $f(x) \sim g(x)$.

1.5. Asymptotic series

We often have the situation that for a function $f(x)$, as $x \to \infty$, say, we have an infinite sequence of O-formulas, each $(n+1)$-th formula being an improvement on the n-th. Frequently the sequence of formulas is of the following type. There is a sequence of functions $\varphi_0, \varphi_1, \varphi_2, \ldots$, satisfying

(1.5.1) $$\varphi_1(x) = o(\varphi_0(x)) \quad (x \to \infty), \quad \varphi_2(x) = o(\varphi_1(x)) \quad (x \to \infty), \ldots,$$

and there is a sequence of constants $c_0, c_1, c_2, \ldots$, such that there is the following sequence of O-formulas for f:

(1.5.2) $$\left\{\begin{array}{ll}
f(x) = O(\varphi_0(x)) & (x \to \infty)\\
f(x) = c_0\varphi_0(x) + O(\varphi_1(x)) & (x \to \infty)\\
f(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + O(\varphi_2(x)) & (x \to \infty)\\
\cdots\cdots\cdots\cdots\cdots\cdots & \\
f(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \ldots + c_{n-1}\varphi_{n-1}(x) + O(\varphi_n(x)) & (x \to \infty)\\
\cdots\cdots\cdots\cdots\cdots\cdots &
\end{array}\right.$$

Obviously, the second formula improves the first one, as

$$c_0\varphi_0(x) + O(\varphi_1(x)) = (c_0 + o(1))\varphi_0(x) = O(\varphi_0(x)) \qquad (x \to \infty).$$

Accordingly, the third formula improves the second one, and so on.

The following notation is used in order to represent the whole set (1.5.2) by a single formula:

(1.5.3) $$f(x) \approx c_0\varphi_0(x) + c_1\varphi_1(x) + c_2\varphi_2(x) + \ldots \qquad (x \to \infty).$$

The right-hand-side is called an *asymptotic series* for $f(x)$, or an *asymptotic expansion* for $f(x)$. The notion is due to Poincaré.

It is easy to see that the c's are uniquely determined when the φ's are given, assuming that such an asymptotic expansion exists. For, assume that (1.5.3) holds, that

$$f(x) \approx d_0\varphi_0(x) + d_1\varphi_1(x) + d_2\varphi_2(x) + \dots \quad (x \to \infty)$$

is also true, and that k is the smallest number for which $c_k \neq d_k$. Then we have, by subtraction

$$0 = (c_k - d_k)\varphi_k(x) + O(\varphi_{k+1}(x)).$$

Dividing by $c_k - d_k$ we find that $\varphi_k(x) = O(\varphi_{k+1}(x))$, which contradicts the fact that $\varphi_{k+1}(x) = o(\varphi_k(x))$.

It can happen that, in (1.5.2), the coefficients $c_0, c_1, c_2, \dots$ are all equal to zero. Then it is conventional to write

$$f(x) \approx 0\cdot\varphi_0(x) + 0\cdot\varphi_1(x) + 0\cdot\varphi_2(x) + \dots \quad (x \to \infty).$$

It means that $f(x) = O(\varphi_n(x))$ $(x \to \infty)$ for all n (but not necessarily uniformly with respect to n). For example, since $e^{-x} = O(x^{-n})$ $(x \to \infty)$ for all n, we write

$$e^{-x} \approx 0\cdot 1 + 0\cdot x^{-1} + 0\cdot x^{-2} + \dots \quad (x \to \infty). \tag{1.5.4}$$

The series occurring in (1.5.3) need not be convergent. At first sight it seems strange that such a sequence, producing sharper and sharper approximations, does not automatically converge. The answer is that convergence means something for some fixed x_0, whereas the O-formulas (1.5.2) are not concerned with $x = x_0$, but with $x \to \infty$. Convergence of the series, for all $x > 0$, say, means that for every individual x there is a statement about the case $n \to \infty$. On the other hand, the statement that the series is the asymptotic expansion of $f(x)$ means that for every individual n there is a statement about the case $x \to \infty$.

Moreover, if the sequence converges, its sum need not be equal to $f(x)$: formula (1.5.4) provides a counterexample. It is even possible to construct functions $f(x), \varphi_0(x), \varphi_1(x), \dots$, such that the series of (1.5.3) converges for all x, but such that the sum of the series does not have the series itself as its own asymptotic series.

A quite simple example of a divergent asymptotic series is the

following one. We consider the function f, defined by

$$f(x) = \int_1^x \frac{e^t}{t}\,dt \tag{1.5.5}$$

(apart from an additional constant this is the so called exponential integral Ei t). Integrating by parts we obtain

$$f(x) = \left[\frac{e^t}{t}\right]_1^x + \int_1^x \frac{e^t}{t^2}\,dt, \tag{1.5.6}$$

where we use the notation $[\varphi(t)]_a^b = \varphi(b) - \varphi(a)$. The first term in (1.5.6) is $x^{-1}e^x - e$, but the second one is of smaller order. Splitting the integral into two parts, we obtain

$$\int_1^{\frac12 x} t^{-2}e^t dt < \int_1^{\frac12 x} e^t dt < e^{\frac12 x},$$

$$\int_{\frac12 x}^{x} t^{-2}e^t dt < \int_{\frac12 x}^{x} (\tfrac12 x)^{-2}e^t dt < 4x^{-2}e^x.$$

Since the $-e$, $e^{\frac12 x}$ and $4x^{-2}e^x$ are all $O(x^{-2}e^x)$, we obtain

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

We can improve the approximation by repeating the procedure. Integrating the integral in (1.5.6) by parts, we get

$$f(x) = \left[\frac{e^t}{t}\right]_1^x + \left[\frac{e^t}{t^2}\right]_1^x + \int_1^x \frac{2e^t}{t^3}\,dt =$$

$$= \left[\frac{e^t}{t}\right]_1^x + \left[\frac{e^t}{t^2}\right]_1^x + \left[\frac{2e^t}{t^3}\right]_1^x + \int_1^x \frac{3!e^t}{t^4}\,dt,$$

and generally $(n = 1, 2, 3, \dots)$

$$f(x) = \left[e^t\left(\frac{1}{t} + \frac{1!}{t^2} + \frac{2!}{t^3} + \dots + \frac{(n-1)!}{t^n}\right)\right]_1^x + n!\int_1^x \frac{e^t}{t^{n+1}}\,dt.$$

The last integral is $O(x^{-n-1}e^x)$ if $x \to \infty$ and if n is fixed. This can again be verified by splitting the integration interval into two parts, viz. $(1, \frac{1}{2}x)$ and $(\frac{1}{2}x, x)$. So for each n, we have

$$e^{-x}f(x) = \frac{1}{x} + \frac{1}{x^2} + \frac{2!}{x^3} + \ldots + \frac{(n-1)!}{x^n} + O\left(\frac{1}{x^{n+1}}\right) \quad (x \to \infty),$$

and it follows that

$$e^{-x}f(x) \approx \frac{1}{x} + \frac{1}{x^2} + \frac{2!}{x^3} + \frac{3!}{x^4} + \ldots \qquad (x \to \infty).$$

The series on the right converges for no single value of x.

A simple, though rather trivial, class of asymptotic series consists of the class of convergent power series. Suppose $f(z)$ is the sum of a convergent power series, say

$$(1.5.7) \qquad f(z) = a_0 + a_1 z + a_2 z^2 + \ldots$$

when $|z| \leq p$, p being any positive number less than the radius of convergence. Then we have

$$(1.5.8) \qquad f(z) \approx a_0 + a_1 z + a_2 z^2 + \ldots \qquad (|z| \to 0).$$

The proof is easy. The series converges at $z = p$, whence the terms $a_n p^n$ are bounded, $|a_n p^n| < A$, say, for all n.

Now for each individual n we have, when $|z| = \frac{1}{2}p$,

$$\left| \sum_{k=n+1}^{\infty} a_k z^k \right| \leq |z|^{n+1} A p^{-n-1} (1 + \tfrac{1}{2} + \tfrac{1}{4} + \ldots),$$

and therefore

$$f(z) = a_0 + a_1 z + \ldots + a_n z^n + O(z^{n+1}) \qquad (|z| < \tfrac{1}{2}p).$$

This implies (1.5.8).

It obviously does not matter whether in this discussion z represents a complex variable, or a real variable, or a real positive variable.

1.6. Elementary operations on asymptotic series

For the sake of simplicity we shall restrict our discussion to asymptotic series of the form

$$(1.6.1) \qquad a_0 + a_1 x + a_2 x^2 + \ldots \qquad (x \to 0),$$

although similar things can be done for several other types.

The series (1.6.1) is a power series (in terms of powers of x), and as long as there is no discussion about its representing anything, we call it a formal power series.

If for these formal power series addition and multiplication are defined in the obvious way, then the set of all formal power series becomes a commutative ring, with $1 + 0 \cdot x + 0 \cdot x^2 + \ldots$ as its unit element (to be denoted by I). If the series $a_0 + a_1 x + \ldots$ and $b_0 + b_1 x + \ldots$ are respresented by A and B, respectively, we define sum and product by

$$A + B = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots,$$
$$AB = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \ldots.$$

If $a_0 \neq 0$, then there is a uniquely determined series C such that $AC = I$. Its coefficients $c_0, c_1, c_2, \ldots$ can be evaluated recursively from the equations

$$a_0 c_0 = 1, \quad a_0 c_1 + a_1 c_0 = 0, \quad a_0 c_2 + a_1 c_1 + a_2 c_0 = 0, \quad \ldots.$$

Furthermore we can define the formal power series that arises from substituting the series B into the series A, provided that $b_0 = 0$. This new series will be denoted by $A(B)$ [1]). It is defined as follows: Let c_{kn} be the coefficient of x^k in the series $a_0 I + a_1 B + a_2 B^2 + \ldots + a_n B^n$. Then it is easily seen that $c_{kk} = c_{k,k+1} = c_{k,k+2} = \ldots$. Writing $c_{kk} = c_k$, we infer that

$$a_0 I + a_1 B + \ldots + a_n B^n =$$
$$= c_0 + c_1 x + \ldots + c_n x^n + c_{n+1,n} x^{n+1} + c_{n+2,n} x^{n+2} + \ldots.$$

We now define

$$A(B) = c_0 + c_1 x + \ldots + c_n x^n + c_{n+1} x^{n+1} + c_{n+2} x^{n+2} + \ldots.$$

So $A(B)$ arises from replacing x in the A-series by B, and combining coefficients afterwards.

A further operation on formal power series is differentiation. The derivative of $A = a_0 + a_1 x + \ldots$ is defined by

$$A' = a_1 + 2a_2 x + 3a_3 x^2 + \ldots$$

that is, by formal term-by-term differentiation.

---

[1]) Note the difference between the notations $A(B)$ and AB. A formula $A(B) = C$ refers to a formal identity $A\{B(x)\} = C(x)$, whereas $AB = D$ simply means that $D(x)$ is the formal product $A(x)B(x)$.

It is well known that if A and B are power series with a positive radius of convergence, these formal operations directly correspond to the same operations on the sums $A(x)$ and $B(x)$ of those series. For example, if $A(B) = C$, then the series C has a positive radius of convergence, and inside the circle of that radius we have $A\{B(x)\} = C(x)$.

When speaking about asymptotic series instead of power series, we have a similar situation, apart from the fact that some extra care is necessary in the case of differentiation. Assume that $A(x)$ and $B(x)$ are functions defined in a neighbourhood of $x = 0$, having asymptotic expansions

$$A(x) \approx A \quad (x \to 0), \qquad B(x) \approx B \quad (x \to 0).$$

Note that $A(x)$ stands for the function, and that A stands for the formal series $a_0 + a_1x + \ldots$. It was already remarked in sec. 1.5 that the coefficients of the series A are uniquely determined by the function $A(x)$, provided that $A(x)$ has an asymptotic series.

Now it is not difficult to show that

$$(1.6.2) \qquad A(x) + B(x) \approx A + B \qquad (x \to 0),$$

$$(1.6.3) \qquad A(x)B(x) \approx AB \qquad (x \to 0),$$

and if $a_0 \neq 0$,

$$(1.6.4) \qquad \{A(x)\}^{-1} \approx A^{-1} \qquad (x \to 0)$$

(A^{-1} stands for the solution of $A^{-1} \cdot A = I$). Furthermore, if $b_0 = 0$, the composite function $A(B(x))$ is defined for all sufficiently small values of x, and

$$(1.6.5) \qquad A\{B(x)\} \approx A(B) \qquad (x \to 0).$$

Formula (1.6.2) is trivial. We shall prove (1.6.3). Writing $AB = C$, we have, for each n,

$$\begin{aligned} A(x) &= a_0 + \ldots + a_nx^n + O(x^{n+1}) \qquad (x > 0), \\ B(x) &= b_0 + \ldots + b_nx^n + O(x^{n+1}) \qquad (x > 0), \end{aligned}$$

and so

$$A(x)B(x) = (a_0 + \ldots + b_nx^n)\,(b_0 + \ldots + b_nx^n) + O(x^{n+1}) \;(x \to 0).$$

Now

$$(a_0 + \ldots + a_nx^n)\,(b_0 + \ldots + b_nx^n) - (c_0 + \ldots + c_nx^n)$$

is a linear combination of $x^{n+1}, x^{n+2}, \ldots, x^{2n}$, and so it is $O(x^{n+1})$ if $x \to 0$. It follows that

$$A(x)B(x) = c_0 + \ldots + c_nx^n + O(x^{n+1}) \qquad (x \to 0),$$

and this proves (1.6.3).

Similar proofs can be given for (1.6.4) and (1.6.5). Actually, (1.6.4) can be considered as a special case of (1.6.5), as $A^{-1} = P(Q)$, with $P = a_0^{-1}(1 + x + x^2 + \ldots)$, $Q = a_0^{-1}(a_0 - A)$.

Suppose that $f(x)$ satisfies

$$f(x) \approx a_0 + a_1x + a_2x^2 + \ldots \qquad (x \to 0),$$

and that $\int_0^x f(t)dt$ exists for all sufficiently small values of x. Then term-by-term integration is legitimate:

$$(1.6.6) \qquad \int_0^x f(t)dt \approx a_0x + \tfrac{1}{2}a_1x^2 + \tfrac{1}{3}a_2x^3 + \ldots \qquad (x \to 0).$$

This is easily proved. If n is given, then there are constants A and a, such that

$$|f(t) - a_0 - a_1t - \ldots - a_{n-1}t^{n-1}| < A|t|^n \qquad (|t| < a),$$

so if $|x| < a$, we have

$$\left| \int_0^x f(t)dt - a_0x - \tfrac{1}{2}a_1x^2 - \ldots - \frac{1}{n}\,a_nx^n \right| < \frac{A}{n+1}\,|x|^{n+1},$$

and (1.6.6) follows.

When we consider differentiation, the situation appears to be different. If $A(x)$ has the asymptotic development $A(x) \sim A$ $(x \to 0)$, then $A'(x)$ does not necessarily exist. If it exists, it does not necessarily have an asymptotic expansion.

For example, we have

$$e^{-1/x}\sin(e^{1/x}) \approx 0 + 0 \cdot x + 0 \cdot x^2 + \ldots \qquad (x > 0,\ x \to 0),$$

but the derivative

$$x^{-2}e^{-1/x}\sin(e^{1/x}) - x^{-2}\cos(e^{1/x})$$

has no such asymptotic expansion.

Term-by-term differentiation of an asymptotic power series is legitimate, however, if it can be shown that the derivative of the function also has an asymptotic expansion (in the form of a formal power series). Assuming that

$$f(x) \approx a_0 + a_1x + a_2x^2 + \dots \qquad (x \to 0),$$
$$f'(x) \approx b_0 + b_1x + b_2x^2 + \dots \qquad (x \to 0),$$

we shall prove that $b_k = (k+1)a_{k+1}$ $(k = 0, 1, 2, \dots)$. Considering, for some $n \geq 1$,

$$g_n(x) = f(x) - (b_0x + \tfrac{1}{2}b_1x^2 + \dots + n^{-1}\, b_{n-1}x^n),$$

we have $g_n'(x) = O(x^n)$ $(x \to 0)$. From the mean value theorem of the differential calculus it follows that $g_n(x) - g_n(0) = O(x^{n+1})$. Since n is arbitrary, we infer that

$$f(x) \approx f(0) + b_0x + \tfrac{1}{2}b_1x^2 + \tfrac{1}{3}b_2x^3 + \dots \qquad (x \to 0).$$

The formula $b_k = (k+1)a_{k+1}$ now follows from the fact that the coefficients in an asymptotic series are uniquely determined.

1.7. Asymptotics and Numerical Analysis

The object of asymptotics is to derive O-formulas and o-formulas for functions in cases where it is difficult to apply the definition of the function for very large (or for very small) values of the variable. It even occurs that the definition of a function is so difficult, even for "normal" values of the variable, that it is easier to find asymptotic information than any other type of information.

As was already stressed in sec. 1.1, neither O-formulas nor o-formulas have, as they stand, any direct value for numerical purposes. However, in almost all cases where such formulas have been derived, it is possible to retrace the proof, replacing all O-formulas by definite estimates involving explicit numerical constants.

That is, at every stage of the procedure we indicate definite numbers or functions with certain properties, where the asymptotic formulas only stated the existence of such numbers or functions.

In most cases, the final estimates obtained in this way are rather weak, with constants a thousand times, say, greater than they could be. The reason is, of course, that such estimates are obtained by means of a considerable number of steps, and in each step a factor 2 or so is easily lost. Quite often it is possible to reduce such errors by a more careful examination.

But even if the asymptotic result is presented in its best possible explicit form, it need not be satisfactory from the numerical point of view. The following dialogue between Miss N.A., a Numerical Analyst, and Dr A.A., an Asymptotic Analyst, is typical in several respects.

N.A.: I want to evaluate my function $f(x)$ for large values of x, with a relative error of at most 1%.
A.A.: $f(x) = x^{-1} + O(x^{-2}) \qquad (x \to \infty)$.
N.A.: I am sorry, but I don't understand.
A.A.: $|f(x) - x^{-1}| < 8x^{-2} \qquad (x > 10^4)$.
N.A.: But my value of x is only 100.
A.A.: Why did not you say so? My evaluations give

$$|f(x) - x^{-1}| < 57000x^{-2} \qquad (x \geq 100).$$

N.A.: This is no news to me. I know already that $0 < f(100) < 1$.
A.A.: I can gain a little on some of my estimates. Now I find that

$$|f(x) - x^{-1}| < 20x^{-2} \qquad (x \geq 100).$$

N.A.: I asked for 1%, not for 20%.
A.A.: It is almost the best thing I possibly can get. Why don't you take larger values of x?
N.A.: !!!. I think it's better to ask my electronic computing machine.
Machine: $f(100) = 0{\cdot}01137\ 42259\ 34008\ 67153$.
A.A.: Haven't I told you so? My estimate of 20% was not very far from the 14% of the real error.
N.A.: !!!...!.

Some days later, Miss N.A. wants to know the value of $f(1000)$. She now asks her machine first, and notices that it will require a month, working at top speed. Therefore, she returns to her Asymptotic Colleague, and gets a fully satisfactory reply.

1.8. Exercises. 1. Show that

$$\int_1^x (1 + t^{-1})^t dt = ex - \tfrac{1}{2}e \log x + O(1) \qquad (x > 1).$$

(Hint: First show that $e^{-1}(1 + t^{-1})^t = 1 - \tfrac{1}{2}t^{-1} + O(t^{-2}) \quad (t \geq 1)$).

2. Show that

$$(x + 1 + O(x^{-1}))^x = ex^x + O(x^{x-1}) \qquad (x \to \infty).$$

3. Prove that, for each n ($n = 1, 2, 3, \ldots$),

$$\int_0^x (\log y)^n dy = O(x(\log x)^n) \qquad (x \to 0).$$

4. Show that

$$t\int_2^\infty e^{-xt}(\log x)^{-1}dx \approx \sum_0^\infty c_n(\log t)^{-n-1} \qquad (t > 0,\ t \to 0),$$

with the coefficients $c_n = -\int_0^\infty e^{-y}(\log y)^n dy$.
(Hint: Use $\int_0^{2t} e^{-y}(\log y)^n dy = O(\int_0^{2t} (\log y)^n dy)$, and apply the result of the previous exercise to this integral).

5. Prove that

$$(xe^{2(x-n)})^n = O(e^{x^2+x}) \qquad (x > 0)$$

holds uniformly with respect to n ($n = 1, 2, 3, \ldots$).
(Hint: Determine the maximum of $(xe^{2(x-n)})^n e^{-x-x^2}$ for $x > 0$, if n is fixed).

6. Prove the following uniform estimate:

$$\sum_{1\le n\le x} \left|\int_0^u e^{i(n-v)t}dt\right| = O(\log(1 + ux)) + O(u) \quad (x \ge 1, u > 0, -\infty < v < \infty).$$

The summation runs over all integers n with $1 \le n \le x$; x need not be an integer.
(Hint: $|\int_0^u e^{ipt}dt| \le \min(u, 2|p|^{-1})$ if p is real, and $u > 0$).

CHAPTER 2

IMPLICIT FUNCTIONS

2.1. Introduction

Let x be given as a function of t by some equation

$$f(x, t) = 0,$$

where, if the equation has more than one root, it is somehow indicated, for each value of t, which one of the roots has to be chosen. Let this root be denoted by $x = \varphi(t)$. The problem is to determine the asymptotic behaviour of $\varphi(t)$ as $t \to \infty$.

We shall only discuss a few examples, since little can be said in general. In general, the question is rather vague, for what we really want is the asymptotic behaviour of $\varphi(t)$ expressed in terms of elementary functions, or at least in terms of explicit functions. In this respect it is, of course, essential what functions are considered as elementary. If no one had ever introduced logarithms, the question about the asymptotic behaviour of the positive solution of the equation $e^x - t = 0$ (as $t \to \infty$) would have been a hopeless problem. But as soon as one considers logarithms as useful functions, the problem vanishes entirely.

In many cases occurring in practice, it is possible to express the asymptotic behaviour of an implicit function in terms of elementary functions. For the sake of curiosity we mention one case where it is quite unlikely that such an elementary expression exists, although it may be difficult to show the contrary. If x is given by

$$x(\log x)^t - t^{2t} = 0,\ x > 1,$$

then we can easily verify that $x = e^{t\varphi(t)}$, where $\varphi(t)$ is the solution of $\varphi e^\varphi = t$. When $t \to \infty$ we have for φ an asymptotic expansion (see sec. 2.4), which involves errors of the type $(\log t)^{-k}$, for k arbitrary but fixed. This means that we have an asymptotic formula for log x, but not for x itself. That is, we do not posses an elementary function $\psi(t)$ such that $x/\psi(t)$ tends to 1 as $t \to \infty$. This would require a formula for $\varphi(t)$ with an error term of $o(t^{-1})$, and it is unlikely that such a formula could ever be found.

In most cases where asymptotic formulas can be obtained, it turns out to be quite easy. Usually it depends on expansions in terms of some small parameter, ordinarily in connection with the Lagrange inversion formula. That formula belongs to complex function theory, but the same results can often be obtained by real function methods. Quite often iteration methods can be applied, but sometimes they fail in a peculiar way (see sec. 2.7).

2.2. The Lagrange inversion formula [1]

Let the function $f(z)$ be analytic in some neighbourhood of the point $z = 0$ of the complex plane. Assuming that $f(0) \neq 0$, we consider the equation

(2.2.1) $$w = z/f(z),$$

where z is the unknown. Then there exist positive numbers a and b, such that for $|w| < a$ the equation has just one solution in the domain $|z| < b$, and this solution is an analytic function of w:

(2.2.2) $$z = \sum_{k=1}^{\infty} c_k w^k \qquad (|w| < a),$$

where the coefficients c_k are given by

(2.2.3) $$c_k = \frac{1}{k!}\left\{\left(\frac{d}{dz}\right)^{k-1} (f(z))^k\right\}_{z=0}.$$

A generalization gives the value of $g(z)$, where g is any function of z, analytic in a neighbourhood of $z = 0$:

(2.2.4) $$g(z) = g(0) + \sum_{1}^{\infty} d_k w^k,$$

$$d_k = (k!)^{-1}(d/dz)^{k-1}\{g'(z)\,(f(z))^k\}_{z=0}\,.$$

Formula (2.2.2), usually quoted as the Lagrange inversion formula, is a special case of a more general theorem on implicit func-

[1] See E. T. Whittaker and G. N. Watson, Modern Analysis, 4th ed., Cambridge 1946, § 7.32.

tions [1]: If $f(z, w)$ is an analytic function of both z and w, in some region $|z| < a_1$, $|w| < b_1$, if $f(0, 0) = 0$ and if $\partial f/\partial z$ does not vanish at the point $z = w = 0$, then there are positive numbers a and b, such that, for each w in the domain $|w| < a$, the equation $f(z, w) = 0$ has just one solution z in the domain $|z| < b$, and this solution can be represented as a power series $z = \sum_{k=1}^{\infty} c_k w^k$.

We shall not reproduce proofs of these theorems here, although in sec. 2.4 (see (2.4.6)) a special case of a slightly more general problem will be treated in detail.

2.3. Applications

Some asymptotic problems on implicit functions admit a direct application of the Lagrange formula. For example, consider the positive solution of the equation

(2.3.1) $$xe^x = t^{-1},$$

when $t \to \infty$. As t^{-1} tends to zero, we apply the Lagrange formula (2.2.2) to the equation $ze^z = w$, so that $f(z) = e^{-z}$. It results that there are constants $a > 0$ and $b > 0$, such that for $|w| < a$ there is only one solution z satisfying $|z| < b$, viz.

$$z = \sum_{k=1}^{\infty} (-1)^{k-1} k^{k-1} w^k / k!$$

(actually, the series converges if $|w| < e^{-1}$). So it is clear that if $t > a^{-1}$, there is one and only one solution in the circle $|x| < b$. But as xe^x increases from 0 to ∞ if x increases from 0 to ∞, the equation (2.3.1) has a positive solution, and this one cannot exceed b if t is sufficiently large. So if t is large enough, the positive solution we are looking for, is given by

(2.3.2) $$x = \sum_{k=1}^{\infty} (-1)^{k-1} k^{k-1} t^{-k} / k!,$$

and this power series also serves as asymptotic development (see sec. 1.5).

Our second example considers the positive solution of

(2.3.3) $$x^t = e^{-x}$$

[1] See L. Bieberbach, Funktionentheorie, vol. 1, 3e Aufl., Leipzig-Berlin 1930, Kap. VII § 8.

when $t \to \infty$. The function x^t is increasing if $x > 0$, and e^{-x} is decreasing. We notice that x^t is small in the interval $0 \leq x \leq 1$ unless x is very close to 1, so that it is clear from the graphs of x^t and e^{-x} that there is just one positive root, close to 1, and tending to 1 as $t \to \infty$.

We now put $x = 1 + z$, $t^{-1} = w$, and try to get an equation of the form (2.2.1). From $x^t = e^{-x}$ we obtain the equation

$$z/f(z) = w, \text{ where } f(z) = -z(1+z)/(\log(1+z)).$$

The function $f(z)$ is analytic at $z = 0$: $f(z) = -1 + c_1 z + \ldots$. It follows that

$$x = 1 - t^{-1} - c_1 t^{-2} + \ldots$$

solves the equation (2.3.3), if t is large enough. As in the previous example, the fact that there is just one positive solution, tending to unity if $t \to \infty$, guarantees that the positive solution is represented by the power series, if t is sufficiently large.

Our third example is stated in a somewhat different form. Consider the equation

(2.3.4) $$\cos x = x \sin x.$$

We observe from the graphs of the functions x and $\operatorname{cotg} x$, that there is just one root in every interval $n\pi < x < (n+1)\pi$ ($n = 0, \pm 1, \pm 2, \ldots$). Denoting this root by x_n, we ask for the behaviour of x_n as $n \to \infty$. As $\operatorname{cotg}(x_n - \pi n) = x_n \to \infty$, we have $x_n - \pi n \to 0$. Putting $x = \pi n + z$, $(\pi n)^{-1} = w$, we find $\cos z = (w^{-1} + z)\sin z$, and so

$$w = z/f(z), \quad f(z) = z(\cos z - z \sin z)/\sin z,$$

where $f(z)$ is analytic at $z = 0$, and $f(0) = 1$. Therefore z is a power series in terms of powers of w, and we infer that $z = w + c_2 w^2 + c_3 w^3 + \ldots$. Consequently we have, if n is large enough,

$$x_n = \pi n + (\pi n)^{-1} + c_2(\pi n)^{-2} + \ldots.$$

As a consequence of the fact that $f(z)$ is an even function of z, we notice that $c_2 = c_4 = c_6 = \ldots = 0$.

2.4. A more difficult case

We take the equation

(2.4.1) $$xe^x = t$$

which has, when $t > 0$, just one solution $x > 0$, as the function xe^x increases from 0 to ∞ when x increases from 0 to ∞. This solution being simply denoted by x, we ask for the behaviour of x as $t \to \infty$.

It is now more difficult than in the previous examples to transform the equation into the Lagrange type. We shall proceed by an iterative method. We write (2.4.1) in the form

(2.4.2) $$x = \log t - \log x.$$

Once we have some approximation to x, we can substitute it on the right-hand-side of (2.4.2), and we obtain a new approximation, better than the former. Note that an error Δ in the value of x gives an error of about Δ/x in the value of $\log x$.

We must have something to start with. As t tends to infinity, we may assume $t > e$, and then we have $x > 1$. For, $0 < x \leq 1$ would imply $\log t - \log x \geq \log t > \log e = 1$, whereas the left-hand-side of (2.4.2) would be ≤ 1.

From $x > 1$ it follows, by (2.4.2), that $x = \log t - \log x < \log t$. So we start with

$$1 < x < \log t.$$

It follows that $\log x = O(\log \log t)$, and so, by (2.4.2),

$$x = \log t + O(\log \log t).$$

The formula refers to $t \to \infty$, and the same thing holds for all other O-formulas in this section.

Taking logarithms, we infer that

$$\log x = \log \log t + \log\{1 + O(\log \log t/\log t)\} =$$
$$= \log \log t + O(\log \log t/\log t).$$

Inserting this into (2.4.2), we get a second approximation

(2.4.3) $$x = \log t - \log \log t + O(\log \log t/\log t).$$

Again taking logarithms here, and inserting the result into (2.4.2), we get the third approximation

$$x = \log t - \log\{\log t - \log\log t + O(\log\log t/\log t)\} =$$
$$= \log t - \log\log t + \frac{\log\log t}{\log t} + \tfrac{1}{2}\left(\frac{\log\log t}{\log t}\right)^2 + O\left(\frac{\log\log t}{(\log t)^2}\right).$$

We shall carry out two further steps. Abbreviating

$$\log t = L_1, \qquad \log\log t = L_2,$$

we obtain

$$\log x = L_2 + \log\{1 - L_2L_1^{-1} + L_2L_1^{-2} + \tfrac{1}{2}L_2^2L_1^{-3} + O(L_2L_1^{-3})\},$$

and so, the term $O(L_2L_1^{-3})$ absorbing all terms $L_2^pL_1^{-q}$ with $q > 3$,

$$\begin{aligned} x = L_1 - L_2 - \{&- L_2L_1^{-1} + L_2L_1^{-2} + \tfrac{1}{2}L_2^2L_1^{-3} + O(L_2L_1^{-3})\} + \\ &+ \tfrac{1}{2}\{- L_2L_1^{-1} + L_2L_1^{-2}\}^2 - \tfrac{1}{3}(L_2L_1^{-1})^3 = \\ &= L_1 - L_2 + L_2L_1^{-1} + \{\tfrac{1}{2}L_2^2 - L_2\}L_1^{-2} + \\ &\quad + \{- \tfrac{1}{3}L_2^3 - \tfrac{3}{2}L_2^2 + O(L_2)\}L_1^{-3}. \end{aligned}$$

The next step can be verified to give

$$\begin{aligned} (2.4.4) \qquad x = L_1 - L_2 + L_2L_1^{-1} &+ \{\tfrac{1}{2}L_2^2 - L_2\}L_1^{-2} + \\ &+ \{- \tfrac{1}{3}L_2^3 - \tfrac{3}{2}L_2^2 + L_2\}L_1^{-3} + \\ &+ \{\tfrac{1}{4}L_2^4 - \tfrac{11}{6}L_2^3 + 3L_2^2 + O(L_2)\}L_1^{-4}. \end{aligned}$$

From these formulas we get the impression that there is an asymptotic series

$$\begin{aligned} (2.4.5) \qquad x \approx L_1 - L_2 &+ L_2P_0(L_2)L_1^{-1} + \\ &+ L_2P_1(L_2)L_1^{-2} + L_2P_2(L_2)L_1^{-3} + \dots, \end{aligned}$$

where $P_k(L_2)$ is a polynomial of degree k ($k = 0, 1, 2, \dots$). This can be proved to be the case, by a careful investigation of the process which led to (2.4.4) and to further approximations of that type. We shall not do this here, as we can show, by a different method, a much stronger assertion: if t is sufficiently large, x can be represented as the sum of a convergent series of this type.

We shall need Rouché's theorem [1]), which reads as follows. Let the bounded domain D have as its boundary a closed Jordan curve C. Let the functions $f(z)$ and $g(z)$ be analytic both in D and on C, and assume that $|f(z)| < |g(z)|$ on C (so automatically $g(z) \neq 0$ on C).

[1]) See E. C. Titchmarsh, Theory of Functions, 2nd ed., Oxford 1939, § 3.42.

Then $f(z) + g(z)$ has in D the same number of zeros as $g(z)$, all zeros counted according to their multiplicity. A proof can be given as follows. It has to be shown that the argument of $f + g$ increases by the same multiple of 2π as the argument of g if z runs through C. This follows from the fact that the real part of $(f + g)/g$ is positive on C (since $|f| < |g|$), whence the argument of $(f + g)/g$ cannot increase by any of the values $\pm 2\pi, \pm 4\pi, \dots$.

Our method for dealing with (2.4.2) is modelled after the usual proof of the Lagrange theorem. For abbreviation, we put

$$x = \log t - \log\log t + v,$$
$$(\log t)^{-1} = \sigma, \quad (\log\log t)/\log t = \tau$$

and we obtain from (2.4.2) that

$$(2.4.6) \qquad e^{-v} - 1 - \sigma v + \tau = 0.$$

For the time being, we ignore the relation that exists between σ and τ, and we shall consider them as small independent complex parameters. We shall show that there exist positive numbers a and b, such that, if $|\sigma| < a$, $|\tau| < a$, the equation (2.4.6) has just one solution in the domain $|v| < b$, and that this solution is an analytic function of both σ and τ in the region $|\sigma| < a$, $|\tau| < a$.

Let δ be the lower bound of $|e^{-z} - 1|$ on the circle $|z| = \pi$. Then δ is positive, and $e^{-z} - 1$ has just one root inside that circle, viz. $z = 0$. Now choose the positive number a equal to $\delta/2(\pi + 1)$. Then we have

$$|\sigma z - \tau| < \tfrac{1}{2}\delta \quad (|\sigma| < a, |\tau| < a, |z| = \pi).$$

A consequence is that $|e^{-z} - 1| > |\sigma z - \tau|$ on the circle $|z| = \pi$. So by Rouché's theorem, the equation $e^{-z} - 1 - \sigma z + \tau = 0$ has just one root inside the circle. Denoting this root by v, we have, in virtue of the Cauchy theorem,

$$(2.4.7) \qquad v = \frac{1}{2\pi i}\int \frac{- e^{-z} - \sigma}{e^{-z} - 1 - \sigma z + \tau} \cdot z\,dz,$$

where the integration path is the circle $|z| = \pi$, taken in the positive direction.

For every z on the integration path $|\sigma z| + |\tau|$ is less than $\tfrac{1}{2}|e^{-z} - 1|$, so that we have the following development

$$(2.4.8)\qquad (e^{-z}-1-\sigma z+\tau)^{-1} = \\ =\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}(e^{-z}-1)^{-k-m-1}z^k\sigma^k\tau^m(-1)^m\frac{(m+k)!}{m!k!},$$

converging absolutely and uniformly when $|z|=\pi$, $|\sigma|<a$, $|\tau|<a$. So in (2.4.7) we can integrate termwise, and v appears as the sum of an absolutely convergent double power series (powers of σ and τ). We notice that all terms not containing τ vanish. For, in (2.4.8) the terms with $m=0$ give rise to integrals

$$(2\pi i)^{-1}\int -(e^{-z}+\sigma)\,(e^{-z}-1)^{-k-1}z^k\cdot z\,dz,$$

which vanish by virtue of the regularity of the integrand at $z=0$.

Our result is that, if $|\sigma|<a$, $|\tau|<a$, (2.4.6) has just one solution v satisfying $|v|<\pi$, and this solution can be written as

$$(2.4.9)\qquad v=\tau\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}c_{km}\sigma^k\tau^m,$$

where the c_{km} are constants.

We now return to the special values of σ and τ, viz. $\sigma=(\log t)^{-1}$, $\tau=\log\log t/\log t$. For t sufficiently large, we have $|\sigma|<a$, $|\tau|<a$. Moreover, the solution of (2.4.6) which we actually want to have, is small: (2.4.3) shows that $v=O(\log\log t/\log t)$. It follows that it coincides with the solution (2.4.9) if t is large. The final result is that if t is large enough,

$$(2.4.10)\qquad x=\log t-\log\log t+ \\ +\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}c_{km}(\log\log t)^{m+1}(\log t)^{-k-m-1},$$

and the series is absolutely convergent for all large values of t. Needless to say, this series can be rearranged into the form (2.4.5).

2.5. Iteration methods

The previous section gave a typical example of the rôle of iteration in asymptotics. In the next sections we shall discuss some further aspects of asymptotic iteration. The subject does not entirely fall under the heading "implicit functions", and therefore our reflections will be somewhat more general.

Let $f(t)$ be a function whose asymptotic behaviour is required as $t\to\infty$. Usually it is quite important to have a reasonable con jecture about this behaviour before we start proving anything. And usually, the better the approximation we guess, the easier it is to prove that it is an approximation indeed.

Let $\varphi_0(t)$, $\varphi_1(t)$, ... be a sequence of functions and assume that, for each separate k, the asymptotic behaviour of $\varphi_k(t)$ is known. Assume that we have reasons to believe that the behaviour of $\varphi_0(t)$ is, in some sense, an approximation to the one of $f(t)$. Moreover assume that there is a procedure that transforms φ_0 into φ_1, φ_1 into φ_2, etc., and that there are reasons to believe that this procedure turns any good approximation into a better approximation. What we hope for is this: it might happen that, for some k, φ_k is so close to f, that we may be able to prove this fact, in some specified sense. It may even happen that we are able to use the procedure itself for proving things. Namely, if we are able to show that (i) if φ_n is an approximation in some n-th sense, then automatically φ_{n+1} is an approximation in some $(n+1)$-th sense, and if moreover (ii) for some k it can be proved that φ_k is an approximation in the k-th sense. The process which led to (2.4.4) provides a simple example. In section 2.4 we were so fortunate to have useful information right from the start: $0<x<\log t$, so that there was no need for guesswork. But quite often there is no such easy first step. For example, if we would again consider (2.4.1), but now with complex values of t, the first step would already be more difficult. In order to be specific, we assume that $\operatorname{Im} t=1$, and that we want to have a solution x with $\operatorname{Re} x\to\infty$, $\operatorname{Im} x\to 0$. Now $x=O(\log t)$ would be a conjecture, and so would be its consequences (2.4.3) and (2.4.4). But at the moment we have reached $x=\log t-\log\log t+o(1)$, we can put $x-\log t+\log\log t=v$, and the discussion of (2.4.6) can be applied. Only then we get to definite results.

This example of iterating conjectures so as to reach a stage, sooner or later, where things can be proved, is too simple to be very fortunate. For, it is not very difficult to prove $x=O(\log t)$ right at the start, using the Rouché theorem. On the other hand, it is easy to imagine slightly more complicated examples, where the application of the Rouché theorem would be very troublesome indeed.

The method of iteration of conjectures also occurs in numerical analysis. There the object to be approximated is not an asymptotic behaviour, but just a number. We shall consider things of that type in sec. 2.6, and compare them to asymptotic problems in sec. 2.7.

2.6. Roots of equations

We want to approximate a special root ξ of some equation $f(x) = 0$. To this end Newton's method usually gives very good results. It consists of taking a rough first approximation x_0 and constructing the sequence $x_1, x_2, x_3, \ldots$ by the formula [1]

$$x_{n+1} = x_n - f(x_n)/f'(x_n). \tag{2.6.1}$$

Its meaning is, that x_{n+1} is the root of the linear function whose graph is the tangent at P_n of the graph of $f(x)$, where P_n denotes the point with coordinates $(x_n, f(x_n))$.

Usually the situation is as follows: There is an interval J, containing ξ as an inner point, having the property that if x_0 belongs to J, then $x_1, x_2, \ldots$ all belong to J and the sequence converges to ξ.

A sufficient condition for the existence of J is, for instance, that $f(x)$ has a continuous second derivative throughout some neighbourhood of ξ and that $f'(\xi) \neq 0$. In that case the process converges very rapidly, for then (2.6.1) guarantees that $x_{n+1} - \xi$ is at most of the order of the square of $x_n - \xi$.

Quite often very little is known about the function $f(x)$, that is, for every special x the value of $f(x)$ can be found, but in larger x-intervals there is not much information about lower and upper bounds of $f(x)$, $f'(x)$, etc. Usually such information can be obtained in very small intervals. In order to find a root of the equation $f(x) = 0$, we then simply choose some number x_0, more or less at random, and we construct $x_1, x_2, \ldots$ by Newton's iteration process. If this sequence shows the tendency to converge, nothing as yet has been proved, as convergence cannot be deduced from a finite number of observations. But it may happen that sooner or later we arrive at a small interval J, where so much information can be obtained about $f(x)$, that it can be proved that the further x_j's remain in J and converge to a point of J, that this limit is a root of $f(x) = 0$,

[1]) See C. Jordan, Calculus of Finite Differences, 2nd ed., New York 1947, § 150.

and that there are no other roots inside J. What we then have achieved is not the exact value of a root, but a small interval in which there is one; moreover we have a procedure to find smaller and smaller intervals to which it belongs. Therefore it is a perfectly happy situation from the point of view of the numerical analyst.

There are also less favourable possibilities, several of which we mention here:

(i) The sequence $x_0, x_1, \ldots$ diverges to infinity.
(ii) It converges to a root, but not to the one we want to approximate.
(iii) It keeps oscillating.
(iv) It converges to the root we have in mind, but we are unable to prove it.

2.7. Asymptotic iteration

Now returning to asymptotic problems about implicit functions, we notice that the Newton method works quite well in small-parameter cases like those of sec. 2.3 or the one of (2.4.6). Needless to say, the root is no longer a number, but a function of t, and we are looking for asymptotic information about this function.

There are two different questions. The first one is whether the Newton method gives a sequence of good approximations.

A far more difficult question is whether we can prove that these approximations are approximations indeed. We shall not discuss this second question, in fact we only discuss examples that have been extensively studied before, so that the asymptotic behaviour is precisely known.

First take the equation (2.3.1), viz. $xe^x = t^{-1}$. We consider $\varphi_0 = 0$ as the first rough approximation to the root. Applying the Newton formula (2.6.1), with $f(x) = xe^x - t^{-1}$, we obtain

$$\varphi_{n+1} = (\varphi_n{}^2 + t^{-1}e^{-\varphi_n})\,(\varphi_n + 1)^{-1},$$

and so, putting $t^{-1} = \varepsilon$,

$$\varphi_1 = \varepsilon,$$
$$\varphi_2 = \varepsilon - \varepsilon(e^\varepsilon - 1)e^{-\varepsilon}(1 + \varepsilon)^{-1} = \varepsilon - \varepsilon^2 + \tfrac{3}{2}\,\varepsilon^3 + O(\varepsilon^4) \qquad (\varepsilon \to 0).$$

Hence φ_2 differs from the true root x (see (2.3.2)) by an amount $O(\varepsilon^4)$. It is not difficult to show, in virtue of (2.3.2), that φ_k differs from x only by $O(\varepsilon^{2k})$.

We next discuss the equation (2.4.1), and we shall apply Newton's method at a stage where we have not yet reached the small-parameter case. Then we shall notice phenomena that did not arise in sec. 2.6.

Observing that the positive root of $xe^x = t$ is small compared to t, we might think $\varphi_0 = 0$ to be a reasonable starting point. We have

$$\varphi_{n+1} = (\varphi_n{}^2 + te^{-\varphi_n})(\varphi_n + 1)^{-1},$$

whence

$$\begin{aligned} \varphi_1 &= t, \\ \varphi_2 &= t - 1 + O(t^{-1}) \qquad (t \to \infty), \\ \varphi_3 &= t - 2 + O(t^{-1}) \qquad (t \to \infty), \end{aligned}$$

and so on. It is clear that this leads us nowhere. None of the φ_k's have any asymptotic resemblance to the true root x, which is $\log t - \log\log t + o(1)$.

The same thing happens if we start with $\varphi_0 = \log t$, which is already a quite reasonable approximation, as $x = \log t + o(\log t)$ $(t \to \infty)$ (see (2.4.3)). Then we again obtain $\varphi_n = \log t - n + o(1)$. It is not difficult to show that we always have $\varphi_n = \varphi_0 - n + o(1)$, as soon as we start with a function φ_0 which is such that $\varphi_0 e^{\varphi_0}/t$ tends to infinity when $t \to \infty$.

Next assume that we try $\varphi_0 = \log t - \log\log t + a_0$ for some constant a_0 (admittedly, this example is not very natural, as no one would try this before trying $\varphi_0 = \log t - \log\log t$). Then we easily calculate that

$$\varphi_n = \log t - \log\log t + a_n + O((\log t)^{-1}), \text{ where } a_{n+1} = a_n + e^{-a_n} - 1.$$

It can be shown (see Ch. 8) that a_n tends to zero quite rapidly. However, not a single φ_k of this sequence gives an approximation essentially better than $\log t - \log\log t + O(1)$.

In some sense $\log t - \log\log t$ is the limit of this sequence φ_0, $\varphi_1, \varphi_2, \ldots$. If we now start the Newton method anew, with $\varphi_0{}^* = \log t - \log\log t$, we suddenly get much better approximations. Actually it means that we consider the small-parameter case (2.4.6), starting with zero as a first approximation to v.

We leave it at these casual remarks; our main aim was to stress the fact that in many asymptotic problems it is of vital importance to start with a good conjecture or a good first approximation.

2.8. Exercises. 1. Show that the equation $\sin x = (\log x)^{-1}$ has just one root x_n in the interval $2\pi n < x_n < 2\pi n + \frac{1}{2}\pi$ $(n = 1, 2, 3, \ldots)$, and that

$$x_n = 2\pi n + (\log 2\pi n)^{-1} + O((\log 2\pi n)^{-3}) \quad (n \to \infty).$$

2. Let $f(t)$ be positive, and assume that

$$e^{tf(t)} = f(t) + t + O(1) \quad (0 < t < \infty).$$

Show that

$$f(t) = t^{-1}\log t + O(t^{-2}) \quad (t \to \infty).$$

3. Show that the positive solution of $e^x + \log x = t$ equals, for large values of t:

$$x = \log t + \frac{\log\log t}{t} P\left(\frac{\log\log t}{t\log t}, \frac{1}{t\log t}, \frac{\log\log t}{t}\right),$$

where $P(\lambda, \mu, \sigma)$ is a multiple power series in the variables λ, μ, σ, convergent for all small values of $|\lambda|, |\mu|, |\sigma|$.