

# **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

*No part of this book may be reproduced  
in any form by print, microfilm or any  
other means without written permission  
from the publisher*

**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

<i>Preface</i> . . . . .	v
CH. 1. INTRODUCTION . . . . .	1
1.1. What is asymptotics? . . . . .	1
1.2. The O-symbol . . . . .	3
1.3. The o-symbol. . . . .	10
1.4. Asymptotic equivalence . . . . .	10
1.5. Asymptotic series . . . . .	11
1.6. Elementary operations on asymptotic series : . .	14
1.7. Asymptotics and Numerical Analysis . . . . .	18
1.8. Exercises . . . . .	19
CH. 2. IMPLICIT FUNCTIONS . . . . .	21
2.1. Introduction . . . . .	21
2.2. The Lagrange inversion formula . . . . .	22
2.3. Applications . . . . .	23
2.4. A more difficult case . . . . .	25
2.5. Iteration methods . . . . .	28
2.6. Roots of equations . . . . .	30
2.7. Asymptotic iteration . . . . .	31
2.8. Exercises . . . . .	33
CH. 3. SUMMATION . . . . .	34
3.1. Introduction . . . . .	34
3.2. Case <i>a</i> . . . . .	34
3.3. Case <i>b</i> . . . . .	36
3.4. Case <i>c</i> . . . . .	37
3.5. Case <i>d</i> . . . . .	38

3.6.	The Euler-Maclaurin sum formula . . . . .	40
3.7.	Example . . . . .	42
3.8.	A remark . . . . .	42
3.9.	Another example . . . . .	43
3.10.	The Stirling formula for the $I$ -function in the complex plane . . . . .	46
3.11.	Alternating sums . . . . .	49
3.12.	Application of the Poisson sum formula . . . . .	52
3.13.	Summation by parts . . . . .	56
3.14.	Exercises . . . . .	58
 CH. 4. THE LAPLACE METHOD FOR INTEGRALS . . . . .		60
4.1.	Introduction . . . . .	60
4.2.	A general case. . . . .	63
4.3.	Maximum at the boundary . . . . .	65
4.4.	Asymptotic expansions . . . . .	66
4.5.	Asymptotic behaviour of the $I$ -function . . . . .	69
4.6.	Multiple integrals . . . . .	71
4.7.	An application . . . . .	72
4.8.	Exercises . . . . .	75
 CH. 5. THE SADDLE POINT METHOD . . . . .		77
5.1.	The method. . . . .	77
5.2.	Geometrical interpretation . . . . .	79
5.3.	Peakless landscapes . . . . .	82
5.4.	Steepest descent . . . . .	83
5.5.	Steepest descent at end-point . . . . .	86
5.6.	The second stage . . . . .	86
5.7.	A general simple case . . . . .	87
5.8.	Path of constant altitude . . . . .	89
5.9.	Closed path . . . . .	90
5.10.	Range of a saddle point . . . . .	91
5.11.	Examples. . . . .	93
5.12.	Small perturbations . . . . .	96
5.13.	Exercises . . . . .	101

CH. 6. APPLICATIONS OF THE SADDLE POINT METHOD . . . . .	102
6.1. The number of class-partitions of a finite set . . . . .	102
6.2. Asymptotic behaviour of $d_n$ . . . . .	104
6.3. Alternative method . . . . .	108
6.4. The sum $S(s, n)$ . . . . .	109
6.5. Asymptotic behaviour of $P$ . . . . .	112
6.6. Asymptotic behaviour of $Q$ . . . . .	115
6.7. Conclusions about $S(s, n)$ . . . . .	118
6.8. A modified Gamma Function . . . . .	119
6.9. The entire function $G_0(s)$ . . . . .	123
6.10. Conclusions about $G(s)$ . . . . .	131
6.11. Exercises . . . . .	133
CH. 7. INDIRECT ASYMPTOTICS . . . . .	134
7.1. Direct and indirect asymptotics . . . . .	134
7.2. Tauberian theorems . . . . .	137
7.3. Differentiation of an asymptotic formula . . . . .	139
7.4. A similar problem . . . . .	141
7.5. Karamata's method . . . . .	143
7.6. Exercises . . . . .	147
CH. 8. ITERATED FUNCTIONS . . . . .	148
8.1. Introduction . . . . .	148
8.2. Iterates of a function . . . . .	148
8.3. Rapid convergence . . . . .	151
8.4. Slow convergence . . . . .	153
8.5. Preparation . . . . .	154
8.6. Iteration of the sine function . . . . .	157
8.7. An alternative method . . . . .	160
8.8. Final discussion about the iterated sine . . . . .	164
8.9. An inequality concerning infinite series . . . . .	166
8.10. The iteration problem . . . . .	169
8.11. Exercises . . . . .	175

CH. 9. DIFFERENTIAL EQUATIONS . . . . .	176
9.1. Introduction . . . . .	176
9.2. A Riccati equation. . . . .	177
9.3. An unstable case . . . . .	184
9.4. Application to a linear second-order equation . .	186
9.5. Oscillatory cases . . . . .	189
9.6. More general oscillatory cases . . . . .	195
9.7. Exercises . . . . .	198
INDEX . . . . .	199

## 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:

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

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<sup>1)</sup>

$$(1.1.2) \quad \lim_{x \rightarrow \infty} \pi(x) \left/ \frac{x}{\log x} \right. = 1.$$

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

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

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

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

When proving  $f(n) \rightarrow 1$ , one usually produces, hidden or not, information of the form<sup>\*</sup> (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  $f(n) \rightarrow 1$ , 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.

---

<sup>1)</sup> 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<sup>1)</sup>. 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  $\{f(n)\}$ :

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

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

$$(1.2.2) \quad f(n) \rightarrow 1 \quad (n \rightarrow \infty).$$

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

$$(1.2.3) \quad \left\{ \begin{array}{l} \text{There exists a number } A \text{ (independent of } n \text{) such that} \\ |f(n) - 1| \leq An^{-1} \quad (n = 1, 2, 3, \dots). \end{array} \right.$$

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

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

There are some minor differences between the various definitions

<sup>1)</sup> 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  $\varphi$  are real or complex functions defined on  $S$ , then the formula

$$(1.2.5) \quad f(s) = O(\varphi(s)) \quad (s \in S),$$

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

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

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:

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

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  $x$  ( $x \rightarrow \infty$ ), but by obvious modifications we get similar notations for cases like

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

The formula

$$(1.2.7) \quad f(x) = O(\varphi(x)) \quad (x \rightarrow \infty)$$

means that there exists a real number  $a$  such that

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

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

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

Examples:

$$\begin{array}{ll} x^2 = O(x) \quad (x \rightarrow 0); & x = O(x^2) \quad (x \rightarrow \infty); \\ e^{-x} = O(1) \quad (x \rightarrow \infty); & (\log x)^6 = O(x^4) \quad (x \rightarrow \infty); \\ (\log x)^{-1} = O(1) \quad (x \rightarrow \infty); & (\sin x^{-1})^{-1} = O(x) \quad (x \rightarrow \infty). \end{array}$$

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  $\varphi$  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  $f/\varphi$  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:

$$Q(x) + O(x^2) = O(x) \quad (x \rightarrow 0).$$

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

$$f(x) = O(x) \quad (x \rightarrow 0), \quad g(x) = O(x^2) \quad (x \rightarrow 0),$$

we have

$$f(x) + g(x) = O(x) \quad (x \rightarrow 0).$$

Analogously,

$$O(x) + O(x^3) = O(x^3) \quad (x \rightarrow \infty),$$

$$e^{O(1)} = O(1) \quad (-\infty < x < \infty),$$

$$e^{O(x)} = e^{O(x^2)} \quad (x \rightarrow \infty).$$

We also write things like

$$e^x = 1 + x + O(x^2) \quad (x \rightarrow 0),$$

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

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

This means that for any function  $f(x)$  with  $f(x) = O(1) \quad (0 < x < \infty)$  we can split  $x^{-1}f(x)$  into two parts  $g(x)$  and  $h(x)$ , such that  $g(x) = O(1)$ ,  $h(x) = O(x^{-2}) \quad (0 < x < \infty)$ . 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  $f(x) = O(1) \quad (0 < x < \infty)$ ,  $g(x) = O(x^2) \quad (0 < x < \infty)$ . 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,  $O(x) = O(x^2)$  ( $x \rightarrow \infty$ ) is correct, but  $O(x^2) = O(x)$  ( $x \rightarrow \infty$ ) 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  $\varphi$  and  $\psi$  be functions such that  $\varphi(x) = O(\psi(x))$  ( $x \rightarrow \infty$ ) is true and  $\psi(x) = O(\varphi(x))$  ( $x \rightarrow \infty$ ) is false. If a third function  $f$  satisfies

$$(1.2.9) \quad f(x) = O(\varphi(x)) \quad (x \rightarrow \infty),$$

then obviously it also satisfies

$$(1.2.10) \quad f(x) = O(\psi(x)) \quad (x \rightarrow \infty).$$

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,

$$2x + x \sin x = O(x) \quad (x \rightarrow \infty)$$

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

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

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

$$(1.2.11) \quad e^{-x} = O(x^{-m}) \quad (x \rightarrow \infty)$$

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  $e^{-x} = O(x^{-m-1})$  ( $x \rightarrow \infty$ ) 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

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

For, we have

$$\begin{aligned}|(f+g)^k| &\leq (|f| + |g|)^k \leq \{2 \max(|f|, |g|)\}^k \leq \\&\leq 2^k \max(|f|^k, |g|^k) \leq 2^k (|f|^k + |g|^k).\end{aligned}$$

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

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

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

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

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

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

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

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

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

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

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,

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

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  $x \rightarrow \infty$ , 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

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

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

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

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

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

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

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

### 1.3. The o-symbol

The expression

$$(1.3.1) \quad f(x) = o(\varphi(x)) \quad (x \rightarrow \infty)$$

means that  $f(x)/\varphi(x)$  tends to 0 when  $x \rightarrow \infty$ . 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

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

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  $x \rightarrow \infty$ , if the quotient  $f(x)/g(x)$  tends to unity. Our notation is

$$f(x) \sim g(x) \quad (x \rightarrow \infty).$$

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

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:

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

When asking for the "asymptotic behaviour" of a given function  $f(x)$  as  $x \rightarrow \infty$ , 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  $x \rightarrow \infty$ , 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  $\varphi_0, \varphi_1, \varphi_2, \dots$ , satisfying

$$(1.5.1) \quad \varphi_1(x) = o(\varphi_0(x)) \quad (x \rightarrow \infty), \quad \varphi_2(x) = o(\varphi_1(x)) \quad (x \rightarrow \infty), \dots,$$

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

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

Obviously, the second formula improves the first one, as

$$c_0 \varphi_0(x) + O(\varphi_1(x)) = (c_0 + o(1)) \varphi_0(x) = O(\varphi_0(x)) \quad (x \rightarrow \infty).$$

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:

$$(1.5.3) \quad f(x) \approx c_0 \varphi_0(x) + c_1 \varphi_1(x) + c_2 \varphi_2(x) + \dots \quad (x \rightarrow \infty).$$

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  $\varphi$ 's are given, assuming that such an asymptotic expansion exists. For, assume that (1.5.3) holds, that

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

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

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

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

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

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  $x \rightarrow \infty$ . Convergence of the series, for all  $x > 0$ , say, means that for every individual  $x$  there is a statement about the case  $n \rightarrow \infty$ . 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  $x \rightarrow \infty$ .

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

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

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

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

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

$$\begin{aligned} \int_1^{4x} t^{-2} e^t dt &< \int_1^{4x} e^t dt < e^{4x}, \\ \int_{4x}^x t^{-2} e^t dt &< \int_{4x}^x (\frac{1}{2}x)^{-2} e^t dt < 4x^{-2} e^x. \end{aligned}$$

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

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

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

$$\begin{aligned} f(x) &= \left[ \frac{e^t}{t} \right]_1^x + \left[ \frac{e^t}{t^2} \right]_1^x + \int_1^x \frac{2e^t}{t^3} dt = \\ &= \left[ \frac{e^t}{t} \right]_1^x + \left[ \frac{e^t}{t^2} \right]_1^x + \left[ \frac{2e^t}{t^3} \right]_1^x + \int_1^x \frac{3!e^t}{t^4} dt, \end{aligned}$$

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  $x \rightarrow \infty$  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

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

and it follows that

$$e^{-x}f(x) \approx \frac{1}{x} + \frac{1}{x^2} + \frac{2!}{x^3} + \frac{3!}{x^4} + \dots \quad (x \rightarrow \infty).$$

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

$$(1.5.7) \quad f(z) = a_0 + a_1z + a_2z^2 + \dots$$

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

$$(1.5.8) \quad f(z) \approx a_0 + a_1z + a_2z^2 + \dots \quad (|z| \rightarrow 0).$$

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  $|z| = \frac{1}{2}p$ ,

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

and therefore

$$f(z) = a_0 + a_1z + \dots + a_n z^n + O(z^{n+1}) \quad (|z| < \frac{1}{2}p).$$

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

$$(1.6.1) \quad a_0 + a_1x + a_2x^2 + \dots \quad (x \rightarrow 0),$$

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  $1 + 0 \cdot x + 0 \cdot x^2 + \dots$  as its unit element (to be denoted by  $I$ ). If the series  $a_0 + a_1x + \dots$  and  $b_0 + b_1x + \dots$  are represented by  $A$  and  $B$ , respectively, we define sum and product by

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

$$AB = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \dots.$$

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

$$a_0c_0 = 1, \quad a_0c_1 + a_1c_0 = 0, \quad a_0c_2 + a_1c_1 + a_2c_0 = 0, \quad \dots.$$

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  $A(B)$ <sup>1)</sup>. It is defined as follows: Let  $c_{kn}$  be the coefficient of  $x^k$  in the series  $a_0I + a_1B + a_2B^2 + \dots + a_nB^n$ . Then it is easily seen that  $c_{kk} = c_{k,k+1} = c_{k,k+2} = \dots$ . Writing  $c_{kk} = c_k$ , we infer that

$$a_0I + a_1B + \dots + a_nB^n =$$

$$= c_0 + c_1x + \dots + c_nx^n + c_{n+1,n}x^{n+1} + c_{n+2,n}x^{n+2} + \dots.$$

We now define

$$A(B) = c_0 + c_1x + \dots + c_nx^n + c_{n+1,n}x^{n+1} + c_{n+2,n}x^{n+2} + \dots.$$

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  $A = a_0 + a_1x + \dots$  is defined by

$$A' = a_1 + 2a_2x + 3a_3x^2 + \dots$$

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

<sup>1)</sup> 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

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

Note that  $A(x)$  stands for the function, and that  $A$  stands for the formal series  $a_0 + a_1x + \dots$ . 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

$$(1.6.2) \quad A(x) + B(x) \approx A + B \quad (x \rightarrow 0),$$

$$(1.6.3) \quad A(x)B(x) \approx AB \quad (x \rightarrow 0),$$

and if  $a_0 \neq 0$ ,

$$(1.6.4) \quad \{A(x)\}^{-1} \approx A^{-1} \quad (x \rightarrow 0)$$

( $A^{-1}$  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

$$(1.6.5) \quad A\{B(x)\} \approx A(B) \quad (x \rightarrow 0).$$

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

$$A(x) = a_0 + \dots + a_n x^n + O(x^{n+1}) \quad (x > 0),$$

$$B(x) = b_0 + \dots + b_n x^n + O(x^{n+1}) \quad (x > 0),$$

and so

$$A(x)B(x) = (a_0 + \dots + a_n x^n)(b_0 + \dots + b_n x^n) + O(x^{n+1}) \quad (x \rightarrow 0).$$

Now

$$(a_0 + \dots + a_n x^n)(b_0 + \dots + b_n x^n) - (c_0 + \dots + c_n x^n)$$

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

$$A(x)B(x) = c_0 + \dots + c_n x^n + O(x^{n+1}) \quad (x \rightarrow 0),$$

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  $P = a_0^{-1}(1 + x + x^2 + \dots)$ ,  $Q = a_0^{-1}(a_0 - A)$ .

Suppose that  $f(x)$  satisfies

$$f(x) \approx a_0 + a_1 x + a_2 x^2 + \dots \quad (x \rightarrow 0),$$

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

$$(1.6.6) \quad \int_0^x f(t)dt \approx a_0 x + \frac{1}{2} a_1 x^2 + \frac{1}{3} a_2 x^3 + \dots \quad (x \rightarrow 0).$$

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

$$|f(t) - a_0 - a_1 t - \dots - a_{n-1} t^{n-1}| < A |t|^n \quad (|t| < a),$$

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

$$\left| \int_0^x f(t)dt - a_0 x - \frac{1}{2} a_1 x^2 - \dots - \frac{1}{n} a_n x^n \right| < \frac{A}{n+1} |x|^{n+1},$$

and (1.6.6) follows.

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

For example, we have

$$e^{-1/x} \sin(e^{1/x}) \approx 0 + 0 \cdot x + 0 \cdot x^2 + \dots \quad (x > 0, x \rightarrow 0),$$

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

$$\begin{aligned} f(x) &\approx a_0 + a_1x + a_2x^2 + \dots & (x \rightarrow 0), \\ f'(x) &\approx b_0 + b_1x + b_2x^2 + \dots & (x \rightarrow 0), \end{aligned}$$

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

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

we have  $g_n'(x) = O(x^n)$  ( $x \rightarrow 0$ ). 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

$$f(x) \approx f(0) + b_0x + \frac{1}{2}b_1x^2 + \frac{1}{3}b_2x^3 + \dots \quad (x \rightarrow 0).$$

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.:  $f(x) = x^{-1} + O(x^{-2}) \quad (x \rightarrow \infty)$ .

N.A.: I am sorry, but I don't understand.

A.A.:  $|f(x) - x^{-1}| < 8x^{-2} \quad (x > 10^4)$ .

N.A.: But my value of  $x$  is only 100.

A.A.: Why did not you say so? My evaluations give

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

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

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

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:  $f(100) = 0.01137\ 42259\ 34008\ 67153$ .

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

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

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

2. Show that

$$\cdot (x + 1 + O(x^{-1}))^x = ex^x + O(x^{x-1}) \quad (x \rightarrow \infty).$$

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

$$\int_0^x (\log y)^n dy = O(x(\log x)^n) \quad (x \rightarrow 0).$$

4. Show that

$$t \int_0^\infty e^{-xt} (\log x)^{-1} dx \approx \sum_0^\infty c_n (\log t)^{-n-1} \quad (t > 0, t \rightarrow 0),$$

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

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

holds uniformly with respect to  $n$  ( $n = 1, 2, 3, \dots$ ).

(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:

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

The summation runs over all integers  $n$  with  $1 \leq n \leq x$ ;  $x$  need not be an integer.

(Hint:  $|\int_0^u e^{tp} dt| \leq \min(u, 2|p|^{-1})$  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  $t \rightarrow \infty$ .

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  $t \rightarrow \infty$ ) 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

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

then we can easily verify that  $x = e^{t\varphi(t)}$ , where  $\varphi(t)$  is the solution of  $\varphi e^\varphi = t$ . When  $t \rightarrow \infty$  we have for  $\varphi$  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 possess an elementary function  $\psi(t)$  such that  $x/\psi(t)$  tends to 1 as  $t \rightarrow \infty$ . 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<sup>1)</sup>

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

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

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  $w$ :

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

where the coefficients  $c_k$  are given by

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

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

$$(2.2.4) \quad g(z) = g(0) + \sum_1^{\infty} d_k w^k,$$

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

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

---

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, *Modern Analysis*, 4th ed., Cambridge 1946, § 7.32.

tions<sup>1)</sup>: 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

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

when  $t \rightarrow \infty$ . As  $t^{-1}$  tends to zero, we apply the Lagrange formula (2.2.2) to the equation  $ze^x = w$ , so that  $f(x) = e^{-x}$ . 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  $\infty$  if  $x$  increases from 0 to  $\infty$ , 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

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

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

Our second example considers the positive solution of

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

<sup>1)</sup> See L. BIEBERBACH, Funktionentheorie, vol. 1, 3e Aufl., Leipzig-Berlin 1930, Kap. VII § 8.

when  $t \rightarrow \infty$ . 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  $t \rightarrow \infty$ .

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  $z = 0$ :  $f(z) = -1 + c_1 z + \dots$ . It follows that

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

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  $t \rightarrow \infty$ , 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

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

We observe from the graphs of the functions  $x$  and  $\cotg x$ , that there is just one root in every interval  $n\pi < x < (n+1)\pi$  ( $n = 0, \pm 1, \pm 2, \dots$ ). Denoting this root by  $x_n$ , we ask for the behaviour of  $x_n$  as  $n \rightarrow \infty$ . As  $\cotg(x_n - n\pi) = x_n \rightarrow \infty$ , we have  $x_n - n\pi \rightarrow 0$ . Putting  $x = n\pi + z$ ,  $(n\pi)^{-1} = w$ , 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  $z = w + c_2 w^2 + c_3 w^3 + \dots$ . Consequently we have, if  $n$  is large enough,

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

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

## 2.4. A more difficult case

We take the equation

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

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

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

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

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  $\Delta$  in the value of  $x$  gives an error of about  $\Delta/x$  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  $\leq 1$ .

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

Taking logarithms, we infer that

$$\begin{aligned} \log x &= \log \log t + \log(1 + O(\log \log t / \log t)) = \\ &= \log \log t + O(\log \log t / \log t). \end{aligned}$$

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

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

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

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

We shall carry out two further steps. Abbreviating

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

we obtain

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

and so, the term  $O(L_2 L_1^{-3})$  absorbing all terms  $L_2^q L_1^{-q}$  with  $q > 3$ ,

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

The next step can be verified to give

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

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

$$(2.4.5) \quad x \approx L_1 - L_2 + L_2 P_0(L_2) L_1^{-1} +$$

$$+ L_2 P_1(L_2) L_1^{-2} + L_2 P_2(L_2) L_1^{-3} + \dots,$$

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 <sup>1)</sup>, 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  $C$ ).

<sup>1)</sup> 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  $2\pi$  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

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

and we obtain from (2.4.2) that

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

For the time being, we ignore the relation that exists between  $\sigma$  and  $\tau$ , 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  $\sigma$  and  $\tau$  in the region  $|\sigma| < a$ ,  $|\tau| < a$ .

Let  $\delta$  be the lower bound of  $|e^{-z} - 1|$  on the circle  $|z| = \pi$ . Then  $\delta$  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

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

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,

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

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  $\frac{1}{2}|e^{-z} - 1|$ ,

so that we have the following development

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

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  $\sigma$  and  $\tau$ ). We notice that all terms not containing  $\tau$  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

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

where the  $c_{km}$  are constants.

We now return to the special values of  $\sigma$  and  $\tau$ , 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,

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

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  $t \rightarrow \infty$ . Usually it is quite important to have a reasonable conjecture 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  $\varphi_0(t), \varphi_1(t), \dots$  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  $\varphi_0$  into  $\varphi_1$ ,  $\varphi_1$  into  $\varphi_2$ , 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  $k$ ,  $\varphi_k$  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  $\varphi_n$  is an approximation in some  $n$ -th sense, then automatically  $\varphi_{n+1}$  is an approximation in some  $(n+1)$ -th sense, and if moreover (ii) for some  $k$  it can be proved that  $\varphi_k$  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 guess-work. 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  $\operatorname{Re} x \rightarrow \infty$ ,  $\operatorname{Im} x \rightarrow 0$ . 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  $\xi$  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  $x_1, x_2, x_3, \dots$  by the formula<sup>1)</sup>

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

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  $\xi$  as an inner point, having the property that if  $x_0$  belongs to  $J$ , then  $x_1, x_2, \dots$  all belong to  $J$  and the sequence converges to  $\xi$ .

A sufficient condition for the existence of  $J$  is, for instance, that  $f(x)$  has a continuous second derivative throughout some neighbourhood of  $\xi$  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  $x_1, x_2, \dots$  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  $f(x) = 0$ ,

---

<sup>1)</sup> 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  $x_0, x_1, \dots$  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  $t^{-1} = \varepsilon$ ,

$$\varphi_1 = \varepsilon,$$

$$\varphi_2 = \varepsilon - \varepsilon(e^\varepsilon - 1)e^{-\varepsilon}(1 + \varepsilon)^{-1} = \varepsilon - \varepsilon^2 + \frac{3}{2}\varepsilon^3 + O(\varepsilon^4) \quad (\varepsilon \rightarrow 0).$$

Hence  $\varphi_2$  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  $\varphi_k$  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

$$\varphi_1 = t,$$

$$\varphi_2 = t - 1 + O(t^{-1}) \quad (t \rightarrow \infty),$$

$$\varphi_3 = t - 2 + O(t^{-1}) \quad (t \rightarrow \infty),$$

and so on. It is clear that this leads us nowhere. None of the  $\varphi_k$ '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  $x = \log t + o(\log t)$  ( $t \rightarrow \infty$ ) (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  $\varphi_0$  which is such that  $\varphi_0 e^{\varphi_0}/t$  tends to infinity when  $t \rightarrow \infty$ .

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  $\varphi_k$  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  $\varphi_0, \varphi_1, \varphi_2, \dots$ . 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  $2\pi n < x_n < 2\pi n + \frac{1}{2}\pi$  ( $n = 1, 2, 3, \dots$ ), and that

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

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

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

Show that

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

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  $\lambda, \mu, \sigma$ , convergent for all small values of  $|\lambda|, |\mu|, |\sigma|$ .

## CHAPTER 3

### SUMMATION

#### 3.1. Introduction

We shall consider sums of the type  $\sum_{k=1}^n a_k(n)$ , where both the terms and the number of terms depend on  $n$ . We ask for asymptotic information about the value of the sum for large values of  $n$ . In many applications,  $a_k(n)$  is independent of  $n$ , and actually several of our examples will be of this type, but the methods by which those examples are tackled, are by no means restricted to this case.

It is of course difficult to say anything in general. The asymptotic problem can be difficult, especially in cases where the  $a_k$  are not all of one sign, and where  $\sum_1^n a_k(n)$  can be much smaller than  $\sum_1^n |a_k(n)|$ . On the other hand, there is a class of routine problems arising in many parts of analysis, and to which a large part of this chapter is devoted: the cases where all  $a_k(n)$  are of one sign and where moreover the  $a_k(n)$  "behave smoothly". We shall not attempt to define what smoothness of behaviour is, but we merely give a number of examples. These fall under four headings *a*, *b*, *c*, *d*, according to the location of the terms which give the main contribution to the sum. The major contribution can come from:

- a.* a comparatively small number of terms at the end, or at the beginning.
- b.* a single term at the end or at the beginning.
- c.* a comparatively small number of terms somewhere in the middle.

There is also a case *d*, where there is not such a small group of terms whose sum dominates the sum of all others.

#### 3.2. Case a

Our first example concerns the behaviour of the sum  $s_n = \sum_{k=1}^n k^{-3}$ . A first approximation to  $s_n$  is the sum  $S = \sum_1^\infty k^{-3}$  of

the infinite series, and the error term is  $-\sum_{n+1}^{\infty} k^{-3}$ . For this last sum we easily obtain the estimate  $O(n^{-2})$ , e.g. by

$$(3.2.1) \quad \sum_{n+1}^{\infty} k^{-3} < \sum_{n+1}^{\infty} \int_{k-1}^k t^{-3} dt = \int_n^{\infty} t^{-3} dt = \frac{1}{2} n^{-2},$$

and therefore

$$(3.2.2) \quad s_n = S + O(n^{-2}) \quad (n \rightarrow \infty).$$

Results of this type are quite satisfactory for many analytical purposes; it should be noted, however, that from the point of view of numerical analysis nothing has been achieved by (3.2.2), unless we know the value of  $S$  from other sources of information. The numerical analyst would prefer to evaluate  $\sum_1^m k^{-3}$  for some suitably chosen value of  $m$ , and to estimate  $\sum_{m+1}^{\infty} k^{-3}$ .

Formula (3.2.2) can be improved by refinement of the argument used in (3.2.1), i.e. comparison of the sum with the integral. We shall return to this technique in secs. 3.5 and 3.6.

Our next example is  $\sum_1^n 2^k \log k$ . In this sum there is a relatively small number of terms at the end whose total contribution is large compared to the sum of all others. If we omit the last  $[\log n]$  terms ( $[\log n]$  denotes the largest integer  $\leq \log n$ ), the sum of the remaining terms is less than

$$\sum_1^{n-[\log n]} 2^k \log n \leq 2^{n+1-\log n} \log n = 2^{n+1} n^{-\log 2} \log n,$$

and this is much smaller than the  $n$ -th term.

We notice that  $\log k$  shows but little variation when  $k$  runs through the indices of the  $[\log n]$  significant terms. We therefore expand  $\log k$  in terms of powers of  $(n - k)/n$ , and in doing this we can easily admit the range  $\frac{1}{2}n < k \leq n$ . We shall be satisfied with

$$\log k = \log(n - h) = \log n - hn^{-1} + O(h^2 n^{-2}) \quad (n \rightarrow \infty)$$

which holds uniformly in  $h$  ( $0 < h < \frac{1}{2}n$ ). We now evaluate

$$\sum_{1 \leq k \leq \frac{1}{2}n} 2^k \log k = O(2^{\frac{1}{2}n} \log n),$$

$$\sum_{\frac{1}{2}n < k \leq n} 2^k \log n = 2^{n+1} \log n + O(2^{\frac{1}{2}n} \log n),$$

$$\sum_{\frac{1}{2}n < k \leq n} 2^k hn^{-1} = n^{-1} 2^n \sum_{h=1}^{\infty} 2^{-h} h + O(2^{\frac{1}{2}n}),$$

$$\sum_{\frac{1}{2}n < k \leq n} 2^k O(h^2 n^{-2}) = O(2^n n^{-2}) \cdot \sum_{h=1}^{\infty} 2^{-h} h^2.$$

The main error term is  $O(2^n n^{-2})$ ; the terms involving  $2^{kn}$  are much smaller than this one. Our result is

$$2^{-n} \sum_1^n 2^k \log k = 2 \log n - n^{-1} \sum_1^\infty h \cdot 2^{-h} + O(n^{-2}),$$

and it is not difficult to extend our argument in order to obtain an asymptotic series in terms of powers of  $n^{-1}$ :

$$2^{-n} \sum_{k=1}^n 2^k \log k - 2 \log n \approx c_1 n^{-1} + c_2 n^{-2} + \dots \quad (n \rightarrow \infty),$$

with  $c_j = -j^{-1} \sum_{h=1}^\infty h^j 2^{-h}$ .

### 3.3. Case b

We are often confronted with sums of positive terms, where each term has at least the order of magnitude of the sum of all previous terms. Our example is  $s_n = \sum_{k=1}^n k!$ . Dividing by the last term, we find that

$$\frac{s_n}{n!} = 1 + \frac{1}{n} + \frac{1}{n(n-1)} + \frac{1}{n(n-1)(n-2)} + \dots + \frac{1}{n!} \dots$$

If we stop after the 5th term, say, we neglect  $n - 5$  terms, each one of which is at most  $(n-5)!/n!$ , and so the error is  $O(n^{-4})$ . But the 5th term itself is  $O(n^{-4})$ , and therefore

$$\frac{s_n}{n!} = 1 + \frac{1}{n} + \frac{1}{n(n-1)} + \frac{1}{n(n-1)(n-2)} + O(n^{-4}) \quad (n \rightarrow \infty).$$

If we so wish, we can expand these terms into powers of  $n^{-1}$ :

$$\frac{s_n}{n!} = 1 + \frac{1}{n} + \frac{1}{n^2} + \frac{2}{n^3} + O(n^{-4}) \quad (n \rightarrow \infty).$$

Replacing the number 5 by an arbitrary integer, we easily find that there is an asymptotic expansion

$$(3.3.1) \quad s_n/n! \approx c_0 + c_1 n^{-1} + c_2 n^{-2} + \dots \quad (n \rightarrow \infty).$$

This series is not convergent, that is to say, the series  $c_0 + c_1 x + c_2 x^2 + \dots$  does not converge unless  $x = 0$ .

This can be shown as follows. The series  $c_0 + c_1 x + c_2 x^2 + \dots$

arises as the formal sum of the power series for the functions

$$1, \frac{x^2}{1-x}, \frac{x^3}{(1-x)(1-2x)}, \frac{x^4}{(1-x)(1-2x)(1-3x)}, \dots$$

and each one of these has non-negative coefficients. So if  $k$  is any positive integer, the coefficients of  $c_0 + c_1x + c_2x^2 + \dots$  exceed those of the series for  $x^{k+1}/(1-x)(1-2x)\dots(1-kx)$ . The latter series diverges at  $x = k^{-1}$ , and therefore  $c_0 + c_1x + c_2x^2 + \dots$  diverges at  $x = k^{-1}$ . Since  $k$  is arbitrary, we infer that the radius of convergence of  $c_0 + c_1x + c_2x^2 + \dots$  is zero.

There is usually no reason to try to obtain an explicit formula for the coefficients of a divergent asymptotic series. For practical purposes only a few terms of the asymptotic series will be needed, and for nearly all theoretical purposes the mere existence of an asymptotic series is already a satisfactory result. So it is only for the sake of curiosity that we mention that  $c_{k+1} = k!d_k$  ( $k = 0, 1, 2, \dots$ ), where the  $d_k$  are the coefficients in

$$\exp(ex - 1) = \sum_0^{\infty} d_k x^k.$$

We leave the proof to the reader [Hint: first prove, e.g. by induction, that

$$\int_0^{\infty} e^{-y/x} \frac{(ey - 1)^k}{k!} dy = \frac{x^{k+1}}{(1-x)(1-2x)\dots(1-kx)} \quad (0 < x < k^{-1});$$

then notice that the coefficient of  $x^{j+1}$  on the left equals  $j!$  times the coefficient of  $y^j$  in  $(ey - 1)^k/k!$ ].

### 3.4. Case c

A typical example is

$$s_n = \sum_{k=1}^n a_k(n), \quad a_k(n) = 2^{2k} \{n!/k!(n-k)!\}^2.$$

We have  $a_{k+1}(n)/a_k(n) = \{2(n-k)/(k+1)\}^2$ . Hence the maximal term occurs at the first value of  $k$  for which  $2(n-k) < (k+1)$ , that is, at about  $k = 2n/3$ .

We notice that in this case, contrary to our previous examples, the sum is large compared to the value of the maximal term. For, if

we move  $k$  in either direction, starting from the maximal term, then  $a_k(n)$  decreases rather slowly ( $n$  is considered to be fixed). It can be shown by various methods, e.g. by the Stirling formula for the factorials, that the number of terms which exceed  $\frac{1}{2} \max a_k(n)$ , is of the order of  $n^{\frac{1}{2}}$ . If, however,  $|k - 2n/3|$  is much greater than  $n^{\frac{1}{2}}$ , then  $a_k$  is very small compared to the maximum, and also the total contribution of these terms is relatively small. Therefore we have to focus our attention on regions of the type  $|k - 2n/3| < An^{\frac{1}{2}}$ .

By application of the Stirling formula,  $a_k(n)$  can be successfully approximated in this region, and then sums are obtained resembling the one discussed in sec. 3.9. We leave it at these brief remarks.

### 3.5. Case d

As a first example we take  $a_k = k^{\frac{1}{2}}$ . The ideal technique for dealing with a case as smooth as this one is given by the Euler-Maclaurin sum formula. Nevertheless we shall start with a more elementary method, which can be applied in less regular cases as well.

There are two steps. First approximate  $a_k$  by a sequence  $u_k$  which is such that  $\sum_{k=1}^n u_k$  is explicitly known; the approximation has to be strong enough for  $\sum_{k=1}^{\infty} (a_k - u_k)$  to converge. The second step deals with  $\sum_{k=1}^n (a_k - u_k)$ . The first approximation to this sum is, as in sec. 3.2, the infinite sum  $S = \sum_{k=1}^{\infty} (a_k - u_k)$ , and we have

$$(3.5.1) \quad s_n = \sum_{k=1}^n a_k = \sum_{k=1}^n u_k + S + \sum_{k=n+1}^{\infty} (u_k - a_k).$$

In the last sum we try to approximate  $u_k - a_k$  by a sequence  $v_k$ , such that  $\sum_{k=n+1}^{\infty} v_k$  is explicitly known, and such that the error term  $\sum_{k=n+1}^{\infty} (u_k - a_k - v_k)$  is known to be small. This procedure can be continued.

The weak point in the procedure is that in general there is hardly any information about the value of  $S$ . The situation is not as serious as in (3.2.2), for in (3.5.1) the major contribution is not  $S$ , but the sum  $\sum_1^n u_k$ , whose value is known.

In our example  $a_k = k^{\frac{1}{2}}$  we can obtain a first approximation to the sum  $s_n$  by taking the integral  $\int_0^n t^{\frac{1}{2}} dt = \frac{2}{3} n^{\frac{3}{2}}$ . If we now try to take  $u_k$  such that  $\sum_1^n u_k = \frac{2}{3} n^{\frac{3}{2}}$ , we still fail. For

$$(3.5.2) \quad k^{\frac{1}{2}} - \left\{ \frac{2}{3} k^{\frac{3}{2}} - \frac{2}{3} (k-1)^{\frac{3}{2}} \right\}$$

is not yet the  $k$ -th term of a convergent series. On expanding  $(1 - k^{-1})^{3/2}$  into powers of  $k^{-1}$  by the binomial series, we find that the expression (3.5.2) is  $\frac{1}{2}k^{-\frac{1}{2}} + O(k^{-\frac{5}{2}})$ , and  $\sum k^{-\frac{1}{2}}$  diverges. But we can again approximate the partial sums of  $\sum k^{-\frac{1}{2}}$  by an integral, viz.  $2n^{\frac{1}{2}}$ . If we now take new  $u_k$ 's, viz.

$$u_k = U_k - U_{k-1}, \quad U_k = \frac{2}{3}k^{3/2} + \frac{1}{2}k^{\frac{1}{2}},$$

we easily obtain that

$$(3.5.3) \quad u_k - a_k = k^{-\frac{3}{2}}/48 + O(k^{-\frac{5}{2}}) \quad (k \rightarrow \infty),$$

whence  $\sum_{k=1}^{\infty} (u_k - a_k)$  converges.

In the second step we have to approximate  $u_k - a_k$  by a sequence  $v_k$ . We take  $v_k = V_{k-1} - V_k$ , where

$$V_k = k^{-\frac{1}{2}}/24, \quad \sum_{n+1}^{\infty} v_k = V_n,$$

as suggested by the integral

$$\int_n^{\infty} (t^{-\frac{3}{2}}/48) dt = n^{-\frac{1}{2}}/24.$$

We so obtain

$$u_k - a_k - v_k = O(k^{-\frac{5}{2}}),$$

and, by (3.5.1),

$$(3.5.4) \quad \sum_{k=1}^n k^{\frac{1}{2}} = \frac{2}{3}n^{3/2} + \frac{1}{2}n^{\frac{1}{2}} + S + \frac{1}{24}n^{-\frac{1}{2}} + O(n^{-\frac{3}{2}}) \quad (n \rightarrow \infty).$$

The term  $O(n^{-\frac{3}{2}})$  can be replaced by an asymptotic series, for the process can be carried on and we can get as many terms as we please. To this end it is, of course, necessary to refine (3.5.3). This is easily done, for  $(u_k - a_k)k^{\frac{3}{2}}$  can be expanded into powers of  $k^{-1}$ , and the expansion converges if  $k > 1$ .

We next ask for the value of  $S$ . We obviously have

$$(3.5.5) \quad S = \sum_{k=1}^{\infty} \left\{ k^{\frac{1}{2}} - \frac{2}{3}k^{3/2} - \frac{1}{2}k^{\frac{1}{2}} + \frac{2}{3}(k-1)^{3/2} + \frac{1}{2}(k-1)^{\frac{1}{2}} \right\} = \\ = \lim_{n \rightarrow \infty} \left\{ \sum_{k=1}^n k^{\frac{1}{2}} - U_n \right\},$$

but it is possible to derive a simpler expression.

The method depends on analyticity properties, and therefore it is not generally applicable.

We first generalize (3.5.4) by introducing a complex parameter  $z$ .

Instead of (3.5.4) we obtain, by the same method,

$$(3.5.6) \quad \sum_{k=1}^n k^{-z} = \frac{n^{1-z}}{1-z} + \frac{1}{2} n^{-z} + S(z) + O(n^{-z-1}) \quad (n \rightarrow \infty),$$

if  $\operatorname{Re} z > -1$ ,  $z \neq 1$ . Here  $S(z)$  is the sum of a convergent series, analogous to (3.5.5). Furthermore, it is not hard to show that this sum is an analytic function of  $z$  in the region  $\operatorname{Re} z > -1$ ,  $z \neq 1$ . If  $\operatorname{Re} z > 1$ , it represents the Riemann zeta function <sup>1)</sup>  $\zeta(z) = \sum_1^\infty n^{-z}$ , as can be seen from (3.5.6) by making  $n \rightarrow \infty$ . Therefore  $S(z) = \zeta(z)$  in the whole region.

Especially, the value of (3.5.5) is

$$S(-\frac{1}{2}) = \zeta(-\frac{1}{2}) = -\zeta(\frac{3}{2})/4\pi.$$

The latter equality follows from the functional equation  $\zeta(1-s) = 2^{1-s}\pi^{-s}T(s)\cos\frac{1}{2}\pi s\zeta(s)$ .

### 3.6. The Euler-Maclaurin sum formula

Our considerations of sec. 3.5 were meant to demonstrate a method, rather than giving the shortest way to deal with  $\sum_1^n k^z$ . It seems that the shortest and most efficient way of dealing with such cases depends on the Euler-Maclaurin sum formula. It is incorporated in many textbooks of advanced analysis <sup>2)</sup>, and therefore we omit its proof.

The basic formula is

$$(3.6.1) \quad \left\{ \begin{aligned} & \frac{g(0) + g(1)}{2} - \int_0^1 g(x)dx = \\ & = \{g'(1) - g'(0)\} \frac{B_2}{2!} + \{g'''(1) - g'''(0)\} \frac{B_4}{4!} + \dots + \\ & + \{g^{(2m-1)}(1) - g^{(2m-1)}(0)\} \frac{B_{2m}}{(2m)!} - \int_0^1 g^{(2m)}(x) \frac{B_{2m}(x)}{(2m)!} dx. \end{aligned} \right.$$

Here  $m$  represents any integer  $\geq 1$ , and  $g$  is a function having  $2m$

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, *Modern Analysis*, 4th ed., Cambridge 1946, ch. 13.

<sup>2)</sup> Ibid., § 7.21. Our notation for the Bernoulli numbers and polynomials is slightly different, however.

continuous derivatives in the interval  $0 \leq x \leq 1$ . The  $B$ 's are the Bernoulli numbers, defined by

$$ze^{zt}/(e^z - 1) = \sum_0^{\infty} B_n z^n/n! \quad (|z| < 2\pi),$$

whence

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{4}, B_3 = B_5 = B_7 = \dots = 0,$$

$$B_4 = -\frac{1}{30}, B_6 = \frac{1}{42}, B_8 = -\frac{1}{30}, B_{10} = \frac{5}{66},$$

$$B_{12} = -\frac{691}{2730}, B_{14} = \frac{7}{8}, B_{16} = -\frac{3617}{510}, \dots$$

Finally, the  $B_n(t)$  denote the Bernoulli polynomials, defined by

$$ze^{zt}/(e^z - 1) = \sum_0^{\infty} B_n(t) z^n/n!.$$

If we write down (3.6.1) for the functions  $g(x) = f(x+1)$ ,  $g(x) = f(x+2)$ , ...,  $g(x) = f(x+n-1)$ , respectively, and if we add the results, many terms cancel out, and we get the Euler-Maclaurin sum formula. We write it in the form

$$(3.6.2) \quad \left\{ \begin{array}{l} f(1) + \dots + f(n) = \int_1^n f(x) dx + C + \frac{1}{2}f(n) + \\ \quad + \frac{B_2}{2!} f'(n) + \frac{B_4}{4!} f''(n) + \dots + \\ \quad + \frac{B_{2m}}{(2m)!} f^{(2m-1)}(n) - \int_1^n f^{(2m)}(x) \frac{B_{2m}(x - [x])}{(2m)!} dx. \end{array} \right.$$

The function  $f$  is assumed to have  $2m$  continuous derivatives if  $x \geq 1$ . The symbol  $[x]$  has the usual meaning of the largest integer  $\leq x$ .  $B_{2m}(x - [x])$  is the value of the  $2m$ -th Bernoulli polynomial at  $x - [x]$ . The number  $C$  is independent of  $n$ :

$$C = \frac{1}{2}f(1) - B_2 f'(1)/2! - \dots - B_{2m} f^{(2m-1)}(1)/(2m)!.$$

It is known that <sup>1)</sup>

$$B_{2m}(x - [x]) = 2(2m)! (2\pi)^{-2m} (-1)^{m+1} \sum_{k=1}^{\infty} k^{-2m} \cos(2k\pi x)$$

if  $m = 1, 2, 3, \dots$ , whence it follows that

$$|B_{2m}(x - [x])| \leq |B_{2m}| = 2(2m)! (2\pi)^{-2m} \sum_{k=1}^{\infty} k^{-2m}.$$

<sup>1)</sup> See W. ROGOSINSKI, Fourier Series, Chelsea, New York, ch. 2, § 4.

This gives a satisfactory estimate for the remainder term in (3.6.1).

If  $f(x)$  is such that  $\int_0^\infty |f^{(2m)}(x)|dx$  converges, we immediately have an asymptotic formula:

$$(3.6.3) \quad f(1) + \dots + f(n) = \int_1^n f(x)dx + S + \frac{1}{2}f(n) + \\ + \sum_{k=1}^m B_{2k} f^{(2k-1)}(n)/(2k)! + O(\int_n^\infty |f^{(2m)}(x)|dx) \quad (n \rightarrow \infty),$$

where  $m$  is a fixed positive integer, and

$$S = C - \int_1^\infty f^{(2m)}(x) \frac{B_{2m}(x - [x])}{(2m)!} dx.$$

### 3.7. Example

Let  $z$  be a complex number, and  $f(x) = x^{-z} \log x$ . Then (3.6.3) can be applied if  $2m > 1 - \operatorname{Re} z$ . It results that

$$\sum_{k=1}^n k^{-z} \log k = \int_1^n x^{-z} \log x dx + C(z) + \frac{1}{2}n^{-z} \log n + R(n; z),$$

where  $C(z)$  depends on  $z$  only, and  $R(n; z)$  has an asymptotic expansion

$$R(n; z) \approx \frac{B_2}{2!} (n^{-z} \log n)' + \frac{B_4}{4!} (n^{-z} \log n)''' + \dots \quad (n \rightarrow \infty).$$

The accents denote differentiation with respect to  $n$ , evaluated as if  $n$  were a continuous variable.

As in sec. 3.5,  $C(z)$  can be determined by an analyticity argument; we obtain  $C(z) = -(1-z)^{-2} - \zeta'(z)$ . The special case  $z = 0$  gives the Stirling formula for  $\log n!$ , as<sup>1)</sup>  $\zeta'(0) = -\frac{1}{2} \log 2\pi$ . It should be remarked, however, that there are many other methods for determining the value of  $C(0)$  (see sec. 3.10).

### 3.8. A remark

Roughly speaking, the Euler-Maclaurin method does not work if the largest term,  $f(n)$ , say, is not small compared to the sum  $f(1) +$

---

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, *Modern Analysis*, 4th ed., Cambridge 1946, § 13.21.

$+ \dots + f(n)$ . In that case one cannot expect the order of  $f^{(2m)}(n)$  to be lower than the one of  $f(n)$ , and so the Euler-Maclaurin formula does not give anything better than  $f(1) + \dots + f(n) = O(f(n))$ . One can illustrate this by the example  $\sum_1^n k!$  of sec. 3.3.

### 3.9. Another example

The Euler-Maclaurin method can also be applied to sums  $\sum_{k=1}^n a_k(n)$  where the terms depend both on  $k$  and  $n$ . There is, however, no point in passing from (3.6.2) to (3.6.3) in that case, for then  $S$  will depend on  $n$ . An unspecified constant may often be tolerated in an asymptotic formula, but having an unspecified function of  $n$  just means having no formula at all. There are some cases, however, where

$$\int_1^n f^{(2m)}(x) B_{2m}(x - [x])/(2m)! dx$$

raises no difficulties, e.g. when  $\int_1^n |f^{(2m)}(x)| dx$  is relatively small.

As an example we take

$$s_n = \sum_{k=-n}^n e^{-k^2\alpha/n},$$

where  $\alpha$  is a positive constant. The Euler-Maclaurin formula gives, if  $f(x) = e^{-x^2\alpha/n}$ ,

$$(3.9.1) \quad s_n = \int_{-n}^n f(x) dx + \frac{1}{2}f(n) + \frac{1}{2}f(-n) + B_2\{f'(n) - f'(-n)\}/2! + \dots + B_{2m}\{f^{(2m-1)}(n) - f^{(2m-1)}(-n)\}/(2m)! + R_n,$$

where

$$(3.9.2) \quad R_m = - \int_{-n}^n f^{(2m)}(x) \frac{B_{2m}(x - [x])}{(2m)!} dx,$$

whence

$$|R_m| \leq \frac{|B_{2m}|}{(2m)!} \int_{-n}^n |f^{(2m)}(x)| dx.$$

We have

$$\int_{-n}^n f(x) dx = \int_{-\infty}^{\infty} f(x) dx + \varepsilon_n = (\pi n/\alpha)^{1/2} + \varepsilon_n,$$

where the error term  $\epsilon_n$  is  $O(e^{-bn})$  with some positive constant  $b$ . Such errors are called *exponentially small*.

The other terms of (3.9.1), apart from  $R_n$ , are all exponentially small because of the fact that every derivative of  $f(x)$  is of the type  $P(x)f(x)$ , where  $P(x)$  is a polynomial. So everything depends on how well  $R_m$  can be estimated.

On substituting  $x = y(n/2\alpha)^{\frac{1}{2}}$  we obtain

$$(3.9.3) \quad \int_{-\infty}^{\infty} |f^{(2m)}(x)| dx = (2\alpha/n)^{m-\frac{1}{2}} \int_{-\infty}^{\infty} |(d/dy)^{2m} e^{-\frac{1}{2}y^2}| dy,$$

and it follows that  $|R_m| < C_m n^{\frac{1}{2}-m}$ , where  $C_m$  is positive and independent of  $n$ . Hence we have, for every  $m$ ,

$$(3.9.4) \quad s_n = (\pi n/\alpha)^{\frac{1}{2}} + O(n^{\frac{1}{2}-m}) \quad (n \rightarrow \infty).$$

In the case of  $s_n$  we accidentally have direct information from another source, viz. a theta function transformation formula, which gives a very good estimate. It is therefore interesting to compare this one to the result of the Euler-Maclaurin method.

For convenience we discuss the infinite sum instead of the finite one. (The difference between the two is exponentially small). Writing down the analogue of (3.9.1) for  $\sum_{-N}^N e^{-k^2\alpha/n}$ , and making  $N \rightarrow \infty$ , we obtain

$$(3.9.5) \quad S_n = \sum_{k=-\infty}^{\infty} e^{-k^2\alpha/n} = \\ = (\pi n/\alpha)^{\frac{1}{2}} - \int_{-\infty}^{\infty} f^{(2m)}(x) B_{2m}(x - [x])/(2m)! dx,$$

where again  $f(x) = e^{-x^2\alpha/n}$ .

We denote the integral by  $R^*$ ; it follows from (3.9.5) than  $R^*$  does not depend on  $m$ . What we shall call here the Euler-Maclaurin method consists of estimating

$$(3.9.6) \quad |R^*| \leq \frac{|B_{2m}|}{(2m)!} \int_{-\infty}^{\infty} |f^{(2m)}(x)| dx$$

and choosing  $m$  such that the right-hand-side is minimal.

The theta transformation formula gives <sup>1)</sup>

$$\sum_{k=-\infty}^{\infty} e^{-k^2 x/n} = (\pi n/\alpha)^{\frac{1}{2}} \sum_{k=-\infty}^{\infty} e^{-k^2 \pi^2 n/\alpha},$$

and therefore

$$(3.9.7) \quad -R^* = 2(\pi n/\alpha)^{\frac{1}{2}} e^{-\pi^2 n/\alpha} + O(n^{\frac{1}{2}} e^{-4\pi^2 n/\alpha}) \quad (n \rightarrow \infty).$$

We shall now investigate whether (3.9.6) can give anything as strong as this. It gives immediately that, for every  $m$ , we have  $R^* = O(n^{1-m})$ . For  $m$  fixed, this is very much weaker than (3.9.7), but by deriving uniform estimates and by taking  $m$  to be a suitable function of  $n$ , we can obtain a better result. There is of course no hope of proving (3.9.7), but it is interesting to see that the Euler-Maclaurin method can still show that  $R^* = O(ne^{-\pi^2 n/\alpha})$ , restricting the losses to the factor  $n^{\frac{1}{2}}$ .

If we adopt the definition

$$H_k(y) = (-1)^k e^{\frac{1}{2}y^2} (d/dy)^k e^{-\frac{1}{2}y^2}$$

for the Hermite polynomials, the integrand on the right-hand-side of (3.9.3) equals  $e^{-\frac{1}{2}y^2} |H_{2m}(y)|$ .

Using the integral representation

$$\begin{aligned} H_{2m}(y) &= (2\pi)^{-\frac{1}{2}} e^{\frac{1}{2}y^2} \left( \frac{d}{dy} \right)^{2m} \int_{-\infty}^{\infty} e^{-\frac{1}{2}v^2 + ivy} dv = \\ &= (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} (iv)^{2m} e^{-\frac{1}{2}v^2 + ivy + \frac{1}{2}y^2} dv = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} (y - iu)^{2m} e^{-\frac{1}{2}u^2} du, \end{aligned}$$

we infer that

$$\int_{-\infty}^{\infty} |f^{(2m)}(x)| dx \leq (2\alpha/n)^{m-\frac{1}{2}} (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(v^2 + u^2)} (y^2 + u^2)^m du dv.$$

Introducing polar coordinates in the  $u$ - $y$ -plane we easily find that the repeated integral equals  $2^{m+1}\pi m!$ .

The factor  $|B_{2m}|/(2m)!$  occurring in (3.9.6) is equal to  $2(2\pi)^{-2m}\zeta(2m)$  (see sec. 3.6), and therefore it is less than  $C(2\pi)^{-2m}$ ,

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, Modern Analysis, 4th ed., Cambridge 1946, § 21.51

A simple direct proof is obtained by taking  $f(x) = \exp(-\beta x^2)$ ,  $a = 0$  ( $\beta = \alpha/n$ ) in Poisson's formula (3.12.1).

where  $C$  is an absolute constant. It now follows from (3.9.6) that

$$|R^*| < C(2\pi)^{-2m}(2\alpha/n)^{m-\frac{1}{2}}(2\pi)^{-\frac{1}{2}2m+1}\pi m!$$

Using Stirling's formula for  $m!$  we infer that there is an absolute constant  $C_1$  such that for all  $m$  and  $n$

$$(3.9.8) \quad |R^*| < C_1(\alpha n/\pi^2 n e)^m (n m/2\alpha)^{\frac{1}{2}}.$$

It is now the right moment to fix the value of  $m$ . The minimum of  $(\alpha t/\pi^2 n e)^t$  is easily seen to be attained at  $t = \pi^2 n/\alpha$ , and the value is  $e^{-\pi^2 n/\alpha}$ . However,  $m$  has to be an integer, and so we shall take  $m = m_0 = [\pi^2 n/\alpha]$ . In order to analyse the difference it makes, we put

$$\psi(\rho) = \rho \log(\alpha\rho/\pi^2 e)$$

whose minimum is  $-\pi^2/\alpha$ , attained at  $\rho = \rho_0 = \pi^2/\alpha$ . We have  $\psi'(\rho_0) = 0$ , and hence

$$\psi(m_0/n) = \psi(\rho_0 + O(n^{-1})) = -\pi^2/\alpha + O(n^{-2}).$$

If we now choose  $m = m_0$ , (3.9.8) becomes

$$R^* = O(ne^{-\pi^2 n/\alpha}) \quad (n \rightarrow \infty).$$

### 3.10. The Stirling formula for the $\Gamma$ -function in the complex plane

In the following, the Euler-Maclaurin formula will play about the same rôle as it did in (3.9.5). We shall have a sum containing a parameter  $z$ . For a fixed value of  $z$  we shall increase the number of terms indefinitely, and only afterwards we allow  $|z|$  to tend to infinity (in (3.9.5) that parameter was  $n$ ).

Let  $z$  be a real or complex number, not lying on the negative part of the real axis, and not equal to zero. We shall apply the Euler-Maclaurin formula to the sum

$$S_n(z) = \sum_{k=1}^n \log(z + k - 1),$$

where the logarithms are given their principal values (imaginary parts of the logarithms absolutely less than  $\pi$ ). With an arbitrary integer  $m \geq 1$  we obtain

$$S_n(z) = \frac{1}{2} \log z + \frac{1}{2} \log(z+n-1) + \int_1^n \log(z+x-1) dx + \\ + \sum_{k=1}^m \{(z+n-1)^{1-2k} - z^{1-2k}\} (2k)^{-1} (2k-1)^{-1} B_{2k} + \\ + \int_1^n (z+x-1)^{-2m} (2m)^{-1} B_{2m}(x-[x]) dx.$$

With  $z$  fixed we easily get an asymptotic formula with an error term  $o(1)$ :

$$(3.10.1) \quad S_n(z) = (z - \frac{1}{2}) \log n - (z - \frac{1}{2}) \log z + n \log n + \\ + z - n - \rho(z) + o(1) \quad (n \rightarrow \infty),$$

where

$$(3.10.2) \quad \rho(z) = \sum_{k=1}^m z^{1-2k} (2k)^{-1} (2k-1)^{-1} B_{2k} - \\ - \int_0^\infty (z+x)^{-2m} (2m)^{-1} B_{2m}(x-[x]) dx.$$

As  $\rho(z)$  does not depend on  $n$ , it is obvious from (3.10.1) that it does not depend on  $m$  either. This fact can also be shown from (3.10.2), integrating by parts.

Taking the difference  $S_n(z) - S_n(1)$ , we obtain, applying (3.10.1) twice,

$$(3.10.3) \quad S_n(z) - S_n(1) = (z-1) \log n - (z-\frac{1}{2}) \log z + z-1 + \\ + \rho(1) - \rho(z) + o(1) \quad (n \rightarrow \infty).$$

This difference is connected with Euler's product formula for  $\Gamma(z)$ :

$$(3.10.4) \quad \Gamma(z) = \lim_{n \rightarrow \infty} n^{z-1} n! / z(z+1)(z+2) \cdots (z+n-1).$$

Taking logarithms, we find:

$$(3.10.5) \quad \log \Gamma(z) = \lim_{n \rightarrow \infty} \{(z-1) \log n + S_n(1) - S_n(z)\}.$$

Now (3.10.3) produces a useful identity:

$$(3.10.6) \quad \log \Gamma(z) = (z - \frac{1}{2}) \log z - z + \rho(z) + 1 - \rho(1).$$

It should be noted that  $\log \Gamma(z)$  does not necessarily represent the principal value of the logarithm. As a matter of fact the right-hand-side depends continuously on  $z$ , and is real if  $z > 0$  ( $\log z$  is given its principal value). Therefore, (3.10.6) represents the

analytic continuation of  $\log \Gamma(z)$  throughout the region  $|\arg z| < \pi$ , starting from real values on the positive real axis.

It is easy to derive an asymptotic formula from (3.10.6), when  $|z| \rightarrow \infty$ . Let  $\delta$  be a positive constant ( $\delta < \pi$ ), and let  $R_\delta$  denote the part of the complex plane which is defined by  $|\arg z| < \pi - \delta$ . Let  $m$  be a given integer  $\geq 1$ . Then  $B_{2m}(x - [x])$  is bounded, whence

$$\begin{aligned} |(z + x)^{-2m} (2m)^{-1} B_{2m}(x - [x]) dx| &< C \int_0^\infty |z + x|^{-2m} dx = \\ &= C |z|^{-2m+1} \int_0^\infty |y + (z|z|^{-1})|^{-2m} dy, \end{aligned}$$

$C$  not depending on  $z$ . Now  $|y + (z|z|^{-1})|$  represents the distance from the point  $-z$  to a point of the unit circle, belonging to the region  $R_\delta$ . It is easy to see geometrically that this distance is at least  $|y + e^{i(\pi-\delta)}|$ . Since  $\int_0^\infty |y + e^{i(\pi-\delta)}|^{-2m} dy$  converges, we infer that the integral in (3.10.2) is  $O(|z|^{1-2m})$ . As  $m$  is arbitrary, we obtain an asymptotic series for  $\rho(z)$ , and (3.10.6) gives

$$(3.10.7) \quad \begin{aligned} \log \Gamma(z) - (z - \frac{1}{2}) \log z + z &\approx 1 - \rho(1) + \\ &+ \sum_{k=1}^{\infty} z^{1-2k} (2k)^{-1} (2k-1)^{-1} B_{2k} \quad (|\arg z| < \pi - \delta, |z| \rightarrow \infty). \end{aligned}$$

It remains to be shown that the constant  $1 - \rho(1)$  equals  $\frac{1}{2} \log(2\pi)$ . We know already that  $1 - \rho(1)$  is real, so it suffices to show that  $e^{1-\rho(1)} = (2\pi)^{\frac{1}{2}}$ . We quote a number of possible methods. (i) In sec. 3.7 it was done with the aid of the Riemann  $\zeta$ -function, but this is certainly not the most elementary way. (ii) We can use the functional equation  $\Gamma(z)\Gamma(-z) = -\pi(z \sin z)^{-1}$ , and make  $z$  tend to infinity along the imaginary axis. (iii) We can use the functional equation<sup>1)</sup>

$$\Gamma(z)\Gamma(z + \frac{1}{2}) = \pi^{\frac{1}{2}} 2^{1-2z} \Gamma(2z),$$

making  $z \rightarrow +\infty$ . (iv) We can evaluate  $\rho(1)$  by evaluation of the integral  $\int_0^\infty (1+x)^2 B_2(x - [x]) dx$ . (v) We can use the beginning of sec. 4.5, with  $z \rightarrow +\infty$ .

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, Modern Analysis, 4th ed., Cambridge 1946, § 1215.

### 3.11. Alternating sums

An alternating sum is a sum of the type  $\sum (-1)^k f(k)$ , where the  $f(k)$  are positive. We usually expect such sums to be small, that is to say, much smaller than the sum of the absolute values of the terms.

We can of course write

$$\sum_{k=0}^{2m+1} (-1)^k f(k) = \sum_{k=0}^m f(2k) - \sum_{k=0}^m f(2k+1),$$

and investigate both sums on the right. Usually these sums will be about equal, whence it is desirable to study them quite closely in order to have sufficient information about their difference.

In most cases, however, the easiest thing to do, is to take pairs of terms together:

$$\sum_{k=0}^{2m+1} (-1)^k f(k) = \sum_{k=0}^m \{f(2k) - f(2k+1)\},$$

and these terms  $f(2k) - f(2k+1)$  will usually be small.

As an example we take the infinite sum

$$(3.11.1) \quad S(t) = \sum_{k=0}^{\infty} (-1)^k f(k), \quad f(x) = (x^2 + t^2)^{-\frac{1}{2}},$$

and we ask for the asymptotic behaviour of  $S(t)$  as  $t \rightarrow \infty$ . The function  $f(x)$  is decreasing, and tends to zero as  $x \rightarrow \infty$ . Therefore the series converges, and we have, by a well-known theorem on alternating series,  $0 < S(t) < f(0)$ . Thus a rough first result is that  $S(t) = O(t^{-1})$ .

We next write

$$S(t) = \sum_{k=0}^{\infty} \{f(2k) - f(2k+1)\}.$$

We shall, of course, compare the difference  $f(2k+1) - f(2k)$  with  $f'(2k)$ , and after that, we shall compare the sum  $-\sum_0^{\infty} f'(2k)$  with the integral  $-\frac{1}{2} \int_0^{\infty} f'(x) dx$  (the factor  $\frac{1}{2}$  arises because  $2k$  only runs through the even numbers). We can carry out these two operations at the same time, comparing

$$f(2k) - f(2k+1) \quad \text{with} \quad -\frac{1}{2} \int_{2k}^{2k+2} f'(x) dx.$$

Using the Taylor series, we can express both in terms of  $f(2k)$ ,  $f'(2k), \dots$ . If we stop the Taylor developments at the terms in-

volving  $f''$ , i.e. if we apply the formula

$$\varphi(a+h) - \varphi(a) = h\varphi'(a) + \int_a^{a+h} (a+h-x)\varphi''(x)dx,$$

then we obtain

$$f(2k+1) - f(2k) = f'(2k) + \int_{2k}^{2k+1} (2k+1-x)f''(x)dx,$$

$$\frac{1}{2} \int_{2k}^{2k+2} f'(x)dx = f'(2k) + \frac{1}{2} \int_{2k}^{2k+2} (2k+2-x)f''(x)dx.$$

On subtraction we find

$$f(2k) - f(2k+1) + \frac{1}{2} \int_{2k}^{2k+1} f'(x)dx = \frac{1}{2} \int_{2k}^{2k+2} (1 - |x - 2k - 1|)f''(x)dx,$$

and so

$$(3.11.2) \quad |f(2k) - f(2k+1) + \frac{1}{2} \int_{2k}^{2k+2} f'(x)dx| \leq \frac{1}{2} \int_{2k}^{2k+2} |f''(x)|dx.$$

In our case we have  $f(x) \rightarrow 0$  when  $x \rightarrow \infty$ , and therefore

$$\sum_{k=0}^{\infty} \int_{2k}^{2k+2} f'(x)dx = \int_0^{\infty} f'(x)dx = -f(0).$$

It follows that

$$(3.11.3) \quad |S(t) - \frac{1}{2}f(0)| \leq \frac{1}{2} \int_0^{\infty} |f''(x)|dx.$$

We have  $f''(x) = (2x^2 - t^2)/(x^2 + t^2)^{3/2}$ . We transform the integral, substituting  $x = yt$ :

$$\int_0^{\infty} |f''(x)|dx = t^{-2} \int_0^{\infty} |1 - 2y^2|/(1 + y^2)^{3/2} dy,$$

and the latter integral is easily seen to be convergent. Therefore, (3.11.3) gives

$$(3.11.4) \quad S(t) = \frac{1}{2}t^{-1} + O(t^{-2}) \quad (t \rightarrow \infty).$$

The process which led to (3.11.2), can of course be continued: in the next step we use the Taylor expansions up to the terms involving  $f'''(x)$ . And in order to eliminate the term involving  $f'''(2k)$ , we subtract a suitable multiple of  $\int_{2k}^{2k+2} f''(x)dx$ , in the same way as we eliminated  $-f'(2k)$  by subtracting  $-\frac{1}{2} \int_{2k}^{2k+2} f'(x)dx$ . This time we

use the formula

$$\varphi(a+h) = \varphi(a) + h\varphi'(a) + \frac{1}{2}h^2\varphi''(a) + \int_a^{a+h} \frac{1}{2}(a+h-x)^2\varphi'''(x)dx,$$

and we obtain

$$(3.11.5) \quad |f(2k) - f(2k+1) + \frac{1}{2} \int_{2k}^{2k+2} f'(x)dx - \frac{1}{2} \int_{2k}^{2k+2} f''(x)dx| \leq \\ \leq C \int_{2k}^{2k+2} |f'''(x)|dx.$$

As  $\int_0^\infty f''(x)dx = -f'(0) = 0$ , we now obtain, in the same way as above,

$$(3.11.6) \quad S(t) = \frac{1}{2}t^{-1} + O(t^{-3}) \quad (t \rightarrow \infty).$$

There is no term with  $t^{-2}$ , and in the next steps of the procedure it will turn out that the coefficients of  $t^{-3}, t^{-4}, \dots$  vanish also. In order to show this, it is easier to put the series in the following form:

$$S(t) = \frac{1}{2}t^{-1} + \frac{1}{2} \sum_{k=-\infty}^{\infty} (-1)^k f(k).$$

Applying (3.11.5) to  $\sum_{-\infty}^{\infty}$ , we obtain, as  $t \rightarrow \infty$ ,

$$S(t) - \frac{1}{2}t^{-1} = -\frac{1}{2} \int_{-\infty}^{\infty} f'(x)dx - \frac{1}{2} \int_{-\infty}^{\infty} f''(x)dx + O(\int_{-\infty}^{\infty} f'''(x)dx).$$

Since  $f(x) \rightarrow 0, f'(x) \rightarrow 0, f''(x) \rightarrow 0, \dots$  as  $x \rightarrow \pm \infty$ , we have

$$\int_{-\infty}^{\infty} f'(x)dx = \int_{-\infty}^{\infty} f''(x)dx = \int_{-\infty}^{\infty} f'''(x)dx = \dots = 0.$$

Furthermore it is easily seen, by substitution of  $x = yt$ , that

$$\int_{-\infty}^{\infty} |f^{(m)}(x)|dx = O(t^{-m}) \quad (t \rightarrow \infty)$$

for every fixed  $m > 0$ . Now it is sufficient to have only a general idea about the continuation of the process which led to (3.11.4) and (3.11.6) in order to see that

$$(3.11.7) \quad S(t) \approx \frac{1}{2}t^{-1} + 0 \cdot t^{-2} + 0 \cdot t^{-3} + 0 \cdot t^{-4} + \dots \quad (t \rightarrow \infty).$$

It may be remarked that the general formula of which (3.11.2), (3.11.5) are special cases, is related in a trivial way to the Boole sum formula<sup>1)</sup>, which we shall not discuss here.

<sup>1)</sup> See C. JORDAN, Calculus of Finite Differences, 2nd ed., New York 1947, § 112.

A second remark is that (3.11.7) can also be derived by the Euler-Maclaurin sum formula, applied separately to  $\sum f(2k)$  and to  $\sum f(2k + 1)$ .

With (3.11.7) we have the same situation as in sec. 3.9. We expect  $S(t) - \frac{1}{2}t^{-1}$  to be exponentially small, and by a careful inspection of the above argument, including estimates holding uniformly in  $m$  and  $t$ , we might be able to show this, although the formulas become quite awkward. But even then we would have only an upper estimate for  $S(t) - \frac{1}{2}t^{-1}$ , and no asymptotic formula, like the one we shall derive in sec. 3.12.

### 3.12. Application of the Poisson sum formula

The formula reads

$$(3.12.1) \quad \sum_{k=-\infty}^{\infty} f(k+a) = \sum_{\nu} e^{2\pi\nu ia} \int_{-\infty}^{\infty} e^{-2\pi\nu iy} f(y) dy,$$

where  $a$  is a real number,  $f(x)$  is Riemann integrable over any finite interval, and

$$\sum_{\nu} \text{ denotes } \lim_{N \rightarrow \infty} \sum_{\nu=-N}^N .$$

The following set of conditions is easily seen to be sufficient<sup>1)</sup>:

- (i)  $\sum_{k=-\infty}^{\infty} f(k+x)$  converges uniformly for  $0 \leq x \leq 1$ .
- (ii) The function  $\phi(x) = \sum_{k=-\infty}^{\infty} f(k+x)$ , which has period 1, satisfies the Fourier conditions (that is,  $\phi(x)$  is the sum of its Fourier series), at least at  $x = a$ .

For, condition (i) enables us to carry out the following operation with the Fourier coefficients of  $\phi(x)$ . If  $\nu$  is any integer, then the  $\nu$ -th Fourier coefficient of  $\phi$  equals

$$\begin{aligned} \int_0^1 e^{-2\pi\nu iy} \phi(y) dy &= \int_0^1 \sum_{k=-\infty}^{\infty} e^{-2\pi\nu iy} f(k+y) dy = \sum_{k=-\infty}^{\infty} \int_0^1 e^{-2\pi\nu iy} f(y) dy = \\ &= \sum_{k=-\infty}^{\infty} \int_k^{k+1} e^{-2\pi\nu iy} f(y) dy = \int_{-\infty}^{\infty} e^{-2\pi\nu iy} f(y) dy. \end{aligned}$$

<sup>1)</sup> See for other sets of sufficient conditions: E. C. TITCHMARSH, Fourier Integrals, 2nd ed., Oxford 1948, ch. 2.

The last step is legitimate, as (i) implies that

$$\int\limits_k^{k+t} e^{-2\pi ty} f(y) dy \rightarrow 0$$

when  $k \rightarrow \pm \infty$ , uniformly in  $0 \leq t \leq 1$ .

The following set of conditions can be shown to imply (i) and (ii):

(iii)  $\sum_{-\infty}^{\infty} f(k+a)$  converges,

(iv)  $f'(x)$  exists ( $-\infty < x < \infty$ ),

(v)  $\sum_{-\infty}^{\infty} f'(k+x)$  converges uniformly in  $0 \leq x \leq 1$ .

For, (iii)+(v) imply (i) (apply the mean value theorem to finite sums  $\sum_N^M f(k+x)$ , with  $M$  and  $N$  either both positive and large or both negative and large), and (v) shows that  $\phi(x)$  is differentiable everywhere, whence it satisfies the Fourier conditions.

Another set of sufficient conditions is (iii)+(vi)+(vii), where (vi)  $f(x)$  has bounded variation<sup>1)</sup> over  $-\infty < x < \infty$ ,

(vii)  $\lim_{h \rightarrow 0} \{f(x+h) + f(x-h)\} = 2f(x)$ , at least for all  $x$  of the form  $a+n$ , where  $n$  is any integer.

We remark that from (iii)+(vi) one can deduce (i)<sup>2)</sup>, as well as the fact that  $\phi(x)$  has bounded variation over  $0 \leq x \leq 1$ ; (vi)+(vii) can be used to show that

$$\lim_{h \rightarrow 0} \{\phi(a+h) + \phi(a-h)\} = 2\phi(a).$$

This formula, in combination with the fact that  $\phi(x)$  has bounded variation, leads again to (ii).

As it is not our present aim to develop Fourier theory here, we leave it at these brief remarks.

We shall apply the Poisson formula to the sum

$$(3.12.2) \quad S_1(t) = \sum_{k=-\infty}^{\infty} f(k), \quad f(x) = e^{\pi i x}(x^2 + t^2)^{-\frac{1}{2}},$$

which is related to (3.11.1) by the formula

$$(3.12.3) \quad S(t) = \frac{1}{2}t^{-1} + \frac{1}{2}S_1(t).$$

The number  $a$  occurring in (3.12.1) will be given the special value 0 here, and, in applying (3.12.1) to (3.12.2),  $t$  is considered to be a fixed positive number.

<sup>1)</sup> See for definition of bounded variation: E. C. TITCHMARSH, Theory of Functions, 2nd ed., Oxford 1939, § 11.4.

<sup>2)</sup> Ibid., § 13.232.

The condition (vi) is not satisfied, but the set (iii)+(iv)+(v) is. Condition (iii) was already checked in the beginning of sec. 3.11, and (iv) is trivial. In order to show (v), we write

$$f'(x) = \pi i e^{\pi i x} (x^2 + t^2)^{-\frac{1}{2}} - x e^{\pi i x} (x^2 + t^2)^{-\frac{3}{2}}.$$

We take two positive integers  $N, M$ , where  $t < N < M$ , and a real number  $x$  in  $0 \leq x \leq 1$ . Then we consider

$$\begin{aligned} & \sum_{k=N}^M f'(x+k) = \\ & = e^{\pi i x} \sum_{k=N}^M (-1)^k \{ \pi i ((x+k)^2 + t^2)^{-\frac{1}{2}} - (x+k)((x+k)^2 + t^2)^{-\frac{3}{2}} \}. \end{aligned}$$

The numbers  $\{(x+k)^2 + t^2\}^{-\frac{1}{2}}$  form a decreasing sequence of  $(M-N+1)$  positive numbers, and the same thing holds for  $(x+k)\{(x+k)^2 + t^2\}^{-\frac{3}{2}}$ . (The function  $y(y^2 + t^2)^{-\frac{3}{2}}$  decreases from  $y = 2^{-\frac{1}{2}}t$  onward). We now use the following well-known fact: If any sequence  $a_N, \dots, a_M$  satisfies  $a_N > a_{N+1} > \dots > a_M > 0$ , then we have

$$|\sum_{k=N}^M (-1)^k a_k| \leq a_N.$$

It follows that

$$|\sum_{k=N}^M f'(x+k)| \leq (N^2 + t^2)^{-\frac{1}{2}} + N(N^2 + t^2)^{-\frac{3}{2}} < 2N^{-1}.$$

As this holds uniformly in  $x$  ( $0 \leq x \leq t$ ), we infer that, for  $t$  fixed,  $\sum_1^\infty f'(x+k)$  converges uniformly in  $0 \leq x \leq 1$ . The same thing can be said about  $\sum_{-\infty}^0$ , and so we have proved (v).

We can now apply (3.12.1) to (3.12.2); the result is that

$$(3.12.4) \quad S_1(t) = \sum_{y=-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-2\pi y t} e^{\pi i y} (y^2 + t^2)^{-\frac{1}{2}} dy,$$

and so we have to study, for  $b = \pm \pi, \pm 3\pi, \pm 5\pi, \dots$ , the integral

$$\int_{-\infty}^{\infty} e^{btz} (z^2 + 1)^{-\frac{1}{2}} dz.$$

On substitution of  $y = iz$  we observe that this integral is a function of  $bt$ , to be denoted by  $\varphi$ :

$$(3.12.5) \quad \varphi(bt) = \int_{-\infty}^{\infty} e^{btz} (z^2 + 1)^{-\frac{1}{2}} dz,$$

and by substitution of  $z = -w$  we infer that  $\varphi$  is an even function. The integral is a Bessel function of zero order, of second kind and of

imaginary argument, and in the standard notation<sup>1)</sup>

$$\varphi(bt) = 2K_0(bt).$$

We shall, however, not explicitly use the theory of Bessel functions here.

Using Cauchy's theorem, the integral (3.12.5) can be transformed. Assuming  $b > 0$ , we deform the integration path  $(-\infty, \infty)$  into a new one, consisting of the following parts: (i) The line  $(-\infty, -R)$ , where  $R$  is a large positive number; (ii) The circular arc  $z = Re^{i\theta}$  ( $\pi \geq \theta \geq \frac{1}{2}\pi$ ); (iii) The segment  $z = is$  ( $R \geq s \geq 1 + \delta$ ), where  $\delta$  is a small positive number; (iv) The circle  $z = i + \delta e^{i\theta}$  ( $\frac{1}{2}\pi \leq \theta \leq \frac{5}{2}\pi$ ); (v) The segment  $z = is$  ( $1 + \delta \leq s \leq R$ ); (vi) The arc  $z = Re^{i\theta}$  ( $\frac{1}{2}\pi \geq \theta \geq 0$ ); (vii) The interval  $(R, \infty)$ . Along (iii), the function  $(1 + z^2)^{-\frac{1}{2}}$  has to be interpreted as  $i(s^2 - 1)^{-\frac{1}{2}}$ , where  $(s^2 - 1)^{\frac{1}{2}}$  is the positive root, and along (v) it has to be interpreted as  $-i(s^2 - 1)^{-\frac{1}{2}}$ . Now making  $R \rightarrow \infty$ ,  $\delta \rightarrow 0$  we infer that

$$\varphi(bt) = 2 \int_1^\infty e^{-bts} (s^2 - 1)^{-\frac{1}{2}} ds = 2e^{-bt} \int_0^\infty e^{-btu} (u^2 + 2u)^{-\frac{1}{2}} du.$$

(The latter integral is obtained from the former by the substitution  $s = u + 1$ ). We have, if  $t > 1$ ,  $b \geq \pi$

$$\int_0^\infty e^{-btu} (u^2 + 2u)^{-\frac{1}{2}} du < \int_0^\infty e^{-\pi u} u^{-\frac{1}{2}} du = 1,$$

and so  $|\varphi(bt)| < 2e^{-bt}$  ( $b \geq \pi$ ,  $t \geq 1$ ). By (3.12.4) we now have

$$|S_1(t) - 2\varphi(\pi t)| \leq |2\varphi(3\pi t) + 2\varphi(5\pi t) + \dots| \leq 2e^{-3\pi t}/(1 - e^{-2\pi t}),$$

and therefore

$$(3.12.6) \quad S_1(t) = 2\varphi(\pi t) + o(e^{-\pi t}) \quad (t \rightarrow \infty).$$

It remains to find the asymptotic behaviour of  $\varphi(\pi t)$ . To this end we write, putting  $u = x^2$ ,

$$\varphi(\pi t) = 2e^{-\pi t} \int_0^\infty e^{-\pi tu} (u^2 + 2u)^{-\frac{1}{2}} du = 2e^{-\pi t} \int_{-\infty}^\infty e^{-\pi tx^2} (x^2 + 2)^{-\frac{1}{2}} dx,$$

in order to be able to apply the method of sec. 4.1. The result is

<sup>1)</sup> See G. N. WATSON, Theory of Bessel Functions, 2nd ed., Cambridge 1952, § 6.16.

(cf. the derivation of (4.1.10))

$$\varphi(\pi t) \approx 2e^{-\pi t} (2\pi t)^{-\frac{1}{2}} \sum_{n=0}^{\infty} d_n \Gamma(n + \frac{1}{2}) \pi^{-n} t^{-n} \quad (t \rightarrow \infty),$$

where the coefficients  $d_n$  are those of

$$(2+x^2)^{-\frac{1}{2}} = 2^{-\frac{1}{2}} \sum_{n=0}^{\infty} d_n x^{2n} \quad (|x^2| < 2).$$

For every  $M$  the term  $O(e^{-3\pi t})$  in (3.12.6) is  $O(e^{-\pi t} t^{-\frac{1}{2}-M})$ , and therefore  $S_1(t)$  has, apart from the factor 2, the same asymptotic power series expansion as  $\varphi(\pi t)$ . So our final result is, as  $d_n = (-1)^n 2^{-3n} (2n)!/(n!)^2$ :

$$S(t) - \frac{1}{2}t^{-1} = \frac{1}{2}S_1(t) \approx e^{-\pi t} \sum_{n=0}^{\infty} (-1)^n t^{-n-\frac{1}{2}} 2^{1-5n} \pi^{-n} \{(2n)!\}^2 \{n!\}^{-3} \quad (t \rightarrow \infty).$$

### 3.13. Summation by parts

We often meet the question of the asymptotic behaviour, as  $n \rightarrow \infty$ , of a sum  $a_1 b(1) + \dots + a_n b(n)$ , where the behaviour of  $a_1 + \dots + a_n$  is known and where the function  $b(x)$  behaves smoothly. Then we can usually apply the formula for summation by parts:

$$(3.13.1) \quad a_1 b_1 + \dots + a_n b_n = A_1(b_1 - b_2) + A_2(b_2 - b_3) + \dots + A_{n-1}(b_{n-1} - b_n) + A_n b_n,$$

where  $A_k = a_1 + a_2 + \dots + a_k$ , and  $b_k$  is an abbreviation for  $b(k)$ . It has some formal advantages to write the formula in terms of integrals. We assume, for simplicity, that  $b(x)$  has a continuous derivative, and we put  $A(x) = \sum_{1 \leq m \leq x} a_m$  (i.e.  $A(x) = 0$  if  $x < 1$ ,  $A(x) = a_1 + \dots + a_k$  if  $k \leq x < k+1$  ( $k = 1, 2, \dots$ )). Then (3.13.1) becomes

$$(3.13.2) \quad a_1 b(1) + \dots + a_n b(n) = A(n) b(n) - \int_1^n A(x) b'(x) dx.$$

As an example we consider the sum  $\sum_{k=1}^n \sin(kt) \log k$ , where  $t$  is a real constant. We put  $\sin kt = a_k$ ,  $\log x = b(x)$ . For every  $n$  we have

$$\sum_{k=1}^n \sin kt = \operatorname{Im} \sum_{k=1}^n e^{ikt} = \operatorname{Im} \{(e^{i(n+1)t} - e^{it})/(e^{it} - 1)\}.$$

Therefore, there is a number  $C > 0$ , not depending on  $n$ , such that

$|\sum_1^n \sin kt| < C$ . (If  $e^{it} - 1 = 0$ , our argument fails, but the result still holds, for then  $\sin kt = 0$  for all  $k$ ). Putting  $A(x) = \sum_{1 \leq k \leq x} \sin kt$ , we have, by (3.13.2),

$$\sum_{k=1}^n \sin(kt) \log k = A(n) \log n - \int_1^n A(x)x^{-1}dx.$$

Since  $|A(x)| < C$  for all  $x$ , we now easily derive that

$$\sum_{k=1}^n \sin(kt) \log k = O(\log n) \quad (t \text{ fixed}, n \rightarrow \infty).$$

We shall discuss a second example, taken from the theory of primes. We take  $a_n = \log n$  if  $n$  is a prime number, and  $a_n = 0$  otherwise. Then  $A(x)$  is the function usually denoted by  $\vartheta(x)$ , and we can write

$$\vartheta(x) = \sum_{p \leq x} \log p.$$

It is a fundamental and far from trivial result of the theory of primes that, for each  $m$  ( $m = 1, 2, 3, \dots$ ), we have<sup>1)</sup>

$$(3.13.3) \quad \vartheta(x) = x + O(x(\log x)^{-m}) \quad (x \rightarrow \infty).$$

Now many other sums involving primes, as  $\sum_{p \leq x} p^{-1}$ ,  $\sum_{p \leq x} p^2$ ,  $\sum_{p \leq x} 1$ , can be dealt with. We consider  $\sum_{p \leq x} 1$ , i.e. the number of primes not exceeding  $x$ . This number is usually denoted by  $\pi(x)$ . We have, by (3.13.2), taking  $b(x) = (\log x)^{-1}$ ,

$$\begin{aligned} \pi(n) &= a_2 b(2) + \dots + a_n b(n) = \\ &= (\log n)^{-1} \vartheta(n) - \int_2^n \vartheta(x) d(\log x)^{-1}. \end{aligned}$$

(We have replaced the lower limit 1 by 2, since  $(\log x)^{-1}$  is singular at  $x = 1$ ; it makes no difference, as  $\vartheta(1) = 0$ ). We compare this with

$$\int_2^n (\log x)^{-1} dx = [x(\log x)^{-1}]_2^n - \int_2^n x d(\log x)^{-1}.$$

On subtraction we obtain, using (3.13.3),

$$\begin{aligned} \pi(n) - \int_2^n (\log x)^{-1} dx &= (\log n)^{-1} O(n(\log n)^{-m}) + \\ &\quad + \int_2^n O(x(\log x)^{-m}) (\log x)^{-2} x^{-1} dx \quad (n \rightarrow \infty). \end{aligned}$$

<sup>1)</sup> See A. E. INGHAM, The Distribution of Primes, Cambridge 1932, p. 12 and p. 63.

The integral on the right can be written as

$$\int_2^{n^{\frac{1}{k}}} O(1) dx + \int_{n^{\frac{1}{k}}}^n O((\log n^{\frac{1}{k}})^{m-2}) dx = O(n(\log n)^{-m+2}) \quad (n \rightarrow \infty),$$

and therefore

$$\pi(n) - \int_2^n (\log x)^{-1} dx = O(n(\log n)^{-m+1}) \quad (n \rightarrow \infty).$$

The integral on the left can easily be expanded in the form of an asymptotic series (cf. (1.5.5)), and we infer that

$$\begin{aligned} \pi(n) \approx & n \log^{-1} n + n \log^{-2} n + 2! n \log^{-3} n + \\ & + 3! n \log^{-4} n + \dots \quad (n \rightarrow \infty). \end{aligned}$$

Meanwhile we notice that (3.13.3) is an example of the situation described in secs. 3.9 and 3.11. Again there is an asymptotic expansion with zero coefficients:

$$e^{-x}\theta(e^x) - 1 \approx 0 \cdot x^{-1} + 0 \cdot x^{-2} + 0 \cdot x^{-3} + \dots \quad (x \rightarrow \infty),$$

but in this case the question as to whether the left hand side is exponentially small, is still unsolved.

**3.14. EXERCISES.** 1. Show (e.g. by the method of summation by parts) that

$$\sum_{n=1}^{\infty} n^{-1} e^{-nt} = -\frac{1}{2} \log t + \frac{1}{2}\gamma + O(t^{\frac{1}{2}}) \quad (0 < t < 1),$$

where  $\gamma = -\int_0^{\infty} e^{-x} \log x dx$  is Euler's constant.

2. Show that

$$\sum_{n=0}^{\infty} \log(1 - be^{-nt}) \approx c_{-1}t^{-1} + c_0 + c_1t + c_2t^2 + \dots \quad (t > 0, t \rightarrow 0).$$

Here  $b$  is a constant,  $0 < b < 1$ , and

$$c_k = (-1)^k (B_{k+1}/(k+1)!) \sum_{n=1}^{\infty} b^n n^{k-1} \quad (k = -1, 0, 1, 2, \dots).$$

3. Show that

$$\begin{aligned} \sum_{n=1}^{\infty} \log(1 - e^{-nt}) \approx & -\pi^2/(6t) - \frac{1}{2} \log t + \frac{1}{2} \log(2\pi) + \frac{1}{24}t + \\ & + 0 \cdot t^2 + 0 \cdot t^3 + \dots \quad (t > 0, t \rightarrow 0). \end{aligned}$$

(Hint: Apply the Euler-Maclaurin formula to  $\sum_0^N f(n)$ , where

$$f(x) = \log((1 - e^{-xt})/xt)).$$

4. Let  $s_n$  be defined by

$$s_n = \frac{1}{\log n} - \frac{1}{\log(n+1)} + \frac{1}{\log(n+2)} - \dots$$

Show that

$$s_n = \frac{1}{2}(\log n)^{-1} + O(n^{-1} \log^{-2} n) \quad (n \rightarrow \infty).$$

5. Derive from (3.13.3) that

$$\sum_p e^{-pt} \approx - \sum_{n=0}^{\infty} t^{-1} (\log t)^{-n-1} \int_0^{\infty} e^{-y} (\log y)^n dy \quad (t > 0, t \rightarrow 0),$$

if the summation variable  $p$  runs through all prime numbers (cf. sec. 1.8, exerc. 4).

## CHAPTER 4

### THE LAPLACE METHOD FOR INTEGRALS

#### 4.1. Introduction

We shall consider integrals over real intervals, where both the integration interval and the integrand may depend on a parameter  $t$ , and we shall ask for the asymptotic behaviour of the integral as  $t \rightarrow \infty$ . We can, of course, extend the interval to the whole line  $(-\infty, \infty)$ , by defining the integrand to be zero in all points outside the original interval. So we have to deal with

$$(4.1.1) \quad F(t) = \int_{-\infty}^{\infty} \varphi(x, t) dx \quad (t \rightarrow \infty).$$

It often occurs that the graph of  $\varphi(x, t)$ , considered as a function of  $x$ , has somewhere a sharp peak, and that the contribution of some neighbourhood of the peak is almost equal to the whole integral, when  $t$  is large. Then we can try to approximate  $\varphi$  in that neighbourhood by simpler functions, for which the integral can be evaluated. This idea is due to Laplace. The advantage is that we only need to approximate in a relatively small interval.

As an example of the method we roughly sketch how to deal with the integral

$$(4.1.2) \quad F(t) = \int_{-\infty}^{\infty} e^{-tx^2} \log(1 + x + x^2) dx.$$

If  $t$  is large, the integrand is very small unless  $x$  is very close to zero. The function  $\log(1 + x + x^2)$  can be successfully approximated by simpler functions if  $-\frac{1}{2} < x < \frac{1}{2}$ , say, and therefore we first try to prove that  $\int_{-\infty}^{-\frac{1}{2}}$  and  $\int_{\frac{1}{2}}^{\infty}$  are very small. Next we remark that

$$\log(1 + x + x^2) = \log \frac{1 - x^3}{1 - x} = x + \frac{1}{2}x^2 - \frac{3}{8}x^3 + O(x^4) \quad (-\frac{1}{2} < x < \frac{1}{2}).$$

So we are led to approximate  $F(t)$  by

$$(4.1.3) \quad \int_{-\frac{1}{t}}^{\frac{1}{t}} e^{-tx^2} (x + \frac{1}{2}x^2 - \frac{1}{3}x^3) dx.$$

It may be remarked that the terms  $x$  and  $-2x^3/3$  give no contribution at all to this integral, these functions being odd.

As  $e^{-tx^2}$  is very small if  $x > \frac{1}{t}$  or  $x < -\frac{1}{t}$ , it is possible to show that it does not make much difference if in (4.1.3) the integration is taken over  $(-\infty, \infty)$  instead of over  $(-\frac{1}{t}, \frac{1}{t})$ .

In the above argument the idea of comparing  $(-\infty, \infty)$  with  $(-\frac{1}{t}, \frac{1}{t})$  occurred twice. It is, however, possible to present the method in such a way that this cutting-off is suppressed entirely (though it is not always practical to do so): Put

$$\log(1 + x + x^2) - x - \frac{1}{2}x^2 + \frac{1}{3}x^3 = R(x).$$

Then we can show that  $\int_{-\infty}^{\infty} e^{-tx^2} R(x) dx$  is small, by virtue of the estimate

$$R(x) = O(x^4) \quad (-\infty < x < \infty).$$

Putting  $tx^2 = y$ , we infer that

$$\int_{-\infty}^{\infty} e^{-tx^2} x^4 dx = t^{-5/2} \int_0^{\infty} e^{-y} y^{3/2} dy = O(t^{-5/2}) \quad (0 < t < \infty).$$

As  $\int_{-\infty}^{\infty} e^{-tx^2} x^2 dx = t^{-3/2} \int_0^{\infty} e^{-y} y^{1/2} dy = \frac{1}{2}t^{-3/2}\pi^{1/2}$  ( $t > 0$ ), it follows that

$$(4.1.4) \quad F(t) = \frac{1}{4}t^{-3/2}\pi^{1/2} + O(t^{-5/2}) \quad (t \rightarrow \infty).$$

The integrals of the type  $\int_{-\infty}^{\infty} e^{-tx^2} x^k dx$  will occur quite often; for future reference we give some formulas here. If  $k$  is an odd positive integer the integral vanishes as the integrand is odd:

$$(4.1.5) \quad \int_{-\infty}^{\infty} e^{-tx^2} x^k dx = 0 \quad (k = 1, 3, 5, \dots).$$

If  $k$  is even, the substitution  $tx^2 = y$  leads to a gamma integral:

$$(4.1.6) \quad \int_{-\infty}^{\infty} e^{-tx^2} x^{2n} dx = t^{-n-\frac{1}{2}} \Gamma(n+\frac{1}{2}) = t^{-n-\frac{1}{2}} \frac{(2n)!}{n!2^{2n}} \pi^{1/2} \quad (n = 0, 1, 2, \dots).$$

These formulas are valid if  $t > 0$ , but also if  $t$  is a complex number with a positive real part.

Sometimes we shall need the following estimate, both for  $k$  odd and  $k$  even:

$$(4.1.7) \quad \int_{-\infty}^{\infty} |e^{-tx^k} x^k| dx = O\{(\operatorname{Re} t)^{-\frac{1}{2}(k+1)}\} \quad (\operatorname{Re} t > 0).$$

This estimate is not uniform with respect to  $k$  ( $k = 0, 1, 2, \dots$ ).

The simpler formulas involving  $e^{-tx} x^k$  will also be frequently applied. If  $k = 0, 1, 2, \dots$ , we have

$$(4.1.8) \quad \int_0^{\infty} e^{-tx} x^k dx = t^{-k-1} k! \quad (\operatorname{Re} t > 0),$$

$$(4.1.9) \quad \int_0^{\infty} |e^{-tx} x^k| dx = O\{(\operatorname{Re} t)^{-k-1}\} \quad (\operatorname{Re} t > 0),$$

and again the estimate does not hold uniformly with respect to  $k$ .

It is quite easy to replace (4.1.4) by an asymptotic series. To this end we remark that

$$\begin{aligned} \log(1 + x + x^2) &= (x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \dots) - (x^3 + \frac{1}{2}x^6 + \dots) = \\ &= c_1 x + c_2 x^2 + c_3 x^3 + \dots \quad (-1 < x < 1). \end{aligned}$$

So, if  $M$  is arbitrary but fixed, we have

$$\begin{aligned} \log(1 + x + x^2) - (c_1 x + \dots + c_{2M-1} x^{2M-1}) &= O(x^{2M}) \\ &\quad (-\frac{1}{2} < x < \frac{1}{2}), \end{aligned}$$

and in the regions  $\frac{1}{2} \leq x < \infty$  and  $-\infty < x \leq -\frac{1}{2}$  this formula is still valid. Using (4.1.5) and (4.1.6) we infer that

$$F(t) = \sum_{n=1}^{M-1} c_{2n} t^{-n-\frac{1}{2}} \Gamma(n + \frac{1}{2}) + O(t^{-M-\frac{1}{2}}) \quad (t \rightarrow \infty),$$

and we have the asymptotic series

$$(4.1.10) \quad F(t) \approx \sum_{n=1}^{\infty} c_{2n} t^{-n-\frac{1}{2}} \Gamma(n + \frac{1}{2}) \quad (t \rightarrow \infty).$$

We started above by expressing the Laplace method in terms of sharp peaks. It is, however, by no means essential that the peak be sharp, nor that its localization on the  $x$ -axis be independent of  $t$ . For, both width and localization of the peak can be controlled simply by a substitution  $x = ay + b$  of the integration variable, where both  $a$  and  $b$  may depend on  $t$ . The only thing that matters is

that there is an interval  $J$  such that  $\int_{-\infty}^{\infty} - f_J$  is small compared to  $f_J$ , and that the integrand can be approximated by simpler functions throughout  $J$ .

In many simple cases the Laplace method can be replaced by the following argument. It changes the general point of view, but usually it hardly changes the details to be carried out.

Assume that, after some substitution  $x = \xi(y)$ , the integral (4.1.1) becomes  $\omega(t) \int_{-\infty}^{\infty} g(y, t) dy$ , and that  $g(y) = \lim_{t \rightarrow \infty} g(y, t)$  exists. Furthermore assume that the approximation of  $g(y, t)$  to  $g(y)$  is strong enough to guarantee that  $\int_{-\infty}^{\infty} g(y, t) dy$  tends to  $\int_{-\infty}^{\infty} g(y) dy$ . Then we immediately have the asymptotic equivalence

$$F(t) \sim \omega(t) \int_{-\infty}^{\infty} g(y) dy \quad (t \rightarrow \infty),$$

and further approximations can often be obtained from a closer study of the difference  $g(y, t) - g(y)$ .

## 4.2. A general case

We shall consider

$$F(t) = \int_{-\infty}^{\infty} e^{th(x)} dx,$$

where  $h(x)$  is a real and continuous function. We shall assume that it has an absolute maximum at  $x = 0$ . Without loss of generality we assume that  $h(0) = 0$ . Moreover, we require  $h(x) < 0$  for all  $x \neq 0$ , and even that there exist positive numbers  $b$  and  $c$  such that

$$h(x) \leq -b \quad \text{if } |x| \geq c.$$

This condition will certainly be satisfied if, as happens in most applications,  $h(x)$  tends to  $-\infty$  if  $x \rightarrow \pm \infty$ .

We require that the integral converges for all sufficiently large values of  $t$ . For simplicity we assume that it converges if  $t = 1$ , so that  $\int_{-\infty}^{\infty} e^{-h(x)} dx$  converges.

Finally, we assume that the derivative  $h'(x)$  exists in some neighbourhood of  $x = 0$ , that  $h''(0)$  exists, and that  $h''(0) < 0$ .

As  $h$  is maximal at  $x = 0$ , we infer that  $h'(0) = 0$ .

It follows from our assumptions that there exists, to any  $\delta > 0$ , a positive number  $\eta(\delta)$  such that  $h(x) \leq -\eta(\delta)$  both in  $-\infty < x \leq -\delta$  and in  $\delta \leq x < \infty$ . This is trivial if  $\delta \geq c$ ; if  $\delta \leq c$  we

observe that the maximum of the continuous function  $h(x)$  in the closed interval  $\delta \leq x \leq c$  is attained at some point of that interval, so that its value is negative (and similarly for  $-c \leq x \leq -\delta$ ).

The contribution of the intervals  $-\infty < x \leq -\delta$  and  $\delta \leq x < \infty$  satisfies

$$(4.2.1) \quad \int_{-\infty}^{-\delta} + \int_{\delta}^{\infty} < e^{-(t-1)\eta(\delta)} \int_{-\infty}^{\infty} e^{th(x)} dx \quad (t > 1).$$

In a neighbourhood of  $x = 0$  we shall approximate the exponent  $h(x)$  by  $\frac{1}{2}x^2h''(0)$ . Given  $\varepsilon$  ( $0 < 3\varepsilon < |h''(0)|$ ), we can determine  $\delta > 0$  such that

$$(4.2.2) \quad |h(x) - \frac{1}{2}x^2h''(0)| \leq \varepsilon x^2 \quad (-\delta \leq x \leq \delta).$$

In order to show this we put  $h(x) - \frac{1}{2}x^2h''(0) = \varphi(x)$ , whence  $\varphi(0) = \varphi'(0) = \varphi''(0) = 0$ . The fact that  $\varphi''(0)$  exists, implies that  $\varphi'$  exists in a neighbourhood of 0. Moreover,  $x^{-1}(\varphi'(x) - \varphi'(0)) \rightarrow 0$  when  $x \rightarrow 0$ , and therefore we have  $\varphi'(x) = o(x)$  ( $x \rightarrow 0$ ). If  $|x|$  is sufficiently small, we can apply the mean value theorem to  $\varphi$ , and it results that  $\varphi(x) - \varphi(0) = x\varphi'(\theta x)$  with some  $\theta$  ( $0 < \theta < 1$ ). So  $\varphi(x) = x\varphi'(\theta x) = o(x^2)$  ( $x \rightarrow 0$ ). From this (4.2.2) follows.

We can now deal with the integral from  $-\delta$  to  $+\delta$ :

$$\int_{-\delta}^{\delta} e^{tx^2(h''(0)-2\varepsilon)} dx < \int_{-\delta}^{\delta} e^{th(x)} dx < \int_{-\delta}^{\delta} e^{tx^2(h(0)+2\varepsilon)} dx.$$

All three integrals  $\int_{-\delta}^{\delta}$  differ from the corresponding integrals  $\int_{-\infty}^{\infty}$  by an amount  $O(e^{-t\alpha})$ , where  $\alpha$  is positive and independent of  $t$ , though the value of  $\alpha$  depends on the choice of  $\delta$ . For the middle one this is expressed by (4.2.1), for the other two it can be established in the same way. Using (4.1.6) (with  $n = 1$ ), we infer that

$$\int_{-\infty}^{\infty} e^{th(x)} dx < (2\pi)^{\frac{1}{2}}(-h''(0) - 2\varepsilon)^{-\frac{1}{2}} t^{-\frac{1}{2}} + O(e^{-t\alpha}) \quad (t > 1).$$

It follows that

$$(4.2.3) \quad \int_{-\infty}^{\infty} e^{th(x)} dx < (2\pi)^{\frac{1}{2}}(-h''(0) - 3\varepsilon)^{-\frac{1}{2}} t^{-\frac{1}{2}}$$

for all sufficiently large values of  $t$ . Needless to say, there is a similar

estimate below. The number  $\varepsilon$  being arbitrary, it follows that

$$(4.2.4) \quad \int_{-\infty}^{\infty} e^{th(x)} dx \sim (2\pi)^{1/2} (-th''(0))^{-1/2} \quad (t \rightarrow \infty).$$

If the restriction  $h(0) = 0$  is dropped, we of course get an extra factor  $e^{th(0)}$  on the right-hand-side of (4.2.4).

### 4.3. Maximum at the boundary

In sec. 4.2 the maximum of the integrand occurred at an inner point of the interval, and owing to our differentiability assumptions we inferred that  $h' = 0$  at that point. If, however, the maximum of  $h$  is attained at the left end-point, say, then  $h'$  will usually be negative at this point.

So assume that we have to deal with

$$\int_0^{\infty} e^{th(x)} dx,$$

where  $h(x)$  is real and continuous, that  $h(x)$  attains its maximum at  $x = 0$ , that  $h'(0)$  exists and that  $h'(0) < 0$ . Moreover we assume that  $h(x) < h(0)$  ( $x > 0$ ), that  $h(x) \rightarrow -\infty$  if  $x \rightarrow \infty$ , and that  $\int_0^{\infty} e^{th(x)} dx$  converges.

We can now repeat the analysis of sec. 4.2 in a somewhat simpler form. Instead of (4.2.2) we have to use an inequality of the type

$$|h(x) - h(0) - xh'(0)| \leq \varepsilon x \quad (0 \leq x \leq \delta),$$

which directly follows from the fact that  $(h(x) - h(0))/x \rightarrow h'(0)$  when  $x \rightarrow 0$ .

It is quite easy to verify by the same method, now using (4.1.8), that

$$\int_0^{\infty} e^{th(x)} dx \sim (-th'(0))^{-1} e^{th(0)} \quad (t \rightarrow \infty).$$

Especially in this simple case it is quite easy to see that the formula remains true if  $t$  is a complex variable, with  $\operatorname{Re} t \rightarrow \infty$  instead of  $t \rightarrow \infty$ . The same remark applies to all integrals discussed in this chapter, if one only replaces error terms  $O(t^{-k})$  by the corresponding terms  $O((\operatorname{Re} t)^{-k})$ , and  $O(e^{-at})$  by  $O(e^{-a\operatorname{Re} t})$ .

#### 4.4. Asymptotic expansions

It is clear that in the case of the integrals discussed in secs. 4.2 and 4.3 more asymptotic information (as  $x \rightarrow 0$ ) about  $h(x)$  leads to more asymptotic information about the integral (as  $t \rightarrow \infty$ ). We shall restrict ourselves to the case of sec. 4.2, the case of sec. 4.3 being analogous. We already gave an example in sec. 4.1, but we shall now discuss a more general and more difficult case.

For simplicity we shall assume that  $h(x)$  is, in some interval  $-\delta \leq x \leq \delta$ , the sum of a convergent power series

$$h(x) = a_2x^2 + a_3x^3 + \dots,$$

and that  $a_2 < 0$ . We shall discuss the integral

$$(4.4.1) \quad F(t) = \int_{-\infty}^{\infty} g(x)e^{th(x)}dx,$$

where  $g(x)$  is an integrable function which is, as far as the interval  $-\delta \leq x \leq \delta$  is concerned, equal to the sum of a convergent power series

$$g(x) = b_0 + b_1x + b_2x^2 + \dots$$

We may assume that both power series are still absolutely convergent if  $|x| = \delta$ , for otherwise we can take a slightly smaller  $\delta$ .

We need some rough estimate expressing that the contributions of the intervals  $(-\infty, -\delta)$  and  $(\delta, \infty)$  are negligible. We shall assume that, for each positive integer  $M$ , we have

$$(4.4.2) \quad \int_{-\infty}^{-\delta} g(x)e^{th(x)}dx = O(t^{-M}), \quad \int_{\delta}^{\infty} g(x)e^{th(x)}dx = O(t^{-M})$$

if  $t \rightarrow \infty$ , not bothering about how such information can be obtained. And we shall assume the existence of a positive number  $\eta$  such that

$$(4.4.3) \quad h(x) \leq -\eta x^2 \quad (-\delta \leq x \leq \delta).$$

(Actually (4.4.3) can be proved, though with a smaller  $\delta$  which is, however, as good as this one.)

We shall give two methods for obtaining an asymptotic series for  $F$ . One will be described here, the second method will be explained in sec. 4.5.

In the first method we consider  $\exp(ta_2x^2)$  as the main factor of

the integrand. The remaining factor

$$g(x) \exp(tx^3(a_3 + a_4x + a_5x^2 + \dots))$$

can be expanded as a double power series in the two arguments  $tx^3$  and  $x$ , convergent for  $|x| \leq \delta$ , and for all values of  $tx^3$ .

We shall denote this double series by

$$P(tx^3, x) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_{mn}(tx^3)^m x^n.$$

Its coefficients  $c_{mn}$  are independent of  $t$  and  $x$ . We want to approximate  $P$  uniformly by its partial sums, and therefore we restrict  $tx^3$  to some finite interval. That is to say, we shall use the power series only if  $|x| \leq t^{-1/3}$ . We abbreviate  $t^{-1/3} = \tau$ . We may assume that  $t > \delta^{-3}$ , whence  $\tau \leq \delta$ .

We first notice that the intervals  $(\tau, \delta)$  and  $(-\delta, -\tau)$  can be neglected. It is not difficult to show that, if  $\eta > 0$ ,  $\eta$  fixed,

$$(4.4.4) \quad \int_{\tau}^{\infty} e^{-\eta tx^2} dx = O\{\exp(-\eta t^{1/3})\} \quad (t > 1).$$

For, as  $\eta t(x^2 - \tau^2) > \eta t\tau(x - \tau) > \eta(x - \tau)$  ( $x > \tau$ ), we have

$$\int_{\tau}^{\infty} e^{-\eta t(x^2 - \tau^2)} dx < \int_{\tau}^{\infty} e^{-\eta(x - \tau)} dx = \eta^{-1},$$

and (4.4.4) follows. More generally, for each integer  $N \geq 0$  we have

$$(4.4.5) \quad \int_{\tau}^{\infty} e^{-\eta tx^2} x^N dx = O\{\exp(-\frac{1}{2}\eta t^{1/3})\} \quad (t > 1).$$

The factor  $x^N$  gives no difficulties. For, if  $x \geq \tau$  we have  $x \leq tx^2$ , and so  $x^N = O(e^{\frac{1}{2}\eta x}) = O\{\exp(\frac{1}{2}\eta tx^2)\}$ . So (4.4.5) follows from (4.4.4), replacing  $\eta$  by  $\frac{1}{2}\eta$  in that formula.

From (4.4.3) and (4.4.4) and from the fact that  $g(x)$  is bounded in  $-\delta \leq x \leq \delta$  it follows that

$$(4.4.6) \quad \int_{\tau}^{\delta} g(x) e^{th(x)} dx + \int_{-\delta}^{-\tau} g(x) e^{th(x)} dx = O\{\exp(-\eta t^{1/3})\} \quad (t > \delta^{-3}).$$

In the remaining interval  $-\tau \leq x \leq \tau$  we approximate  $P$  by a partial sum. We choose a positive integer  $A$ , and we write

$$P_A(tx^3, x) = \sum_{m \geq 0, n \geq 0, m+n \leq A} c_{mn}(tx^3)^m x^n.$$

Then we have, if  $|x| < \tau$ ,

$$(4.4.7) \quad P - P_A = O((tx^3)^{A+1}) + O(x^{A+1}),$$

uniformly with respect to  $x$  and  $t$ . This step requires some explanation. If a double power series  $\sum_{m \geq 0, n \geq 0} c_{mn} z^m w^n$  converges for  $|z| < 2R$ ,  $|w| < 2S$ , then we have, the terms of a convergent series being bounded,  $c_{mn} = O(R^{-m} S^{-n})$ . Now we easily estimate, if  $|z| < \frac{1}{2}R$ ,  $|w| < \frac{1}{2}S$ ,

$$\begin{aligned} \sum_{m \geq 0, n \geq 0, m+n > A} c_{mn} z^m w^n &= O(\sum |z/R|^m |w/S|^n) = \\ &= O\left(\sum_{k=A+1}^{\infty} (|z/R| + |w/S|)^k\right) = O\{(|z/R| + |w/S|)^{A+1}\} = \\ &\quad = O\{(|z| + |w|)^{A+1}\} = O(|z|^{A+1}) + O(|w|^{A+1}). \end{aligned}$$

For the last step, cf. (1.2.12). It should be stressed that the estimates are not uniform with respect to  $A$ .

As to (4.4.7) we remark that  $P$  is continuous if  $-\delta \leq x \leq \delta$ ,  $|tx^3| \leq 1$ , so that the formula just proved for a possibly smaller region can be extended to this larger range (cf. sec. 1.2).

We have, by (4.4.5), if  $A$  is fixed,

$$\int_{-\infty}^{\infty} \int_{-\tau}^{\tau} P_A \exp(ta_2 x^2) dx = O(t^A \exp(\frac{1}{2}a_2 t^{1/3})) \quad (t \rightarrow \infty),$$

$a_2$  being negative. Hence, by (4.4.2), (4.4.5) (4.4.6) and (4.4.7), we have, for each positive integer  $M$ ,

$$\begin{aligned} (4.4.8) \quad \int_{-\infty}^{\infty} g(x) e^{th(x)} dx - \int_{-\infty}^{\infty} P_A \exp(ta_2 x^2) dx &= O(t^{-M}) + \\ &+ O\left\{ \int_{-\infty}^{\infty} \exp(ta_2 x^2) (|tx^3|^{A+1} + |x|^{A+1}) dx \right\} \quad (t \rightarrow \infty). \end{aligned}$$

For these integrals we refer to (4.1.5) and (4.1.6). By virtue of (4.1.7), the last  $O$ -term in (4.4.8) is easily seen to be  $O(t^{-\frac{1}{2}A-1})$ . So we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} g(x) e^{th(x)} dx &= \\ &= \sum_{\substack{m \geq 0, n \geq 0 \\ m+n \leq A}} c_{mn} \varepsilon_{m+n} t^{-\frac{1}{2}(m+n+1)} (-a_2)^{-\frac{1}{2}(3m+n+1)} \Gamma\left\{\frac{1}{2}(3m+n+1)\right\} + \\ &\quad + O(t^{-\frac{1}{2}A-1}) + O(t^{-M}) \quad (t \rightarrow \infty), \end{aligned}$$

where  $\varepsilon_{m+n}$  denotes 1 if  $m+n$  is even, and 0 if  $m+n$  is odd.

As  $A$  and  $M$  are arbitrary, we have an asymptotic series

$$(4.4.9) \quad \int_{-\infty}^{\infty} g(x) e^{th(x)} dx \approx \sum_{v=0}^{\infty} d_v t^{-\frac{1}{2}v} \quad (t \rightarrow \infty).$$

where

$$d_\nu = (-a_2)^{-\nu-\frac{1}{2}} \sum_{m=0}^{2\nu} c_{m, 2\nu-m} (-a_2)^{-m} \Gamma(m + \nu + \frac{1}{2}).$$

It is easily seen that the main term  $d_0 t^{-\frac{1}{2}}$  equals

$$\{-2\pi/th''(0)\}^{\frac{1}{2}} g(0).$$

We assumed above that  $h$  and  $g$  were the sums of convergent power series in some neighbourhood of  $x = 0$ . It is not difficult to show that the results also apply to the case that

$h(x) \approx a_2 x^2 + a_3 x^3 + \dots, \quad g(x) \approx b_0 + b_1 x + b_2 x^2 + \dots \quad (x \rightarrow 0)$   
in the sense of asymptotic series.

#### 4.5. Asymptotic behaviour of the $\Gamma$ -function

We shall apply the result of sec. 4.4 to the problem of the behaviour of  $\Gamma(t+1)$  as  $t \rightarrow \infty$ . We wish to keep our discussions entirely independent of sec. 3.10, where the Euler-Maclaurin method was applied to the infinite product of the  $\Gamma$ -function. Presently we shall restrict ourselves to the case that  $t$  is real.

We start from the integral representation

$$(4.5.1) \quad \Gamma(t+1) = \int_0^\infty e^{-u} u^t du \quad (t > -1).$$

The integrand has a peak, but the localization of the peak is not fixed: the maximum of  $e^{-u} u^t$  occurs at  $u = t$ , and the maximal value is  $e^{-t} t^t$ . Therefore we introduce a substitution  $u = t + y$ , taking  $y$  as the new integration variable, and we have to investigate the integrand in the neighbourhood of  $y = 0$ . The neighbourhoods that matter, are rather large in this case. It can be seen from the analysis below that they are intervals of lengths exceeding  $t^{\frac{1}{2}}$ . This fact, however, does not influence the method in any respect, and if we carry out the further substitution  $y = tx$ , this is only done in order to obtain some minor simplifications in the formulas.

On substituting  $u = t(1+x)$  in (4.5.1) the integral becomes simply

$$\Gamma(t+1) = e^{-t} t^{t+1} \int_{-1}^{\infty} \{e^{-x}(1+x)\}^t dx.$$

The function  $e^{-x}(1+x)$  has its maximum at  $x = 0$ , and putting

$e^{-x}(1+x) = e^{h(x)}$  we have

$$h(x) = -\frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \dots$$

So  $h(x)$  satisfies the conditions of sec. 4.2 with  $h''(0) = -1$ . Therefore, by (4.2.4),

$$\Gamma(t+1) \sim e^{-t} t^{t-\frac{1}{2}} (2\pi)^{\frac{1}{2}} \quad (t \rightarrow \infty).$$

This is Stirling's formula.

The method of sec. 4.4 leads to an asymptotic series for the function  $\Gamma(t+1)t^{-t-\frac{1}{2}}e^t$ . We shall now explain a modification of that method, which works out quite simply in many cases. We introduce a new integration variable  $z$  by  $\frac{1}{2}z^2 = -h(x)$ , or, more precisely,

$$z = x(1 - \frac{3}{4}x + \frac{2}{4}x^2 - \dots)^{\frac{1}{2}},$$

where the principal value of the root is chosen. In a certain neighbourhood of the origin the Lagrange inversion formula (see sec. 2.2) gives  $x$  as a power series in  $z$ :

$$x = z + c_2 z^2 + c_3 z^3 + \dots$$

It does not matter how small that neighbourhood is, as it is independent of  $t$ , whence the integral can be restricted to that neighbourhood (cf. (4.2.1)). So we infer that there exist positive numbers  $\delta$  and  $c$ , both independent of  $t$ , such that

$$\Gamma(t+1) = e^{-t} t^{t+\frac{1}{2}} \left\{ \int_{-\delta}^{\delta} e^{-\frac{1}{2}z^2 t} (1 + 2c_2 z + 3c_3 z^2 + \dots) dz + O(e^{-ct}) \right\}$$

if  $t \rightarrow \infty$ , and the method of sec. 4.1 leads to

$$(4.5.2) \quad (2\pi)^{-\frac{1}{2}} \Gamma(t+1) e^{t-t-\frac{1}{2}} \approx \\ \approx 1 + 3c_2 \cdot \frac{2}{t} \cdot \frac{2!}{1!2^2} + 5c_3 \left( \frac{2}{t} \right)^2 \frac{4!}{2!2^4} + \dots \quad (t \rightarrow \infty).$$

We can compare this result with (3.10.7), which can be written in the form

$$(4.5.3) \quad \log\{(2\pi)^{-\frac{1}{2}} \Gamma(t+1) e^{t-t-\frac{1}{2}}\} \approx \\ \approx \frac{B_2}{1 \cdot 2t} + \frac{B_4}{3 \cdot 4t^3} + \frac{B_6}{5 \cdot 6t^5} + \dots, \quad (t \rightarrow \infty).$$

It follows that, if we denote the formal power series in (4.5.2) and (4.5.3) by  $P(t^{-1})$  and  $Q(t^{-1})$ , respectively, we have formally  $e^{Q(z)} = P(z)$ . It is, however, by no means easy to verify this directly.

#### 4.6. Multiple integrals

The Laplace method can easily be carried over to multiple integrals. We shall consider

$$F(t) = \int_{-1}^1 \dots \int_{-1}^1 \exp\{th(x_1, \dots, x_n)\} dx_1 \dots dx_n,$$

where  $h$  is continuous in the cube  $-1 \leq x_1 \leq 1, \dots, -1 \leq x_n \leq 1$ . We assume that  $h(0, \dots, 0) = 0$ , but  $h(x_1, \dots, x_n) < 0$  in all other points of the cube. Moreover we assume that all second order derivatives of  $h$  exist, and are continuous in a neighbourhood of the origin, and that the maximum at  $(0, \dots, 0)$  is of elliptic type. What we actually need is this:

$$(4.6.1) \quad h(x_1, \dots, x_n) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \\ + o(x_1^2 + \dots + x_n^2) \quad (x_1^2 + \dots + x_n^2 \rightarrow 0),$$

where the quadratic form  $\sum \sum a_{ij} x_i x_j$  ( $a_{ij} = a_{ji}$ ) is positive definite.

The procedure of sec. 4.2 can now be repeated, with obvious alterations. Omitting details, we only mention the result

$$(4.6.2) \quad F(t) \sim At^{-\frac{1}{2}n} \quad (t \rightarrow \infty),$$

where

$$A = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp(-\frac{1}{2} \sum \sum a_{ij} x_i x_j) dx_1 \dots dx_n.$$

It is well-known that

$$(4.6.3) \quad A = (2\pi)^{\frac{1}{2}n} D^{-\frac{1}{2}},$$

where  $D$  is the determinant of the matrix  $(a_{ij})$ . This is usually derived by an orthogonal substitution in the integration variables, in such a manner that the matrix gets the diagonal form, and then the integral becomes the product of  $n$  single integrals.

If  $h$  admits a development into powers of  $x_1, \dots, x_n$ , we can obtain an asymptotic series for  $F(t)$ , in the same way as it was done in sec. 4.4.

In many cases, especially for theoretical purposes, it is easier

to apply an analogy of the method used in sec. 4.5. We shall give only a brief description.

If  $y$  is a positive number, then  $\varphi(y)$  will denote the volume of that part of the cube  $-1 \leq x_1 \leq 1, \dots, -1 \leq x_n \leq 1$ , the points of which satisfy  $h(x_1, \dots, x_n) \geq -\frac{1}{2}y^2$ . Then we have

$$F(t) = \int_0^\infty e^{-t y^2} d\varphi(y),$$

and so the problem has been reduced to a question about a single integral. Usually  $\varphi(y)$  will be differentiable, and  $\varphi'(y) \sim ny^{n-1}D^{-1}V_n$  ( $y \rightarrow 0$ ), where  $V_n$  is the volume of the unit sphere in  $n$ -dimensional Euclidean space. For the main term we now obtain

$$F(t) \sim nD^{-1}V_n \int_0^\infty e^{-t y^2} y^{n-1} dy = nD^{-1}V_n \cdot 2^{-1+\frac{1}{2}n} \Gamma(\frac{1}{2}n) t^{-\frac{1}{2}n} \quad (t \rightarrow \infty).$$

As  $V_n = \pi^{\frac{1}{2}n} / \Gamma(\frac{1}{2}n + 1)$ , this gives the same result as (4.6.2).

#### 4.7. An application

We shall discuss an instructive example of the multidimensional Laplace method. We consider the sum

$$(4.7.1) \quad S(s, n) = \sum_{k=0}^{2n} (-1)^{k+n} \binom{2n}{k}^s,$$

where  $s$  and  $n$  are positive integers. It is well-known that  $S(1, n) = 0$ ,  $S(2, n) = (2n)!/(n!)^2$ , and a formula of Dixon<sup>1)</sup> gives  $S(3, n) = (3n)!/(n!)^3$ . One of course expects similar formulas for larger values of  $s$ , but no such formula is known. A simple method to decide on the existence of such a formula is to determine the asymptotic behaviour of  $S(s, n)$  as  $n \rightarrow \infty$  ( $s$  fixed) and to investigate whether this corresponds to the behaviour of multiplicative combinations of factorials. It will turn out that the asymptotic formula for  $S(s, n)$  involves  $(\cos \pi/2s)^{2ns}$ . The number  $(\cos \pi/2s)^{2s}$  is rational if  $s = 2$  or  $3$ . If  $s > 3$ , however, this is no longer true, and it follows that  $(\cos \pi/2s)^{2sn}$  does not occur in the Stirling formulas for  $n!$ ,  $2n!$ ,  $3n!$ ,  $\dots$ . Therefore, we cannot expect simple extensions of the Dixon formula if  $s > 3$ .

Properly speaking, the discussion of  $S(s, n)$  belongs to Ch. 3. We

<sup>1)</sup> See W. N. BAILEY, Generalized Hypergeometric Series, Cambridge 1935, p. 13.

are, however, in the situation described in sec. 3.11: the sum is exponentially small compared to the largest term (i.e. the term with  $k = n$ ). This fact is easily verified in the cases  $s = 1, 2, 3$ , and for general  $s$  it follows from our final result (4.7.4). (We notice that the term with  $k = n$ , which we denote by  $t_n$ , is asymptotically  $(2^{2n}(n\pi)^{-\frac{1}{2}})^s$ ). This means, roughly, that the Euler-Maclaurin method (in the version of sec. 3.11, because of the alternating signs) gives a result of the type

$$S/t_n \approx 0 + 0 \cdot n^{-1} + 0 \cdot n^{-2} + \dots,$$

and possibly (by the method of sec. 3.10) that  $S/t_n$  is exponentially small, but we will not be satisfied with a mere upper estimate. Moreover, in this case, the terms are, considered as functions of the summation variable  $k$ , quite awkward, and the Euler-Maclaurin analysis becomes involved. For these reasons it is worth while to try other explicit expressions for  $S$ . One possibility is used below, another one (not restricted to the case that  $s$  is an integer) will be used in sec. 6.4.

It is easily seen that the sum  $S(s, n)$  is equal to the coefficient of  $z_1^0 z_2^0 \dots z_r^0$  in the product

$$(-1)^n (1 + z_1)^{2n} (1 + z_2)^{2n} \dots (1 + z_r)^{2n} \{1 - (z_1 \dots z_r)^{-1}\}^{2n},$$

where  $r = s - 1$ .

As  $S(1, n) = 0$  is trivial, we henceforth assume  $s \geq 2$ ,  $r \geq 1$ .

By Cauchy's formula we have

$$S(r+1, n) = (-1)^n (2\pi i)^{-r} \int \dots \int (1 + z_1)^{2n} \dots (1 + z_r)^{2n} \cdot \{1 - (z_1 \dots z_r)^{-1}\}^{2n} (z_1^{-1} dz_1 \dots z_r^{-1} dz_r),$$

where the integrals are taken along the unit circles in the complex  $z$ -planes.

On substituting  $z_j = \exp(2i\varphi_j)$  we obtain

$$(4.7.2) \quad S(r+1, n) = 2^{2rn+2n} \pi^{-r} \cdot$$

$$\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \dots \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} \{ \cos \varphi_1 \dots \cos \varphi_r \sin(\varphi_1 + \dots + \varphi_r) \}^{2n} d\varphi_1 \dots d\varphi_r,$$

and to this multiple integral we can apply the Laplace method. We put

$$G(\varphi_1, \dots, \varphi_r) = \cos \varphi_1 \dots \cos \varphi_r \sin(\varphi_1 + \dots + \varphi_r).$$

and our first question concerns the extreme points of  $G$ . As  $G = 0$  on the boundary of the cube

$$-\frac{1}{2}\pi \leq \varphi_1 \leq \frac{1}{2}\pi, \dots, -\frac{1}{2}\pi \leq \varphi_r \leq \frac{1}{2}\pi,$$

whereas  $G$  takes both positive and negative values inside the cube, the boundary can be neglected. As to the inner points, we remark that  $G$  has continuous partial derivatives, and so we only need to consider points where  $\partial G / \partial \varphi_1 = \dots = \partial G / \partial \varphi_r = 0$ . Excluding points where  $G = 0$ , we have, if  $j = 1, \dots, r$ ,

$$(4.7.3) \quad \partial G / \partial \varphi_j = \{-\tan \varphi_j + \cot(\varphi_1 + \dots + \varphi_r)\} \cdot G.$$

Hence our condition implies that  $\tan \varphi_1 = \dots = \tan \varphi_r$ . The  $\varphi_j$  being restricted to the interval  $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$  it follows that all  $\varphi_j$  are equal,  $\varphi_1 = \dots = \varphi_r = \alpha$ , say. We obtain  $\cot r\alpha = \tan \alpha$ , and so  $\alpha + r\alpha = \frac{1}{2}\pi + k\pi$ , where  $k$  is an integer. In other words  $\alpha = \nu\pi/2s$ , where  $s = r + 1$ , and  $\nu$  is an odd integer,  $|\nu| < s$ . The value of  $G$  in such a point is

$$G(\alpha, \dots, \alpha) = (\cos \alpha)^r \sin(r\alpha) = \pm (\cos \alpha)^s.$$

So there are two absolute maxima of  $G^2$ , corresponding to  $\nu = +1$  and  $\nu = -1$ . These are  $\alpha = \beta$  and  $\alpha = -\beta$ , respectively, where  $\beta = \pi/2s$ . It is sufficient to consider only one of them,  $\alpha = +\beta$ , say. For, the integral in (4.7.2) can be split into two equal parts, according to  $\varphi_1 + \dots + \varphi_r > 0$  or  $< 0$ .

We shall write, in a neighbourhood  $\Omega$  of  $(\beta, \dots, \beta)$ ,

$$G(\varphi_1, \dots, \varphi_r) = G(\beta, \dots, \beta) \exp h(\beta + x_1, \dots, \beta + x_r),$$

and we have to deal with

$$2 \int \dots \int \exp(2nh(\beta + x_1, \dots, \beta + x_r)) dx_1 \dots dx_r,$$

the integral being extended over a neighbourhood  $\Omega'$  of  $(0, \dots, 0)$ . As  $G$  has continuous partial derivatives of all orders, we have a multiple Taylor expansion for  $h$  (cf.(4.6.1)). As  $G$  is maximal at  $x_1 = \dots = x_r = 0$ , and as  $h = 0$  at that point, the constant term and all linear terms vanish:

$$h(\beta + x_1, \dots, \beta + x_r) = -\frac{1}{2} \sum_{i=1}^r \sum_{j=1}^r a_{ij} x_i x_j + \dots,$$

where  $a_{ij} = -(\partial/\partial \varphi_i)(\partial/\partial \varphi_j)(\log G)$ , evaluated at  $x_1 = \dots = x_r = 0$ .

From (4.7.3) we infer

$$\begin{aligned} a_{ij} &= (\partial/\partial\varphi_i) \{\tan \varphi_j - \cot(\varphi_1 + \dots + \varphi_r)\} = \\ &= \delta_{ij} \cos^2 \varphi_j + \sin^{-2}(\varphi_1 + \dots + \varphi_r) = (\delta_{ij} + 1) \cos^2(\pi/2s), \end{aligned}$$

for at  $\varphi_1 = \dots = \varphi_r = \pi/2s$  we have  $\sin(\varphi_1 + \dots + \varphi_r) = \sin(r\pi/2s) = \cos(\pi/2s)$ . Here  $\delta_{ij}$  is the Kronecker symbol.  $\delta_{ij} = 1$  if  $i = j$ ,  $\delta_{ij} = 0$  if  $i \neq j$ . The determinant of the matrix  $(1 + \delta_{ij})$  ( $i, j = 1, \dots, r$ ) has elements 2 in the main diagonal, and all other elements are 1. Its value equals  $s$  (the order of the matrix is  $r$ ), which easily can be shown by induction. It can also be derived from eigenvalue theory: the numbers 1 and  $r + 1$  are obviously eigenvalues, and as subtraction of the unit matrix from the given matrix leads to a matrix of rank 1, the multiplicity of the eigenvalue 1 equals  $r - 1$ . Therefore, there are no other eigenvalues. The determinant equals the product of the eigenvalues, whence the determinant equals  $r + 1$ .

The matrix  $(1 + \delta_{ij})$  is positive definite, for it is the matrix of the quadratic form

$$x_1^2 + \dots + x_r^2 + (x_1 + \dots + x_r)^2.$$

We are now in a position to apply (4.6.2) and (4.6.3), and the result is that  $S(s, n)$  is asymptotically equivalent to

$$2^{2rn+2n}\pi^{-r}2\cdot(2\pi)^{1/2}D^{-1/2}\cdot(2n)^{-1/2}\cdot\{G(\beta, \dots, \beta)\}^{2n},$$

where  $D = s \cos^{-2r}(\pi/2s)$ ,  $G(\beta, \dots, \beta) = \cos^s(\pi/2s)$ . It results that

$$(4.7.4) \quad S(s, n) \sim \{2 \cos(\pi/2s)\}^{2ns+s-1} 2^{2-s} (\pi n)^{s/2} (s-1)_s^{-1}$$

if  $n \rightarrow \infty$  and if  $s$  is fixed ( $s = 2, 3, \dots$ ).

As a verification we take  $s = 3$ . Then we find

$$S(3, n) \sim 3^{3n+1/2} (2\pi n)^{-1} \quad (n \rightarrow \infty),$$

and since

$$(3n)!/(n!)^3 \sim (3n)^{3n+1/2} (2\pi)^{1/2} e^{-3n} \{n^{n+1/2} (2\pi)^{1/2} e^{-n}\}^{-3},$$

this is in accordance with Dixon's formula  $S(3, n) = (3n)!/(n!)^3$ .

#### 4.8. EXERCISES. 1. Prove the asymptotic equivalence

$$\int_0^\pi x^n \sin x \, dx \sim \pi^{n+2} n^{-2} \quad (n \rightarrow \infty).$$

2. Show that

$$\int_0^1 e^{xx^n} (1+x^2)^{-n} dx \sim (\frac{1}{2}\pi)^{\frac{1}{2}} n^{-\frac{1}{2}} 2^{-n} \quad (n \rightarrow \infty).$$

3. Prove the asymptotic equivalence

$$\sum_{k=0}^n \binom{n}{k} k! n^{-k} \sim (\frac{1}{2}\pi n)^{\frac{1}{2}} \quad (n \rightarrow \infty), .$$

and if  $\lambda$  is a constant,  $0 < \lambda < 1$ , show that

$$\sum_{k=0}^n \binom{n}{k} k! n^{-k} \lambda^k = (1-\lambda)^{-1} - n^{-1} \lambda^2 (1-\lambda)^{-3} + O(n^{-2}) \quad (n \rightarrow \infty).$$

(Hint:  $k! n^{-k-1} = \int_0^\infty e^{-nx} x^k dx$ ).

4. Show that there is an asymptotic expansion

$$\int_0^\infty \int_0^\infty e^{-n(x^2+y^2)} (1-x-y)^n dx dy \approx c_2 n^{-2} + c_3 n^{-3} + c_4 n^{-4} + \dots \quad (n \rightarrow \infty),$$

with  $c_2 = 1$ ,  $c_3 = -7$ .

## CHAPTER 5

### THE SADDLE POINT METHOD

#### 5.1. The method

The saddle point method, due to B. Riemann and P. Debye<sup>1)</sup>, is one of the most important and most powerful methods in asymptotics. Its object is to obtain useful approximations to integrals

$$(5.1.1) \quad F(t) = \int_P \varphi(z) dz,$$

where  $P$  is an integration path in the complex plane, and  $\varphi(z)$  is analytic in some neighbourhood of  $P$ . We assume that both the integration path and the function  $\varphi$  depend on a real parameter  $t$ , and we want to study the asymptotic behaviour of  $F(t)$  as  $t \rightarrow \infty$ .

Any special application of the saddle point method consists of two stages.

(i) The stage of exploring, conjecturing and scheming, which is usually the most difficult one. It results in choosing a new integration path, made ready for application of (ii).

(ii) The stage of carrying out the method. Once the path has been suitably chosen, this second stage is, as a rule, rather a matter of routine, although it may be complicated. It essentially depends on the Laplace method of Ch. 4.

The first stage, however, is usually quite difficult, especially in those cases where the new path has to depend on  $t$ . Most authors

<sup>1)</sup> B. RIEMANN, Gesammelte Mathematische Werke, 2e Aufl., Leipzig 1892, pp. 424–430; P. DEBYE, Math. Ann. 67 (1909), pp. 535–558.

Some books which give expositions of the saddle point method, with examples, are: R. COURANT und D. HILBERT, Methoden der Mathematischen Physik, vol. 1, New York 1955; H. and B. S. JEFFREYS, Methods of Mathematical Physics, 3rd ed., Cambridge 1956; E. T. COPSON, The Asymptotic Expansion of a Function Defined by a Definite Integral or a Contour Integral, Admiralty Computing Service, London 1946; A. ERDÉLYI, Asymptotic Expansions, Dover Publ., 1956.

dealing with special applications do not go into the trouble of explaining what arguments led to their choice of the path. The main reason is that it is always very difficult to say why a certain possibility is tried and others are discarded, especially since this depends on personal imagination and experience.

In the present exposition we shall try to give, in each example, some arguments accounting to some extent for the choice of the path, although many readers may find these quite unsatisfactory. There are, of course, general arguments, to be explained in secs. 5.2–5.5, which form the basis of the method, but in special applications these generalities give only partial answers.

The general idea of the saddle point method can easily be grasped in the following way: Assume, for a moment, that we are not interested in the value of  $F(t)$ , but that we only want to find a good upper estimate for  $|F(t)|$ . We have, of course, if the path has finite length  $l_P$ ,

$$(5.1.2) \quad |F(t)| \leq \int_P |\varphi(z)| \cdot |dz| \leq l_P \cdot \max_P |\varphi(z)|,$$

where  $\max_P |\varphi(z)|$  is the maximum of  $|\varphi(z)|$  along  $P$ . It may be possible, however, to obtain a better estimate by taking a different path. By Cauchy's theorem, the path  $P$  in (5.1.1) may be replaced by other paths  $C$ , having the same endpoints as  $P$ , provided that  $C$  can be continuously deformed into  $P$  without leaving the domain of analyticity of  $\varphi$ . We shall call these paths  $C$  admissible.

We now wish to determine  $C$  such that the value of

$$(5.1.3) \quad l_C \max_C |\varphi(z)|$$

is minimal. This  $C$  can of course depend on  $t$ .

The role of  $l_C$  is, as a rule, quite unimportant. In the first place, we may remark that the estimation (5.1.2) is a rough one. Along the largest part of the path the value of  $|\varphi(z)|$  may be much smaller than the maximum, so that only a small part of the path may count. Secondly, we are thinking of applications where  $\varphi(z)$  behaves rather violently, at any rate if  $t$  is large: very large at some places, and very small at other places. Therefore, small variations of the path may result in large variations of  $\max_C |\varphi(z)|$ , whereas the value of  $l_C$  hardly changes.

Finally it may be remarked that if  $\varphi(z)$  does not behave violently, the saddle point method has not much of a chance to succeed.

For these reasons we may expect that the value (5.1.3) is pretty close to its minimum if we choose from the set of all admissible paths the one,  $C$  say, for which

$$(5.1.4) \quad \max_C |\varphi(z)|$$

is minimal. It is of course not generally true that a minimizing path  $C$  exists, but commonly it does.

Usually it will turn out that we are so fortunate that the path  $C$ , chosen this way in order to obtain an upper bound for  $|F(t)|$ , is at the same time a good path for the evaluation of  $F(t)$  itself. That is, we can parametrize the path and write  $\int_C \varphi(z) dz$  as an integral along a real interval, to which we can try to apply the Laplace method (Ch. 4). If the Laplace method fails, all we can say is that the problem was really no case for the saddle point method.

The above statement is, of course, very rough. There are many paths giving the same value of  $\max_C |\varphi(z)|$ , but not all are suitable for application of the Laplace method. We shall have to adjust the path  $C$  in such a way that the parts of the path where  $|\varphi(z)|$  is close to its maximum, become small in length. This can be achieved by the so-called method of steepest descent (sec. 5.4).

If  $C$  is chosen according to that method, the maximum of  $|\varphi(z)|$  will be attained at a few isolated points only, and as a rule only in one single point. These isolated points are either end-points of the path, or they turn out to be saddle points, i.e. points where the derivative  $\varphi'(z)$  vanishes. The saddle points are usually easy to find, and they form the basis for the construction of the path  $C$ .

As we said before, it may occur that the point of  $C$  where  $|\varphi(z)|$  is maximal, is one of the end-points, and that there is no question of saddle points. We shall then still speak about the saddle point method, as the general aspect, both of problem and method, is of the same type as in the saddle point case.

## 5.2. Geometrical interpretation

In order to make things clear we shall give a geometrical illustration. We put  $z = x + iy$ , and we consider the surface in three-dimensional  $x$ - $y$ - $w$ -space, whose equation is  $w = |\varphi(x+iy)|$ . We ex-

clude the trivial case that  $\varphi$  is constant. It is easily seen that at points where  $\varphi'(x + iy) = 0$ , and at no other points, the tangent plane is horizontal, by virtue of the formula

$$\begin{aligned}\frac{\partial|\varphi|}{\partial x} - i \frac{\partial|\varphi|}{\partial y} &= |\varphi| \left\{ \frac{\partial \operatorname{Re} \log \varphi}{\partial x} - i \frac{\partial \operatorname{Re} \log \varphi}{\partial y} \right\} = \\ &= |\varphi| \frac{d \log \varphi}{dz} = \frac{|\varphi|}{\varphi} \frac{d\varphi}{dz}.\end{aligned}$$

(For simplicity, we shall forget a moment about points where  $\varphi(x + iy) = 0$ ). By the maximum modulus principle, there are no maxima nor minima, apart from minima at points where  $\varphi = 0$ , so that the points where the tangent plane is horizontal, are saddle points. These have the property that their neighbourhoods on the surface are partly above and partly below the level of the saddle point itself.

Let us imagine a man who wants to move from spot  $A$  to spot  $B$  in some mountain district, and whose physical condition makes it desirable to avoid the higher altitudes as much as possible. On the other hand, he has no objection whatsoever against walking, nor against climbing. He therefore tries to do the same thing as we want to do on our surface  $w = |\varphi(x + iy)|$ : he wants to take a path such that the maximum altitude is as low as possible.

If there happens to be a path of which  $A$  is the highest level, his problem is solved. It is clear that no path leading from  $A$  to  $B$  has a maximum altitude below the one of  $A$ . The same remark applies to  $B$ .

On the other hand it may occur that no such path exists, so that every path leading from  $A$  to  $B$  contains altitudes above those of  $A$  and  $B$ . Unless our man is a mathematician, it will be immediately clear to him that if there exists a path which solves his minimum problem, the highest point of that path will be a saddle point, that is, in his terminology, the highest point of a pass. A mathematician will be able to prove it (also assuming the existence of a solution of the minimum problem) under some continuity conditions, which are amply satisfied in our case  $w = \varphi(x + iy)$ . If the surface of the earth were not a sphere, but an infinite plane, the man would readily understand that the existence of a minimizing path is not guaran-

teed. It might be possible that by making wider and wider detours the maximum altitude of the path could be reduced and reduced, without every attaining a minimum. A sufficient, though not necessary, condition for the existence of a minimizing path is, that there exists a large circle, containing both  $A$  and  $B$  in its interior, such that every point of its circumference is higher than all points on the straight line segment  $AB$ .

We shall now describe a method for finding the best possible path for the mountaineer. Let  $h_0$  be the highest of the altitudes of  $A$  and  $B$ . Now for every number  $h \geq h_0$  we construct the region  $R_h$ , consisting of all points whose altitudes are  $\leq h$ . Then both  $A$  and  $B$  belong to  $R_h$ , and the question is whether  $A$  and  $B$  can be connected by a path entirely belonging to  $R_h$ . If the answer is negative, the highest point of his minimal path will certainly have an altitude  $> h$ . If the answer is affirmative, there is a path whose maximum altitude is  $\leq h$ . So his problem is solved if he knows the smallest value of  $h$  for which the answer to the above question is affirmative. But how to find this smallest value?

To his non-mathematical mind the following will be obvious: if  $h$  is this smallest value of  $h$ , and if  $P$  is a path from  $A$  to  $B$  entirely in  $R_h$ , then the highest point of  $P$  will be a saddle point. For otherwise, this highest point could be circumvented by making some detour.

It follows that the minimum value of  $h$  only needs to be sought amongst the altitudes of the various saddle points. And even without these discussions the first thing our mountaineer would do, would be to look up the altitudes of the various cols in the neighbourhood, and try whether he can do with some of the lowest.

It is clear that the highest saddle point will be crossed by our mountaineer, i.e. passed in such a way that in each neighbourhood of the saddle point, and on both sides of the path, there are points above the level of the saddle point.

The method described above simplifies the problem in so far that the unknown quantity no longer is a path, but a number. Unfortunately, it is often very difficult to carry out this procedure in our mathematical saddle point problems. However, we can easily look up the altitudes of the surrounding saddle points and try to use the lowest one. In the next section it will be shown that we often obtain a definite answer in this way.

### 5.3. Peakless landscapes

The landscape we are especially interested in, viz.  $w = |\varphi(x+iy)|$ , has a simple property: by the maximum modulus theorem we know that there are no peaks. Assuming that there are no singularities either, this property has the following consequence: if a *closed* path crosses a saddle point, and if this saddle point is not a double point of the path, then this saddle point is *not* the highest point of the path. (This has to be taken in the sense that there are other points on the path which are definitely higher). In order to show this it is sufficient to restrict ourselves to closed paths without double points, for otherwise we can easily find a closed part of the path, without double points, still crossing the saddle point (mathematically we are still on the level of the mountaineer). Now by the maximum modulus theorem a closed path contains in its interior no points higher than the highest point of the path. So in every neighbourhood of this highest point there are, on one side of the path, no points higher than this point itself. Therefore, this highest point cannot be a saddle point crossed by the path (see the definition of "crossing" at the end of sec. 5.2).

This closed path theorem has a consequence which is of great practical value for our minimum problem. If (in a landscape without peaks and without singularities) a path, without double points, leading from  $A$  to  $B$  crosses a saddle point, and if this saddle point is a highest point of the path (in the weak sense that no point of the path is higher), then the path solves the minimum problem, that is, any other path from  $A$  to  $B$  has at least one point of at least the same altitude as the saddle point just mentioned. The proof is easy. If we go from  $A$  to  $B$  along this path, and go back from  $B$  to  $A$  along any other path, we have described a closed path to which the previous theorem can be applied.

This means that the minimum path can be immediately recognized as such. If we conjecture that some path solves the minimum problem, this conjecture can be tested by looking at this path only, and it is no longer necessary to investigate all other paths.

If we now include the case that there is a path of which either  $A$  or  $B$  is the highest point, we can formulate the following rule for the solution of the minimum problem: *Find a path, from  $A$  to  $B$ ,*

*without double points, the highest point of which is either an end-point or a saddle point which is crossed by the path.*

The additional words “which is crossed by the path” are clearly necessary in order to avoid that our mountaineer goes up to some very high pass and after reaching the saddle point, goes back in the direction he came from.

If there are singularities, we have to take care that the minimal path hangs together with the path originally given, in the sense that the first path can be continuously deformed into the second one without ever leaving the regularity domain. Or in other words, we have to take care that Cauchy’s theorem can be applied.

However, it is often useful to violate this provision about singularities. Of course, we then have to take certain residues into account, but these are, as a rule, easily determined with great precision.

Various statements made thus far could be stated more precisely, and could be proved more rigorously. But for the present purpose it is not necessary to do so, as these matters only play a part at what we called, in sec. 5.1, the first stage of the procedure. Rigorous proofs will only be needed at the second stage, where calculations are based upon one special path, and then we have nothing to do with the question why just that path was chosen.

#### 5.4. Steepest descent

Let us consider a path, the highest point of which is a saddle point (crossed by this path). We shall examine a neighbourhood of the saddle point.

It is slightly easier to work with the logarithm of  $\varphi(z)$ , to be denoted by  $\psi(z)$ , so that  $\varphi(z) = e^{\psi(z)}$ . As the saddle point, to be denoted by  $\zeta$ , is the highest point of the path, we have  $\varphi'(\zeta) \neq 0$ . As we are further only interested in a neighbourhood of  $\zeta$ , we can select for  $\psi(z)$  any branch of the logarithm, and we have no difficulties due to the multi-valuedness.

It follows from  $\varphi(\zeta) \neq 0$  that the conditions  $\varphi'(\zeta) = 0$  and  $\psi'(\zeta) = 0$  are equivalent.

So far we have taken  $|\varphi(z)|$  to be the altitude of the landscape. If we replace it by  $\operatorname{Re} \psi(z)$ , which is a monotonic function of  $|\varphi(z)|$ , there is no change in any of the arguments of the preceding sections.

Moreover,  $\operatorname{Re} \psi(z)$  is single-valued (at least if  $\varphi(z)$  is single-valued).

Unless  $\psi(z)$  is a constant, at least one of the higher derivatives  $\psi''(\zeta)$ ,  $\psi'''(\zeta)$ , ... differs from zero. Let  $k$  be the smallest positive integer such that  $\psi^{(k)}(\zeta) \neq 0$ . We shall assume that  $k = 2$ . The cases  $k > 2$  are only slightly more difficult, and as they hardly ever occur in applications we shall disregard them.

The situation in a small neighbourhood of  $\zeta$  is mainly determined by the value of  $\psi''(\zeta)$ , since

$$\psi(z) = \psi(\zeta) + \frac{1}{2}\psi''(\zeta) \cdot (z - \zeta)^2 + \dots$$

We shall define the *axis* of the saddle point  $\zeta$  as the straight line in the complex  $z$ -plane defined by

$$\psi''(\zeta) (z - \zeta)^2 \text{ real and } \leq 0.$$

This is a line passing through  $\zeta$ . The *argument* of the axis is  $\frac{1}{2}\pi - -\frac{1}{2}\arg \psi''(\zeta)$ . (If a line  $l$  is parallel to the line connecting 0 and  $\alpha$ , then we say that  $\arg \alpha$  is the argument of  $l$ . The argument of  $l$  is uniquely determined apart from an additional multiple of  $\pi$ .)

The line defined by  $\psi''(\zeta) (z - \zeta)^2 \geq 0$  is clearly perpendicular to the axis.

The curves where  $\operatorname{Re} \psi(z) = \operatorname{Re} \psi(\zeta)$ , intersect the axis at  $\zeta$ , making angles of  $\pm \frac{1}{4}\pi$  with the axis. These curves are drawn in Fig. 1 as heavy lines. The curves where  $\operatorname{Im} \psi(z) = \operatorname{Im} \psi(\zeta)$  are drawn as dotted lines. Their tangents at  $\zeta$  are the axis of the saddle point and the line perpendicular to it, respectively. The axis itself is denoted by  $a$ . The curves  $\operatorname{Re} \psi(z) = \operatorname{Re} \psi(\zeta)$  divide the neighbourhood of  $\zeta$  into 4 parts. The two of these which do not contain the axis, are hatched. In these regions we have  $\operatorname{Re} \psi(z) > \operatorname{Re} \psi(\zeta)$ , whence it follows that our integration path, of which  $\zeta$  is the highest point, does not enter into these regions. The integration path has to cross the saddle point, so that it connects the two non-hatched regions.

There is a limit case, where the integration path is exactly one of the lines with  $\operatorname{Re} \psi(z) = \operatorname{Re} \psi(\zeta)$ , and where  $\zeta$  is not the only highest point of the path, all points of at least a part of the path having the same altitude. There is a method specially devoted to this case, called the "method of stationary phase". We shall see,

however, that in our case of analytic functions this case can always be avoided (sec. 5.8), and we shall not further discuss it.

In the general case, where the integration path uses, apart from  $\zeta$  itself, only inner points of the non-hatched regions, we can always deform it such that its tangent at  $\zeta$  coincides with the axis  $a$ , and such that, in the cases of violent behaviour of  $\operatorname{Re} \psi(z)$  we are interested in, the value of  $\operatorname{Re} \psi(z)$  is in all points of the path very much smaller than  $\operatorname{Re} \psi(\zeta)$ , apart from a small segment of the path around  $\zeta$ . If we start at the saddle point, and go in one of the two directions

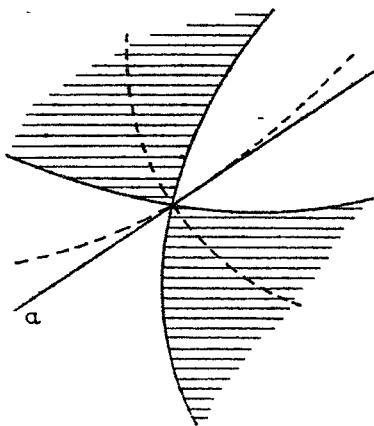


Fig. 1.

of the axis, the function  $\operatorname{Re} \psi(z)$  decreases. It is easily verified that this decrease is stronger than the decrease of  $\operatorname{Re} \psi(z)$  in any other direction. Therefore, the directions of the axis are called directions of *steepest descent*.

The use of these directions of steepest descent is not strictly essential for the saddle point method. We can usually take any other curve connecting the non-hatched regions, provided that the angle it makes with the axis, is less than  $\frac{1}{4}\pi$ , and does not tend to  $\frac{1}{4}\pi$  if the parameter  $t$  tends to infinity.

### 5.5. Steepest descent at end-point

Suppose that we have a path from  $A$  to  $B$ , the highest point of which is  $A$ . In the general case we have  $\psi'(A) \neq 0$ . We shall not consider what happens if  $\psi'(A) = 0$ , for then we have a saddle

point at  $A$ , and so things can be discussed according to sec. 5.4.

If  $\psi'(A) \neq 0$ , the value of  $\operatorname{Re} \psi(z)$  in small neighbourhoods of  $A$  is to a large extent determined by the value of  $\psi'(A)$ , since

$$\psi(z) = \psi(A) + (z - A)\psi'(A) + \dots$$

We shall again define the axis as the set defined by

$$(z - A)\psi'(A) \text{ real and } \leq 0,$$

and in the present case this is a half-line through  $A$ .

Perpendicular to this line is the curve along which we have  $\operatorname{Re} \psi(z) = \operatorname{Re} \psi(A)$ . At one side of this curve, the side of the axis, we have  $\operatorname{Re} \psi(z) < \operatorname{Re} \psi(A)$ , at the other side we have  $\operatorname{Re} \psi(z) > \operatorname{Re} \psi(A)$ . The path under discussion certainly does not enter into the latter region, the highest point of the path being  $A$  itself.

The direction of the axis is the direction of steepest descent, and we shall preferably take our path starting in this direction.

### 5.6. The second stage

Suppose that we have found a curve that minimizes  $\max_C \operatorname{Re} \psi(z)$ , along the principles expounded in sec. 5.3, and that we have modified it so as to show steepest descent at end-points and saddle points. Furthermore assume that the behaviour of  $\operatorname{Re} \psi(z)$  is violent, in the sense that it is, at least relatively, very large at some of the saddle points or end-points, and that at these points  $|\psi''(\zeta)|$  is large as far as saddle points are concerned, and  $|\psi'(\zeta)|$  is large if end-points are concerned. Then on our path we have large values of  $|e^{\psi(z)}|$  in small neighbourhoods of some of the saddle points or end-points, and in all other points of the path its value is negligible compared to these. Not all saddle points or end-points need to be important, for it may happen that at some of them the value of  $|e^{\psi(z)}|$  is negligible compared to its value at some of the others. Accordingly, in these insignificant points there is no need for steepest descent, and if saddle points are concerned, it is even not necessary to draw the path exactly through these saddle points.

For the final evaluation of asymptotic expressions for our integral we now immediately apply the Laplace method of Ch. 4. At the end-points the integrand is, roughly speaking, of the type  $\exp(-cs)$  and at the saddle points it is of the type  $\exp(-cs^2)$ . In both cases  $c$

is a constant whose real part is large, and  $s$  is a real variable used for parametrization of the path in the neighbourhood of the significant point under consideration.

### 5.7. A general simple case

We consider a simply connected region  $D$  in the complex plane, and two functions  $g(z)$ ,  $h(z)$ , both independent of  $t$ . The functions  $g$  and  $h$  are analytic functions of  $z$  for all  $z$  inside  $D$ . The points  $A$ ,  $B$  are in  $D$  and independent of  $t$ . We want to discuss the asymptotic behaviour of

$$F(t) = \int_A^B g(z) e^{th(z)} dz$$

as  $t > 0$ ,  $t \rightarrow \infty$ . Assume that there is a point  $\zeta \in D$  where  $h'(\zeta) = 0$ ,  $h''(\zeta) \neq 0$ . Properly speaking,  $\zeta$  is not a saddle point of  $get^h$ , but of  $e^{th}$ ; nevertheless it will turn out that  $\zeta$  can be used for the problem about  $get^h$  as well.

The landscape of  $|e^{th(z)}|$  around  $\zeta$  is of the type illustrated in fig. 1, p. 85. If  $\delta$  is a number independent of  $t$ , and  $0 < \delta < \frac{1}{4}\pi$ , then we can find  $\rho > 0$  such that there are two opposite sectors of the circle with centre  $\zeta$  and radius  $\rho$ , with apertures  $\frac{1}{2}\pi - 2\delta$ , both symmetrical with respect to the axis of the saddle point, in which  $|e^{th(z)}| < |e^{th(\zeta)}|$ , or  $\operatorname{Re} h(z) < \operatorname{Re} h(\zeta)$ . In formulas, these sectors can be described by

- (i)  $0 < |z - \zeta| < \rho$ ,  $|\arg(z - \zeta) + \frac{1}{2}\pi + \frac{1}{2}\arg h''(\zeta)| < \frac{1}{4}\pi - \delta$ ,
- (ii)  $0 < |z - \zeta| < \rho$ ,  $|\arg(z - \zeta) - \frac{1}{2}\pi + \frac{1}{2}\arg h''(\zeta)| < \frac{1}{4}\pi - \delta$ .

For, in both sectors we have

$$|\arg(-(z - \zeta)^2 h''(\zeta))| < \frac{1}{2}\pi - 2\delta,$$

and so

$$\operatorname{Re}\{-(z - \zeta)^2 h''(\zeta)\} > |z - \zeta|^2 \sin 2\delta;$$

consequently

$$\begin{aligned} \operatorname{Re} h(z) - \operatorname{Re} h(\zeta) &= \frac{1}{2} \operatorname{Re}\{h''(\zeta)(z - \zeta)^2\} + O(|z - \zeta|^3) < \\ &< \operatorname{Re}\{-\frac{1}{2}|z - \zeta|^2 \sin 2\delta\} + O(|z - \zeta|^3). \end{aligned}$$

So this is negative if  $\rho$  is sufficiently small.

Obviously, both  $\delta$  and  $\rho$  can be chosen independently of  $t$ . Now

assume that  $A_1$  is a point in sector (i),  $B_1$  in sector (ii), both independent of  $t$ , and that the integration path from  $A_1$  to  $B_1$  remains inside the domain  $D$ . Then we can replace the path by a new one: First connect  $A_1$  to a point  $A_2$  on the axis, inside (i); then cross the saddle point along the axis, from  $A_1$  to a point  $B_2$  ( $B_2$  inside (ii)); finally connect  $B_2$  to  $B_1$  by a path inside (ii). Along the paths from  $A_1$  to  $A_2$  and from  $B_1$  to  $B_2$ ,  $\operatorname{Re}(h(z) - h(\zeta))$  has a negative upper bound  $-c$ , and therefore the contribution of these parts to the integrals is  $O(e^{-ct} \cdot e^{t\operatorname{Re} h(\zeta)})$ . The contribution of the integral from  $A_2$  to  $B_2$  can be evaluated by application of the Laplace method. We parametrize the path by

$$z = \zeta + \alpha x, \quad a \leq x \leq b \quad (-\rho < a < 0 < b < \rho, \quad \alpha = \exp(\frac{1}{2}\pi i - \frac{1}{2}i \arg h''(\zeta))),$$

and the integral from  $A_2$  to  $B_2$  becomes

$$\alpha \int_a^b g(\zeta + \alpha x) e^{th(\zeta + \alpha x)} dx.$$

We have thus obtained an integral that was discussed extensively in sec. 4.4, as

$$h(\zeta + \alpha x) = h(\zeta) + \frac{1}{2}h''(\zeta)\alpha^2 x^2 + \dots, \quad \frac{1}{2}h''(\zeta)\alpha^2 < 0.$$

The results of that section can be applied immediately. We find an asymptotic series of the form (cf. (4.4.9))

$$(5.7.1) \quad \int_{A_2}^{B_2} g(z) e^{th(z)} dz \approx e^{th(\zeta)} t^{-\frac{1}{2}} \sum_0^\infty d_n t^{-n} \quad (t \rightarrow \infty).$$

If  $g(\zeta) \neq 0$ , the main term is easily evaluated:

$$(5.7.2) \quad \begin{aligned} \int_{A_2}^{B_2} g(z) e^{th(z)} dz &= \\ &= (2\pi)^{\frac{1}{2}} \alpha^{-\frac{1}{2}} |h''(\zeta)|^{-\frac{1}{2}} g(\zeta) e^{th(\zeta)} \{1 + O(t^{-1})\} \quad (t \rightarrow \infty). \end{aligned}$$

We note that  $\alpha$  is the complex number with modulus 1 whose argument corresponds to the direction on the axis from (i) to (ii).

The right-hand-side of (5.7.1) will be referred to as the *contribution* of the saddle point  $\zeta$ . It is a trivial but important remark that the contribution of a saddle point depends on the direction in which it is crossed. If we reverse the direction, the contribution is of course multiplied by  $-1$ .

As the integrals from  $A_1$  to  $A_2$  and from  $B_2$  to  $B_1$  are exponentially small compared to  $e^{th(\zeta)}$ , formulas (5.7.1) and (5.7.2) remain true if we replace  $A_2$  by  $A_1$  and  $B_2$  by  $B_1$ .

The question whether the integral along the original path from  $A$  to  $B$  can be represented asymptotically by the contribution of the saddle point, is of a different type. It cannot be answered by studying small neighbourhoods of  $\zeta$ . The answer is affirmative if  $A$  can be linked to  $A_1$ , and  $B$  to  $B_1$ , in such a way that on those connecting paths the maximum of  $\operatorname{Re} h(z)$  is less than  $\operatorname{Re} h(\zeta)$ , for then the contribution of these paths is again negligible.

It is, of course, not necessary to use  $A_1$ ,  $B_1$  as intermediates between  $A$ ,  $B$  and  $A_2$ ,  $B_2$ . In the above presentation it was done for the sake of a minor simplification. The price we have to pay for this, is exponentially small, and therefore we need not bother about the strict necessity of this step.

If the path from  $A$  to  $B$  has  $\zeta$  as its highest point, and if the tangent at  $\zeta$  makes an angle  $< \frac{1}{4}\pi$  with the axis, then we can take  $A_1$  and  $B_1$  on the path itself, sufficiently close to  $\zeta$ .

A similar, but simpler, discussion can be given for the contribution of an end-point to our integral  $\int g(z)e^{th(z)}dz$ . We only state the result: If  $g(A) \neq 0$ ,  $h'(A) \neq 0$ , and if the path starts from  $A$  in a direction in which  $\operatorname{Re} h(A)$  decreases, then the first term of the contribution of a neighbourhood of  $A$  equals (cf. sec. 4.3)

$$g(A)e^{th(A)}(-th'(A))^{-1}.$$

### 5.8. Path of constant altitude

We again consider integrals

$$\int_A^B \varphi(z)dz = \int_A^B e^{th(z)}dz.$$

If the points  $A$  and  $B$  are connected by a path, all points of which have the same altitude in the landscape  $w = \operatorname{Re} \psi(z)$ , then this path automatically solves the minimum problem (for no other path from  $A$  to  $B$  has a maximum altitude below the one of  $A$ ). We shall show that this path can always be slightly deformed so as to give a path having only a discrete number of highest points.

To this end we consider parts of the path, whose end-points are

either  $A$  or  $B$  or else saddle points, not containing saddle points as inner points. Obviously the whole path can be divided into such parts. Let  $A_k A_{k+1}$  be such a part of the path, and let this arc be given in the parametric representation  $z = f(s)$ ,  $0 \leq s \leq 1$ , where  $f(s)$  is continuously differentiable,  $f'(s) \neq 0$  ( $0 \leq s \leq 1$ ) and  $f(0) = A_k$ ,  $f(1) = A_{k+1}$ . As there are no saddle points on the arc between  $A_k$  and  $A_{k+1}$ , we have

$$\frac{d\psi}{ds} = \psi'(z) \cdot f'(s) \neq 0 \quad (0 < s < 1)$$

As  $\operatorname{Re} \psi$  was constant along the arc, we now know that  $i d\psi/ds$  is real, continuous and  $\neq 0$  ( $0 < s < 1$ ). It follows that its sign is constant, say

$$i d\psi/ds \quad \text{real and } < 0 \quad (0 < s < 1).$$

It now follows from the Cauchy-Riemann equations that the derivative of  $\operatorname{Re} \psi(z)$  in a direction perpendicular to the arc, is also negative, provided that the positive direction on the normal is pointing to the left bank of the arc. It follows that on the left bank  $\operatorname{Re} \psi(z)$  has values less than the constant value which it has along the arc, so instead of left bank we may speak about lower bank.

We now describe a new path from  $A_k$  to  $A_{k+1}$ . At  $A_k$  and  $A_{k+1}$  it goes, preferably by steepest descent, into the lower bank, and further it proceeds along the lower bank at a small distance of the arc. Here "small" means: sufficiently small in order to guarantee that we are below the level of the original path.

We can do the same thing for other parts of the path  $A$ ,  $B$ . Sometimes the lower bank will be on the left, sometimes it will be on the right. At a saddle point, the new path goes up to the level of the saddle point and descends again on the other side. (The last few words are not strictly true. If we have a higher-order saddle point, and if the order of the first non-vanishing derivative is odd, then we have to descend at the bank we came from).

The asymptotic behaviour of the integral from  $A$  to  $B$  is now given by the sum of the contributions of the points  $A = A_0$ ,  $A_1$ , ...,  $A_n = B$ .

In sec. 5.11 a worked example can be found.

### 5.9. Closed path

Instead of for a curve leading from a point  $A$  to a point  $B$ , the

problem about (5.1.1) can also be proposed for a closed path. We of course do not assume that the integrand is analytic everywhere inside the path, for then the integral is trivially zero.

Considering a closed path, we of course get no contributions from end-points. If the path can be deformed into another closed path crossing just one saddle point, and if this saddle point is higher than all other points of the path, then this path is fit for application of the techniques of what we called the second stage.

If we have a closed path of constant altitude, it need not solve the minimum problem (the argument given at the beginning of sec. 5.8 essentially depended on the fact that  $A$  and  $B$  were fixed beforehand). For example, if  $\varphi(z) = z^{-2}$ , any circle whose centre is in the origin, is a curve of constant altitude, and none of them solves the minimum problem.

### 5.10. Range of a saddle point

It is often quite difficult to determine a saddle point exactly. However, for asymptotic purposes it is not always necessary to take a path exactly through the saddle point. (In sec. 5.7 we actually did not use the exact saddle point of  $g(z)e^{th(z)}$ , but an approximation to it, viz. the saddle point of  $e^{th(z)}$ ). In order to get an idea of what deviations from the saddle point are possible, we shall introduce the notion of the “*range*” of a saddle point.

If  $\zeta$  is a saddle point of the function  $\psi$ , then the range of  $\zeta$  is a circular neighbourhood of  $\zeta$ , consisting of all  $z$ -values which are such that  $|\psi''(\zeta)(z - \zeta)^2|$  is not very large.

This is obviously not a proper mathematical definition, however useful it may be. The word “range” is as unmathematical as words like “small”, “large compared to”. We might easily give it a definite meaning, but that would be quite an arbitrary one. And, as the word only plays a rôle in what we have called the first stage, we need not be very precise.

Occasionally, we shall also use the word “range” for the radius of the circular neighbourhood mentioned above.

We have to remember that everything we are doing, depends on the parameter  $t$ , although the letter  $t$  was not explicitly written in our formulas. Especially, the saddle point  $\zeta$  may depend on  $t$ . But even if the saddle point is fixed, the range may still depend on  $t$ .

E.g., if  $\psi(z) = -tz^2$ , then  $\zeta = 0$  is a fixed saddle point, and its range is of the order of  $t^{-\frac{1}{2}}$ .

If we have to deal with an integral  $\int e^{\psi(z)} dz$ , and if  $\zeta$  is a saddle point, then it is very important to know whether in the formula

$$(5.10.1) \quad \psi(z) = \psi(\zeta) + \frac{1}{2}\psi''(\zeta)(z - \zeta)^2 + \frac{1}{3}\psi'''(\zeta)(z - \zeta)^3 + \dots$$

the sum of the terms

$$(5.10.2) \quad \frac{1}{3}\psi'''(\zeta)(z - \zeta)^3 + \dots$$

is or is not small compared to the term  $\frac{1}{2}\psi''(\zeta)(z - \zeta)^2$ , when  $z$  lies in the range of the saddle point. (Needless to say, everything depends on the parameter  $t$ , and "small" has to be interpreted in the sense of a  $\alpha$ -formula as  $t \rightarrow \infty$ ). If it is small, we are in a position to apply the technique of sec. 4.4 and, as far as the contribution of the saddle point  $\zeta$  is concerned, the integral can be successfully compared to

$$(5.10.3) \quad \int_L \exp\{\psi(\zeta) + \frac{1}{2}\psi''(\zeta)(z - \zeta)^2\} dz = \\ = (2\pi)^{\frac{1}{2}} \alpha |\psi''(\zeta)|^{-\frac{1}{2}} e^{\psi(\zeta)}.$$

Here  $L$  is the axis of the saddle point, extended to infinity in both directions, and the sense in which  $L$  is taken corresponds with the sense in which our integration path crosses the saddle point. The number  $\alpha$  has absolute value 1, and its argument indicates the direction on  $L$  (cf. 5.7.2).

In the special case that  $\psi(z) = th(z)$ ,  $h(z)$  independent of  $t$ ,  $h'(\zeta) = 0$ ,  $h''(\zeta) \neq 0$  (see sec. 5.7), we have an example where inside the range the terms of third and higher order are small compared to the second order terms. For then (5.10.1) has a positive radius of convergence  $R$ , say, where  $R$  does not depend on  $t$ . The range has the order of  $t^{-\frac{1}{2}}$ , for  $|\psi''(\zeta)(z - \zeta)^2|$  is large if  $|z - \zeta|$  is much larger than  $t^{-\frac{1}{2}}$ . Furthermore, (5.10.2) converges if  $|z - \zeta| < R$ , and so its sum is  $O(t(z - \zeta)^3)$  if  $|z - \zeta| < R_1$  (where  $R_1$  is a constant,  $0 < R_1 < < R$ ). It follows that the sum of (5.10.2) is  $O\{t^{-\frac{1}{2}} \cdot (\psi''(\zeta)(z - \zeta)^2)\}$  if  $|z - \zeta| = O(t^{-\frac{1}{2}})$ . This means that inside the range the second order term dominates all other terms if  $t \rightarrow \infty$ .

If, on the other hand, the sum (5.10.2) is not small compared to  $\frac{1}{2}\psi''(\zeta)(z - \zeta)^2$  throughout the range of the saddle point  $\zeta$ , it is difficult to say anything in general. Usually it means that there are other significant saddle points in the range of  $\zeta$ , or that there are

even singularities of  $\psi$  in that range. We shall discuss a few examples in sec. 5.12.

### 5.11. Examples

The examples we shall give in the final sections of this chapter, will be somewhat artificial in two respects. First, they did not arise from practical problems, but were just designed to illustrate some aspects of the saddle point method. Secondly, in each of these examples only some of the typical difficulties of the saddle point method will occur, whereas in most applications occurring in practice, almost all possible difficulties occur in one and the same problem. We shall give some of these more complicated problems in the next chapter.

Our first example is

$$f(t) = \int_0^\infty \exp(t(z + iz - z^3)) dz,$$

which is of the type of the integrals considered in sec. 5.7.

Even in a simple case like this, it is not easy to get an adequate survey over the landscape. Fortunately, the problem can be solved almost blindfolded.

We put  $z + iz - z^3 = h(z)$ . The saddle points are the solutions of  $h'(z) = 1 + i - 3z^2 = 0$ . So there are two of them,  $\zeta$  and  $-\zeta$ , where  $\zeta = 2^{1/4} \cdot 3^{-1/2} \cdot e^{\pi i/8}$ . At first sight it seems unlikely that  $-\zeta$  needs to be considered, and therefore we turn our attention to  $+\zeta$ .

The axis of this saddle point (see sec. 5.4) has the argument  $-\pi/16$ . Therefore, the straight line  $l$  connecting 0 and  $\zeta$  cuts the axis under an angle  $3\pi/16$ , which is less than  $\pi/4$ . This means that  $l$  is a suitable path in a neighbourhood of the saddle point. Fortunately, this line also serves our other purposes: it turns out that no other point of the line is higher than the saddle point itself. For, if we describe the line by  $z = e^{\pi i/8}x$ ,  $0 \leq x < \infty$ , the function  $h(z)$  becomes

$$h(z) = h(e^{\pi i/8}x) = e^{3\pi i/8}(2^{\frac{1}{4}}x - x^3),$$

and  $2^{\frac{1}{4}}x - x^3$  is maximal at  $x = x_0 = 2^{\frac{1}{4}}3^{-\frac{1}{2}}$ . Therefore, the same thing applies to  $\operatorname{Re} h(z)$ .

On the line  $l$  we have  $\operatorname{Re} h(z) < \operatorname{Re} h(\zeta) - x$  for  $x$  sufficiently large, and therefore the contribution of the part from  $x_0 + 1$ , say,

to  $\infty$  is exponentially small compared to the contribution of the saddle point.

It is of course a routine matter to show that the integral along the real axis from 0 to  $\infty$  is equal to the integral from 0 to  $e^{\pi i/8} \cdot \infty$  along  $l$ .

In order to evaluate the contribution of the saddle point, we evaluate

$$h(\zeta) = (1+i)\zeta - \zeta^3 = \frac{2}{3}(1+i)\zeta = 2^{7/4}3^{-3/4}e^{3\pi i/8},$$

$$h''(\zeta) = -6\zeta = -2^{5/4}3^{1/2}e^{3\pi i/8}.$$

As we cross the saddle point from left to right, the number  $\alpha$  (see sec. 5.7) equals  $\alpha = e^{-\pi i/16}$ . So by (5.7.2) the first term of the contribution equals

$$(2\pi)^{1/4}e^{-\pi i/16}t^{-1/4}(2^{5/4}3^{3/4})^{-1/4}e^{th(\zeta)} = e^{-\pi i/16}2^{-1/8}3^{-1/4}e^{it}e^{th(\zeta)}.$$

This gives at the same time the asymptotic behaviour of  $f$ , and so

$$f(t) \sim e^{-\pi i/16}2^{-1/8}3^{-1/4}e^{it}t^{-1/4}\exp\{2^{7/4}3^{-3/4}e^{3\pi i/8}t\} \quad (t \rightarrow \infty).$$

Our next example is

$$F(t) = \int_{-\infty}^{\infty} \varphi(z, t) dz, \quad \varphi(z, t) = e^{it(3z-z^3)},$$

and it is meant to illustrate sec. 5.8.

Although  $\varphi$  has absolute value 1 for all real values of  $z$ , the integral will be proved to converge. So  $\int_a^\infty$  tends to 0 if  $a \rightarrow \infty$  ( $t$  fixed). But it is not true that  $\int_a^\infty$  is exponentially small if  $a$  is fixed and  $t \rightarrow \infty$ . Therefore it is not advisable to apply the method of sec. 5.8 to  $F_a = \int_a^\infty$  and to make  $a \rightarrow \infty$  afterwards. To that end we would need a formula for  $F_a$  holding uniformly in  $t$  and  $a$ .

Therefore we prefer to replace the whole path  $(-\infty, \infty)$  by a new infinite path  $P$ , before we start making  $t \rightarrow \infty$ .

The saddle points are  $z = -1$  and  $z = +1$ . We first want to know what the lower bank is (the words lower and upper bank refer to the magnitude of the integrand, and not to lower and upper half-plane). Taking  $z = x + iy$ , we have  $\operatorname{Re}(it(3z - z^3)) = t(3y(x^2 - 1) - y^3)$ . This is negative for small positive values of  $y$  if  $-1 < x < 1$ , and it is negative for small negative values of  $y$  if  $x > 1$  or  $x < -1$  (we of course assume  $t > 0$ ). So between

$x = -1$  and  $x = +1$  the lower bank lies in the upper half-plane, and outside that interval it lies in the lower half-plane.

According to sec. 5.8, we now construct a path crossing the saddle point 1 from north-west to south-east, from  $1 - \varepsilon + i\delta$  to  $1 + \varepsilon - i\delta$ , say, where  $\varepsilon$  and  $\delta$  are small positive numbers. These numbers are chosen such that the saddle point is the highest point of that path. The saddle point  $-1$  can be crossed by a similar path from  $-1 - \varepsilon - i\delta$  to  $-1 + \varepsilon + i\delta$ . Finally we link the points  $-1 + \varepsilon + i\delta$  and  $+1 - \varepsilon + i\delta$  by a straight line, and we link the points  $-1 - \varepsilon - i\delta$  and  $1 + \varepsilon - i\delta$  to  $-\infty$  and  $+\infty$ , respectively, by lines parallel to the real axis. This defines our modified path  $P$ .

The integral along  $P$  is easily seen to converge, as

$$|\varphi(x - i\delta)| = |\varphi(1 + \varepsilon - i\delta)| \cdot \exp(-3\delta t(x^2 - (1 + \varepsilon)^2)) \leq \\ \leq |\varphi(1 + \varepsilon - i\delta)| \exp(-6\delta t(x - 1 - \varepsilon)) \quad (x > 1 + \varepsilon),$$

and a similar estimate holds on the part of  $P$  extending to  $-\infty$ . Therefore we have

$$\left| \int_{1+\varepsilon-i\delta}^{\infty-i\delta} \varphi dz \right| \leq |\varphi(1 + \varepsilon - i\delta)| / 6\delta t,$$

and  $|\varphi(1 + \varepsilon - i\delta)|$  is exponentially small, as  $\operatorname{Re}(i(3z - z^3)) < 0$  at  $z = 1 + \varepsilon - i\delta$ . It follows that the asymptotic behaviour of the integral along  $P$  just consists of the contributions of the saddle points 1 and  $-1$ .

We still have the question whether  $f_P = f_{-\infty}$ . When investigating this, we can consider  $t$  as a positive constant. If  $b$  is a large positive number we have, by Cauchy's theorem,

$$f = \int_{-b}^b - \int_{b-i\delta}^{\infty-i\delta} - \int_{-\infty-i\delta}^{-b-i\delta} + \int_{b-i\delta}^b - \int_{-b-i\delta}^{-b},$$

where the integration paths, apart from  $P$ , are straight horizontal or vertical lines. It follows from the convergence of  $f_P$  that the second and the third integral on the right-hand-side tend to 0 as  $b \rightarrow \infty$ . The same thing is true for the fourth and for the fifth, as

$$\left| \int_{b-i\delta}^b \right| \leq \int_0^\delta \exp(-t(3u(b^2 - 1) - u^3)) du < \\ < C(\delta, t) \int_0^\delta \exp(-3tu(b^2 - 1)) du < C(\delta, t)/3t(b^2 - 1),$$

and this tends to 0 as  $b \rightarrow \infty$  ( $C(\delta, t)$  is independent of  $b$ ). This proves that  $\int_{-\infty}^b$  tends to  $\int_P$  as  $b \rightarrow \infty$ . So we have established the convergence of  $\int_{-\infty}^\infty$ , and we have shown that  $\int_{-\infty}^\infty = \int_P$ .

The contributions of the saddle points are easily calculated anyway their first order terms. At  $z = -1$  the axis is exactly south-west to north-east, and the number  $\alpha$  (see sec. 5.7) equals  $e^{\pi i/4}$ . Putting  $h(z) = i(3z - z^3)$ , we have  $h''(z) = -6iz$ . Therefore  $|h''(-1)| = 6$ . Now (5.7.2) gives the first term of the contribution of  $-1$ , viz.

$$(2\pi)^{1/2} e^{\pi i/4} t^{-1/2} \cdot 6^{-1/2} e^{-2it}.$$

Similarly, we find for the first term of the contribution of the saddle point at  $z = 1$  exactly the complex conjugate of this expression. Our final result is that

$$F(t) = 2(\pi/3t)^{1/2} \cos(2t - \pi/4) + O(t^{-3/2}) \quad (t \rightarrow \infty).$$

In a case like this we cannot state that the first term gives the asymptotic behaviour, as the error term is not always small compared to the main term. In other words, it is not true that  $F(t) \sim 2(\pi/3t)^{1/2} \cos(2t - \pi/4)$ . The latter formula would imply that for large values of  $t$ ,  $F(t)$  vanishes at exactly the same places where  $\cos(2t - \pi/4)$  vanishes, and this is not necessarily so.

## 5.12. Small perturbations

We shall now discuss some examples illustrating the notion of the range of a saddle point (sec. 5.10). We start from the integral

$$(5.12.1) \quad \int_{-\infty}^{\infty} e^{-tz^2} dz,$$

whose saddle point is  $z = 0$ . The circle  $|z| \leq t^{1/2}$  can be considered as its range. The notion of range has some importance when one wants to discuss integrals obtained from (5.12.1) by small perturbations. For example, in the integral

$$(5.12.2) \quad \int_{-\infty}^{\infty} (1 + iz)^{1/2} e^{-tz^2} dz$$

the factor  $(1 + iz)^{1/2}$  behaves quite smoothly within the range of the saddle point  $z = 0$  of (5.12.1). Admittedly it becomes large far outside the range, but there it is quite innocent compared to the very

small factor  $e^{-tz^2}$ , so that the contribution of these parts of the integration path is negligible. This means that, although  $z = 0$  is not a saddle point of (5.12.2), the integration path  $-\infty < z < \infty$  of (5.12.2) can still successfully be used for the calculation of the asymptotic behaviour of (5.12.2). Actually this is what we did in sec. 4.4, and it is not necessary to repeat those details here.

Our next example is slightly more complicated, the extra factor depending on  $t$ . We put

$$(5.12.3) \quad F_\alpha(t) = \int_{-\infty}^{\infty} e^{-tz^2} \frac{e^z}{1 + t^\alpha z^2} dz,$$

where  $\alpha$  is a positive parameter. For each fixed value of  $\alpha$  the asymptotic behaviour (as  $t \rightarrow \infty$ ) is required. We shall investigate in what respect this can still be considered as a "minor modification" of (5.12.1).

The factor  $e^z$  is harmless: it behaves smoothly within the range. It gets large if  $z$  is positive and large with respect to 1, but then the factor  $\exp(-tz^2)$  is overwhelmingly small.

The factor  $(1 + t^\alpha z^2)^{-1}$  behaves smoothly if  $t^\alpha z^2$  is small, for then it can successfully be expanded into a power series. Also, if  $t^\alpha z^2$  is large, that factor can be expanded as  $(t^\alpha z^2)^{-1}$  times a power series in terms of  $(t^\alpha z^2)^{-1}$ . The dangerous zone is the circular annulus  $R$  in the  $z$ -plane where  $t^\alpha z^2$  is neither large nor small, and actually there are two poles in that annulus, viz.  $z = \pm it^{-\frac{1}{\alpha}}$ . Now there are these possibilities:

- (i)  $0 < \alpha < 1$ . Then the annulus  $R$  lies far outside the range.
- (ii)  $\alpha = 1$ . The annulus  $R$  covers a considerable part of the range.
- (iii)  $\alpha > 1$ . The annulus  $R$  lies inside the range, but is very small compared to the range.

Case (i). Here the state of affairs can be compared to the case of (5.12.2). According to the technique of sec. 4.4, we choose a positive number  $T$  which is large compared to the radius of the range, and small compared to the radii of  $R$ ; we can take  $T = t^\beta$ , where  $\frac{1}{2}\alpha < \beta < \frac{1}{2}$ . The integrals from  $T$  to  $\infty$  and from  $-T$  to  $-\infty$  are easily seen to be  $O(\exp(-\frac{1}{2}t^{1-2\beta}))$ . As  $1 - 2\beta > 0$ , this term is negligible. In the interval  $-T \leq z \leq T$ , the extra factor can be

successfully approximated. Taking the first three terms only, we have, if  $-T \leq z \leq T$ ,

$$\begin{aligned} e^z(1+t^\alpha z^2)^{-1} &= \{1+z+\frac{1}{2}z^2+O(z^3)\}\{1-t^\alpha z^2+O(t^{2\alpha}z^4)\}= \\ &= 1+z+(\frac{1}{2}-t^\alpha)z^2-t^\alpha z^3+O(t^{2\alpha}z^4)+O(z^3). \end{aligned}$$

We now easily find (cf. sec. 4.1)

$$F_\alpha(t) = \int_{-\infty}^{\infty} \{1+z+(\frac{1}{2}-t^\alpha)z^2-t^\alpha z^3\} e^{-tz^2} dz + t^{-\frac{1}{2}} O(t^{2\alpha-\frac{1}{2}}+t^{-\frac{3}{2}}),$$

and therefore,

$$F_\alpha(t) = \pi^{\frac{1}{2}} t^{-\frac{1}{2}} \{1 - \frac{1}{2} t^{\alpha-1} + O(t^{-1}) + O(t^{2\alpha-2})\} \quad (t \rightarrow \infty).$$

Here we dealt with a few terms only, but it is easily seen that the complete asymptotic series has the form

$$t^{-\frac{1}{2}} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} c_{kl} t^{-k} t^{-l(1-\alpha)}.$$

Case (ii). If  $\alpha = 1$ , the function  $e^{-tz^2}$  fails to be a good first approximation to the integrand within the range, the factor  $(1+tz^2)^{-1}$  changing the scene entirely. In order to get a better survey, we first transform the variable by  $z = t^{-\frac{1}{2}}w$ , thus obtaining a case where the range is independent of  $t$ :

$$F_1(t) = t^{-\frac{1}{2}} \int_{-\infty}^{\infty} (1+w^2)^{-1} \exp(-w^2+t^{-\frac{1}{2}}w) dw.$$

The rôle played thus far by the integrals  $\int_{-\infty}^{\infty} \exp(-w^2) w^k dw$ , is now taken over by integrals  $\int_{-\infty}^{\infty} \exp(-w^2) w^k (1+w^2)^{-1} dw$ . Otherwise, the technique is the same as in sec. 4.4. We obtain the asymptotic series

$$F_1(t) \approx \sum_{\nu=0}^{\infty} c_\nu t^{-\nu-\frac{1}{2}} \quad (t \rightarrow \infty),$$

where

$$c_\nu = \{(2\nu)!\}^{-1} \int_{-\infty}^{\infty} e^{-w^2} w^{2\nu} (1+w^2)^{-1} dw.$$

In the present simple case it is not difficult to show that the asymptotic series converges and that its sum equals  $F_1(t)$ , for all  $t > 0$ .

Case (iii).  $\alpha > 1$ . The factor  $(1 + t^\alpha z^2)^{-1}$  gives, in this case, two poles which are very close to the point  $z = 0$ , that is to say, very close compared to the dimensions of the range. And, the integration path passes just between the poles. We shall now shift the integration path downwards, over a distance  $2t^{-\frac{1}{2}\alpha}$ , say, taking the residue of the lower pole into account. The effect is, that the path has hardly been altered if we compare the shift to the size of the range, whereas on the other hand  $(1 + t^\alpha z^2)^{-1}$  can be expanded into powers of  $z^{-2}$  in all points of the new path. The residue at  $-it^{-\frac{1}{2}\alpha}$  equals

$$r(t) = \frac{1}{2}it^{-\frac{1}{2}\alpha} \exp(t^{1-\alpha} - it^{-\frac{1}{2}\alpha}),$$

and we obtain

$$F_\alpha(t) = -2\pi i r(t) + \int_P \exp(-tz^2 + z) \cdot (1 + t^\alpha z^2)^{-1} dz,$$

the integration path  $P$  being the straight line from  $-2i \cdot t^{-\frac{1}{2}\alpha} - \infty$  to  $-2i \cdot t^{-\frac{1}{2}\alpha} + \infty$ . On this path the expansion

$$(1 + t^\alpha z^2)^{-1} = t^{-\alpha} z^{-2} - t^{-2\alpha} z^{-4} + t^{-3\alpha} z^{-6} - \dots$$

converges uniformly (as  $|t^\alpha z^2| \geq 4$  on the path). The function  $e^z$  can also be expanded into powers of  $z$ , and within the range  $z$  is small. Therefore, it is quite easy to obtain an asymptotic series for  $F_\alpha(t) + 2\pi i r(t)$ . But the situation is even more favourable; we are able to prove that the asymptotic series is convergent for all positive values of  $t$ , and also that its sum equals  $F_\alpha(t) + 2\pi i r(t)$ . To this end we write

$$\begin{aligned} F_\alpha(t) + 2\pi i r(t) &= \\ &= \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \exp(-tz^2) (z^k/k!) (-1)^n t^{-(n+1)\alpha} z^{-2(n+1)} dx, \end{aligned}$$

where  $z = x - 2it^{-\frac{1}{2}\alpha}$ , and  $x$  is the new integration variable.

We want to carry out termwise integration here, and this has to be justified. We want to prove that

$$(5.12.4) \quad \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} f_{kn}(x) dx = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} f_{kn}(x) dx,$$

where

$$f_{kn}(x) = \exp(-tz^2) (z^k/k!) (-1)^n t^{-(n+1)\alpha} z^{-2(n+1)},$$

and  $z = x - 2it^{-\frac{1}{2}\alpha}$ ;  $t$  is a positive constant. It is not a difficult case, since there is absolute convergence. We have

$$\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} |f_{kn}(x)| = |\exp(-tz^2)| \cdot e^{|z|} \cdot (t^{\alpha}|z|^2 - 1)^{-1},$$

since  $t^{\alpha}|z|^2 \geq t^{\alpha}|2it^{-\frac{1}{2}\alpha}|^2 \geq 4$ . By virtue of the overwhelming power of the factor  $|\exp(-tz^2)|$  it is easily seen that

$$\int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} |f_{kn}(x)| dx$$

converges. It now follows from the theory of Lebesgue integrals<sup>1)</sup> that (5.12.4) holds, and also that

$$\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} f_{kn}(x) dx$$

converges absolutely. However, it is not difficult to give a more elementary proof, avoiding Lebesgue integrals: If  $h$  is an integer, then the double series  $\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} f_{kn}(x)$  converges uniformly in the interval  $h \leq x \leq h+1$ , whence it follows that

$$\int_h^{h+1} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} f_{kn}(x) dx = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \int_h^{h+1} f_{kn}(x) dx.$$

Writing  $\int_h^{h+1} f_{kn}(x) dx = c_{hkn}$ , we have, by virtue of the absolute convergence of the multiple series on the left,

$$\sum_{h=-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} c_{hkn} = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \sum_{h=-\infty}^{\infty} c_{hkn},$$

and this is equivalent to (5.12.4).

Returning to our original variable  $z$ , we deduce from (5.12.4) that

$$F_{\alpha}(t) + 2\pi i r(t) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} t^{-(n+1)\alpha} (-1)^n (k!)^{-1} \int_P \exp(-tz^2) z^{k-2n-2} dz,$$

and that this double series converges absolutely.

It remains to evaluate the integrals occurring on the right. We substitute  $-tz^2 = w$ ,  $z = -i(w/t)^{\frac{1}{2}}$ , where the principal value of the

<sup>1)</sup> See E. C. TITCHMARSH, Theory of Functions, 2nd ed., Oxford 1939, § 10.9 and § 1.77.

square root is taken. If  $z$  runs through  $P$ , then  $w$  describes a path  $C$  starting at  $-\infty$ , encircling the origin in the positive sense, and leading back to  $-\infty$ . The integral becomes

$$\begin{aligned} \int_P \exp(-tz^2) z^{k-2n-2} dz &= -\frac{1}{2} t^{-1} \int_P \exp(-tz^2) z^{k-2n-3} d(-tz^2) = \\ &= -\frac{1}{2} t^{-1} \int_C e^{w \cdot i} i^{2n+3-k} t^{\frac{1}{2}(2n+3-k)} w^{\frac{1}{2}(k-2n-3)} dw = \\ &= \pi t^{n-\frac{1}{2}(k-1)} i^{2n-k+2} / \Gamma(n + \frac{1}{2}(3-k)), \end{aligned}$$

according to Hankel's formula <sup>1)</sup> for  $(\Gamma(a))^{-1}$ . Our final result is that

$$F_\alpha(t) = \pi t^{-\frac{1}{2}\alpha} \exp(t^{1-\alpha} - it^{-\frac{1}{2}\alpha}), - \pi t^{-\alpha+\frac{1}{2}} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{t^{-n(\alpha-1)-\frac{1}{2}k} i^{-k}}{k! \Gamma(n + \frac{1}{2}(3-k))},$$

and the double series converges absolutely for all  $t > 0$ . As  $\alpha > 1$ , only negative exponents occur in the double series, and therefore the series is at the same time an asymptotic series.

If we expand the function  $\exp(t^{1-\alpha} - it^{-\frac{1}{2}\alpha})$  as a double series, we get again terms of the same type. Combining the two double series, one easily finds

$$F_\alpha(t) = \pi t^{-\frac{1}{2}\alpha} \sum_{h=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{h+m} t^{-\frac{1}{2}m(\alpha-1)-h}}{(2h)! \Gamma(\frac{1}{2}m+1-h)} \quad (t > 0).$$

### 5.13. EXERCISES. 1. Show that

$$\int_{-\infty}^{\infty} e^{itz} (1+x^2)^{-t} dx \sim (\pi(1-c)/t)^{\frac{1}{2}} e^{-ct} (2c)^{-t} \quad (t \rightarrow \infty),$$

where  $c = -1 + 2^{\frac{1}{2}}$

### 2. Show that

$$\int_{-1}^{\infty} (x^3 + 3x - 2i)^{-n} e^{ix} dx \sim 2e(i/4)^n (\pi/3n)^{\frac{1}{2}} \quad (n \rightarrow \infty),$$

where  $n$  runs through the integers.

### 3. Show that

$$\int_{i-\infty}^{i+\infty} e^{-z^2} (1+z)^{-n} dz \sim (\pi/2e)^{\frac{1}{2}} i^{-n} e^{\frac{1}{2}n} (\frac{1}{2}n)^{-\frac{1}{2}n} e^{i\sqrt{2n}} \quad (n \rightarrow \infty),$$

where  $n$  runs through the integers.

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, *Modern Analysis*, 4th ed., Cambridge, 1946, § 12.22.

## CHAPTER 6

### APPLICATIONS OF THE SADDLE POINT METHOD

In Ch. 5 we gave several relatively simple applications of the saddle point method, intended to illustrate the principles of the method with easy examples. Since simple examples of the saddle point method seem to be relatively rare in practice, we shall expose in this chapter three more difficult cases, in order to give an idea of the complications which may occur in such problems.

The first problem covers secs. 6.1—6.3. In this case we have a quite simple integrand. There are infinitely many saddle points, but as the contribution of the main saddle point is very large compared to the contribution of all others, these other saddle points need not even enter into the discussion. An extra difficulty is that the main saddle point cannot be represented explicitly; it is given by a transcendental equation. Fortunately, this equation has already been studied extensively in Ch. 2.

The second problem (secs. 6.4—6.7) is complicated because the integrand contains gamma functions. We have to simplify them by application of the Stirling formula, but this does not work for the whole integration path. The problem of finding a suitable path, however, is not exceedingly difficult in this case.

In the third example (secs. 6.8—6.10) the major difficulty lies in finding a suitable path. The parameter is a complex number in this case, and it is by no means easy to give a suitable path for all possible values of the parameter. The difficulty is overcome by application of conformal mapping, which usually is a very efficient instrument for obtaining a survey of the behaviour of an analytic function in a large area.

#### 6.1. The number of class-partitions of a finite set

Let  $S$  be a finite set. By a class-partition of  $S$  we denote a collec-

tion of non-empty subsets of  $S$ , which are mutually disjoint and whose union is  $S$ . For example, if  $S$  consists of the three elements  $a, b, c$ , then there are five possible class-partitions, viz. (i)  $(a)(b)(c)$ , (ii)  $(ab)(c)$ , (iii)  $(ac)(b)$ , (iv)  $(bc)(a)$ , (v)  $(abc)$ . The number of class-partitions of  $S$  obviously depends only on the number of elements of  $S$ . Now by  $d_n$  we denote the number of class-partitions of a set of  $n$  elements. One easily finds  $d_1 = 1$ ,  $d_2 = 2$ ,  $d_3 = 5$ ,  $d_4 = 15$ . Our problem is to determine the asymptotic behaviour of  $d_n$  as  $n \rightarrow \infty$ .

There is a recurrence relation, expressing  $d_{n+1}$  in terms of  $d_1, \dots, d_n$ :

$$(6.1.1) \quad d_{n+1} = \binom{n}{0} d_0 + \binom{n}{1} d_1 + \dots + \binom{n}{n} d_n \quad (n=0,1,2,\dots),$$

where  $d_0 = 1$ . The proof runs as follows. Let  $S$  have the elements  $a_1, \dots, a_n, a_{n+1}$ . Consider a class-partition of  $S$ , and assume that the subset which contains  $a_{n+1}$ , contains  $k$  further elements ( $0 \leq k \leq n$ ). If we fix  $k$ , there are  $\binom{n}{k} d_{n-k}$  class-partitions of this type. For, the  $k$  further elements just mentioned, can be chosen from the set  $\{a_1, \dots, a_n\}$  in  $\binom{n}{k}$  ways, and the set of the remaining  $n - k$  elements from  $\{a_1, \dots, a_n\}$  admits  $d_{n-k}$  class-partitions. (If  $k = n$ , there are no remaining elements, but the convention  $d_0 = 1$  covers this case). Now carrying out summation with respect to  $k$ , we obtain (6.1.1).

Starting from the recurrence relation (6.1.1), we proceed by the method of generating functions. Putting

$$D(z) = \sum_{n=0}^{\infty} d_n z^n / n!,$$

we deduce that  $D'(z) = e^z \cdot D(z)$ , and, as  $D(0) = 1$ , it follows that  $D(z) = \exp(e^z - 1)$ . Therefore, the  $d_n/n!$  are the coefficients in the expansion

$$(6.1.2) \quad \exp(e^z - 1) = \sum_{n=0}^{\infty} d_n z^n / n!.$$

A more direct way of counting the class-partitions, leading to the same formula (6.1.2), is the following one. Consider a class

partition of the set  $S$ ,  $S$  having  $n$  elements. This is a collection of subsets; let in this collection the number of subsets having  $j$  elements be denoted by  $s_j$ . So  $s_j \geq 0$  ( $j = 1, 2, 3, \dots$ ), and  $n = s_1 + 2s_2 + 3s_3 + \dots$ . We now fix a sequence  $s_1, s_2, \dots$ , satisfying these conditions, and we ask for the number of class-partitions corresponding to this sequence. This number is easily seen to be equal to

$$n! \{ (1!)^{s_1} (2!)^{s_2} (3!)^{s_3} \dots s_1! s_2! s_3! \dots \}^{-1}.$$

It follows that  $d_n/n!$  equals the coefficient of  $z^n$  in the power series development of

$$\sum_{s_1=0}^{\infty} \frac{z^{s_1}}{s_1!} \sum_{s_2=0}^{\infty} \left( \frac{z^2}{2!} \right)^{s_2} \frac{1}{s_2!} \sum_{s_3=0}^{\infty} \left( \frac{z^3}{3!} \right)^{s_3} \frac{1}{s_3!} \dots,$$

and this represents

$$\exp(z) \cdot \exp\left(\frac{z^2}{2!}\right) \cdot \exp\left(\frac{z^3}{3!}\right) \dots = \exp\left(z + \frac{z^2}{2!} + \dots\right) = \exp(e^z - 1).$$

## 6.2. Asymptotic behaviour of $d_n$

We shall study the asymptotic behaviour of the coefficients in (6.1.2) by expressing them via Cauchy's formula for the coefficients of a power series:

$$(6.2.1) \quad 2\pi i d_n/n! = \int_C \exp(e^z) z^{-n-1} dz,$$

where the integration path  $C$  encircles the origin once, in the positive sense. To this integral we shall apply the saddle point method. The saddle points are the roots of  $ze^z = n + 1$ . This equation has one positive solution, discussed in sec. 2.4, but this is not the only solution. Actually it can be shown, for each integer  $k$ , that there is just one saddle point in the horizontal strip

$$(2k - 1)\pi < \operatorname{Im} z < (2k + 1)\pi \quad (k = 0, \pm 1, \pm 2, \dots),$$

provided that  $n$  is sufficiently large.

Let the positive saddle point, i.e. the positive solution of  $ze^z = n + 1$ , be denoted by  $u$ . Fortunately, we are in a position where the other saddle points can be disregarded, that is to say, we are able to find a path through  $u$ , of which  $u$  itself is the highest point.

The axis of the saddle point  $u$  is easily seen to be vertical. Proceeding along the principle that the simplest possibilities should be tried first, we try to take a large part of the path as a vertical line. Along this vertical line through the point  $u$  we have  $|\exp e^z| \leq \exp e^u$ , since  $\operatorname{Re} e^z \leq |e^z| = e^{\operatorname{Re} z} = e^u$  holds on that line. Secondly, the factor  $z^{-n-1}$  is, in absolute value, maximal at the point  $u$ . So a vertical path satisfies the requirement that the integrand should attain its maximal absolute value at the saddle point.

However, a vertical line does not encircle the origin. But if we take a large segment of the vertical line, and complete it to a closed contour by adding a large semi-circle, it does. And, if we make the radius  $R$  of the semi-circle tend to infinity, its contribution to the integral (6.2.1) tends to zero (if  $n > 0$ ), the factor  $z^{-n-1}$  being  $O(R^{-n-1})$ , whereas  $\exp e^z$  is bounded in the half-plane  $\operatorname{Re} z \leq u$ . Therefore, the integral  $\int_C$  in (6.2.1) may be replaced by  $\int_{u-i\infty}^{u+i\infty}$ .

The integrand is

$$\exp(e^z - (n+1)\log z) = \exp(e^z - ue^u \log z).$$

Writing  $z = u + iy$ , we obtain

$$(6.2.2) \quad 2\pi i e_n/n! = \exp(e^u - ue^u \log u) \int_{-\infty}^{\infty} \exp \psi(y) dy,$$

where

$$\psi(y) = e^u[(e^{iy} - 1) - u \log(1 + iyu^{-1})].$$

As  $|\exp \psi(y)| = \exp \operatorname{Re} \psi(y)$ , we have to study

$$\operatorname{Re} \psi(y) = e^u[-1 + \cos y - u \log(1 + y^2u^{-2})].$$

As long as  $y$  is not too large, the terms  $-1 + \cos y$  are predominant. Therefore, there is a maximum at  $y = 0$ , there are further maxima around  $y = \pm 2\pi$ , etc. The influence of these further maxima is very small, because of the large factor  $e^u$  in front. We shall show that in (6.2.2) we can restrict ourselves essentially to the interval  $-\pi < y < \pi$ .

If  $\pi < y < u$ , then we have  $\log(1 + y^2/u^2) > \frac{1}{2}y^2/u^2$ , and therefore

$$\left| \int_u^\infty \exp \psi(y) dy \right| < u \cdot \exp\{-\pi^2 e^u / (4u)\}.$$

If  $y > u$ , then we use  $1 + y^2/u^2 > 2y/u$ , and so we have, putting  $y = ux$ ,

$$\left| \int_u^\infty \exp \psi(y) dy \right| \leq u \int_1^\infty \exp\{-\frac{1}{2}e^{uy} u \log(2x)\} dx.$$

It is easily seen that  $\int_1^\infty (2x)^{-p} dx = O(e^{-tp})$  ( $p > 2$ ), and therefore

$$\int_u^\infty \exp \psi(y) dy = O(u \cdot \exp[-\frac{1}{2}ue^u]) \quad (u \rightarrow \infty).$$

It follows that

$$(6.2.3) \int_{-\infty}^\infty \exp \psi(y) dy - \int_{-\pi}^\pi \exp \psi(y) dy = O\{\exp(-eu/u)\} \quad (u \rightarrow \infty),$$

and now we can direct our attention to the interval  $(-\pi, \pi)$ , where the saddle point at  $y = 0$  gives the main contribution.

As

$$\psi(y) = e^u \left[ -\frac{y^2}{2!} + \frac{(iy)^3}{3!} + \dots - \frac{y^2}{2u} - \frac{(iy)^3}{3u^2} - \dots \right],$$

we find by the Laplace method (Ch. 4),

$$(6.2.4) \int_{-\pi}^\pi \exp \psi(y) dy = (2\pi e^{-u})^{\frac{1}{2}} \{1 + O(u^{-1})\}.$$

In order to get an asymptotic expansion, it seems to be inconvenient to follow the method of sec. 4.4, as  $\psi$  depends on  $u$  in a rather complicated way. We prefer to apply the method used in sec. 4.5, consisting of the introduction of a new integration variable  $w$  such that  $\exp \psi(y)$  is transformed into  $\exp(-w^2 f(u))$  ( $f(u)$  depending on  $u$  only).

So far  $y$  was real; we shall now treat it as a complex variable. If  $|y|$  is small, we define  $w$  by

$$e^{iy} - 1 - u \log(1 + iyu^{-1}) = -\frac{1}{2}w^2(1 + u^{-1}),$$

choosing the root  $w$  which satisfies  $dw/dy = +1$  at  $y = 0$ . Now  $w$  can be written as

$$(6.2.5) \quad w = y + y^2 P(y, u^{-1}),$$

where  $P(y, u^{-1})$  is a power series in the variables  $y$  and  $u^{-1}$ , convergent if both  $y$  and  $u^{-1}$  are sufficiently small. It follows that  $y$  can be solved from (6.2.5), by the technique explained in sec. 2.4

(cf. (2.4.7)):

$$y = y(w) = \frac{1}{2\pi i} \int_C \{\eta + \eta^2 P(\eta, u^{-1}) - w\}^{-1} \eta \frac{d}{d\eta} \{\eta + \eta^2 P(\eta, u^{-1})\} d\eta,$$

where  $C$  is a contour encircling the origin in the positive direction, as long as  $w$  and  $u^{-1}$  are sufficiently small. Therefore,  $y$  is a power series in  $w$  and  $u^{-1}$ . We also need its derivative

$$dy/dw = 1 + w\gamma_1(u^{-1}) + w^2\gamma_2(u^{-1}) + \dots,$$

where  $\gamma_1, \gamma_2, \dots$  are power series in terms of  $u^{-1}$ , simultaneously convergent for large values of  $u$ .

It is not quite sure that the integral (6.2.4) can be transformed successfully this way, but at any rate we can find a positive number  $c$  such that it works for  $\int_{-c}^c$ . In the  $w$ -plane the integration path becomes a curve which crosses the saddle point at  $w = 0$ . The errors made by breaking off the integrals at  $c$  and  $-c$ , and at the corresponding points in the  $w$ -plane, are in no way alarming, as they are of the type  $\exp(-Ae^u)$ , where  $A$  is a positive constant.

Our final result is

$$(2\pi)^{-\frac{1}{2}} U^{\frac{1}{2}} \int_{-\infty}^{\infty} \exp w(y) dy \approx \\ \approx 1 + a_1 y_2(u) U^{-1} + a_2 y_4(u) U^{-2} + \dots \quad (u \rightarrow \infty),$$

where  $U = e^u(1 + u^{-1})$ ,  $a_k = (2k)! (k!)^{-1} 2^{-k}$ .

Using one term only, we get the following expression for  $d_n$ :

$$(6.2.6) \quad d_n = n! e^{-1} (2\pi)^{-\frac{1}{2}} (1 + u^{-1})^{-\frac{1}{2}} \cdot \{1 + O(e^{-u})\} \exp(e^u - ue^u \log u - \frac{1}{2}u).$$

Here  $u$  is related to  $n$ , by the formula

$$ue^u = n + 1, \quad u > 0.$$

The asymptotic behaviour of  $u$ , as  $n \rightarrow \infty$ , was investigated in sec. 2.4. By (2.4.10), which was the solution of the equation  $xe^x = t$ , we now have

$$u = \log t - \log \log t + (\log \log t)(\log t)^{-1} Q(\sigma, \tau),$$

where  $Q(\sigma, \tau)$  is a double power series, and we abbreviate

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

However, if we approximate  $u$  by taking a finite number of terms of  $Q(\sigma, \tau)$ , the error introduced in (6.2.6) becomes considerable and much of the accuracy obtained in (6.2.6) is lost. Approximating  $u$  by  $\log t - \log \log t$ , we find that

$$\begin{aligned} e^u - ue^u \log u - \frac{1}{2}u &= t(u^{-1} - \log u) + O(\log t) = \\ &= t \left\{ -\log \log t + \frac{\log \log t}{\log t} + \frac{1}{\log t} + \frac{1}{2} \left( \frac{\log \log t}{\log t} \right)^2 \right\} + \\ &\quad + O \left( \frac{t \log \log t}{(\log t)^2} \right). \end{aligned}$$

Here  $t$  is still  $n+1$ , but it is easy to see that when replacing  $n+1$  by  $n$ , we make an error which is much smaller than the error already involved. Furthermore, we use a rough estimate for  $n!$ :

$$\log n! = n \log n - n + O(\log n),$$

and we find that

$$(6.2.7) \quad \begin{aligned} \frac{\log d_n}{n} &= \log n - \log \log n - 1 + \frac{\log \log n}{\log n} + \\ &+ \frac{1}{\log n} + \frac{1}{2} \left( \frac{\log \log n}{\log n} \right)^2 + O \left( \frac{\log \log n}{(\log n)^2} \right) \quad (n \rightarrow \infty). \end{aligned}$$

It is quite easy to replace the  $O$ -term by an asymptotic series, with terms of the form  $(\log \log n)^k (\log n)^{-m}$ .

### 6.3. Alternative method

We briefly indicate an entirely different method by which the asymptotic behaviour of the sequence  $d_n$ , dealt with in sec. 6.2, can be obtained. Starting from (6.1.2), we expand  $\exp ex$  as  $\sum_{k=0}^{\infty} ekx/k!$ , and in each term we expand  $e^{kx}$  into its power series. So we get an absolutely convergent double series, in which the order of summation may be changed, and (6.1.2) gives

$$d_n = e^{-1} \sum_{k=0}^{\infty} k^n / k!.$$

This sum can be tackled by the methods of Ch. 3. The index  $k_{\max}$  of the maximal term lies close to  $e^u$ , where  $u$  is again the so-

lution of the equation  $ue^u = n + 1$ . For,  $k^n/k!$  roughly equals

$$(2\pi)^{-1/2} k^{-3/2} \exp((n+1)\log k - k \log k + k),$$

and the function  $(n+1)\log x - x \log x + x$  is maximal at  $x = e^u$ . The second derivative of this function is  $-(n+1)x^{-2} - x^{-1}$ , and it follows that, roughly speaking, it is only an interval  $|k - k_{\max}| < n^{1/2}$  that gives a substantial contribution to the sum. In this interval the sum can be replaced by an integral, if we only carry out corrections according to the Euler-Maclaurin sum formula.

#### 6.4. The sum $S(s, n)$

In sec. 4.7 we obtained the asymptotic behaviour of  $S(s, n)$ , defined by

$$(6.4.1) \quad S(s, n) = \sum_{k=0}^{2n} (-1)^{k+n} \binom{2n}{k}^s,$$

if  $s$  is a fixed integer  $> 1$ , and  $n \rightarrow \infty$ . The method was definitely restricted to this case, as  $s - 1$  occurred as the number of dimensions of a euclidean space. In the present section we shall study this sum for general real values of  $s$  ( $s$  fixed,  $n \rightarrow \infty$ ). It should be admitted that this is not a very natural question, as non-integral powers of binomial coefficients do not frequently occur in mathematics. The main reason for its discussion here is, that it is a difficult problem with various interesting aspects.

If  $s = 2, 3, \dots$  we have, by (4.7.4),

$$(6.4.2) \quad S(s, n) \sim \{2 \cos(\pi/2s)\}^{2ns+s-1} 2^{2-s} (\pi n)^{1/(1-s)} s^{-1}$$

if  $n \rightarrow \infty$ . This formula is definitely false if  $s = 0$ , for  $S(0, n)$  equals  $(-1)^n$ . If  $s$  is negative, it does not hold either. Then we are in the situation that the first few terms, and the last few, are prevailing. If  $s < 0$ , we have the asymptotic series

$$S(s, n) \approx 2 \cdot (-1)^n \left\{ 1 - \binom{2n}{1}^s + \binom{2n}{2}^s - \dots \right\} \quad (n \rightarrow \infty),$$

or, more explicitly, for  $h$  and  $s$  fixed ( $s < 0$ ), and  $n \rightarrow \infty$ ,

$$(6.4.3) \quad S(s, n) = 2 \cdot (-1)^n \sum_{k=0}^h (-1)^k \binom{2n}{k}^s + O(n^{(h+1)s}).$$

(In order to prove this, we have to consider all terms for which  $k < h + (-s)^{-1}$ , and to remark that the sum of the remaining terms with  $k \leq n$  is at most  $n$  times the first neglected term).

If  $s > 0$ , there is no such trivial way. We shall then replace the sum by an integral, by application of the residue theorem:

$$(6.4.4) \quad S(s, n) = \int_C \left( \frac{\Pi(2n)}{\Pi(n+z)\Pi(n-z)} \right)^s \frac{dz}{2i \sin \pi z}.$$

Here the integration path  $C$  is a curve which encircles the points  $-n, -n+1, \dots, -1, 0, 1, 2, \dots, n$  just once (in the positive sense), but does not encircle the points  $\pm(n+1), \pm(n+2), \dots$ . We might take for  $C$  a rectangle, with vertices  $\pm(n+\frac{1}{2}) \pm pi$ , where  $p$  is some positive number. The function  $\Pi(z)$  is Gauss' extension of the factorial ( $\Pi(k) = k!$ ); it is slightly more convenient to operate with  $\Pi(z)$  than with the gamma function, which is related to  $\Pi$  by the formula  $\Pi(z) = \Gamma(z+1)$ .

The function  $\Pi(n+z)\Pi(n-z)$  has poles at the points  $\pm(n+1), \pm(n+2), \dots$ , and therefore the integrand is, unless  $s$  is an integer, a multi-valued function. However, if we cut the  $z$ -plane along the real axis from  $n+1$  to  $+\infty$  and from  $-n-1$  to  $-\infty$ , the integrand is single-valued in the remaining domain  $D$  of the  $z$ -plane. Needless to say, the value of the  $s$ -th power occurring in the integrand, is given its positive value at  $z=0$ . Furthermore we can take care that the path  $C$  lies entirely inside the domain  $D$ . Now it is easy to prove (6.4.4): the residue at  $z=h$  ( $h=0, \pm 1, \dots, \pm n$ ) is

$$\{(2n)!/(n+h)!(n-h)!\}^s \cdot (-1)^h \cdot (2\pi i)^{-1},$$

and on replacing  $h-n$  by  $k$ , the integral turns out to be equal to  $S(s, n)$ .

Before starting any serious work with the integral (6.4.4) we make some observations.

(i) The integrand is an odd function of  $z$ . Therefore, if  $C$  is symmetric with respect to the origin, it is sufficient to consider only half the integration path, and to multiply the result by 2.

(ii) If  $z$  is somewhere in the right half-plane, then the absolute value of the integrand decreases on replacing  $z$  by  $z+1$ . For, by this operation, this absolute value is multiplied by the factor

$|((n-z)/(n+z+1))|^s$ . This makes it plausible that something like a saddle point can be expected on the upper, and also on the lower part, of the imaginary axis.

(iii) The integrand is far too complicated for calculations of exact saddle points. However, on a large part of the path  $C$  we have a reasonable approximation to the integrand, by the Stirling formula. Precisely, if  $\delta$  is a positive number ( $0 < \delta < \pi$ ), and if  $R_\delta$  denotes the sector  $|\arg z| < \pi - \delta$ , then we infer from (3.10.7) (with  $1 - \rho(1) = \frac{1}{2} \log 2\pi$ ) that

$$(6.4.5) \quad \Pi(z) = (2\pi)^{\frac{1}{2}} z^{s+\frac{1}{2}} e^{-z} \{1 + O(|z^{-1}|)\} \quad (z \in R_\delta),$$

where  $z^{s+\frac{1}{2}} = e^{(s+\frac{1}{2})\log z}$ , with the principal value of the logarithm.

Formula (6.4.5) means that the integrand of (6.4.4) can be approximated by elementary functions as long as  $z$  stays sufficiently far away from the boundary of  $D$ , i.e. from the half lines  $(-\infty, -n-1)$  and  $(n+1, \infty)$ . However, we are not in a position to apply this to the whole path  $C$ , as  $C$  has to cross the real axis between  $n$  and  $n+1$ , and this is *not* far from the boundary of  $D$ .

(iv) It will not be difficult to find a useful approximation to the integrand, also by the Stirling formula, for values of  $z$  which are not too far from  $z = n$ . Therefore, we shall have to work with two different approximations, in different regions. This gives, of course, some difficulty in fitting the respective parts of the path together. This is not so much an essential difficulty, but rather a technical one, caused by the relative complexity of the integrand.

(v) The difficulty just stated can be overcome by making the connection between the two regions far away from the main battle field, viz. at  $+\infty$  and  $-\infty$ , respectively. Remark (ii) suggests that a retreat to  $+\infty$  or to  $-\infty$  is comparatively easy.

Following the above suggestions, we shall split the problem into two parts. Let  $N$  be an integer  $> n$ , and let  $\rho$  be a positive number. We define  $P_N$  and  $Q_N$  by

$$P_N = \int_{-N-\frac{1}{2}+ip}^{N+\frac{1}{2}+ip} \Omega dz, \quad Q_N = \int_{(N+\frac{1}{2})}^{\infty} \Omega dz.$$

Here  $\Omega$  is an abbreviation for the integrand of (6.4.4). In the case of  $P_N$ , the integration path is a straight line. The number  $\rho$  is a positive constant. The integration path of  $Q_N$  starts at  $N+\frac{1}{2}$ , pro-

ceeds through the lower half-plane, crosses the real axis between  $n$  and  $n + 1$ , and then leads back to  $N + \frac{1}{2}$  through the upper half-plane. Therefore, this path lies, apart from its end-points, inside the domain  $D$ . We shall show, if  $n$  is fixed, but sufficiently large, that

$$\lim_{N \rightarrow \infty} P_N = P, \quad \lim_{N \rightarrow \infty} Q_N = Q$$

both exist, and that

$$(6.4.6) \quad S(s, n) = -2P + 2Q.$$

These statements will easily follow from the fact that the integrand  $\Omega$  is  $O(|z|^{-2})$  in the domain described by

$$(6.4.7) \quad |\operatorname{Im} z| \leq p, \quad |z - k| < \frac{1}{4} \quad (k = 0, \pm 1, \pm 2, \dots),$$

provided that  $n$  is fixed but sufficiently large (it suffices that  $(2n+1)s \geq 2$ ). For, by the functional equation

$$\Pi(z) \cdot \Pi(-z) = \pi z / \sin \pi z,$$

we have

$$(6.4.8) \quad \Omega = \frac{1}{2i} \left( \frac{(-1)^n (2n)! \Pi(z-n)}{\pi(z-n) \Pi(z+n)} \right)^s (\sin \pi z)^{s-1},$$

whence it follows that

$$(6.4.9) \quad \Omega = O(|z|^{-(2n+1)s} |\sin \pi z|^{s-1}) \quad (|z| > 2n)$$

if  $s$  and  $n$  are fixed ( $s > 0, n > 0$ ). This implies that  $\Omega = O(|z|^{-2})$  if  $z$  lies in the region just considered, provided that  $(2n+1)s \geq 2$ .

As to formula (6.4.6) we refer to remark (i), which shows that  $\frac{1}{2}S(s, n)$  equals  $-P_N + Q_N$  plus an integral from  $N + \frac{1}{2}$  to  $N + \frac{1}{2} + ip$  and a similar integral in the lower half-plane. The latter integrals tend to zero as  $N \rightarrow \infty$ , since  $\Omega = O(|z|^{-2})$  ( $n$  fixed).

There are now two separate problems, viz. the asymptotic behaviour of  $P$  and  $Q$ , respectively, as  $n \rightarrow \infty$ .

## 6.5. Asymptotic behaviour of $P$

We shall first deal with  $P = \int_{ip-\infty}^{ip+\infty} \Omega dz$ , where  $p$  is a positive number. Clearly,  $P$  is independent of  $p$  (cf. 6.4.9). We shall approximate  $\Omega$  with the aid of Stirling's formula. However, a slight extension of (6.4.5) is necessary, as our values  $n+z$  and  $n-z$  do not always belong to the sector  $R_s$ . If we use the relation  $\Pi(z)\Pi(-z)$

$= \pi z / \sin \pi z$ , the behaviour of  $\Pi(z)$  in the quadrant  $\operatorname{Im} z > 0$ ,  $\operatorname{Re} z < 0$  can be deduced from its behaviour in the opposite quadrant. By a careful discussion of the arguments of  $z^{s+\frac{1}{2}}$  and  $(-z)^{-s+\frac{1}{2}}$  (the principal values are related by  $(-z)^\nu = e^{-\pi i \nu} z^\nu$ ) we find that

$$\Pi(z)\{1 - e^{2\pi iz}\} = (2\pi)^{\frac{1}{2}} z^{s+\frac{1}{2}} e^{-z}\{1 + O(z^{-1})\} \quad (\operatorname{Re} z < 0, \operatorname{Im} z > 0).$$

It now easily follows that we have, if  $p_0$  is positive and fixed,

$$\Pi(z) = (2\pi)^{\frac{1}{2}} z^{s+\frac{1}{2}} e^{-z}\{1 + O(|z^{-1}|) + O(e^{-2\pi|\operatorname{Im} z|})\} \quad (|\operatorname{Im} z| > p_0).$$

From this formula we can immediately deduce an estimate for  $\Omega$  (i.e. the integrand of (6.4.4)) by some trivial calculations:

$$(6.5.1) \quad \Omega = -2^{2ns}(\pi n)^{-\frac{1}{2}s}(1 - \zeta^2)^{-\frac{1}{2}s}.$$

$$\begin{aligned} &\cdot \exp \left[ n \left\{ -s \log(1 - \zeta^2) - s\zeta \log \frac{1 + \zeta}{1 - \zeta} + \pi i \zeta \right\} \right] \cdot \\ &\cdot \{1 + O(n^{-1} + n^{-1}|1 + \zeta|^{-1} + n^{-1}|1 - \zeta|^{-1} + e^{-2\pi n |\operatorname{Im} \zeta|})\} \\ &\quad (n > 1, \operatorname{Im} n\zeta > p). \end{aligned}$$

Here  $s$  and  $p_0$  are positive and fixed, and  $\zeta$  stands for  $z/n$ .

The integrand is, roughly speaking, of the type considered in sec. 5.7, if we put

$$(6.5.2) \quad \psi(\zeta) = -s \log(1 - \zeta^2) - s\zeta \log \frac{1 + \zeta}{1 - \zeta} + \pi i \zeta.$$

We remark that  $s$  is a fixed positive constant, and that the multi-valued functions  $\log(1 + \zeta)$  and  $\log(1 - \zeta)$  are given their principal values. Disregarding the  $O$ -terms of (6.5.1) for the moment, we start looking for saddle points of  $\psi$ . We have

$$(6.5.3) \quad \psi'(\zeta) = \pi i - s \log\{(1 + \zeta)/(1 - \zeta)\},$$

and therefore  $\psi'(\eta) = 0$  if  $\eta = i \tan(\pi/2s)$ . So we observe that there is a saddle point in the upper half-plane if and only if  $s > 1$ . First assuming  $s > 1$ , we take for integration path the infinite straight horizontal line through the saddle point  $\eta$ . As we have

$$\begin{aligned} \psi''(\eta) &= -2s(1 - \eta^2)^{-1} = \\ &= -2s(1 + \tan^2(\pi/2s))^{-1} = -2s \cos^2(\pi/2s) < 0, \end{aligned}$$

this line coincides with the axis of the saddle point.

Our path can be described by  $\zeta = \eta + x$  ( $-\infty < x < \infty$ ).

Fortunately, the saddle point  $\eta$  is "the highest point" of the path, as far as  $\psi$  is concerned.  $\operatorname{Re} \psi(\zeta)$  decreases if  $x$  increases from 0 to  $\infty$ , and also if  $x$  runs from 0 to  $-\infty$ . For, by (6.5.3) we have

$$\operatorname{Re} \frac{d\psi}{dx} = -s \log \left| \frac{1+\zeta}{1-\zeta} \right| < 0 \quad (x > 0),$$

as  $|1+\zeta| > |1-\zeta|$  if  $x > 0$ . If  $x < 0$ , we have some thing similar.

By sec. 5.7 we have

$$(6.5.4) \quad \int_{\eta-1}^{\eta+1} (1-\zeta^2)^{-\frac{1}{2}s} \exp\{n\psi(\zeta)\} d\zeta = \\ = (2\pi)^{\frac{1}{2}} n^{-\frac{1}{2}} |\psi''(\eta)|^{-\frac{1}{2}} \exp\{n\psi(\eta)\} (1-\eta^2)^{-\frac{1}{2}s} \{1 + O(n^{-1})\}.$$

We have

$$\begin{aligned} \psi''(\eta) &= -2s(\cos(\pi/2s))^2, \\ \exp\{n\psi(\eta)\} &= (\cos(\pi/2s))^{2ns}, \\ (1-\eta^2)^{-\frac{1}{2}s} &= (\cos(\pi/2s))^s, \end{aligned}$$

and therefore the value of (6.5.4) is reduced to

$$(6.5.5) \quad (\pi/ns)^{\frac{1}{2}} (\cos(\pi/2s))^{2ns+s-1} \{1 + O(n^{-1})\}.$$

Now returning to  $\Omega$ , we remark that on our path we have  $\operatorname{Im} n\zeta = n \tan(\pi/2s)$ , and this exceeds  $p_0$  if  $n$  is sufficiently large. So (6.5.1) can be applied. Moreover, the O-terms in (6.5.1) can be reduced to one term  $O(n^{-1})$ . It is not difficult to show that

$$\int_{\eta-1}^{\eta+1} |(1-\zeta^2)^{-\frac{1}{2}s} \exp(n\psi(\zeta)) \cdot O(n^{-1})| d\zeta = O\{n^{-\frac{3}{2}s} \exp(n\psi(\eta))\},$$

and therefore these O-terms result in a correction of the same order as the O-term in (6.5.5).

The fact that  $\operatorname{Re} \psi(\zeta)$  decreases (if  $x \rightarrow \infty$ ), combined with the occurrence of the factor  $(1-\zeta^2)^{-\frac{1}{2}s}$ , makes that hardly any complication is caused by the circumstance that the path has infinite length, so, fortunately, we need not go into the trouble of investigating  $\psi(\zeta)$  as  $x \rightarrow \infty$ . Actually we infer that

$$\left| \int_{\eta+1}^{\eta+\infty} \Omega d\zeta \right| < C 2^{2ns} (\pi n)^{-\frac{1}{2}s} |\exp(n\psi(\eta+1))| \int_{\eta+1}^{\eta+\infty} |1-\zeta^2|^{-\frac{1}{2}s} d\zeta$$

with some positive number  $C$  (independent of  $n$ ). The latter integral is easily seen to converge by virtue of our assumption  $s > 1$ .

Needless to say, the integral  $\int_{\eta-\infty}^{\eta+\infty} \Omega dz$  can be estimated in the same way.

Collecting the various results, we obtain our final formula for  $P$ :

$$(6.5.6) \quad P = \int_{\eta-\infty}^{\eta+\infty} \Omega dz = n \int_{\eta-\infty}^{\eta+\infty} \Omega d\zeta = \\ = -\{2 \cos(\pi/2s)\}^{2ns+s-1} 2^{1-s} (\pi n)^{\frac{1}{2}(1-s)} s^{-\frac{1}{2}} \{1 + O(n^{-1})\},$$

valid if  $s$  is fixed and  $> 1$ .

It may be remarked that the factor  $\{1 + O(n^{-1})\}$  can be replaced by an asymptotic series  $\sum_0^\infty c_k n^{-k}$ . In order to obtain this series it is necessary, of course, to use the Stirling formula in the form of an asymptotic series.

The saddle point  $\eta$  moves to  $i\infty$  when  $s$  decreases to 1. This suggests that, if  $s \leq 1$ , an estimate for the integral can be obtained by shifting the whole horizontal path upwards to infinity. In fact we can show that, if  $0 < s \leq 1$ ,  $n > \frac{1}{2}(s^{-1} - 1)$ , the integral  $P$  vanishes. For, if  $0 < s \leq 1$  and  $\operatorname{Im} z > \phi_0$  ( $\phi_0$  positive and fixed) we have  $|\sin \pi z|^{s-1} = O(1)$ . Now we deduce from (6.4.9) that  $\Omega = O(|z|^{-\lambda})$  ( $|z| > 2n$ ,  $\operatorname{Im} z > \phi_0$ ) with a constant  $\lambda > 1$ , if  $(2n+1)s > 1$ ,  $0 < s \leq 1$ ,  $n$  and  $s$  fixed. It follows that

$$P = \int_{i\phi-\infty}^{i\phi+\infty} \Omega dz = O(\phi^{1-\lambda}) \quad (\phi \rightarrow \infty),$$

in this case. Since  $P$  does not depend on  $\phi$ , we obtain that  $P = 0$  ( $0 < s \leq 1$ ,  $(2n+1)s > 1$ ).

## 6.6. Asymptotic behaviour of $Q$

We next turn our attention to  $Q$ , defined in sec. 6.4. We assume throughout that  $s > 0$ . Of course, we expect a saddle point on the real axis between  $n$  and  $n+1$ , close to  $n+1$ . If we replace, in the integrand,  $\Pi(n+z)$  by its Stirling approximation, such a saddle point turns out to exist, but only if  $0 < s < 1$ . Moreover, the singularity at  $n+1$  can be shown to lie within the range of this saddle point. Therefore the ordinary saddle point analysis does not apply in the case of our integral for  $Q$ .

In order to study  $\Omega$  for values of  $z$  near  $n+1$ , we use (6.4.8). For a first orientation, we remark that the only factor in (6.4.8) depending "heavily" on  $n$ , apart from the trivial multiplier  $\{(2n)!\}^s$ ,

is the factor  $\{\Pi(z+n)\}^{-s}$ . To this factor we apply the following direct consequence of Stirling's formula:

$$(6.6.1) \quad (2n)!/\Pi(x+2n) = (2n)^{-x}\{1 + O(n^{-1})\} \quad (x \text{ fixed}, n \rightarrow \infty),$$

and we take  $z = n+x$ .

Because of the factor  $(2n)^{-x}$  it seems to be sensible to pull the integration path as far as possible to the right. But needless to say, we cannot pull it over the branch point at  $z = n+1$ .

We shall therefore deform the integration path of the integral  $Q$  by making a path following the real axis from  $\infty$  to  $n+1$ , taking the values of the multi-valued functions which correspond to their values in the lower half-plane, and back from  $n+1$  to  $\infty$  with the values from the upper half-plane. (Notice that the integrand is one-valued in the domain  $D$ , described in sec. 6.4). This path has to be provided with small semi-circles around  $n+2, n+3, \dots$  and with a full circle around  $n+1$  in the well-known fashion, in order to circumvent the singularities, but, as commonly happens in such cases, these circles can be removed by making their radii tend to zero. Furthermore it should be remarked that the integral was originally not defined as an integral from  $\infty$  to  $\infty$ , but as the limit of an integral  $Q_N$  leading from  $N+\frac{1}{2}$  to  $N+\frac{1}{2}$ , where  $N$  runs through the integers. However, it is easily shown that this makes no difference, as  $\Omega = O(|z|^{-2})$  in the region (6.4.7).

According to (6.4.8), we write  $\Omega$  in the form

$$\Omega = (-1)^n \pi^{-s} (2i)^{-1} \{(2n)! \Pi(z-n-1)/\Pi(z+n)\}^s (\sin \pi(z-n))^{s-1}.$$

The function  $\Pi(z-n-1)/\Pi(z+n)$  is single-valued in the half-plane  $\operatorname{Re} z > n$ , and positive if  $z$  is real,  $z > n$ . The behaviour of  $\{\sin \pi(z-n)\}^{s-1}$  is also easy to describe. It is positive if  $n < z < n+1$ . Its argument increases by  $\pi(s-1)$  if we pass the branch point  $n+k$  ( $k = 1, 2, 3, \dots$ ) by a semicircle in the lower half-plane from  $n+k-\delta$  to  $n+k+\delta$ , and for the similar thing in the upper half-plane we find  $-\pi(s-1)$ . So we obtain, again writing  $z = n+x$ ,

$$(6.6.2) \quad Q = (-1)^{n+1} \pi^{-s} \int_1^\infty A dx,$$

where

$$A = \{(2n)! \Pi(x-1)/\Pi(x+2n)\}^s \cdot |\sin \pi x|^{s-1} \cdot \sin\{\pi(s-1)[x]\},$$

and  $[x]$  denotes the largest integer  $\leq x$ . A first consequence is that  $Q = 0$  if  $s$  is an integer  $> 0$ , for then  $\sin(\pi(s-1)[x])$  vanishes identically.

It will turn out that the main contribution to the integral is given by values of  $x$  close to 1. In the interval  $1 \leq x \leq 2$ , say, we can use (6.6.1), which is easily seen to hold uniformly with respect to  $x$  in that interval. It results that  $\int_1^2 Adx$  can be written as

$$\int_1^2 (2n)^{-xs} \{ \Pi(x-1) \}^s |\sin \pi x|^{s-1} \{ 1 + O(n^{-1}) \} \sin \{ \pi(s-1) \} dx.$$

This integral is of the type of those discussed in sec. 4.3,  $\log 2n$  playing the rôle of the parameter  $t$  which occurs in that section. Writing  $x = 1 + y$ , and comparing the integral with

$$\int_0^\infty e^{-ys} \log 2n y^{s-1} dy = \Gamma(s) (s \log 2n)^{-s},$$

we easily obtain that

$$(6.6.3) \quad \int_1^2 Adx = \pi^{s-1} \Gamma(s) (2ns \log 2n)^{-s} \cdot \{ \sin \pi(s-1) + O((\log n)^{-1}) \}.$$

Instead of the term  $O((\log n)^{-1})$  we can get an asymptotic series in terms of powers  $(\log n)^{-k}$  ( $k = 1, 2, 3, \dots$ ).

The remaining integral  $\int_2^\infty$  requires some careful attention, as (6.6.1) does not hold uniformly with respect to  $x$  in the interval  $2 \leq x < \infty$ . On the other hand a rough estimate will do : we shall show that  $\int_2^\infty$  is  $O(n^{-2s})$ , so that it amply vanishes into the  $O$ -term of (6.6.3).

Let  $K$  be an integer  $> s-1$ . We shall show that there exists a constant  $C_1$  (depending neither on  $x$  nor on  $n$ ), such that

$$(6.6.4) \quad \Pi(x+2n)/\{(2n)! \Pi(x-1)\} > C_1 x^K n^2 \quad (x \geq 2, n \geq K).$$

The left-hand-side can be written as

$$x \left( 1 + \frac{x}{1} \right) \left( 1 + \frac{x}{2} \right) \dots \left( 1 + \frac{x}{2n} \right),$$

and this is greater than

$$x^K \{ (K-1)! \}^{-1} \prod_{h=K}^{2n} (1 + 2h^{-1}) \geq C_2 x^K \prod_{h=1}^{2n} (1 + 2h^{-1}),$$

with  $C_2^{-1} = (K - 1)! \prod_{k=1}^{K-1} (1 + 2h^{-1})$ . The product  $\prod_1^{2n}$  can be compared to  $\prod(1 + 2h^{-1} + h^{-2})$ :

$$\begin{aligned} \prod_1^{2n} (1 + 2h^{-1}) &= \prod_1^{2n} (1 + h^{-1})^2 \cdot \prod_1^{2n} \{1 - (h + 1)^{-2}\} \geq \\ &\geq (2n + 1)^2 \prod_1^{\infty} \{1 - (h + 1)^{-2}\}. \end{aligned}$$

The latter product being convergent, we have proved (6.6.4), with

$$C_1 = 4C_2 \prod_1^{\infty} \{1 - (h + 1)^{-2}\}.$$

We can now give a satisfactory upper bound for the integral from 2 to  $\infty$ :

$$\left| \int_2^{\infty} A dx \right| < n^{-2s} \int_2^{\infty} C_1^{-s} x^{-Ks} |\sin \pi x|^{s-1} dx.$$

The integral on the right being convergent (as  $K > s-1$ ), and independent of  $n$ , we have

$$\left| \int_2^{\infty} A dx \right| < C_3 n^{-2s} \quad (C_3 \text{ independent of } n).$$

Comparing this to (6.6.3) we observe that  $n^{-2s}$  can be absorbed by the O-term  $n^{-s}(s \log n)^{-s}O((\log n)^{-1})$ . Therefore, our final result for  $Q$  (cf. (6.6.2)) is, if  $s > 0$ ,

$$(6.6.5) \quad Q = (-1)^s \pi^{-1} \Gamma(s) (2ns \log 2n)^{-s} \{\sin \pi s + O((\log n)^{-1})\}.$$

### 6.7. Conclusions about $S(s, n)$

Thus we have obtained the asymptotic behaviour of  $S(s, n)$ , for all real values of  $s$ . If  $s < 0$ , we have (6.4.3); if  $s = 0$  we have  $S(0, n) = (-1)^n$ ; if  $0 < s < 1$  we use  $S = 2Q - 2P$  (see (6.4.6)), with  $P = 0$ ; if  $s = 1$  we have  $S(1, n) = 0$ ; if  $1 < s < \frac{3}{2}$ ,  $P$  is much smaller than  $Q$ , as  $2 \cos(\pi/2s) < 1$  in that case (cf. (6.5.6) and (6.6.5)); if  $s > \frac{3}{2}$ ,  $P$  is much larger than  $Q$  (if  $s = \frac{3}{2}$ ,  $P$  and  $Q$  are almost of the same order, although  $Q$  is still the smaller of the two). We give a list of the results:

$$s < 0: \quad S(s, n) = 2 \cdot (-1)^n + O(n^s).$$

$$s = 0: \quad S(0, n) = (-1)^n.$$

$$0 < s < \frac{3}{2}: \quad S(s, n) = 2(-1)^n \pi^{-1} \Gamma(s) (2ns \log(2n))^{-s} \cdot \{\sin \pi s + O((\log n)^{-1})\}.$$

$$s = 1: \quad S(1, n) = 0.$$

$$s \geq \frac{3}{2}: \quad S(s, n) = 2^{2-s} \{2 \cos(\pi/2s)\}^{2ns+s-1} \cdot (\pi n)^{\frac{1}{2}(1-s)} s^{-\frac{1}{2}} \{1 + O(n^{-1})\}.$$

In the cases  $s < 0$  and  $s > \frac{3}{2}$  the  $O$ -term can be replaced by an asymptotic series in terms of powers of  $n^{-1}$ ; the same thing can be done if  $0 < s < \frac{3}{2}$ , but then in terms of powers of  $(\log n)^{-1}$ . If  $s = \frac{3}{2}$  the asymptotic series is more complicated, as both  $P$  and  $Q$  give their contributions:

$$S\left(\frac{3}{2}, n\right) \approx 2 \cdot 3^{-1/2} \cdot \pi^{-1/4} \{n^{-1/4} + c_1 n^{-5/4} + c_2 n^{-3/4} (\log n)^{-3/2} + c_3 n^{-3/2} (\log n)^{-5/2} + \dots\};$$

the terms of the development of  $P$  are, from the third term onwards, negligible compared to the development of  $Q$ .

### 6.8. A modified Gamma function

We shall discuss an example where the problem of finding a suitable integration path is quite difficult. This difficulty is mainly caused by the circumstance that the real variable  $t$ , which occurred thus far in our saddle point problems, is replaced by a complex variable  $s$ , and we want to find the asymptotic behaviour of the integral for all complex values of  $s$ , when  $|s| \rightarrow \infty$ . The integration path will therefore depend both on  $|s|$  and on  $\arg s$ , and it is the dependence on  $\arg s$  which gives the major trouble. We shall meet these difficulties by application of conformal mapping.

The function to be considered is defined by

$$(6.8.1) \quad G(s) = \int_0^\infty e^{-P(u)} u^{s-1} du \quad (\operatorname{Re} s > 0),$$

where  $P(u)$  is a polynomial

$$P(u) = u^N + \alpha_{N-1} u^{N-1} + \dots + \alpha_1 u + \alpha_0.$$

The degree  $N$  is a fixed positive integer, and the coefficients  $\alpha_{N-1}, \dots, \alpha_0$  are fixed complex numbers. In the special case that  $\alpha_{N-1} = \dots = \alpha_0 = 0$  the function  $G(s)$  becomes  $N^{-1} \Gamma(s/N)$  and,

therefore the complex Stirling formula (cf. (6.4.5)) will form a special case. This special case can, of course, be derived much easier than the formula for  $G(s)$ . For example, with the gamma function it is sufficient to discuss the half-plane  $\operatorname{Re} s > 0$ , because of the functional equation  $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$ . Moreover, it can be done by entirely different methods as well (see, for example, sec. 3.10).

As in the case of the  $\Gamma$ -integral, it is easily seen that the integral (6.8.1) converges only if  $\operatorname{Re} s > 0$ . But it is not difficult to show that  $G(s)$  can be continued analytically over the whole plane, except for single poles at the points  $s = 0, -1, -2, \dots$  (however, for exceptional sets of coefficients  $\alpha$ , some of these points can be regular points of the function). The possibility of this continuation is a well-known consequence of the fact that  $e^{-P(u)}$  is analytic at  $u = 0$ . The argument is as follows. There is a power series expansion

$$e^{-P(u)} = a_0 + a_1 u + a_2 u^2 + \dots,$$

valid for all values of  $u$ . So if  $k$  is any integer  $\geq 0$ , we have

$$e^{-P(u)} = a_0 + a_1 u + a_2 u^2 + \dots + a_k u^k + R(u),$$

$$R(u) = O(u^{k+1}) \quad (|u| < 1),$$

and therefore

$$\begin{aligned} G(s) &= \int_1^\infty e^{-P(u)} u^{s-1} du + \int_0^1 R(u) u^{s-1} du + \\ &\quad + \sum_{j=0}^k a_j (s+j)^{-1} \quad (\operatorname{Re} s > 0). \end{aligned}$$

The first integral is analytic for all complex values of  $s$ , and the second one is analytic in the half-plane  $\operatorname{Re} s > -k - 1$ . This shows the analytic continuation of  $G(s)$  throughout that half-plane, and, as  $k$  is arbitrary, it solves the problem for the whole plane.

A second method for establishing the analytic continuation depends on the functional equation

$$\begin{aligned} sG(s) &= N G(s+N) + (N-1)\alpha_{N-1} G(s+N-1) + \\ &\quad + \dots + \alpha_1 G(s+1). \end{aligned}$$

which is, if  $\operatorname{Re} s > 0$ , easily derived from (6.8.1) by partial integration.

A third method is closely connected to our way of attacking the asymptotic problem, and we shall postpone it for a moment (see (6.8.8)).

First we want to get rid of the multi-valued function  $u^{s-1}$  in the integrand. Performing the substitution  $u = e^z$ , we obtain

$$(6.8.2) \quad G(s) = \int_{-\infty}^{\infty} \exp\{-P(e^z) + sz\} dz \quad (\operatorname{Re} s > 0).$$

The integral converges if  $\operatorname{Re} s > 0$ , but if we alter the integration path, it can be used for other parts of the  $s$ -plane as well. We choose some small positive number  $\delta$ , and we define the path  $C_\delta$ , consisting of two half lines

- (i) the half line described by  $z = ix e^{i\delta}$ ,  $\infty > x \geq 0$ ,
- (ii) the positive real axis  $z = x$ ,  $0 \leq x < \infty$ .

It is not difficult to show that the integral along  $C_\delta$  is equal to  $G(s)$  if  $\operatorname{Re} s > 0$ ,  $\operatorname{Im} s \geq 0$ , and that the integral is an analytic function of  $s$  in the half-plane defined by  $-\delta < \arg s < \pi - \delta$ . As  $\delta$  is arbitrary, this furnishes the analytic continuation over the whole upper half-plane, but it does not give the behaviour of  $G(s)$  on the negative real axis.

The asymptotic behaviour of the integral along  $C_\delta$  can be tackled by saddle point analysis. The saddle points are the solutions of

$$(6.8.3) \quad e^z P'(e^z) = s.$$

If  $|s|$  is large, the solutions of (6.8.3) are easily localized. We write  $s = |s|e^{i\theta}$ , with  $-\frac{1}{2}\pi < \theta < \frac{3}{2}\pi$ . Now in every horizontal strip  $S_k$  ( $k = 0, \pm 1, \pm 2, \dots$ ), defined by

$$(6.8.4) \quad S_k: |\operatorname{Im} z - (\theta + 2k\pi)/N| \leq \pi/N$$

there lies just one root of (6.8.3), close to  $z = z_k$ , where

$$z_k = N^{-1}\{\log(|s|/N) + \theta i + 2k\pi i\}.$$

This we observe when replacing  $e^z P'(e^z)$  by its first term, viz.  $Ne^{Nz}$ , and applying the Rouché theorem (see sec. 2.4) to the strip just mentioned, with the functions  $e^z P'(e^z) - s$  and  $Ne^{Nz} - s$ . The difference between the two is, in absolute value, smaller than

$|Ne^{Nz} - s|$  for all values of  $z$  on the boundary, provided that  $|s|$  is sufficiently large. (The Rouché theorem requires a bounded domain so that it is necessary to approximate the strip by a long horizontal rectangle). As  $Ne^{Nz} - s$  has just one root, viz.  $z_k$ , in the  $k$ -th strip, the same thing holds for  $e^z P'(e^z) - s$ .

Moreover, if  $|s|$  is large enough, the roots of (6.8.3) can be expanded into powers of  $s^{-1/N}$ ,

$$(6.8.5) \quad \zeta_k = z_k + c_{k1}s^{-1/N} + c_{k2}s^{-2/N} + \dots,$$

where  $\zeta_k$  denotes the root in the  $k$ -th strip. The series in (6.8.5) converges absolutely for all large values of  $|s|$ . We do not go into details of the proof of (6.8.5), but refer to a similar case in sec. 2 (see (2.4.7)).

Our integration path  $C_8$  leads to  $+\infty$  through the strip  $S_0$ . We have to move the path such that it leads over the saddle point  $\zeta_0$ . However, we cannot keep the path entirely inside the strip  $S_0$ ; as  $C_8$  starts at  $ie^{i\theta} \cdot \infty$ , it has to cross the strips  $S_1, S_2, \dots$ . And, it is to be feared that (in order to avoid values of the integrand greater than its value at  $\zeta_0$ ) the crossings have to be made quite close to the saddle points  $\zeta_1, \zeta_2, \zeta_3, \dots$ . Actually this makes our problem awkward to deal with. It could easily be handled if the problem were restricted to the case that  $|s|$  tends to infinity with  $\arg s$  fixed, or with  $\arg s$  restricted to some small interval. Under such circumstances the problem would be of the type of the one in sec. 6.2, where the infinite collection of saddle points did not cause much trouble. Even so a certain amount of non-elegant calculations would be involved. And, as we are interested in the whole upper half-plane, we should have to divide it into some smaller sectors, and in each sector the calculations would be different.

Fortunately there is a much simpler way out, by virtue of the fact that the exponents in  $P(e^z) = e^{Nz} + \alpha_1 e^{(N-1)z} + \dots$  are integers, which implies that  $P(e^z)$  is periodic mod  $2\pi i$ . We shall first define new paths  $L_k$  ( $k = 0, \pm 1, \pm 2, \dots$ ).

$L_k$  consists of three parts:

- (i) The half-line  $z = 2k\pi N^{-1} + x$  ( $\infty > x \geq 0$ ),
- (ii) The segment  $z = ix$  ( $2k\pi/N \leq x \leq 2(k+1)\pi/N$ ),
- (iii) The half-line  $z = 2(k+1)\pi N^{-1} + x$  ( $0 \leq x < \infty$ ).

Let the function  $G_k(s)$  be defined by

$$(6.8.6) \quad G_k(s) = \int_{L_k} \exp\{-P(e^z) + sz\} dz.$$

Obviously,  $G_k(s)$  is analytic for all  $s$ . And, as  $L_{k+N}$  is obtained from  $L_k$  by shifting it in vertical direction over a distance  $2\pi$ , we have

$$(6.8.7) \quad G_{k+N}(s) = e^{2\pi i s} G_k(s) \quad (k = 0, \pm 1, \pm 2, \dots).$$

The function  $G(s)$  can be expressed in terms of  $G_0, \dots, G_{N-1}$ . Assume, for a moment, that  $\operatorname{Re} s > 0$ , so that  $G(s)$  is represented by (6.8.2). Obviously we have

$$\int_{2\pi i - \infty}^{2\pi i + \infty} = e^{2\pi i s} \int_{-\infty}^{\infty},$$

in analogy to (6.8.7). Furthermore, it can be shown that

$$\int_{2\pi i - \infty}^{2\pi i + \infty} - \int_{-\infty}^{\infty} = G_0 + G_1 + \dots + G_{N-1},$$

by shifting the paths  $L_0, \dots, L_{N-1}$  indefinitely to the left. Therefore, if  $\operatorname{Re} s > 0$ ,

$$(6.8.8) \quad G(s) = -(1 - e^{2\pi i s})^{-1} \{G_0(s) + \dots + G_{N-1}(s)\}.$$

The right-hand-side is analytic for all  $s$ , except for possible poles at  $s = 0, \pm 1, \pm 2, \dots$ . But, of course, we know that  $G$  is regular at  $s = 1, 2, 3, \dots$ . So the possibility of analytic continuation has been proved for the third time.

It will turn out that the asymptotic behaviour of  $G_0, \dots, G_{N-1}$  can be satisfactorily described in the sector  $\delta < \arg s < 2\pi - \delta$ . So (6.8.8) will give us a satisfactory result for  $G(s)$ , except for those  $s$  which are close to the positive real axis. It is, however, quite easy to solve the asymptotic problem for  $s$  in a small sector around the positive real axis,  $|\arg s| < \pi/8$ , say, directly from (6.8.2) by saddle point analysis, and we shall not devote much attention to it.

### 6.9. The entire function $G_0(s)$

For the time being, we shall consider  $G_0(s)$  only. We shall assume that  $\delta < \arg s < 2\pi - \delta$ , with some positive number  $\delta$ . Then the saddle point  $\zeta_0$  of  $\exp(-P(e^z) + sz)$  lies inside the path  $L_0$  if  $|s|$  is

large, for,  $\zeta_0$  is close to  $z_0$  (see (6.8.5)), and  $z_0 = N^{-1}\{\log(|s|/N) + i\theta\}$ , where  $\theta = \arg s$ .

In order to find our way in the darkness, we first take the special case that  $\alpha_{N-1} = \dots = \alpha_1 = 0$ , whence  $\zeta_0$  coincides with  $z_0$ . In this special case we write  $G_0^*$  instead of  $G_0$ .

By the substitution  $z = z_0 + w$ , the saddle point is shifted to the origin. The path is also shifted; it becomes  $(L_0| - z_0)$ , by which we denote the path described by  $z - z_0$ , if  $z$  describes the path  $L_0$ . It follows from Cauchy's theorem that a horizontal shift of the path  $(L_0| - z_0)$  has no influence upon the value of the integral, and therefore we may replace it by  $(L_0| - i\theta/N)$ . This path passes through the saddle point  $w = 0$ .

The integrand becomes

$$\exp(-e^{Nz} + sz) = \exp(s(z_0 - N^{-1})) \cdot \exp\{-sN^{-1}(e^{Nw} - Nw - 1)\},$$

where the splitting has been made so as to make the second factor of the form  $\exp(-\frac{1}{2}sN(w^2 + \dots))$ , for small values of  $w$ . Now  $G_0^*(s)$  becomes

$$(6.9.1) \quad G_0^*(s) = \exp(s(z_0 - N^{-1})).$$

$$\int_{(L_0| - i\theta/N)} \exp\{-sN^{-1}(e^{Nw} - Nw - 1)\} dw.$$

At this point we apply a conformal mapping in order to get a clear idea about the behaviour of  $e^{Nw} - Nw - 1$ . We consider this function in the strip  $|\text{Im } w| < 2\pi/N$ . The path  $(L_0| - i\theta/N)$  lies, for all  $\theta$  satisfying  $\delta < \theta < 2\pi - \delta$ , inside this strip.

Needless to say,  $e^{Nw} - Nw - 1$  cannot give a conformal mapping of any region containing  $w = 0$ , the function having a double zero here. Instead, we consider the function

$$\xi(w) = \{2(e^{Nw} - Nw - 1)\}^{\frac{1}{2}},$$

where in a neighbourhood of  $w = 0$  the sign of the square root has been chosen such that  $\xi(w) = Nw + \dots$

We shall show that the strip  $|\text{Im } w| < 2\pi/N$  is mapped conformally onto a region  $S$ , which is obtained from the complete  $\xi$ -plane by deleting two hyperbolic arcs in the left half-plane. These arcs can be described by

$$(\text{Re } \xi) \cdot (\text{Im } \xi) = \pm 2\pi, \quad \text{Re } \xi \leq -|\text{Im } \xi|.$$

In order to prove this, we first remark that  $\xi^2$  is positive if  $w > 0$ , so from the choice of the sign of the square root near  $w = 0$  we infer that  $\xi$  itself is positive if  $w > 0$ . Moreover,  $\xi$  tends to  $+\infty$  if  $w$  tends to  $+\infty$ .

We now consider a rectangle with vertices  $\pm M \pm 2\pi i/N$ , where  $M$  is a large positive number, and we analyze what  $\xi$  does if  $w$  proceeds through its boundary. We start at  $w = M$ , where  $\xi$  is positive and large. If  $w$  runs from  $M$  to  $M + 2\pi i/N$ , then  $|\xi|$  is large, and  $\arg \xi$  increases from 0 to almost  $\pi$ . If  $w$  runs through the upper horizontal side of the rectangle, then  $\xi$  runs through the upper hyperbolic arc, up to its end-point  $(2\pi)^{1/2}(i - 1)$  (this is attained if  $w = 2\pi i/N$ ), and back along the same arc to a point far to the left, and close to the real axis. If  $w$  proceeds from  $-M + 2\pi i/N$  to  $-M$ , then  $\xi$  jumps from the hyperbolic arc to the real axis, over a small distance only. If  $w$  proceeds along the lower part of the rectangle, then  $\xi$  runs through the complex conjugate of the path described just now.

Let  $\alpha$  be any point of the region  $S$ . Then it is clear that  $\arg(\xi(w) - \alpha)$  increases by  $2\pi$  if  $w$  runs through the boundary of the rectangle, since the path of  $\xi$  encircles  $\alpha$  just once, provided that  $M$  is large enough. We infer that  $\xi(w)$  attains the value  $\alpha$  just once inside the rectangle<sup>1)</sup>, and, therefore, just once inside the infinite strip.

If  $\alpha$  lies on one of the hyperbolic arcs, then  $\xi(w)$  is never equal to  $\alpha$  if  $w$  lies inside the strip. This can be shown, for instance, by investigating a rectangle with vertices  $\pm M \pm (2\pi - \varepsilon)i/N$ , where  $\varepsilon$  is positive and small.

So we have shown that  $\xi(w)$  gives a one-to-one mapping of the open infinite strip onto the region  $S$ . As  $\xi$  is an analytic function of  $w$ , it is a conformal mapping.

We want to introduce  $\xi$  as a new integration variable in the integral (6.9.1), and therefore we have to investigate  $dw/d\xi$ . Needless to say, this is an analytic function of  $\xi$  throughout the region  $S$ .

As  $\xi^2 = 2(e^{Nw} - Nw - 1)$ , we have  $\xi d\xi = N(e^{Nw} - 1)dw$ . Now  $e^{Nw} - 1$  is, as far as our strip is concerned, close to 0 only if  $w$  is

---

<sup>1)</sup> See E. C. TITCHMARSH, Theory of Functions, 2nd ed., Oxford 1939, § 3.41.

close to either 0, or  $2\pi/N$ , or  $-2\pi/N$ . Therefore we have

$$(6.9.2) \quad dw/d\xi = O(\xi) \quad (|\arg \xi| \leq \frac{3}{4}, |\xi| > 3).$$

It is not even difficult to show that  $dw/d\xi$  is  $O(\xi^{-1})$  in this region, yet  $O(\xi)$  is sufficient for our purpose.

Our integral becomes

$$(6.9.3) \quad G_0^*(s) = \exp(s(z_0 - N^{-1})) \int_C \exp(-\frac{1}{2}sN^{-1}\xi^2) \frac{dw}{d\xi} d\xi,$$

and next we ask what  $C$  is. Analyzing the image of  $(L_0] - i\theta/N)$  under the conformal mapping, we find that  $C$  is a curve starting at  $e^{-i\theta} \cdot \infty$  and tending to  $e^{i\pi - i\theta} \cdot \infty$ , avoiding the hyperbolic arcs. We have, however, considerable freedom in modifying this path. Along the straight line through the origin, with arguments  $-\frac{1}{2}\theta$  and  $\pi - \frac{1}{2}\theta$  (a full line through the origin has *two* arguments!) the expression  $\frac{1}{2}sN^{-1}\xi^2$  is positive. It is easily seen that this line may be replaced by other lines through the origin, whose arguments differ from  $-\frac{1}{2}\theta$  and  $\pi - \frac{1}{2}\theta$  by less than  $\pi/4$ , for then the real part of  $\frac{1}{2}sN^{-1}\xi^2$  is still larger than a constant positive multiple of  $|\xi|^2$ , so that, by virtue of (6.9.2), the convergence is guaranteed. Naturally, in the finite part of the plane a deviation from the straight line may be necessary in order to avoid the hyperbolic arcs.

All such straight lines are, at the same time, satisfactory from the point of view of the saddle point method: on these lines the absolute value of  $\exp(-\frac{1}{2}sN^{-1}\xi^2)$  is maximal at the saddle point  $\xi = 0$ . However, for this purpose, deviations from the straight line, as mentioned above, cannot always be tolerated. In other words, we can only admit straight lines from  $-e^{i\eta} \cdot \infty$  to  $+e^{i\eta} \cdot \infty$ , where  $\frac{1}{4}\pi < \eta < \frac{3}{4}\pi$ . This is no objection: as  $\eta - (\pi - \frac{1}{2}\theta)$  is allowed to lie between  $-\frac{1}{2}\pi$  and  $\frac{1}{2}\pi$ , we can find a satisfactory value for  $\eta$  to any  $\theta$  in the interval  $\delta < \theta < 2\pi - \delta$ . Actually we might take  $\eta = \frac{1}{4}(3\pi - \theta)$ .

It is, however, preferable to have a fixed integration path, not depending on  $\eta$ . This can only be achieved under restriction of  $\theta$ , and therefore we shall consider two different values of  $\eta$ :

- (i)  $\eta = \frac{1}{4}(3\pi - \delta)$ . This can be used as long as  $\frac{1}{2}\delta < \theta < \pi + \frac{1}{2}\delta$ .
- (ii)  $\eta = \frac{1}{4}(\pi + \delta)$ . Can be used if  $\pi - \frac{1}{2}\delta < \theta < 2\pi - \frac{1}{2}\delta$ .

It will turn out that, for reasons of symmetry, we can restrict ourselves to  $\delta < \theta < \pi + \frac{1}{2}\delta$ . So we shall only consider case (i). We

replace, in (6.9.3), the path  $C$  by the straight line through the origin, from  $-e^{i(3\pi-\delta)/4} \cdot \infty$  to  $+e^{i(3\pi-\delta)/4} \cdot \infty$ . This path will be denoted by  $D$ . Along  $D$  we have

$$|\exp(-\frac{1}{2}sN^{-1}\xi^2)| \leq \exp\{-\frac{1}{2}|s| \cdot N^{-1} \cdot |\xi|^2 \cdot \sin(\frac{1}{4}\delta)\}.$$

Now the stage has been set for application of the method of sec. 4.1. The value of  $dw/d\xi$  at  $\xi = 0$  equals  $N^{-1}$  (for,  $\xi = Nw + \dots$ ). Furthermore,  $dw/d\xi$  is an analytic function of  $\xi$  along the line  $D$ , and  $dw/d\xi = O(\xi)$  if  $|\xi| > 3$ .

The integral can be compared with the formula

$$\int_D \exp(-\frac{1}{2}sN^{-1}\xi^2) d\xi = -(2\pi N)^{\frac{1}{2}} s^{-\frac{1}{2}},$$

where  $s^{-\frac{1}{2}}$  is to be interpreted as  $|s|^{-\frac{1}{2}} e^{-i\theta}$ . So finally we obtain

$$(6.9.4) \quad G_0^*(s) = -\exp(s(z_0 - N^{-1})) \left(\frac{2\pi}{Ns}\right)^{\frac{1}{2}} \{1 + O(s^{-1})\},$$

and the  $O$ -term can be replaced by an asymptotic series of the form  $c_1 s^{-1} + c_2 s^{-2} + \dots$ . It should be noted that (6.9.4) holds uniformly in the region  $|s| > 1$ ,  $\delta < \arg s < \pi + \frac{1}{2}\delta$ .

Formula (6.9.4) is not new, of course. We have (notice that  $G_k^*(s) = e^{2\pi i k s/N} G_0^*(s)$  and apply (6.8.8)),

$$G_0^*(s) = -N^{-1}(1 - e^{2\pi i s/N}) \Gamma(s/N).$$

So if we replace  $s/N$  by  $w$ , (6.9.4) reduces to

$$\Gamma(w) = (2\pi)^{\frac{1}{2}} (1 - e^{2\pi i w})^{-1} e^{(w-\frac{1}{2}) \log w - w} \{1 + O(w^{-1})\} \\ (\delta < \arg w < \pi + \frac{1}{2}\delta).$$

This is an easy consequence of the Stirling formula ( $\log w$  is given the value with  $\delta < \operatorname{Im} \log w < \pi + \frac{1}{2}\delta$ ). It was not, however, our purpose to deduce well-known results concerning the gamma function, for which easier methods exist, but rather to develop a technique for  $G_0^*(s)$  which can be modified to a technique for  $G_0(s)$ .

As a preparation to the problem of the asymptotic behaviour of  $G_0(z)$  we shall first investigate the following integral

$$(6.9.5) \quad f(t, \omega) = \int_{-\rho}^{\rho} g(x) \exp(-tx^2 - tQ(x, \omega)) dx.$$

Here  $\rho$  is a fixed positive number,  $Q(x, \omega)$  is a double power series in  $x$  and  $\omega$ :

$$Q(x, \omega) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} c_{nm} x^n \omega^m,$$

absolutely convergent if  $|x| < 2\phi$  and  $|\omega| < b$ , where  $b$  is a fixed positive number. The function  $g(x)$  is assumed to be analytic in the circle  $|x| < 2\phi$ , and it is assumed that  $g(0) \neq 0$ . We want to determine the behaviour of  $f(t, \omega)$  when  $|\omega|$  is small and  $|t| \rightarrow \infty$ , where  $t$  is restricted to a sector  $|\arg t| < \frac{1}{2}\pi - \delta$ .

Searching for saddle points, we investigate the equation

$$(6.9.6) \quad -2ix - tQ'(x, \omega) = 0,$$

the dash indicating differentiation with respect to  $x$ . By the Rouché theorem (see sec. 2.4) this equation has exactly one solution  $x_0$  inside the circle  $|x| \leq \phi$  if  $|\omega|$  is sufficiently small. For then we have  $|Q'(x, \omega)| < 2|x|$  on the boundary of that circle, whence  $2x + Q'(x, \omega)$  has as many roots in the interior as  $2x$  itself. Furthermore,  $x_0$  can be written as a sum of a power series:

$$x_0 = \sum_{k=1}^{\infty} d_k \omega^k,$$

convergent if  $|\omega|$  is sufficiently small.

The straight line from  $-\phi$  to  $+\phi$  is a satisfactory path for applying the saddle point method to the simplified integral  $\int \exp(-tx^2) dx$ . This path makes an angle  $< \frac{1}{2}\pi - \delta$  with the axis of the saddle point in the origin. In the modified integral (6.9.5) the saddle point  $x_0$  is close to the original saddle point, and also the direction of its axis does not differ much from the original axis. So if  $|\omega|$  is small enough, the horizontal path through  $x_0$  makes an angle less than  $\frac{1}{2}\pi - \frac{1}{2}\delta$  with the new axis, and therefore it can be used, at any rate in a neighbourhood of the saddle point. These remarks are, of course, only intended as an orientation; the work remains to be done.

We replace (6.9.5) by

$$(6.9.7) \quad \int_{-p+x_0}^{-p+x_0} + \int_{-p+x_0}^{p+x_0} + \int_{p+x_0}^p.$$

where the integration is taken along straight line segments in all three cases. The first and the last term in (6.9.7) are easily seen to be exponentially small in comparison to the value of  $\exp(-tx^2 - tQ)$  at the saddle point, and therefore they can be neglected. In the middle term of (6.9.7) we carry out the substitution  $x = x_0 + y$ .

Then the integral becomes

$$(6.9.8) \quad f(t, \omega) = \exp(-tx_0^2 - tQ(x_0, \omega)) \cdot \int_{-\rho}^{\rho} g(x_0 + y) \exp(-ty^2 - tQ_1(y, \omega)) dy,$$

where  $Q_1(y, \omega) = Q(x_0 + y, \omega) - Q(x_0, \omega) - yQ'(x_0, \omega)$  is again a double power series

$$Q_1(y, \omega) = \sum_{n=2}^{\infty} \sum_{m=1}^{\infty} \gamma_{nm} y^n \omega^m,$$

convergent if  $|y| < 3\rho/2$  and  $|\omega|$  sufficiently small (if  $|\omega|$  is small enough, this condition on  $y$  implies that  $|y + x_0| < 2\rho$ ).

We can now proceed along various methods. For example, we can expand  $\exp(-tQ_1(y, \omega))$  in terms of powers of  $y$ . Another method is to apply conformal transformation again, introducing a new variable  $z$  by  $y^2 + Q(y, \omega) = z^2$ . Then  $z$  can be solved as a double power series in  $y$  and  $\omega$ , and we get an integral

$$\int Q_1(z, \omega) \exp(-tz^2) dz.$$

Omitting the details, which are all implied in the usual saddle point routine, we state the result:

$$(6.9.9) \quad f(t, \omega) = \pi^{\frac{1}{2}} t^{-\frac{1}{2}} \{1 + \frac{1}{2} Q''(x_0, \omega)\}^{-\frac{1}{2}} \cdot \{g(x_0) + O(t^{-1})\} \exp\{-tx_0^2 - tQ(x_0, \omega)\},$$

uniformly in the region

$$|\arg t| < \frac{1}{2}\pi - \delta, |t| > 1, |\omega| < b_1$$

(for some fixed positive number  $b_1$ ). The term  $O(t^{-1})$  can be replaced by an asymptotic series with terms  $f_k(\omega)t^{-k}$ , where the  $f_k(\omega)$  are power series in terms of powers of  $\omega$ , convergent if  $|\omega| < b_1$ . It should be remarked that also  $g(x_0)$  and  $Q''(x_0, \omega)$  are convergent power series in terms of powers of  $\omega$ .

In order to apply the result about  $f(t, \omega)$  to  $G_0$  it is easier to express the main result about (6.9.5) in words: If  $\omega$  is small enough, there is just one saddle point near  $x = 0$ , and the contribution of this saddle point gives an asymptotic series for  $f(t, \omega)$  if  $t \rightarrow \infty$ .

In the integral (6.8.6) for  $G_0(s)$ , we carry out the same substitutions as we did in the case of  $G_0^*(s)$ . That is,  $z = z_0 + w$ , where

$$Nz_0 = \log(|s|/N) + i\theta, \theta = \arg s, 2(e^{Nw} - Nw - 1) = \xi^2,$$

and we discuss the integral in the  $\xi$ -plane:

$$(6.9.10) \quad G_0(s) = \exp(s(z_0 - N^{-1})).$$

$$\int_D \exp\{-\frac{1}{2}sN^{-1}\xi^2 + (e^{Nz} - P(e^z))\} \frac{dw}{d\xi} d\xi.$$

$D$  is the straight line through the origin, from  $-e^{t(3\pi-\delta)/4} \cdot \infty$  to  $+e^{t(3\pi-\delta)/4} \cdot \infty$ . We shall again restrict  $\theta$  by  $\delta < \theta < \pi + \delta/4$ .

We first investigate the term  $e^{Nz} - P(e^z)$ , which embodies the deviation of  $G_0$  from  $G_0^*$ . It is equal to

$$(6.9.11) \quad e^{Nz_0+Nw} \{-\alpha_{N-1}e^{-z_0-w} - \alpha_{N-2}e^{-2z_0-2w} - \dots\}.$$

It is not difficult to show from the properties of the conformal mapping that  $\operatorname{Re} w$  tends to  $+\infty$  if  $|\xi|$  tends to infinity, provided that  $\xi$  runs along  $D$ . It easily follows, for any positive number  $p$ , that  $e^w = O(\xi^{2/N})$  ( $\xi$  on  $D$ ,  $|\xi| > p$ ). Therefore

$$|e^{Nz} - P(e^z)| \leq c_1 |s\xi^2|^{(N-1)/N} \quad (\xi \text{ on } D, |\xi| > p, |s| > c_2),$$

where the  $c$ 's are suitable positive numbers independent of  $s$  and  $\xi$ . It follows that, if  $p$  is any positive number, we can restrict the integration in (6.9.10) to a segment of  $D$  with length  $2p$ , symmetric with respect to the origin. For, the further parts of  $D$  are easily seen to give a contribution which is exponentially negligible if  $|s| \rightarrow \infty$ .

To the remaining part of  $D$ , with length  $2p$ , we apply the result about (6.9.5). We have, of course, to turn the integration path over an angle  $-\frac{1}{2}(3\pi - \delta)$ , by the substitution  $\xi = xe^{t(3\pi-\delta)/4}$ , and to put  $s = e^{-\frac{1}{2}(3\pi-\delta)t}$ , so that  $t$  is restricted by the condition  $-\frac{1}{2}\pi + \frac{1}{2}\delta < \arg t < \frac{1}{2}\pi - \frac{1}{2}\delta$ , and  $s\xi^2$  becomes  $tx^2$ . Furthermore, if  $p$  is small enough we have, in the circle  $|\xi| < 2p$ ,

$$e^{Nz} - P(e^z) = N^{-1}sQ(x, \omega),$$

where (cf. (6.9.11))  $\omega = s^{-1/N}$ , and  $Q$  is a double power series with  $Q(\xi, 0) = 0$ . For,  $w = N^{-1}\xi + \dots$  and  $e^{-w}, \dots, e^{-(N-1)w}$  are convergent power series in powers of  $\xi$  for small values of  $\xi$ , and  $e^{-z_0} = N^{1/N}\omega$ . (Actually it can be seen from the conformal mapping that the series for  $w$  converges if  $|\xi| < 2\pi^{\frac{1}{N}}$ , so  $p$  can be taken  $= \pi^{\frac{1}{N}}$ ).

It follows from our discussions of (6.9.5) that for  $G_0(s)$  we have an

asymptotic series, which is identical to the contribution of the saddle point. The series equals a certain function multiplied by an asymptotic series  $c_0(\omega) + c_1(\omega)s^{-1} + \dots$ , and the  $c(\omega)$ 's are convergent series in powers of  $\omega = s^{-1/N}$ . Therefore, the series  $c_0(\omega) + \dots$  can be rewritten as an asymptotic series of the form  $c_0 + c_1s^{-1/N} + c_2s^{-2/N} + \dots$ .

We do not evaluate explicitly the contribution of the saddle point in the integral in the  $\xi$ -plane, as it is easier to do so in the original  $z$ -plane (see (6.8.6), with the saddle point  $\zeta_0$ , given by (6.8.5)). It is not difficult to see that the contribution of the saddle point is not affected by the substitutions relating  $z$  to  $\xi$  and  $\xi$  to  $x$ . So our final result is an asymptotic series

$$(6.9.12) \quad G_0(s) \approx -\exp\{-P(e^{\zeta_0}) + s\zeta_0\} \cdot$$

$$\cdot (-2\pi/\psi''(\zeta_0))^\frac{1}{2} \sum_0^\infty c_k s^{-k/N},$$

where  $c_0 = 1$ , and  $\psi(z)$  is the function  $-P(e^z) + sz$ . There is of course the difficulty to determine the sign of the contribution, but this sign is easily derived from the sign in (6.9.4), by a continuity argument. We have

$$\begin{aligned} -\psi''(\zeta_0) &= N^2 e^{Nz} + (N-1)^2 \alpha_{N-1} e^{(N-1)z} + \dots = \\ &= Ns\{1 + a_1 s^{-1/N} + a_2 s^{-2/N} + \dots\}, \end{aligned}$$

so that (6.9.12) can be slightly simplified. Introducing new coefficients  $d_k$  (with  $d_0 = 1$ ), we get

$$(6.9.13) \quad G_0(s) \approx -\left(\frac{2\pi}{Ns}\right)^\frac{1}{2} \exp\{-P(e^{\zeta_0}) + s\zeta_0\} \sum_0^\infty d_k s^{-k/N},$$

and the formula has been proved if  $\delta < \arg s < \pi + \frac{1}{2}\delta$ ,  $|s| \rightarrow \infty$ . The argument of  $s^\frac{1}{2}$  is understood to lie between  $\frac{1}{2}\delta$  and  $\frac{1}{2}\pi + \frac{1}{2}\delta$ .

Formula (6.9.13) obviously generalizes (6.9.4).

## 6.10. Conclusions about $G(s)$

In order to deal with  $G(s)$ , we need, according to (6.8.8), the asymptotic behaviour of  $G_0, G_1, \dots, G_{N-1}$ .

The problem about  $G_k$  ( $k = 1, \dots, N-1$ ) is easily reduced to

the problem for  $G_0$ . It follows from (6.8.6) that

$$(6.10.1) \quad G_k(s) = \exp(2\pi iks/N) \int_{L_0} \exp\{-P(e^{2\pi i k/N} e^z) + sz\} dz,$$

and this integral equals the function  $G_0(s)$ , constructed for the polynomial  $P(e^{2\pi i k/N} u)$  instead of  $P(u)$ . The leading term of the new polynomial is again  $u^N$ .

Replacing  $P(u)$  by  $P(e^{2\pi i k/N})$  can have an influence upon the situation of the saddle point; the difference of the two saddle points can be of the order of  $s^{-1/N}$ . And, as in (6.9.13) there occurs a term  $s\zeta_0$  in the exponent, the influence upon the asymptotic behaviour can be considerable.

It is quite easy to state a simpler but weaker result:

$$(6.10.2) \quad G_k(s) = e^{2\pi iks/N} (2\pi/Ns)^{\frac{1}{2}} \cdot$$

$$\cdot \exp\left\{-\frac{s}{N} + \frac{s}{N} \log \frac{s}{N} + O(s^{(N-1)/N})\right\},$$

where  $\delta < \arg s < \pi + \frac{1}{4}\delta$ , and  $|s| \rightarrow \infty$ . From this we infer that  $G_k$  ( $k > 0$ ) is negligible compared to  $G_0$  as soon as the factor  $\exp(2\pi iks/N)$  beats the  $\exp(O(s^{(N-1)/N}))$ . This is certainly the case if  $s$  is restricted to  $\delta < \arg s < \pi - \delta$ . And, under that assumption, the factor  $(1 - e^{2\pi i s})^{-1}$  of (6.8.8) can be replaced by 1. The relative errors made this way are of the type  $O(e^{-c|s|})$  with some positive  $c$ . So summarizing, we infer, from (6.9.13), that

$$(6.10.3) \quad G(s) \approx (2\pi/Ns)^{\frac{1}{2}} \exp(-P(e^{\zeta_0}) + s\zeta_0) \sum_0^{\infty} d_k s^{-k/N}$$

if  $|s| \rightarrow \infty$ , in the sector  $\delta < \arg s < \pi - \delta$ . We observe that  $d_0 = 1$ .

For the sector  $-\pi + \delta < \arg s < -\delta$  we can obtain the same result, provided that  $\zeta_0$  indicates the saddle point close to  $N^{-1} \log(s/N)$ , where the log has its principal value. In a sector like  $-2\delta < \arg s < 2\delta$  it is quite easy to obtain (6.10.3) again, by direct application of the saddle point method to the integral (6.8.2). In this case we can take the horizontal line through the saddle point as the integration path.

So (6.10.3) has been proved in the sector  $-\pi + \delta < \arg s < \pi - \delta$ . In the special case that  $P(u) = e^{Nu}$ , it reduces to the Stirling formula, which is known to hold in the same sector.

We finally state a rough inequality which is easily deduced from (6.10.2) and (6.8.8). If we delete from the complex plane the half strip described by  $|\operatorname{Im} s| < 1$ ,  $\operatorname{Re} s < 1$ , then we have in the remaining region

$$|G(s)/\Gamma(s/N)| < C_1 \exp(C_2 |s|^{(N-1)/N}).$$

6.11. EXERCISES. 1. If  $\alpha$  is a fixed real number we put

$$S_n = \sum_{j=1}^n (-1)^{n+j} \binom{2n}{n+j} j^{2n+\alpha}$$

Show that there is an asymptotic series

$$n^{-8\alpha/3} S_n / (2n)! \approx c_0 + c_1 n^{-1} + c_2 n^{-2} + \dots \quad (n \rightarrow \infty),$$

with  $c_0 = \frac{1}{2} 3^{-\frac{1}{2}\alpha} / \Gamma(1 + \frac{1}{2}\alpha)$ .

2. If  $\alpha$  is a non-negative even integer, then at most a finite number of the  $c$ 's in the previous problem are different from zero, and the sum of this finite asymptotic series is exactly equal to the expression on the left (Hint: consider the  $(2n + \alpha)$ -th derivative of  $(e^{\frac{1}{2}x} - e^{-\frac{1}{2}x})^{2n}$  at  $x = 0$ ). If  $\alpha$  is even and negative, then  $S_n = 0$  if  $2n + \alpha > 0$ .

3. If  $\delta$  is a positive constant, then show that

$$\int_{-\infty}^{\infty} \exp(-\frac{1}{2}(x-w)^2 - e^w) dw \approx (2\pi/\lambda)^{\frac{1}{2}} \exp(-\frac{1}{2}\lambda^2 - \lambda) \sum_{k=0}^{\infty} b_k \lambda^{-k},$$

in the sector  $|\arg x| < 3\pi/4 - \delta$ , with  $|x| \rightarrow \infty$ . Here  $\lambda$  represents the root of  $\lambda + \log \lambda = x$  ( $|\arg \lambda| < \pi$ ;  $\log \lambda$  has its principal value). The  $b$ 's are constants;  $b_0 = 1$ . (Cf. N. G. de Bruijn, Nederl. Akad. Wetensch. Proc. Ser. A 56 = Indagationes Math. 15, p. 456 (1953)).

## CHAPTER 7

### INDIRECT ASYMPTOTICS

#### 7.1. Direct and indirect asymptotics

The first six chapters of this book (with the exception of some arguments in Ch. 1) have been devoted to asymptotic methods which we shall call *direct* methods. The common features are (i) the function  $f(t)$  whose asymptotic behaviour (as  $t \rightarrow \infty$ ) is required, is represented by some explicit formula in the form of a series or integral, and (ii) this expression is split into parts, some of which are proved to be small, whereas the dominating parts are compared with known functions, and it is shown that their deviations from these known functions are small; (iii) the final result is then obtained from the fact that the absolute value of a sum (i.e. the total error) does not exceed the sum of the absolute value of the terms (i.e. the sum of the absolute values of the partial errors). We remark that these direct methods are always constructive, in the sense that the methods supply the means for replacing O- and o-formulas by definite numerical estimates (cf. sec. 1.7).

Observing that the direct methods essentially depend on the inequality  $|a + b| \leq |a| + |b|$ , we remark that these direct methods work in the real field as well as in the complex field, and actually they can be applied to functions  $f(t)$  whose values belong to a system where each element  $a$  has a "norm"  $|a|$ , such that  $|a + b| \leq |a| + |b|$  holds (we of course assume that the system is an abelian group with respect to addition).

In the remaining chapters of this book we shall come across several methods which are not of this direct type; we shall call them *indirect* methods. It is difficult to describe common features in the way we did above for the direct methods. We shall indicate some aspects here which are, however, by no means common to *all* indirect methods.

(i) Most indirect methods consider real functions only. Usually they essentially depend upon the property that the set of real numbers is a linearly ordered system (i.e. a system where an inequality relation  $<$  is given, such that  $a < b$  and  $b < c$  imply  $a < c$ ), and upon the fact that any bounded monotonic sequence converges.

(ii) Sometimes indirect proofs play a rôle. This may mean that the resulting O- and o-formulas, which are statements as to the existence of a number, or of two numbers, or of a function (see sec. 1.2 and 1.3), are reduced to statements expressing the absurdity of non-existence. In such cases we cannot replace the O- or o-formulas by definite numerical estimates.

However, non-constructivity can be the result of direct proofs as well. In particular it can occur after an application of the theorem that any bounded monotonic sequence converges: If  $a_1 < a_2 < \dots < 1$ , then there exists a number  $\alpha$  such that  $a_n = \alpha + o(1)$  ( $n \rightarrow \infty$ ), but we cannot be more specific as long as we have no further information about the  $a_n$ 's.

(iii) Frequently indirect methods essentially involve proofs by induction. We mention a typical argument: In order to show that a certain function  $f(t)$  satisfies  $f(t) = t^2 + O(t)$  ( $t \rightarrow \infty$ ) we show, for instance, that  $|f(t)| < 10$  ( $2 < t \leq 3$ ) and that, for all  $T \geq 3$ , the assumption

$$|f(t) - t^2| < 80t \quad (T - 1 < t \leq T)$$

leads to

$$|f(t) - t^2| < 80t \quad (T < t \leq T + 1).$$

Now the principle of induction shows that  $|f(t) - t^2| < 80t$  ( $t > 2$ ).

The difficulty with such arguments is the same one as the difficulty of induction proofs in general. It may happen that the induction step fails if we replace  $80t$  by  $100t$ . Or it may happen that the induction step fails for any inequality  $|f(t) - t^2| < At$ , but that it works for some stronger inequality of the type  $|f(t) - t^2| < At^k$ . In other words, it may happen that a stronger assertion is easier to prove than a weaker one.

Needless to say, proofs by induction often occur in cases where we want to know the asymptotic behaviour of a sequence  $\{a_n\}$

which is given by induction, and where we do not possess a formula expressing  $a_n$  in terms of  $n$  (cf. Ch. 8).

(iv) Quite often we shall use formulas expressing the function  $f(t)$  (whose asymptotic behaviour is required), not only in terms of known expressions, but also in terms of  $f$  itself. It may happen, for example, that  $f(t)$  can be expressed, by some integral, in terms of the values of  $f(\tau)$  in the interval  $t < \tau < \infty$ . Such expressions can often be used for determining the behaviour as  $t \rightarrow \infty$ ; they may permit to transform quite rough information about  $f(\tau)$  (in the interval  $t < \tau < \infty$ ) into more definite information about  $f(t)$ .

For example, if we know that the real function  $f(t)$  satisfies the relation

$$f(t) = \cos t^{-1} + \int_t^{\infty} \{\tau^2 + (f(\tau))^2\}^{-1} d\tau \quad (t > 1),$$

then it is easily seen that the integral is  $O(t^{-1})$ . It follows that

$$f(t) = \cos t^{-1} + O(t^{-1}) = 1 + O(t^{-1}) \quad (t > 1).$$

Inserting this into the integral, we get

$$\begin{aligned} \int_t^{\infty} \{\tau^2 + 1 + O(\tau^{-1})\}^{-1} d\tau &= \int_t^{\infty} \{\tau^{-2} - \tau^{-4} + O(\tau^{-5})\} d\tau = \\ &= t^{-1} - \frac{1}{3}t^{-3} + O(t^{-4}), \end{aligned}$$

whence  $f(t) = 1 + t^{-1} - \frac{1}{3}t^{-3} + O(t^{-4}) \quad (t > 1)$ .

The procedure can be carried on, and it is easy to show that  $f(t)$  has an asymptotic series.

(v) Sometimes we have to find the behaviour of a function  $f(t)$ , given by a number of data, one of which is a requirement about the asymptotic behaviour itself. An example: Suppose that we are dealing with a bounded function  $f(t)$  (in  $0 \leq t < \infty$ ) which satisfies a given differential equation. Suppose it turns out that this equation has just one bounded solution. Now again the problem is to transform the rough asymptotic information (the boundedness) into something more definite. Some of the problems in Ch. 9 will be of this type.

(vi) Needless to say, many indirect arguments contain parts which have a direct nature (see, for example, sec. 7.5).

From the above remarks it will be clear that "indirect asymptotics" indicates not so much a method as a lack of general methods. As the further chapters of this book give indirect methods applied in several fields, we give only one type of examples in the present chapter: some of the simpler parts of Tauberian asymptotics<sup>1)</sup>.

Tauberian theorems are usually proved by indirect methods, though this cannot always be said to be essential. However, their counterparts, the so called Abelian theorems, entirely belong to direct asymptotics.

## 7.2. Tauberian theorems

A well-known theorem of Abel reads as follows: If the power series  $\sum_{k=0}^{\infty} a_k x^k$  converges at  $x = 1$ , then it converges throughout the interval  $0 \leq x \leq 1$ , and its sum  $f(x)$  satisfies

$$\lim_{x \rightarrow 1, 0 \leq x < 1} f(x) = f(1).$$

Subtracting the constant value  $f(1)$ , we can put this theorem into the following form: If

$$(7.2.1) \quad a_0 + a_1 + \dots + a_n = o(1) \quad (n \rightarrow \infty),$$

then we have

$$(7.2.2) \quad \sum_{k=0}^{\infty} a_k x^k = o(1) \quad (0 \leq x < 1, x \rightarrow 1).$$

This theorem can be proved by a method belonging to direct asymptotics. If we write

$$A(y) = \sum_{0 \leq k \leq y} a_k,$$

then we have by partial summation (cf. (3.13.2))

$$(7.2.3) \quad \sum_{k=0}^n a_k x^k = A(n)x^n - \log x \int_0^n A(y)x^y dy.$$

Taking a fixed value of  $x$  in the interval  $0 < x < 1$ , we observe that  $A(n)x^n = o(x^n) \rightarrow 0$  as  $n \rightarrow \infty$ , and that  $\int_0^\infty A(y)x^y dy$  converges. It therefore follows from (7.2.3) that

$$(7.2.4) \quad f(x) = -\log x \int_0^\infty A(y)x^y dy.$$

<sup>1)</sup> For a more extensive treatment the reader may consult: G. H. HARDY, Divergent Series, Oxford 1949.

Now split the integral into two parts. If a number  $\varepsilon > 0$  is given, we determine  $b$  such that  $|A(y)| < \frac{1}{2}\varepsilon$  when  $b \leq y < \infty$ . Then we have, splitting  $\int_0^\infty = \int_0^b + \int_b^\infty$ ,

$$|f(x)| \leq |\log x| \int_0^b |A(y)| dy + \frac{1}{2}\varepsilon |\log x| \int_b^\infty x^y dy.$$

The first term is  $< \frac{1}{2}\varepsilon$  if  $x$  is sufficiently close to 1. The second term is at most

$$\frac{1}{2}\varepsilon |\log x| \int_0^\infty x^y dy = \frac{1}{2}\varepsilon.$$

It follows that  $|f(x)| < \varepsilon$  if  $x$  is sufficiently close to 1, so that we have proved (7.2.2).

Formula (7.2.4) can also be written in the form

$$f(x) = \int_0^\infty A(y) x^y dy / \int_0^\infty x^y dy,$$

and this expresses that  $f(x)$  is an average of values of  $A(y)$ , with positive weights. Abel's theorem derives asymptotic information about this average of  $A(y)$  from asymptotic information about  $A(y)$  itself.

The converse of Abel's theorem is not true: (7.2.2) does not imply (7.2.1). It is easy to give a counterexample. If

$$f(x) = \frac{1}{2} - x + x^2 - x^3 + \dots = \frac{1}{2}(1-x)/(1+x) \quad (0 \leq x < 1),$$

then  $f(x) \rightarrow 0$  as  $x \rightarrow 1$ , i.e. (7.2.2) holds. However, (7.2.1) is false in this case:

$$a_0 + a_1 + \dots + a_n = \frac{1}{2} \cdot (-1)^n \neq o(1).$$

It is possible to prove that (7.2.2) implies (7.2.1) by assuming some supplementary condition. The first result in this direction was obtained by A. Tauber, who showed that

$$(7.2.5) \quad a_n = o(n^{-1}) \quad (n \rightarrow \infty)$$

is satisfactory: (7.2.2) and (7.2.5) together imply (7.2.1). It was proved later by Hardy and Littlewood that (7.2.5) can be replaced by the weaker condition that there exists a positive constant  $C$  such

that

$$(7.2.6) \quad a_n > -Cn^{-1} \quad (n = 1, 2, 3, \dots).$$

We shall show the sufficiency of (7.2.6) in sec. 7.5.

The general terminology is modelled after this special case. A theorem which derives asymptotic information about some kind of average of a function from asymptotic information about the function itself, is called an *Abelian* theorem. If one can find a supplementary condition under which the converse of an Abelian theorem holds, then this condition is called a *Tauberian* condition, and the converse theorem is called a Tauberian theorem.

In secs. 7.3 and 7.4 we shall deal with quite simple cases, and a more difficult Tauberian theorem will be proved in sec. 7.5.

### 7.3. Differentiation of an asymptotic formula

Let  $f(x)$  be integrable over every finite interval, and put

$$(7.3.1) \quad F(t) = \int_0^t f(x) dx.$$

Assuming some asymptotic behaviour of  $f(t)$ , say, an equivalence

$$(7.3.2) \quad f(t) \sim t^\alpha \quad (t \rightarrow \infty),$$

where  $\alpha$  is a constant  $\geq 0$ , it is easy to derive an Abelian result about  $F(t)$ , in this case

$$(7.3.3) \quad F(t) \sim (\alpha + 1)^{-1} t^{\alpha+1} \quad (t \rightarrow \infty).$$

The converse is not always true, i.e. formal differentiation of (7.3.3) is not always legitimate. For example, if  $F(t) = t^2 + t \sin t$ , then  $F'(t) = (2 + \cos t)t + \sin t$ , so  $F(t) \sim t^2$ , but not  $F'(t) \sim 2t$ .

In order to be able to deduce (7.3.2) from (7.3.3), we need a Tauberian condition, and as such we take

$$(7.3.4) \quad f(t) \text{ is real and non-decreasing} \quad (0 \leq t < \infty).$$

We shall prove the Tauberian theorem that (7.3.3) and (7.3.4) together imply (7.3.2).

Let  $\epsilon$  be a positive number. By (7.3.3) we can take  $T$  such that

$$|F(t) - (\alpha + 1)^{-1} t^{\alpha+1}| < \epsilon t^{\alpha+1} \quad (t \geq T).$$

Let  $p$  and  $t$  be numbers such that  $t \geq T$ ,  $t + p \geq T$ ;  $p$  may be negative. Then we have also

$$|F(t + p) - (\alpha + 1)^{-1}(t + p)^{\alpha+1}| < \varepsilon (t + p)^{\alpha+1},$$

and taking differences we find that

$$(7.3.5) \quad \left| \int_t^{t+p} f(x)dx - \int_t^{t+p} x^\alpha dx \right| < \varepsilon t^{\alpha+1} + \varepsilon (t + p)^{\alpha+1}.$$

First assume that  $p > 0$ . Then it immediately follows from (7.3.4) that  $p f(t) \leq \int_t^{t+p} f(x)dx$ . Moreover, we have  $\int_t^{t+p} x^\alpha dx < p(t + p)^\alpha$ . It follows that

$$f(t) \leq (t + p)^\alpha + 2\varepsilon p^{-1}(t + p)^{\alpha+1}.$$

This gives an upper estimate for  $f(t)$ , and we can still fix  $p$  in order to make it as efficient as possible. Writing  $p = qt$  we get

$$f(t) \leq t^\alpha \{(1 + q)^\alpha + 2\varepsilon q^{-1}(1 + q)^{\alpha+1}\}.$$

It is not necessary to find the exact minimum. We only remark that  $q$  should be small enough to keep  $(1 + q)^\alpha$  close to 1, and that  $q$  should be large with respect to  $\varepsilon$ , in order to keep  $\varepsilon q^{-1}$  small. So we take  $q = \varepsilon^k$ , and we get

$$(7.3.6) \quad f(t) \leq t^\alpha \{(1 + \varepsilon^k)^\alpha + 2\varepsilon^k(1 + \varepsilon^k)^{\alpha+1}\} \quad (t \geq T).$$

A lower bound is obtained by taking  $p < 0$  in (7.3.5). We immediately take  $p = -\varepsilon^k t$ . Assuming  $\varepsilon < \frac{1}{2}$ , and  $t \geq 2T$ , we have  $p + t \geq T$ . Moreover we have

$$|p| f(t) \geq \int_{t+p}^t f(x)dx, \quad \int_{t+p}^t x^\alpha dx \geq |p|(t + p)^\alpha,$$

and it follows from (7.3.5) that

$$(7.3.7) \quad f(t) \geq t^\alpha \{(1 - \varepsilon^k)^\alpha - 2\varepsilon^k\} \quad (t \geq 2T).$$

From (7.3.6) and (7.3.7) we can deduce  $f(t) \sim t^\alpha$  ( $t \rightarrow \infty$ ). For, if  $\varepsilon' > 0$  is given, then  $\varepsilon$  can be chosen such that the factors between  $\{\}$  in (7.3.6) and (7.3.7) lie between  $1 - \varepsilon'$  and  $1 + \varepsilon'$ . With this value of  $\varepsilon$  we can determine  $T$ , and then we have  $|f(t) - t^\alpha| < \varepsilon' t^\alpha$  ( $t \geq 2T$ ). This proves our Tauberian theorem.

### 7.4. A similar problem

We again consider  $F(t) = \int_0^t f(x)dx$ , as in the previous section, but instead of (7.3.2) we assume

$$(7.4.1) \quad f(t) = 2t + O(1) \quad (t \rightarrow \infty).$$

Then we can derive by an Abelian argument

$$(7.4.2) \quad F(t) = t^2 + O(t) \quad (t \rightarrow \infty).$$

We again ask whether the supplementary condition

$$(7.4.3) \quad f(t) \text{ is non-decreasing} \quad (0 \leq t < \infty)$$

is sufficient in order to make (7.4.2) imply (7.4.1). It will turn out that it is not.

Proceeding in the same way as in sec. 7.3, we choose a positive function  $p(t)$  of  $t$ , and we obtain from (7.4.2) and (7.4.3):

$$\begin{aligned} f(t) &\leq p^{-1} \int_t^{t+p} f(x)dx = p^{-1} \{F(t+p) - F(t)\} = \\ &= 2t + p + O(tp^{-1}) + O(1) \quad (t \rightarrow \infty). \end{aligned}$$

The best possible  $O$ -result is obtained by taking  $p$  such that the terms  $p$  and  $tp^{-1}$  are of the same order. So taking  $p = t^{\frac{1}{2}}$ , we obtain  $f(t) \leq 2t + O(t^{\frac{1}{2}})$ . We easily get the corresponding lower estimate, and so

$$(7.4.4) \quad f(t) = 2t + O(t^{\frac{1}{2}}) \quad (t \rightarrow \infty).$$

Roughly speaking this is the best possible result that can be derived from (7.4.2) and (7.4.3). More precisely we shall show that there exists a function  $f(t)$  satisfying (7.4.2) and (7.4.3), which is of the form  $f(t) = 2t + O(t^{\frac{1}{2}})$ , but which does not satisfy  $f(t) = 2t + o(t^{\frac{1}{2}})$ .

A good example can be obtained by graphical arguments. We shall assume that  $|F(t) - t^2| \leq t$  ( $t \geq 0$ ), which means that the graph of the function  $y = F(t)$  in the  $(t, y)$ -plane lies below the graph of  $y = t^2 + t$  and above the graph of  $y = t^2 - t$ , as far as values  $t \geq 0$  are concerned. We shall denote the parabola  $y = t^2 + t$  by  $\pi_1$ , and the parabola  $y = t^2 - t$  by  $\pi_2$ . The condition (7.4.3) means that the graph of  $F(t)$  is convex. We now want to draw a convex curve between  $\pi_1$  and  $\pi_2$  which behaves as irregularly as

possible. By "irregularly" we mean that the deviation of the slope of the graph from the slope of the parabolas is occasionally large.

We therefore choose the graph of  $F$  as follows. We take a sequence of points  $P_0 = (t_0, y_0)$ ,  $P_1 = (t_1, y_1)$ , ... on  $\pi_2$ , such that, for each value of  $k$ , the line connecting  $P_k$  and  $P_{k+1}$  touches  $\pi_1$  in a point somewhere between these two. Now the graph of  $F$  is the broken line  $P_0P_1P_2P_3\dots$ .

It does not matter that  $F$  has no derivative at the vertices  $P_k$ , for  $f$  is not defined as the derivative of  $F$ , but  $F$  is defined as the integral of  $f$ . We can give  $f(t_k)$  any value between the slopes of  $P_{k-1}P_k$  and  $P_kP_{k+1}$ .

The condition that  $P_kP_{k+1}$  touches  $\pi_2$  is, by elementary analytic geometry, easily translated into the relation  $(t_k + t_{k+1} - 2)^2 = 4t_k t_{k+1}$ , whence

$$t_{k+1} = t_k + 2 + (8t_k)^{\frac{1}{2}}.$$

If  $t_0$  is chosen arbitrarily, then  $t_1, t_2, \dots$  can be evaluated successively. Accidentally we are able to give an explicit solution. (If we were not, we should have to study the asymptotic behaviour of  $t_k$ , as  $k \rightarrow \infty$ , and that can be done by methods indicated in Ch. 8). If we take  $t_0 = 0$ , then  $t_1 = 0 + 2 + 0 = 2$ ,  $t_2 = 2 + 2 + 4 = 8$ ,  $t_3 = 8 + 2 + 8 = 18$ ,  $t_4 = 32$ ,  $t_5 = 50$ , ... and it is easy to show that  $t_k = 2k^2$ . Now the slope of  $P_kP_{k+1}$  is easily seen to be  $(2k+1)^2$ , and the point of contact lies at  $t = 2k^2 + 2k$ . We therefore define  $f$  by

$$f(t) = (2k+1)^2 \quad (2k^2 \leq t < 2(k+1)^2, k = 0, 1, 2, \dots).$$

Now obviously  $f(t_k) - 2t_k = 4k+1 > t_k^{\frac{1}{2}}$ , so that  $(f(t) - 2t)t^{-\frac{1}{2}}$  does not tend to zero. The function  $f(t)$  is non-decreasing, and  $F(t) = \int_0^t f(x)dx$  lies between  $t^2 - t$  and  $t^2 + t$  for all  $t$ . This follows from the geometrical argument, but it can of course be verified by integration, which gives

$$\begin{aligned} F(t) &= t^2 + t - (t - 2k^2 - 2k)^2 = \\ &= t^2 - t + (2(k+1)^2 - t)(t - 2k^2) \quad (2k^2 \leq t \leq 2(k+1)^2). \end{aligned}$$

From these formulas it is evident that

$$t^2 - t \leq F(t) \leq t^2 + t \quad (0 \leq t < \infty).$$

We just established that (7.4.3) is not a satisfactory Tauberian condition in order to pass from (7.4.2) to (7.4.1). We shall now consider a much stronger condition: we assume that  $f$  has a non-negative second derivative:

$$(7.4.5) \quad f''(t) \geq 0 \quad (t \geq 0).$$

Moreover assuming (7.4.2), i.e.  $F(t) = t^2 + O(t)$ , we can derive (7.4.1) by some simple arguments. We can even prove more, viz. that there exists a number  $b$  such that  $f(t) = 2t + b + o(1)$ . First we remark that  $f''(t) \geq 0$  means that  $f'(t)$  is non-decreasing. If for some  $t_0$  we had  $f'(t_0) = a$ , where  $a > 2$ , we would have  $f'(t) \geq a$  ( $t > t_0$ ), and this would conflict with  $F(t) = t^2 + O(t)$ . So  $f'(t) \leq 2$  for all  $t$ , and it follows that  $f(t) - 2t$  is non-increasing. If  $t > t_1 > 1$  we have

$$F(t) - t^2 - F(t_1) + t_1^2 = \int_{t_1}^t (f(x) - 2x) dx \leq (t - t_1)(f(t_1) - 2t_1).$$

Let  $A$  be such that  $|F(t) - t^2| < At$  ( $t > 1$ ). Making  $t \rightarrow \infty$  we infer that  $f(t_1) - 2t_1 \geq -A$ . So  $f(t) - 2t$  is non-increasing and bounded below; it follows that  $f(t) - 2t$  tends to a limit when  $t \rightarrow \infty$ . We can also say something about  $f'(t)$ ; as  $f'$  is non-decreasing,  $f' \leq 2$ , and  $f(t) = 2t + b + o(1)$ , it is evident that  $f'(t) = 2 + o(1)$ , and even that  $\int_0^\infty (f'(t) - 2) dt$  converges.

### 7.5. Karamata's method

Let  $a_0 + a_1x + a_2x^2 + \dots$  be a power series, convergent if  $|x| < 1$ , and let some asymptotic information be given about the partial sums  $a_1 + \dots + a_n$ , as  $n \rightarrow \infty$ . If this behaviour is not too irregular, we can deduce, by Abelian arguments, the behaviour of the sum function

$$f(x) = a_0 + a_1x + a_2x^2 + \dots$$

as  $0 < x < 1$ ,  $x \rightarrow 1$ . (A special case of this problem and of its Tauberian counterpart was discussed in sec. 7.2).

Conversely, assume that the asymptotic behaviour of  $f(x)$  (as  $x \rightarrow 1$ ) is known, and that we want to derive information about the partial sums

$$A(n) = \sum_{0 \leq k \leq n} a_k.$$

These can be written as

$$(7.5.1) \quad A(n) = \sum_{k=0}^{\infty} a_k g(e^{-k/n}) \quad (n > 0),$$

where

$$g(x) = \begin{cases} 0 & (0 \leq x < e^{-1}), \\ 1 & (e^{-1} \leq x \leq 1). \end{cases}$$

In Karamata's method this discontinuous function  $g(x)$  is approximated, in some sense to be specified later, by a polynomial  $P(x)$ , with  $P(0) = 0$ . If

$$P(x) = \sum_{j=1}^m p_j x^j,$$

then the sum corresponding with (7.5.1) is

$$(7.5.2) \quad \sum_{k=0}^{\infty} a_k P(e^{-k/n}) = \sum_{j=1}^m p_j f(e^{-j/n}).$$

If  $P$  is a fixed polynomial, then the asymptotic behaviour (as  $n \rightarrow \infty$ ) of the right-hand-side is known.

The method can be applied to a fairly large class of cases. We shall specialize by taking a fixed real number  $\gamma$  and assuming

$$(7.5.3) \quad f(x) = \sum_{k=0}^{\infty} a_k x^k = o((1-x)^{-\gamma}) \quad (0 < x < 1, x \rightarrow 1).$$

Using the Tauberian condition

$$(7.5.4) \quad a_k > -C(k+1)^{\gamma-1} \quad (k = 0, 1, 2, \dots),$$

where  $C$  is a positive constant, we shall prove that

$$(7.5.5) \quad A(n) = \sum_{k=0}^n a_k = o(n^{\gamma}) \quad (n \rightarrow \infty).$$

The special case  $\gamma = 0$  has been announced already in sec. 7.2. This case is usually deduced from the case  $\gamma = 1$  by some auxiliary Tauberian theorems, but Karamata's method is strong enough to cover the case  $\gamma = 0$  as well<sup>1)</sup>.

As to the approximations of  $g(x)$  by polynomials we shall stipulate the following conditions. Let  $h$  be an integer,  $h > 0$ ,  $h > -\gamma$ ;

<sup>1)</sup> See H. WIELANDT, Math. Zeitschr. 56 (1952), pp. 206–207.

in fact we take for  $h$  the smallest integer satisfying these inequalities. Let  $\varepsilon$  be a positive number,  $0 < \varepsilon < \frac{1}{2}$ . Then we want to have polynomials  $P_1(x)$ ,  $P_2(x)$  such that

$$\begin{aligned} P_1(x) &\leq g(x) \leq P_2(x) & (0 \leq x \leq 1), \\ |P_2(x) - P_1(x)| &\leq 2 & (0 \leq x \leq 1), \\ |P_2(x) - P_1(x)| &\leq \varepsilon x & (0 \leq x \leq e^{-1-\varepsilon}), \\ |P_2(x) - P_1(x)| &\leq \varepsilon(1-x)^h & (e^{-1+\varepsilon} \leq x \leq 1). \end{aligned}$$

We do not require anything regarding the degree or the coefficients of these polynomials. The possibility of finding  $P_1$  and  $P_2$  can be shown as follows. We first get rid of the discontinuity at  $x = e^{-1}$ , constructing continuous functions  $g_1(x)$  and  $g_2(x)$  such that

$$\begin{aligned} g_1(x) &\leq g(x) \leq g_2(x) & (0 \leq x \leq 1), \\ g_2(x) &= g_1(x) = g(x) & \text{if } x \leq e^{-1-\varepsilon} \text{ or } x \geq e^{-1+\varepsilon}, \\ g_2(x) - g_1(x) &\leq 1 & \text{if } e^{-1-\varepsilon} \leq x \leq e^{-1+\varepsilon}. \end{aligned}$$

Next we determine a polynomial  $Q(x)$  such that  $Q(0) = g(0)$ ,  $Q(1) = g(1)$ ,  $Q'(1) = g'(1)$ , ...,  $Q^{(h-1)}(1) = g^{(h-1)}(1)$ . Actually,  $Q(x) = 1 - (1-x)^h$  is already suitable, but the explicit form of  $Q$  does not matter. Now the function  $\varphi_2$ , defined by

$$(g_2(x) - Q(x))x^{-1}(1-x)^{-h} = \varphi_2(x)$$

is a continuous function in  $0 \leq x \leq 1$ . By the Weierstrass approximation theorem<sup>1)</sup> we can find a polynomial  $R_2(x)$  such that

$$|\varphi_2(x) + \frac{1}{2}\varepsilon - R_2(x)| < \frac{1}{2}\varepsilon \quad (0 \leq x \leq 1).$$

Putting  $Q + x(1-x)^h R_2 = P_2$ , we observe that

$$g_2(x) \leq P_2(x) \leq g_2(x) + \frac{1}{2}\varepsilon x(1-x)^{-h} \quad (0 \leq x \leq 1).$$

Similarly we construct  $P_1(x)$  such that

$$g_1(x) - \frac{1}{2}\varepsilon x(1-x)^{-h} \leq P_1(x) \leq g_1(x) \quad (0 \leq x \leq 1).$$

Then  $P_1$  and  $P_2$  obviously satisfy all requirements.

By (7.5.1) and (7.5.4) we have, for all positive values of  $n$ ,

---

<sup>1)</sup> See E. C. TITCHMARSH, Theory of Functions, 2nd ed., Oxford 1939, § 13.33.

$$(7.5.6) \quad -A(n) + \sum_{k=0}^{\infty} a_k P_1(e^{-k/n}) \leq \\ \leq \sum_{k=0}^{\infty} C(k+1)^{\gamma-1} (g(e^{-k/n}) - P_1(e^{-k/n})),$$

$$(7.5.7) \quad A(n) - \sum_{k=0}^{\infty} a_k P_2(e^{-k/n}) \leq \\ \leq \sum_{k=0}^{\infty} C(k+1)^{\gamma-1} (P_2(e^{-k/n}) - g(e^{-k/n})),$$

by virtue of the fact that  $g - P_1 \geq 0$ ,  $P_2 - g \geq 0$ . The right-hand-sides are at most

$$\sum_{k=0}^{\infty} C(k+1)^{\gamma-1} (P_2(e^{-k/n}) - P_1(e^{-k/n})).$$

This amount is easily estimated above, by splitting the sum according to  $k \leq n(1-\varepsilon)$ ,  $n(1-\varepsilon) < k \leq n(1+\varepsilon)$ ,  $n(1+\varepsilon) < k < \infty$ . We have, for the first part of the sum

$$\begin{aligned} \sum_{0 \leq k \leq n(1-\varepsilon)} &\leq \sum_{0 \leq k \leq n(1-\varepsilon)} C(k+1)^{\gamma-1} \varepsilon (1 - e^{-k/n})^h \leq \\ &\leq \sum_{0 \leq k \leq n} \varepsilon C(k+1)^{\gamma-1} k^h n^{-h} \leq \varepsilon C D_1 n^\gamma, \end{aligned}$$

where  $D_1$  depends on  $\gamma$ , but not on  $\varepsilon$  and  $n$ . For,  $(k+1)^{\gamma-1} k^h \leq (k+1)^{\gamma+h-1}$ , and  $\gamma+h$  is positive. And if  $\beta$  is a positive number, we have

$$\sum_{k=0}^n (k+1)^{\beta-1} = O(n^\beta) \quad (n > 0).$$

The second part of the sum has at most  $2\varepsilon n + 1$  terms, and so

$$\sum_{n(1-\varepsilon) < k \leq n(1+\varepsilon)} \leq 2C \sum_{n(1-\varepsilon) < k \leq n(1+\varepsilon)} (k+1)^{\gamma-1} \leq 2C(2\varepsilon n + 1) D_2 n^{\gamma-1},$$

where  $D_2$  depends on  $\gamma$  only (we can take  $D_2 = 3^{|\gamma-1|}$ , since  $0 < \varepsilon < \frac{1}{2}$  guarantees that  $n/3 < k+1 < 3n$  if  $n(1-\varepsilon) < k \leq n(1+\varepsilon)$ ).

The third part of the sum is

$$\sum_{k \geq n(1+\varepsilon)} \leq \sum_{k \geq n(1+\varepsilon)} C(k+1)^{\gamma-1} \varepsilon e^{-k/n} \leq \varepsilon C D_3 n^\gamma,$$

where  $D_3$  depends on  $\gamma$  only. This is easily obtained by comparing the sum with the corresponding integral

$$\int_n^{\infty} x^{\gamma-1} e^{-x/n} dx = n^\gamma \int_1^{\infty} y^{\gamma-1} e^{-y} dy.$$

So the sums occurring on the right-hand-sides of (7.5.6) and (7.5.7) give at most  $\varepsilon Cnr(D_1 + 4D_2 + D_3 + 2n^{-1}\varepsilon^{-1}D_2)$ , where the  $D$ 's depend on  $\gamma$  only.

The sum  $\sum_{k=0}^{\infty} a_k P_1(e^{-k/n})$  occurring in (7.5.6), can be estimated by virtue of (7.5.2) and (7.5.3) (it should be noticed that  $P_1(0)=0$ ). As

$$(1 - e^{-j/n})^{-\gamma} \sim j^{-\gamma} n^{\gamma} \quad (n \rightarrow \infty, j \text{ fixed}),$$

the sum is  $O(n^{\gamma})$ . More precisely, we can determine a number  $n_0$ , depending on  $\gamma$ ,  $\varepsilon$  and on the polynomial  $P_1$ , such that the sum is, in absolute value, less than  $\varepsilon n^{\gamma}$ . The same thing holds for the sum with  $P_2$ , occurring in (7.5.7). We can assume that  $n_0$  serves both  $P_1$  and  $P_2$ . We finally obtain, from (7.5.6) and (7.5.7),

$$|A(n)| < \varepsilon n^{\gamma} (CD_1 + 4CD_2 + CD_3 + 2n^{-1}\varepsilon^{-1}CD_2 + 1) \quad (n > n_0).$$

If moreover  $n > \varepsilon^{-1}$ , we infer that  $|A(n)| < \varepsilon n^{\gamma} D_4$ , where  $D_4$  is independent of  $\varepsilon$  and  $n$ . As  $\varepsilon$  is arbitrary, this proves (7.5.5).

If  $\gamma > 0$ , it is easy to prove (cf. sec. 3.5) that

$$\begin{aligned} \sum_{k=0}^{\infty} (k+1)^{\gamma-1} x^k &\sim \Gamma(\gamma) (1-x)^{-\gamma} \quad (0 < x < 1, x \rightarrow 1), \\ \sum_{k=0}^{\infty} (k+1)^{\gamma-1} &\sim \gamma^{-1} n^{\gamma} \quad (n \rightarrow \infty). \end{aligned}$$

So for  $\gamma > 0$  our Tauberian theorem can be generalized as follows. If

$$\begin{aligned} \sum_{k=0}^{\infty} a_k x^k &\sim (1-x)^{-\gamma} \quad (0 < x < 1, x \rightarrow 1), \\ a_k &> -C(k+1)^{\gamma-1} \quad (k = 0, 1, 2, \dots), \end{aligned}$$

then we have

$$\sum_{k=0}^n a_k \sim (\Gamma(\gamma+1))^{-1} n^{\gamma} \quad (n \rightarrow \infty).$$

**7.6. EXERCISES.** 1. Assume  $\alpha \geq 0$ ,  $C > 0$ ,  $f'(t) > -Ct^{\alpha-1}$  ( $t > 1$ ), and  $\int_0^t f(x) dx \sim (\alpha+1)^{-1} t^{\alpha+1}$  ( $t \rightarrow \infty$ ). Show that  $f(t) \sim t^{\alpha}$  ( $t \rightarrow \infty$ ).

2. Assume that

$$\sum_{k=0}^{\infty} a_k x^k \sim -\log(1-x) \quad (0 < x < 1, x \rightarrow 1),$$

where

$$a_k > -C k^{-1} \log k \quad (k = 1, 2, 3, \dots),$$

with some positive constant  $C$ . Show that  $\sum_0^n a_k \sim \log n$  ( $n \rightarrow \infty$ ).

## CHAPTER 8

### ITERATED FUNCTIONS

#### 8.1. Introduction

Many problems in asymptotics can be stated in the following terms: Let a sequence of functions  $F_1, F_2, \dots$  be given, and let  $x_0$  be a number. Now we put

$$x_1 = F_1(x_0), \quad x_2 = F_2(x_1), \quad x_3 = F_3(x_2), \dots,$$

assuming that  $F_1$  is defined at  $x_0$ ,  $F_2$  is defined at  $x_1$ , etc. The problem is to find the asymptotic behaviour of  $x_n$  as  $n$  tends to infinity.

In the larger part of the present chapter we shall discuss only a very special case of the problem, taking all functions  $F_1, F_2, \dots$  to be one and the same function  $f$ . Nevertheless, cases where the  $F_i$  are different, can quite often be tackled by methods devised for this special case. This remark holds for even more general cases. We mention the possibility that  $F_n$  is a function of  $n$  variables instead of one variable, and that  $x_n$  is defined recursively by the formula  $x_n = F_n(x_0, x_1, \dots, x_{n-1})$ . A further generalization is obtained if we replace the  $x$ 's by functions and the  $F$ 's by operators. Many asymptotic problems about the solutions of differential or integral equations fall under this heading.

#### 8.2. Iterates of a function

Until sec. 8.9 we shall take all functions  $F_k$  to be equal to a fixed real function  $f$ . Then we have, if  $x_0$  is any real number,

$$x_1 = f(x_0), \quad x_2 = f(x_1), \quad x_3 = f(x_2), \dots$$

We shall denote by  $f_n$  the  $n$ -th iterate of  $f$ , which is defined by

$$f_1(x_0) = f(x_0), \quad f_{n+1}(x_0) = f\{f_n(x_0)\} = f_n\{f(x_0)\} \quad (n = 0, 1, 2, \dots).$$

Therefore we have  $x_n = f_n(x_0)$  ( $n = 1, 2, \dots$ ).

For a geometrical interpretation of these formulas we refer to Fig. 2. The figure shows the graph of  $y = f(x)$  and the graph of  $y = x$ , and it shows how  $x_1, x_2, x_3, \dots$  can be constructed consecutively if  $x_0$  is given. The reader may take a pleasure in drawing some graphs himself and in discovering what happens to sequences  $x_0, x_1, x_2, \dots$  with various values of  $x_0$ .

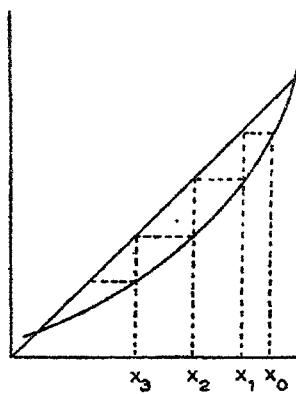


Fig. 2.

In the following discussion we shall assume that  $f$  is defined for all real values of the variable, so that there is no question about the  $f_n$ 's being defined or not.

It is possible that the sequence  $x_0, x_1, x_2, \dots$  tends to a limit  $c$ . If the function  $f$  is continuous at  $c$ , then  $x_{n+1} = f(x_n)$  shows that  $c = f(c)$ . Therefore, if  $f$  is continuous everywhere, the possibilities for  $c$  are restricted to the solutions of the equation  $c = f(c)$ .

Convergence to a point  $c$ , where  $c = f(c)$ , can often be proved in the following way. We show the existence of a neighbourhood  $N$  of  $c$  such that, once some  $x_n$  falls into  $N$ , the sequence  $x_{n+1}, x_{n+2}, \dots$  converges to  $c$ . In such cases it is likely that the asymptotic behaviour of  $x_n$  (as  $n \rightarrow \infty$ ) can be studied in detail, especially if  $f(x)$  is analytic at  $x = c$ .

The problem whether for a given value of  $x_0$  there exists an  $n$  such that  $x_n$  lies in that neighbourhood  $N$  of  $c$ , is of a different nature. It is often quite easy if  $f$  is a real continuous function and  $x$  is a real variable. We shall discuss a general example.

Assume that  $f(x)$  is continuous in the interval  $J$ , defined by  $c \leq x < d$  ( $d$  may also stand for  $+\infty$ ). Furthermore assume that  $f(c) = c$  and  $f(c) \leq f(x) < x$  if  $x \in J$ ,  $x > c$ . Then we have, for any  $x$  in  $J$ , that  $\lim_{n \rightarrow \infty} f_n(x) = c$ . For, our assumptions imply that  $f$  maps  $J$  into itself, and therefore the same thing can be said about  $f_2, f_3, \dots$ . Furthermore  $f(x) \leq x$  ( $x \in J$ ) guarantees that  $x \geq f_1(x) \geq f_2(x) \geq \dots$ . As all  $f_n(x)$  are in  $J$ , the sequence is bounded below. It follows that  $\lim f_n(x)$  exists, and that it tends to a solution of  $x = f(x)$ , which cannot be anything but  $c$ .

A similar discussion applies if  $f(c) = c$ ,  $x < f(x) \leq c$  in an interval  $d < x \leq c$ .

It has to be remarked that  $\lim_{n \rightarrow \infty} f_n(x)$  need not be a continuous function of  $x$ . If, for example, we apply our previous results to the function  $f(x) = x + \sin x$ , we observe that the limit exists for all  $x$ , and if it is denoted by  $\varphi(x)$ , it can be described by  $\varphi(0) = 0$ ,  $\varphi(x) = \pi$  ( $0 < x < 2\pi$ ),  $\varphi(2\pi) = 2\pi$ ,  $\varphi(x) = 3\pi$  ( $2\pi < x < 4\pi$ ),  $\varphi(4\pi) = 4\pi$ ,  $\varphi(x) = 5\pi$  ( $4\pi < x < 6\pi$ ), etc.

However, the situation can be much more complicated than in the cases we discussed just now. If  $f(x)$  is continuous in  $-\infty < x < \infty$ , and if  $f(x) < x$  in an interval  $c < x < d$ , with  $f(c) = c$ , but if  $f(x)$  is not  $\geq f(c)$  throughout that interval, then the behaviour of  $x_n$  (as  $c < x < d$ ,  $n \rightarrow \infty$ ) is no longer exclusively determined by the behaviour of  $f$  in the interval  $c < x < d$ . In such cases the complete discussion of the behaviour of  $x_n$  might be very difficult.

We again turn to the local problem, i.e. the question what happens in small neighbourhoods of a point  $c$  where  $f(c) = c$ .

Without loss of generality we take  $c = 0$  (otherwise consider  $f^*$ , defined by  $f^*(x) = f(x + c) - c$ ; notice that its iterates are given by  $f_n^*(x) = f_n(x + c) - c$ ), and by virtue of the relation  $c = f(c)$  we now have  $f(0) = 0$ .

In order to be able to be more specific we shall assume that  $f$  is analytic at  $x = 0$ :

$$(8.2.1) \quad f(x) = a_1 x + a_2 x^2 + a_3 x^3 + \dots \quad (|x| < p),$$

where  $p$  is some positive number. The coefficients  $a_1, a_2, \dots$  are allowed to be complex numbers, and  $x$  is a complex variable.

The absolute value of the coefficient  $a_1$  is decisive for our problem. If  $|a_1| < 1$  the sequence  $x_0, x_1, x_2, \dots$  converges to 0 indeed,

provided that the starting point  $x_0$  is sufficiently close to 0. Moreover the asymptotic behaviour of  $x_n$  is not difficult to find (see sec. 8.3). Convergence is rapid in this case. If  $0 < |a_1| < 1$ , then  $\log|x_n^{-1}|$  behaves as  $Cn$ , where  $C$  is a positive constant. If  $a_1 = 0$  the convergence is even much faster. If  $|a_1| > 1$  it is easy to see that  $x_n$  does not converge to 0, unless the  $x_n$ 's vanish identically from a certain value of  $n$  onwards. If  $|a_1| = 1$  the problem is more intricate (see sec. 8.5), and if there is convergence it is quite slow.

### 8.3. Rapid convergence

If  $|a_1| < 1$  we are in the fortunate circumstance that the iteration problem for  $f(x)$ , given by (8.2.1), can be solved by a direct method. That is, for  $x_n$  (defined by  $x_n = f_n(x_0)$ ) we can derive a new formula, from which the asymptotic behaviour of  $x_n$  (as  $n \rightarrow \infty$ ) can be obtained, provided that  $|x_0|$  is not too large.

We assume here that  $a_1 \neq 0$ ; the case  $a_1 = 0$  will be considered at the end of this section.

We start from a rough estimate for  $f_n(x)$ . Let  $b$  satisfy  $|a_1| < b < 1$ . Then there exists a number  $p_1$  ( $0 < p_1 < p$ ) such that

$$(8.3.1) \quad 0 < |f_n(x)| < b^n|x| \quad (n = 1, 2, \dots; 0 < |x| < p_1).$$

For, the power series for  $x^{-1}f(x)$  has the value  $a_1$  at  $x = 0$ . Therefore  $p_1$  can be found such that  $|x| < p_1$  implies  $0 < |x^{-1}f(x)| < b$ . So if  $x_0$  satisfies  $0 < |x_0| < p_1$  we have  $0 < |x_1| < b|x_0|$ , and therefore  $0 < |x_1| < p_1$ . In the second step we infer that  $0 < |x_2| < b|x_1|$  and  $0 < |x_2| < p_1$ . By induction we find  $0 < |x_n| < b|x_{n-1}|$  and  $0 < |x_n| < p_1$ . It follows that  $0 < |x_n| < b^n|x_0|$ , and (8.3.1) follows.

We next prove that, if  $x_0$  is fixed ( $0 < |x_0| < p_1$ ), then  $x_n a_1^{-n}$  tends to a limit, which we denote by  $\omega(x_0)$ . We have

$$(8.3.2) \quad \frac{x_{n+1}}{a_1 x_n} = \frac{f(x_n)}{a_1 x_n} = 1 + \frac{a_2}{a_1} x_n + \frac{a_3}{a_1} x_n^2 + \dots,$$

and this is close to 1 if  $n$  is large. Writing  $1 + r_n$  for the right-hand-side of (8.3.2), we infer from (8.3.1) that  $r_n = O(b^n)$ . Consequently, the product  $\prod_{k=0}^{\infty} (1 + r_k)$  converges. As  $\prod_{k=0}^{n-1} (1 + r_k)$  equals  $x_n a_1^{-n} x_0^{-1}$ ,

we infer that  $x_n a_1^{-n}$  tends to a limit  $\omega(x_0)$ , where

$$(8.3.3) \quad \omega(x_0) = x_0 \prod_{n=0}^{\infty} (1 + r_n).$$

As  $f_n(x_0) = f_{n-1}(f(x_0)) = f_{n-1}(x_1)$ , we find that  $x_n a_1^{1-n}$  tends to  $\omega(x_1)$ , and therefore  $\omega(x_1) = a_1 \omega(x_0)$ . That is, the function  $\omega$  satisfies the so-called Schröder equation <sup>1)</sup>

$$(8.3.4) \quad \omega(f(x)) = a_1 \omega(x) \quad (|x| < p_1).$$

It should be remarked that  $\omega(x)$  is analytic inside the circle  $|x| < p_1$ . This follows from the fact that each factor  $1 + r_n$  in the product (8.3.3) is an analytic function of  $x_0$ , and the product converges uniformly for  $x_0$  in that circle. And (8.3.3) shows that  $\omega(0) = 0$ ,  $\omega'(0) = 1$ . We put

$$(8.3.5) \quad \omega(x) = x + d_2 x^2 + d_3 x^3 + \dots \quad (|x| < p_1).$$

The coefficients  $d_2, d_3, \dots$  can be determined recursively from the identity (8.3.4). Once we have determined  $d_2, d_3, \dots, d_{n-1}$ , we can evaluate  $d_n$  by equating the coefficients of  $x^n$  in (8.3.4). This gives a linear equation for  $d_n$ , only containing  $d_2, \dots, d_n$ . In this equation  $d_n$  gets the coefficient  $a_1^n - a_1$ , and this is  $\neq 0$  by our assumption that  $0 < |a_1| < 1$ .

By repeated application of (8.3.4) we obtain  $\omega(x_n) = a_1^n \omega(x_0)$ , and solving this for  $x_n$  we shall get an explicit formula. The Lagrange inversion formula (see sec. 2.2) gives the inverse function  $\Omega$  of  $\omega$ , satisfying  $\Omega(\omega(x)) = x$  in a suitable circle  $|x| < p_2$ ; and

$$\Omega(x) = x + e_2 x^2 + e_3 x^3 + \dots \quad (|x| < p_2).$$

We remark that the coefficients of  $\Omega$  can be evaluated recursively without using the coefficients of  $\omega$ . To this end we can use the identity  $f(\Omega(y)) = \Omega(a_1 y)$ , which follows from (8.3.4) by putting  $\omega(x) = y$ .

If  $p_3 = \min(p_1, p_2)$ , and  $|x_0| < p_3$ , then we have by (8.3.1) that  $|x_n| < p_3$  for all  $n$ . It follows that

$$(8.3.6) \quad \begin{aligned} x_n &= \Omega(a_1^n \omega(x_0)) = \\ &= a_1^n \omega(x_0) + e_2 a_1^{2n} \omega(x_0)^2 + e_3 a_1^{3n} \omega(x_0)^3 + \dots \quad (|x_0| < p_3). \end{aligned}$$

<sup>1)</sup>ANN E. SCHRÖDER, Math. . 3 (1871), pp. 296–322.

This formula gives very satisfactory information about the behaviour of  $x_n$  when  $n \rightarrow \infty$ .

Although it has no direct consequence for asymptotics, we mention that the above formulas can be used for continuous iteration. That is to say, we can define functions  $f_\lambda(x)$  for all  $\lambda \geq 0$ , such that

$$f_\lambda(f_\mu(x)) = f_{\lambda+\mu}(x) \quad (\lambda \geq 0, \mu \geq 0), \quad f_1 = f,$$

and  $f_0$  is the identity ( $f_0(x) = x$ ). If  $\lambda$  is a positive integer, then  $f_\lambda$  is the  $\lambda$ -th iterate of  $f$ . The functions  $f_\lambda$  can be defined by<sup>1)</sup>

$$f_\lambda(x) = \Omega\{a_1^\lambda \omega(x)\} \quad (|x| < p_3).$$

Thus far we have assumed that  $a_1 \neq 0$ . We now briefly indicate what happens if  $a_1 = 0$ . If all coefficients in (8.2.1) vanish, then all  $x_n$  are zero, and there is no problem. So assume that  $a_k$  is the first non-vanishing coefficient, and that  $k > 1$ . Without loss of generality we may take  $a_k = 1$ , for otherwise we can consider the function  $f^*(x) = \alpha^{-1}f(\alpha x)$  where  $\alpha$  is chosen such that  $a_k \alpha^{k-1} = 1$ . So we put

$$f(x) = x^k + a_{k+1}x^{k+1} + a_{k+2}x^{k+2} + \dots$$

The iteration machinery can be controlled by the following formulas:

$$\lim_{n \rightarrow \infty} \{f_n(x)\}^{k-n} = \omega(x), \quad \omega(0) = 0, \quad \omega'(0) = 1.$$

$$\omega(f(x)) = (\omega(x))^k.$$

$$f_n(x) = \Omega\{(\omega(x))^{kn}\} \quad (\Omega = \text{inverse of } \omega).$$

This is only the formal side of the matter, but it is not difficult to give a detailed justification in the way it was done in the case  $0 < |a_1| < 1$ .

#### 8.4. Slow convergence

Our next case is the iteration problem for a function  $f$ , in the form (8.2.1), with  $|a_1| = 1$ . Our methods will be applicable to functions which have an asymptotic series of the type

$$f(x) \approx x + a_k x^k + a_{k+1} x^{k+1} + \dots \quad (x \rightarrow 0, x > 0),$$

with  $k > 1$ ,  $a_k < 0$ . As an example we shall treat the case that

$$f(x) = \sin x = x - x^3/3! + x^5/5! - \dots,$$

<sup>1)</sup> The formula is due to G. KOENIGS, Ann. Sci. Éc. Norm. Sup., (3) 1 (1884), Supplém. pp. 3–41, and 2 (1885), pp. 385–404.

but the general case is not more difficult (see sec. 8.11, exercise 1). As before, we write

$$\sin_1 x = \sin x, \quad \sin(\sin_n x) = \sin_n(\sin x) = \sin_{n+1} x; \quad x_n = \sin_n x_0.$$

If  $0 < x_0 < \pi$ , then we have  $0 < \sin x_0 < x_0$ . Therefore, by induction,  $0 < x_n < \pi$  for all  $n$ , and  $x_0 > x_1 > x_2 > \dots$ . It follows that  $\lim x_n = 0$  (see sec. 8.2). We now raise the question of the asymptotic behaviour.

There is a difference between this case and the cases of sec. 8.3. In sec. 8.3 there was convergence to 0 for all complex  $x_0$  inside a certain circle. In the present case this is no longer true. For example, if we take  $x_0$  purely imaginary,  $x_0 = it_0$  ( $t_0 > 0$ ), then we have  $x_n = it_n$ , where  $t_n = \sinh_n t_0$ . And it is easily seen that  $0 < t_0 < t_1 < t_2 < \dots$ , and that  $t_n$  tends to infinity, no matter how small  $t_0$  was chosen.

In sec. 8.3 we were able to solve the problem by means of certain series in terms of powers of  $x$ , for which it did not matter whether  $x_0$  was real or not. In the present case it seems to matter indeed, and therefore we cannot expect to be able to do much with such power series. Apart from that, the study of complex values of  $x_0$  seems to be difficult, and so we shall confine ourselves to real values of  $x_0$ . It is no essential restriction to assume that  $0 < x_0 < \pi$  ( $x_1 = \sin x_0$  at any rate satisfies  $-\pi < x_1 < \pi$ , and  $\sin_n x_0 = \sin_{n-1} x_1$ ; furthermore, owing to the symmetry, there is no harm in considering positive values only).

We shall give two different methods for the solution of the problem of the asymptotic behaviour of  $\sin_n x_0$ . The first one (secs. 8.5 and 8.6) is quite natural, and uses ideas which are generally applicable in iteration problems; the second method (sec. 8.7 and 8.8) is more effective, but essentially restricted to iteration problems of the type we are dealing with now.

### 8.5. Preparation

The following question will serve as a preparation. Let  $u_1, u_2, \dots$  be a sequence of positive numbers, and assume that

$$(8.5.1) \quad u_{n+1} = u_n - u_n^2 + O(u_n^3) \quad (n \rightarrow \infty).$$

What can be said about the asymptotic behaviour of  $u_n$  as  $n \rightarrow \infty$ ?

In the first place it is clear that nothing can be said if we do not assume something like  $u_n \rightarrow 0$ . For, if  $p$  and  $q$  are fixed positive numbers,  $0 < p < q$ , then any arbitrary sequence of numbers  $\{u_n\}$  with  $p \leq u_n \leq q$  obviously satisfies (8.5.1).

We want to be more specific about the O-term: Let  $A$  be a fixed positive number such that

$$(8.5.2) \quad |u_{n+1} - u_n + u_n^2| \leq Au_n^3 \quad (n = 1, 2, 3, \dots).$$

It is not difficult to show that there exists a number  $p > 0$ , such that, whenever  $0 < u_k < p$  for some value of  $k$ , we automatically get  $u_k > u_{k+1} > u_{k+2} > \dots$ , and  $u_k \rightarrow 0$ . To this end we choose  $p$  such that  $0 < x < p$  implies both  $x - x^2 > Ax^3$  and  $Ax^3 < x^2$  (therefore  $Ap \leq 1$ ). Then it follows from (8.5.2) that  $0 < u_k < p$  implies  $0 < u_{k+1} < u_k < p$ , and so forth. The sequence is decreasing and bounded below, and so it converges to a limit  $c$ , with  $0 \leq c < p$ . From (8.5.2) we infer that  $|c - c + c^2| \leq Ac^3$ . As  $Ap \leq 1$ ,  $0 \leq c < p$ , it follows that  $c = 0$ .

We just learned (under the general assumption that all  $u_n$  are positive) that either all  $u_n$  are  $\geq p$  or  $u_n$  tends monotonically to 0 from a certain  $n$  onward. For the latter case we can deduce a much sharper result:

$$(8.5.3) \quad \text{If } u_n \rightarrow 0, \quad \text{then} \quad u_n = n^{-1} + O(n^{-2} \log n).$$

This can be proved as follows. By some elementary computations we find that there exist positive constants  $K$  and  $N$  such that, for all  $n \geq N$  the following is true: for all  $x$  in the interval

$$(8.5.4) \quad 0 < x < n^{-1} + Kn^{-2} \log n$$

we have

$$(8.5.5) \quad 0 < x - x^2 + Ax^3 < (n+1)^{-1} + K(n+1)^{-2} \log(n+1).$$

(In order to show this, put  $\lambda_n = n^{-1} + Kn^{-2} \log n$ , and show that

$$\lambda_{n+1} - \lambda_n + \lambda_n^2 - A\lambda_n^3 = (1 + K - A)n^{-3} + O(n^{-4} \log^2 n),$$

whence  $\lambda_n - \lambda_n^2 + A\lambda_n^3 < \lambda_{n+1}$  if  $K$  and  $n$  are sufficiently large).

Now let  $k$  be chosen such that  $0 < u_k < N^{-1} + KN^{-2} \log N$ ; this is possible by virtue of the assumption  $u_n \rightarrow 0$ . Then it is easy

to prove by induction that

$$0 < u_{k+m} < (N + m)^{-1} + K(N + m)^{-2} \log(N + m) \quad (m = 0, 1, 2, \dots).$$

Therefore  $u_n < n^{-1} + O(n^{-2} \log n)$ . A similar argument can be used for the lower estimation.

The difficulty lies, of course, not so much in proving a result like (8.5.3) as in guessing what one has to prove. That  $n^{-1}$  is a first approximation may already be guessed by comparing the difference equation  $u_{n+1} - u_n = -u_n^2$  with the differential equation  $u'(t) = -u(t)^2$ , whose solutions are  $u(t) = (t + c)^{-1}$ . It is not so easy to describe how the  $O(n^{-2} \log n)$  can be guessed. It requires some imagination and experience to see that just the term  $Kn^{-2} \log n$  makes it possible to pass from (8.5.4) to (8.5.5).

Accidentally there is a much simpler way to prove (8.5.3). It depends on the substitution  $u_n = v_n^{-1}$ , which transforms the equation (8.5.1) into a more suitable form. There is no obvious reason to expect this beforehand, but it is suggested by the form of (8.5.3). We obtain from (8.5.1)

$$v_{n+1} = v_n(1 - u_n + O(u_n^2))^{-1} = v_n(1 + u_n + O(u_n^2)) \quad (n \rightarrow \infty),$$

and therefore

$$(8.5.6) \quad v_{n+1} - v_n = 1 + O(v_n^{-1}), \quad v_n \rightarrow \infty \quad (n \rightarrow \infty).$$

As  $v_n^{-1} \rightarrow 0$ , we infer that  $v_{n+1} - v_n > \frac{1}{2}$  for all large  $n$ , and therefore  $v_n > n/4$  for all large  $n$ . It follows that  $v_n^{-1} = O(n^{-1})$ . Consequently  $v_{n+1} - v_n = 1 + O(n^{-1})$ , and from this we infer  $v_n = n + O(\log n)$ . As

$$u_n = v_n^{-1} = \{n + O(\log n)\}^{-1} = n^{-1} + O(n^{-2} \log n),$$

we again have (8.5.3).

About (8.5.1) we proved (8.5.3), that is a statement of the type: if the  $u_n$ 's are not too large, then they are very small. Without proof we quote two similar results, whose validity, however, is not restricted to real sequences.

(i) If the sequence  $\{a_n\}$  satisfies

$$(8.5.7) \quad a_{n+1} - a_n = O(n^{-1}) + O(n^{-2}a_n^2), \quad a_n = o(n) \quad (n \rightarrow \infty),$$

then we have  $a_n = O(\log n)$ .

(ii) If the sequence  $\{b_n\}$  satisfies

$$b_{n+1} - b_n = O(n^{-2}b_n^2), \quad b_n = o(n) \quad (n \rightarrow \infty).$$

then we have  $b_n = O(1)$ .

Theorem (i) is related to (8.5.1): assuming that we have already established  $nu_n \rightarrow 1$ ,  $u_n$  satisfying (8.5.1), then the substitution  $u_n = n^{-1} + n^{-2}a_n$  leads to (8.5.7).

### 8.6. Iteration of the sine function

We return to the iteration problem for the sine function. It is assumed that  $0 < x_0 < \pi$ ,  $x_1 = \sin x_0$ ,  $x_2 = \sin x_1$ , etc. We already showed that  $x_n \rightarrow 0$ . We have

$$x_{n+1} = \sin x_n = x_n - \frac{1}{6}x_n^3 + \frac{1}{120}x_n^5 - \dots$$

As the series contains only odd powers of  $x_n$ , the formula can be simplified by putting  $x_n^2 = y_n$ :

$$y_{n+1} = y_n(1 - \frac{1}{6}y_n + \frac{1}{120}y_n^2 - \dots)^2.$$

Writing  $\frac{1}{6}y_n = \frac{1}{6}x_n^2 = z_n$ , we obtain something resembling (8.5.1):

$$(8.6.1) \quad z_{n+1} = z_n(1 - \frac{1}{2}z_n + \frac{3}{40}z_n^2 - \dots)^2 = \\ = z_n - z_n^2 + \frac{2}{5}z_n^3 - \dots$$

As  $z_n \rightarrow 0$ , (8.5.3) gives an asymptotic formula for  $z_n$ , viz.  $z_n = n^{-1} + O(n^{-2} \log n)$ , whence  $\sin_n x_0 = 3^{\frac{1}{2}}n^{-\frac{1}{2}}\{1 + O(n^{-1} \log n)\}^{\frac{1}{2}}$ .

It has to be understood that all O-formulas in this section refer to the set of all positive integers  $n$ , and that they do not hold uniformly with respect to  $x_0$ .

Further results can be found by inserting this result into the equation (8.6.1). Calculations are somewhat easier if we consider  $w_n = z_n^{-1}$  instead of  $z_n$ , the same substitution that gave such an easy success in sec. 8.5. The relation between  $w_n$  and  $x_n$  is

$$w_n = 3x_n^{-2},$$

and we know that  $w_n \rightarrow \infty$ . Moreover,  $w_n$  satisfies the relation

$$(8.6.2) \quad w_{n+1} = w_n + 1 + \frac{3}{5}w_n^{-1} + O(w_n^{-2}).$$

---

<sup>1)</sup> A slightly weaker result, viz.  $n^{\frac{1}{2}} \sin_n x \rightarrow 3^{\frac{1}{2}}$  ( $x$  fixed,  $0 < x < \pi$ ,  $n \rightarrow \infty$ ) was given by G PÓLYA and G SZEGÖ, Aufgaben und Lehrsätze aus der Analysis, vol. 1, New York 1945, Abschn 1, problem nr 173.

It is not difficult to obtain the full development. By differentiation of the well-known series for  $\cot x$  it follows that

$$(8.6.3) \quad (\sin x)^{-2} = \sum_{k=0}^{\infty} (-4)^k (1 - 2k) B_{2k} x^{2k-2} / (2k)!,$$

where the  $B$ 's are the Bernoulli numbers (see sec. 3.6). It follows from  $x_{n+1} = \sin x_n$  that

$$(8.6.4) \quad w_{n+1} = w_n \sum_{k=0}^{\infty} w_n^{-k} (1 - 2k) (-12)^k B_{2k} / (2k)! = \\ = w_n + 1 + \frac{3}{5} w_n^{-1} + \frac{2}{7} w_n^{-2} + \frac{3}{25} w_n^{-3} + \frac{18}{385} w_n^{-4} + \dots$$

Just as in our conclusions about (8.5.6), we infer that  $w_n^{-1} = O(n^{-1})$  and  $w_n = n + O(\log n)$ . Inserting this into (8.6.4), we obtain

$$w_{n+1} - w_n = 1 + \frac{3}{5} n^{-1} + O(n^{-2} \log n).$$

From this we infer, putting  $w_n = n + \frac{3}{5} \log n + t_n$ , that

$$t_{n+1} - t_n = O(n^{-2} \log n),$$

and it follows that  $t_n$  tends to a limit, to be denoted by  $C$ , and that

$$t_n = C + \sum_{k=n}^{\infty} (t_k - t_{k+1}) = C + O\left(\sum_n^{\infty} k^{-2} \log k\right) = \\ = C + O(n^{-1} \log n).$$

We shall write  $t_n = C + r_n$ , whence  $r_n = O(n^{-1} \log n)$ . Substituting  $w_n = n + \frac{3}{5} \log n + C + r_n$  into (8.6.4), we obtain

$$r_{n+1} - r_n = -\frac{9}{25} \frac{\log n}{n^2} + \frac{41 - 42C}{70n^2} + O\left(\frac{\log^2 n}{n^3}\right),$$

and we easily infer by summation that

$$r_n = \frac{9}{25} \frac{\log n}{n} + \frac{-79 + 210C}{350n} + O\left(\frac{\log^2 n}{n^2}\right).$$

This procedure can be continued, and it is possible to show that there exists an asymptotic series for  $r_n$ , of the form

$$\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} c_{kl} (n^{-1} \log n)^k n^{-l}.$$

It follows that  $n^{-1} w_n$  has a similar asymptotic series. Finally we get for  $x_n = 3^4 w_n^{-1}$  an asymptotic series, of which we produce a few

terms here:

$$(8.6.5) \quad x_n = \sin_n x_0 = \\ = 3^{\frac{1}{2}} n^{-\frac{1}{2}} \left\{ 1 - \frac{\frac{3}{10} \log n}{n} - \frac{C}{2n} + \frac{\alpha \log^2 n + \beta \log n + \gamma}{n^2} + O\left(\frac{\log^3 n}{n^3}\right) \right\},$$

where

$$\alpha = \frac{27}{200}, \quad \beta = \frac{9}{20}C - \frac{9}{50}, \quad \gamma = \frac{3}{8}C^2 - \frac{3}{10}C + \frac{79}{700}.$$

The value of  $C$  of course depends on  $x_0$ . It is remarkable that the first two terms of the asymptotic series are independent of  $x_0$ .

In order to find out something about how  $C$  depends on  $x_0$ , we replace  $C$  by  $\Psi(x_0)$  and we consider the following abbreviation of formula (8.6.5):

$$(8.6.6) \quad x_n = 3^{\frac{1}{2}} n^{-\frac{1}{2}} \left\{ 1 - \frac{\frac{3}{10} \log n}{n} - \frac{\Psi(x_0)}{2n} + O(n^{-2} \log^2 n) \right\}.$$

As  $x_{n+1} = \sin x_n = \sin_n (\sin x_0)$ , we have

$$x_{n+1} = 3^{\frac{1}{2}} n^{-\frac{1}{2}} \left\{ 1 - \frac{\frac{3}{10} \log n}{n} - \frac{\Psi(\sin x_0)}{2n} + O(n^{-2} \log^2 n) \right\},$$

and so

$$x_{n+1} - x_n = \frac{1}{2} \cdot 3^{\frac{1}{2}} n^{-\frac{3}{2}} \{ \Psi(x_0) - \Psi(\sin x_0) \} + O(n^{-5/2} \log^2 n).$$

On the other hand, we have

$$x_{n+1} - x_n = \sin x_n - x_n = -\frac{1}{6} x_n^3 + O(x_n^5) = \\ = -\frac{1}{2} \cdot 3^{\frac{1}{2}} n^{-\frac{3}{2}} + O(n^{-5/2} \log n).$$

Therefore, it results that  $\Psi$  satisfies the equation

$$(8.6.7) \quad \Psi(\sin x_0) - \Psi(x_0) = 1.$$

The function  $\Psi$  is uniquely determined by (8.6.6). In fact (8.6.6) implies that

$$\Psi(x_0) = \lim_{n \rightarrow \infty} \{ 2n - \frac{3}{5} \log n - 2x_n n^{3/2} 3^{-\frac{1}{2}} \}.$$

This formula shows that  $\Psi(x) = \Psi(\pi - x)$  ( $0 < x < \pi$ ), and that  $\Psi(x)$  is non-increasing in the interval  $0 < x < \frac{1}{2}\pi$ . (For, if  $x_0$  increases, and  $n$  is fixed, then  $x_n$  increases). Moreover it can be shown from (8.6.7) that  $\Psi(x) \rightarrow \infty$  if  $x \rightarrow 0$  ( $x > 0$ ).

### 8.7. An alternative method

Our second method for dealing with the asymptotic behaviour of  $\sin_n x_0$  has some analogy to the contents of sec. 8.3. It starts from the Schröder equation (8.3.4), or, what is slightly simpler in our case, the so-called Abel equation

$$(8.7.1) \quad \varphi\{f(x)\} - \varphi(x) = 1$$

(the connection between the two is expressed by the relation  $\omega(x) = \alpha_1^{\varphi(x)}$ ).

If  $\varphi$  is known, and if  $f$  satisfies (8.7.1), then the iterates of  $f$  are easily obtained. In analogy with (8.3.6) we find  $f_n$  by solving

$$(8.7.2) \quad \varphi\{f_n(x)\} = n + \varphi(x)$$

for  $f_n(x)$ . We shall give two examples.

If  $\varphi(y) = y^{-1}$ , then  $f(x)$  equals  $x/(1+x)$ , as follows from (8.7.1). Now  $f_n(x)$  can be found from (8.7.2):

$$f_n(x) = x/(1+nx),$$

and this can also be used for continuous iteration (see sec. 8.3).

Next we take  $\varphi(y) = c^{-1}y^{-2}$ , where  $c$  is a constant. This corresponds to the system of iterates:

$$(8.7.3) \quad f_n(x) = x(1+nCx^2)^{-\frac{1}{2}}.$$

In the following, however, we shall discuss (8.7.1) from a different point of view by considering  $f$  as given and  $\varphi$  as unknown. We again take  $f(x) = \sin x$ , and the Abel equation becomes

$$(8.7.4) \quad \psi(\sin x) - \psi(x) = 1.$$

A special solution was obtained at the end of sec. 8.6, but we shall not use this information here.

Restricting ourselves to the interval  $0 < x \leq \frac{1}{2}\pi$ , it is quite easy to describe the general solution of (8.7.4): Choose an arbitrary function  $\psi^*$  in the interval  $1 < x \leq \frac{1}{2}\pi$ , and take  $\psi = \psi^*$  in that interval. In the next interval  $\sin 1 < x \leq 1$  we take  $\psi$  such that (8.7.4) is satisfied. Next we define it in the interval  $\sin^2 1 < x \leq \sin 1$ , such that (8.7.4) again holds. Continuing this process indefinitely we obtain a solution for the interval  $0 < x \leq \frac{1}{2}\pi$ .

It is clear that the asymptotic behaviour (as  $x \rightarrow 0$ ) of this

solution  $\psi(x)$  can be described to a certain extent once the asymptotic behaviour of  $\sin_n x$  is known. We want to go the other way round, however, in the following order: (i) Find an explicit solution. (ii) Determine its asymptotic behaviour. (iii) Apply this to the asymptotic behaviour of  $\sin_n x$ .

Properly speaking we do not require, for our purpose, that the solution is explicit itself, but only that its asymptotic behaviour is explicitly known.

Instead of asking for elementary functions which approximate a solution of (8.7.4), we start with a simpler problem: we ask for elementary functions  $\psi$  for which  $\psi(\sin x) - \psi(x)$  tends rapidly to 1 as  $x \rightarrow 0$ . One of the first functions to try is a monomial  $\psi_1(x) = ax^{-b}$  ( $b > 0$ ). It gives

$$\begin{aligned}\psi_1(\sin x) - \psi_1(x) &= a(\sin x)^{-b} - ax^{-b} = \\ &= ax^{-b}\{(1 - \frac{1}{6}x^2 + \dots)^{-b} - 1\} = \frac{1}{6}abx^{-b+2} + \dots\end{aligned}$$

We want it to approximate 1, and this is achieved by taking  $b = 2$ ,  $a = 3$ . So  $\psi_1(x) = 3x^{-2}$ , and

$$\psi_1(\sin x) - \psi_1(x) = 1 + \frac{1}{5}x^2 + \frac{2}{63}x^4 + \frac{1}{225}x^6 + \dots$$

(the series on the right equals

$$12 \sum_0^\infty (-4)^k (2k+1) B_{2k+2} x^{2k} / (2k+2)! \text{, cf. (8.6.3).}$$

Next we want to modify  $\psi_1$  in order to compensate the term  $x^2/5$ . Therefore we try to find a function  $\chi_1$  such that the difference  $\chi_1(\sin x) - \chi_1(x)$  is approximately equal to  $-x^2/5$ . We notice that a monomial does not work now. It is worth while to try  $\log x$ , for if we replace the difference  $\chi_1(\sin x) - \chi_1(x)$  by  $(\sin x - x)\chi_1'(x)$ , we see that  $\chi_1'(x)$  should be approximately  $6/(5x)$ .

Actually, if we take  $\chi_1(x) = \frac{6}{5} \log x$ , we have

$$\chi_1(\sin x) - \chi_1(x) = \frac{6}{5} \log(\sin x/x) = -\frac{1}{5}x^2 - \frac{1}{150}x^4 + \dots$$

If we take  $\psi_2 = \psi_1 + \chi_1$ , we have

$$\psi_2(\sin x) - \psi_2(x) = 1 + p(x),$$

$$p(x) = (\frac{2}{63} - \frac{1}{150})x^4 + \dots = \frac{3}{\sin^2 x} - \frac{3}{x^2} - 1 + \frac{6}{5} \log \frac{\sin x}{x}.$$

Thus we can go on, by choosing  $\chi_2$  such that  $\chi_2(\sin x) - \chi_2(x)$  equals approximately  $-\dot{p}(x)$ , and so on. In this way we would still be constructing approximate solutions instead of approximations to a solution. Fortunately we are now able to indicate an *exact* solution of the equation  $\chi(\sin x) - \chi(x) = -\dot{p}(x)$ . This equation is satisfied by the function

$$(8.7.5) \quad \chi_2(x) = p(x) + p(\sin x) + p(\sin_2 x) + p(\sin_3 x) + \dots.$$

We only need to show that the series is convergent for all real values of  $x$ . This immediately follows from the formulas  $p(y) = O(y^4)$  ( $y \rightarrow 0$ ) and  $\sin_n x = O(n^{-1})$  ( $n \rightarrow \infty$ ), whence  $p(\sin_n x) = O(n^{-2})$  ( $n \rightarrow \infty$ ).

The formula  $\sin_n x = O(n^{-1})$  can be proved in the way we proved (8.5.3). We shall, however, indicate a short independent proof: If  $\epsilon$  is a positive constant, then the iterates of the function  $f(x) = x(1 + ex^2)^{-\frac{1}{2}}$  can be evaluated explicitly. By (8.7.3) we have  $f_n(x) = x(1 + nex^2)^{-\frac{1}{2}}$ . Furthermore it is not difficult to show that  $\epsilon$  can be chosen such that  $\sin x < f(x) < x$  ( $0 < x < \pi$ ). We then obtain  $\sin_2 x < f(\sin x) < f_2(x)$ ,  $\sin_3 x < f(\sin_2 x) < f(f_2(x)) = f_3(x)$ , etc. It follows that

$$(8.7.6) \quad \sin_n x < x(1 + nex^2)^{-\frac{1}{2}} \quad (0 < x < \pi, \quad n = 1, 2, 3, \dots),$$

and so  $\sin_n x = O(n^{-\frac{1}{2}})$ . The assumption  $0 < x < \pi$  is, of course, no essential restriction.

We just established the convergence of (8.7.5), but we shall need more, viz. an estimate for  $\chi_2(x)$  as  $x \rightarrow 0$ . We shall prove that

$$(8.7.7) \quad \chi_2(x) = O(x^2) \quad (0 < x \leq \frac{1}{2}\pi).$$

We established before that  $p(x) = O(x^4)$  ( $x > 0$ ), and that  $p(\sin_n x) = O(n^{-2})$  ( $x$  fixed,  $n \rightarrow \infty$ ). We can easily be more specific: There are absolute constants  $C_1$  and  $C_2$  such that

$$|p(x)| < C_1 x^4, \quad |p(\sin_n x)| < C_2 n^{-2} \quad (0 < x \leq \frac{1}{2}\pi, \quad n = 1, 2, \dots).$$

It follows that

$$\sum_{k=n+1}^{\infty} p(\sin_k x) < C_2 n^{-1}.$$

Furthermore we have, if  $0 < x \leq \frac{1}{2}\pi$ ,

$$\sum_{k=0}^n p(\sin_k x) < C_1 \sum_{k=0}^n (\sin_k x)^4 < C_1 \sum_{k=0}^n x^4 = C_1 n x^4.$$

So for  $0 < x \leq \frac{1}{2}\pi$  and for all values of  $n$  we have

$$|\chi_2(x)| < C_1 n x^4 + C_2 n^{-1}.$$

We now choose  $n$ : As  $x \leq \frac{1}{2}\pi$ , we have  $3x^{-2} > 1$ . Let  $n$  be the smallest integer exceeding  $3x^{-2}$ . Then  $3x^{-2} < n < 6x^{-2}$ , and so

$$|\chi_2(x)| < (6C_1 + \frac{1}{3}C_2)x^2,$$

which proves (8.7.7). A slightly simpler proof is obtained by using the inequality (8.7.6).

As  $\chi_2$ , given by (8.7.5), exactly satisfies  $\chi_2(\sin x) + \chi_2(x) = -p(x)$ , the function

$$(8.7.8) \quad \psi(x) = 3x^{-2} + \frac{6}{5} \log x + \chi_2(x)$$

is an exact solution of (8.7.4). We have the estimate (8.7.7) for  $\chi_2$ , but it is not difficult to see that there is an asymptotic series

$$(8.7.9) \quad \chi_2(x) \approx c_2 x^2 + c_4 x^4 + c_6 x^6 + \dots \quad (x \rightarrow 0).$$

The leading coefficient equals  $c_2 = 79/1050$ . We shall show the first step of the proof of (8.7.5); all further steps are analogous. We have

$$\chi_2(\sin x) - \chi_2(x) = -p(x) = \alpha x^4 + \dots, \text{ with } \alpha = -79/3150.$$

Furthermore

$$(\sin x)^k - x^k = -kx^{k+2}/6 + \dots \quad (k = 1, 2, \dots);$$

here we use the case  $k = 2$ . Now put  $\chi_3(x) = -3\alpha x^2 + \chi_3(x)$ , then  $\chi_3(\sin x) - \chi_3(x) = -q(x)$  is a power series starting with a term  $\beta x^6$ .

From (8.7.7) we derive the rough estimate  $\chi_3(x) = O(x^2)$ , and we infer that  $\chi_3(\sin_n x) \rightarrow 0$ , if  $n \rightarrow \infty$  ( $x$  fixed). It follows that

$$\begin{aligned} \chi_3(x) &= \sum_{n=1}^{\infty} \{\chi_3(\sin_{n-1} x) - \chi_3(\sin_n x)\} = \\ &= q(x) + q(\sin x) + q(\sin_2 x) + \dots \end{aligned}$$

It can now be derived, in analogy with the proof of (8.7.7), that  $\chi_3(x) = O(x^4)$ . So  $\chi_2(x) = -3\alpha x^2 + O(x^4)$ .

It should be noted that the functions  $p(x)$ ,  $q(x)$ , ... are power series, with a positive radius of convergence, but that it is very

improbable that the same thing could be said about  $\chi_2(x)$  (or about  $\chi_3(x), \dots$ ).

### 8.8. Final discussion about the iterated sine

The asymptotic information about  $\psi$ , obtained in sec. 8.7 (see (8.7.8) and (8.7.9)), will now be applied to the problem about  $\sin_n x$ . To this end we start from the formula

$$(8.8.1) \quad \psi(\sin_n x) = n + \psi(x),$$

which is a direct consequence of (8.7.4). We shall consider  $x$  as a fixed number, and  $n$  as a large integer. Since  $\psi$  is a given function, we can consider (8.8.1) as an equation for the unknown quantity  $\sin_n x$ . Therefore the question to express the asymptotic behaviour of the solution  $\sin_n x$  in terms of the parameter  $n$  is a problem of the type discussed in Ch. 2.

We replace the unknown  $\sin_n x$  by the single letter  $u$ . The equation can be written as

$$(8.8.2) \quad n + \psi(x) \approx 3u^{-2} + \frac{8}{3} \log u + \frac{79}{1650} u^2 + \dots \quad (n \rightarrow \infty).$$

In some respects the question is more difficult than the problems considered in Ch. 2. Firstly, the right-hand-side represents the function  $\psi(u)$  asymptotically, but probably not exactly, and  $\psi(u)$  is probably very singular at  $u=0$ . Secondly, we are not yet sure, that  $\psi$  is monotonic, and therefore the uniqueness of the solution of  $\psi(u) = n + \psi(x)$  is unproved as yet. (This refers to the question whether  $\sin_n x$  depends uniquely on  $\psi(x)$ ; it is of course trivial that  $\sin_n x$  depends uniquely on  $x$ ).

But fortunately we need not bother about the existence and uniqueness of the solution  $u$  of (8.8.2). For,  $u = \sin_n x$  is a well-defined quantity, and the thing we want to do is to obtain asymptotic information about  $u$  in the way we use to handle the solution of an equation. In other words,  $u$  is not virtually unknown, but the fact that links our problem with the problems of Ch. 2 is that the asymptotic behaviour of  $u$  is unknown.

To start with, we know that  $u \rightarrow 0$  (when  $n \rightarrow \infty$ ), and it easily follows from (8.8.2) that  $u \sim (3/n)^{\frac{1}{2}}$ . This will be used repeatedly for estimating rounding-off errors.

In order to eliminate the difficulty that the right-hand-side of

(8.8.2) represents  $\psi(u)$  only asymptotically, we break it off somewhere, after the second term say, and the error made this way is transported to the left-hand-side. At the same time we subtract on both sides  $\frac{6}{5} \log 3$  in order to simplify further calculations. So we have

$$(8.8.3) \quad n + \psi(x) - \rho - \frac{6}{5} \log 3 = 3u^{-2} + \frac{6}{5} \log(3^{-\frac{1}{4}}u).$$

Here  $\rho$  is equal to  $\psi(u) - 3u^{-2} - \frac{6}{5} \log u$ . So by (8.7.8) and (8.7.9) it is asymptotically

$$(8.8.4) \quad \rho \approx c_2 u^2 + c_4 u^4 + \dots \quad (u \rightarrow 0),$$

whence  $\rho = O(u^{-1}) \quad (n \rightarrow \infty)$ .

Replacing the left-hand-side of (8.8.3) by  $y$ , we have

$$y = 3u^{-2} + \frac{6}{5} \log(3^{-\frac{1}{4}}u),$$

and  $y = n + O(1)$ , for  $x$  is a fixed number. Putting  $u = (3/y)^{\frac{1}{4}}$  we get

$$(8.8.5) \quad v^{-2} = 1 + \frac{\frac{3}{4} \log y}{y} - \frac{6}{5y} \log v.$$

From  $u \sim (3/n)^{\frac{1}{4}}$ ,  $y \sim n$  we know that  $v \rightarrow 1$ , whence  $\log v = O(1)$ . Raising both sides of (8.8.5) to the power  $-\frac{1}{2}$ , we get

$$(8.8.6) \quad v = 1 - \frac{1}{2} \left\{ \frac{\frac{3}{4} \log y}{y} - \frac{6}{5y} \log v \right\} + O(y^{-2} \log^2 y),$$

and so  $v = 1 + O(y^{-1} \log y)$ ,  $\log v = O(y^{-1} \log y)$ .

Inserting this into (8.8.6) we get

$$v = 1 - \frac{\frac{3}{16}}{y} \frac{\log y}{y} + O(y^{-2} \log^2 y),$$

whence  $\log v = -\frac{3}{16} y^{-1} \log y + O(y^{-2} \log^2 y)$ . Now we again raise (8.8.5) to the power  $-\frac{1}{2}$ , and we develop it a little further, according to the formula

$$(1 + z)^{-\frac{1}{2}} = 1 - \frac{1}{2}z + \frac{3}{8}z^2 + O(z^3).$$

Then we obtain

$$v = 1 - \frac{1}{2} \left\{ \frac{\frac{3}{4} \log y}{y} - \frac{6}{5y} \log v \right\} + \frac{3}{8} \left( \frac{\frac{3}{4} \log y}{y} \right)^2 + O(y^{-3} \log^3 y).$$

Contenting ourselves with a third-order error, we thus have

$$(8.8.7) \quad v = 1 - \frac{3}{10} \frac{\log y}{y} + \frac{27}{200} \frac{\log^2 y}{y^2} - \frac{9}{50} \frac{\log y}{y^2} + O(y^{-3} \log^3 y).$$

It is not difficult to show that there exists an asymptotic series in terms of powers of  $y^{-1}$  and  $y^{-1} \log y$ . Actually, the equation (8.8.5) can be transformed into an equation of the type (2.4.6) (put  $v = e^{1/w}$ ), and it follows that the asymptotic series even converges.

We next replace, in (8.8.7),  $v$  by  $3^{-\frac{1}{2}} y^{\frac{1}{2}} u$ , and  $y$  by  $n + \psi(x) + \rho - \frac{3}{5} \log 3$ , and we use that  $\rho = O(n^{-1})$ . Then we get  $\sin_n x$  expressed in terms of  $n$ , with an error term  $O(n^{-5/2} \log^2 n)$ :

$$(8.8.8) \quad u = \sin_n x = (3/n)^{-\frac{1}{2}} \left\{ 1 - \frac{3}{10} n^{-1} \log n - \frac{1}{2} (\psi(x) - \frac{3}{5} \log 3) n^{-1} + O(n^{-2} \log^2 n) \right\} \quad (n \rightarrow \infty).$$

Comparing this to (8.6.6) we learn that there is a simple relation

$$\Psi(x) = \psi(x) - \frac{3}{5} \log 3$$

between the special solutions  $\Psi(x)$  (see sec. 8.6) and  $\psi(x)$  (see sec. 8.7) of the Abel equation (8.7.3). Incidentally this shows that  $\psi(x)$  is decreasing ( $0 < x < \pi/2$ ).

Refinements of (8.8.8) can be found as follows. On the strength of (8.8.8) we can improve our formula  $\rho = O(n^{-1})$ . Formula (8.8.4) gives

$$\rho = 3c_2 n^{-1} + O(n^{-2} \log n).$$

Inserting this into (8.8.7), with  $y = n + \Psi(x) - \rho$ , we obtain the next approximation to  $\sin_n x$ , already given in (8.6.5). It is clear that this process can be continued indefinitely, and it follows, for the second time, that there exists an asymptotic series.

Above it was remarked that the asymptotic series for  $v$ , of which (8.8.7) is a beginning segment, is convergent in the ordinary sense. There is no reason, however, to expect this to remain true for the series that expresses  $v$  asymptotically in terms of  $n$ . For in passing from  $y$  to  $n$  we use the series (8.8.4), and that series cannot be expected to be convergent in the ordinary sense.

## 8.9. An inequality concerning infinite series

In sec. 8.10 we shall solve an iteration problem belonging to the general class described in sec. 8.1. The most remarkable feature of

the problem of sec. 8.10 will be the fact that there is an unexpected discontinuity (with respect to the parameter  $x$ ) in the asymptotic behaviour.

As the problem will look quite artificial, we shall show in the present section how it originates from a series problem in a natural way.

If  $\lambda > 1$ ,  $a_n \geq 0$  ( $n = 1, 2, \dots$ ), then we have the inequality

$$\sum_1^{\infty} a_n \leq \lambda^{1/\lambda} \sum_{n=1}^{\infty} n^{-1/\lambda} (a_n^{\lambda} + a_{n+1}^{\lambda} + \dots)^{1/\lambda},$$

due to Copson<sup>1)</sup>. The constant on the right, viz.  $\lambda^{1/\lambda}$ , is not best possible, however. It will be our aim to discuss the question of the best possible constant in the special case  $\lambda = 2$ . Actually, we shall show that there is a constant  $\gamma$  ( $1 < \gamma \leq 2^{\frac{1}{2}}$ ) such that

$$(8.9.1) \quad \sum_1^{\infty} a_n \leq \gamma \sum_{n=1}^{\infty} n^{-\frac{1}{2}} (a_n^2 + a_{n+1}^2 + \dots)^{\frac{1}{2}}$$

for all sequences of non-negative  $a_n$ 's (provided that the series on the right converges), and such that there is a special convergent series  $\sum a_n$  (with positive terms) for which the equality sign holds. The method depends, of course, on application of Cauchy's inequality  $(\sum u_n v_n)^2 \leq \sum u_n^2 \cdot \sum v_n^2$ . Introducing positive numbers  $p_{kn}$ , we have

$$(8.9.2) \quad \sum_{k=n}^{\infty} a_k p_{kn} \leq \left( \sum_{k=n}^{\infty} a_k^2 \right)^{\frac{1}{2}} \cdot \left( \sum_{k=n}^{\infty} p_{kn}^2 \right)^{\frac{1}{2}},$$

whence

$$\sum_{k=1}^{\infty} a_k \sum_{n=1}^k p_{kn} = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} a_k p_{kn} \leq \sum_{n=1}^{\infty} \left( \sum_{k=n}^{\infty} a_k^2 \right)^{\frac{1}{2}} \left( \sum_{k=n}^{\infty} p_{kn}^2 \right)^{\frac{1}{2}}.$$

In order to get an inequality of the required type we want to choose the  $p_{kn}$  such that

$$\sum_{n=1}^k p_{kn} = 1 \quad (k = 1, 2, \dots), \quad \sum_{k=n}^{\infty} p_{kn}^2 = x^2 n^{-1} \quad (n = 1, 2, \dots),$$

where  $x$  is a positive number not depending on  $n$ . If the  $p_{kn}$ 's are chosen that way, then (8.9.1) holds, with  $x$  instead of with  $\gamma$ , and the next problem is to minimize  $x$ .

<sup>1)</sup> See G. H. HARDY, J. E. LITTLEWOOD and G. PÓLYA, Inequalities, Cambridge 1934, theorem 345.

We eventually want to have an inequality for which the equality sign holds for at least one series  $\sum \alpha_k$ . Therefore it will be a good idea to take the  $\phi_{kn}$ 's such that there is a possibility,  $\alpha_k = \alpha_k$ , say, to have the equality sign in (8.9.2) for all  $n$  simultaneously. This leads us to the restriction  $\phi_{kn} = \alpha_k \beta_n$ , where both  $\{\alpha_k\}$  and  $\{\beta_n\}$  are positive sequences. And these should satisfy, for  $n = 1, 2, \dots$ ,

$$\sum_1^n \beta_k = \alpha_n^{-1}, \quad \sum_n^\infty \alpha_k^2 = x^{2n-1} \beta_n^{-2}.$$

For each  $n$ , the second equation contains infinitely many variables. Therefore we put  $\sum_n^\infty \alpha_k^2 = \delta_n$ , and eliminating the  $\beta$ 's we get the equations

$$x \sum_1^n k^{-1} \delta_k^{-1} = (\delta_n - \delta_{n+1})^{-1} \quad (n = 1, 2, \dots),$$

with the supplementary conditions

$$\delta_1 > \delta_2 > \delta_3 > \dots > 0, \quad \lim_{n \rightarrow \infty} \delta_n = 0.$$

Taking differences, we get the equations

$$xn^{-1} = \left(1 - \frac{\delta_{n+1}}{\delta_n}\right)^{-1} - \left(\frac{\delta_{n-1}}{\delta_n} - 1\right)^{-1} \quad (n = 2, 3, \dots),$$

and  $x = (1 - \delta_2 \delta_1^{-1})^{-1}$ . We choose  $\delta_1 = 1$ , and  $\delta_2 = v_1, \delta_3 = v_1 v_2, \delta_4 = v_1 v_2 v_3, \dots$ . The  $v$ 's should lie between 0 and 1, and should satisfy

$$xn^{-1} = (1 - v_n)^{-1} - (v_{n-1}^{-1} - 1)^{-1} \quad (n = 2, 3, \dots),$$

where  $x = (1 - v_1)^{-1}$ . Finally writing  $1 - v_n = u_n^{-2}$ , we obtain the following system of equations:

$$(8.9.3) \quad \begin{cases} u_1 = x, \\ u_2 = 2^{-1}x + (u_1^2 - 1)^{\frac{1}{2}}, \\ u_3 = 3^{-1}x + (u_2^2 - 1)^{\frac{1}{2}}, \\ \dots \end{cases}$$

Now assume that  $x, u_1, u_2, \dots$  are positive numbers satisfying (8.9.3). Then we take

$$\begin{aligned} \phi_{kn} &= \alpha_k \beta_n, & \alpha_k &= u_k^{-1} (v_1 \dots v_{k-1})^{\frac{1}{2}}, \\ \beta_k &= x k^{-1} (v_1 \dots v_{k-1})^{-\frac{1}{2}}, & \delta_k &= v_1 \dots v_{k-1}, \end{aligned}$$

where  $v_k = 1 - u_k^{-2}$ , and  $v_1 \dots v_{k-1}$  denotes 1 if  $k = 1$ . Indeed it can be proved that  $\delta_k \rightarrow 0$  ( $k \rightarrow \infty$ ) (see sec. 8.10), and it follows that (8.9.1) is true (with  $x$  instead of  $\gamma$ ). In order to find a series  $\sum a_n$  for which the equality sign holds, we remark that we have equality in (8.9.2) if we take  $a_n = \alpha_n$ . So the problem is whether the series

$$\sum_1^{\infty} \alpha_k = \sum_1^{\infty} u_k^{-1} (v_1 \dots v_{k-1})^{\frac{1}{2}} \text{ and } \sum_1^{\infty} k^{-\frac{1}{2}} \delta_k = \sum_1^{\infty} k^{-\frac{1}{2}} (v_1 \dots v_{k-1})^{\frac{1}{2}}$$

converge. In sec. 8.10 we shall find out that they do converge if  $x$  has its minimal value  $\gamma$  (the existence of a minimal value is also proved in sec. 8.10). It does not happen, of course, if  $x > \gamma$ . For if (8.9.1) is true, it remains true if  $\gamma$  is enlarged to  $x$ , but then we loose the equality case.

### 8.10. The iteration problem

We shall investigate the solutions of the system (8.9.3). The word "solution" will indicate a sequence of positive numbers  $u_1, u_2, \dots$ , satisfying (8.9.3) for some given value of  $x$ . If  $x$  is given, there is at most one solution, since the  $u$ 's can be evaluated consecutively. We shall show:

- (i) There is a number  $\gamma$  ( $1 < \gamma \leq 2^{\frac{1}{2}}$ ) such that there is a solution if and only if  $x \geq \gamma$ .
  - (ii) For every  $x \geq \gamma$  the sequence  $\{n^{-\frac{1}{2}} u_n\}$  converges, and
- $$\lim n^{-\frac{1}{2}} u_n = x + (x^2 - 1)^{\frac{1}{2}} \quad \text{if } x \text{ fixed, } x > \gamma,$$
- $$\lim n^{-\frac{1}{2}} u_n = \gamma - (\gamma^2 - 1)^{\frac{1}{2}} \quad \text{if } x = \gamma.$$

Proof of (i): The system (8.9.3) has a solution if  $x \geq 2^{\frac{1}{2}}$ . For then we can prove by induction that  $u_1 \geq 2^{\frac{1}{2}}, u_2 \geq 3^{\frac{1}{2}}, u_3 \geq 4^{\frac{1}{2}}, \dots$ . This depends on the inequality

$$(n+1)^{\frac{1}{2}} < ((n^{\frac{1}{2}})^2 - 1)^{\frac{1}{2}} + n^{-\frac{1}{2}} 2^{\frac{1}{2}} \quad (n = 2, 3, \dots),$$

which follows from

$$(n+1)^{\frac{1}{2}} - (n-1)^{\frac{1}{2}} = \frac{1}{2} \int_{n-1}^{n+1} t^{-\frac{1}{2}} dt < (n-1)^{-\frac{1}{2}} \leq (\frac{1}{2}n)^{-\frac{1}{2}}.$$

There is not always a solution of (8.9.3) if  $0 < x < 2^{\frac{1}{2}}$ . If  $x$  is given, we can determine  $u_1, u_2, \dots$  consecutively. It may happen, however, that for a certain  $k$  we get  $u_k < 1$ , and then either  $u_k$

or  $u_{k+1}$  is not positive. We can take  $u_1 = u_1(x) = x$  if  $x > 0$ . So there is a positive solution for both  $u_1$  and  $u_2$  if and only if  $x \geq 1$ . The function  $u_2(x)$  is continuous and increasing if  $x \geq 1$ , and since  $u_2(1) = 2^{-\frac{1}{2}}$ , its value at  $x = 1$  is  $< 1$ . On the other hand  $u_2(x) \rightarrow \infty$  ( $x \rightarrow \infty$ ), since  $u_2(x) \geq 2^{-\frac{1}{2}}x$ . So there is a number  $c_2 > 1$  with  $u_2(c_2) = 1$ . Hence there is a positive solution for  $u_1, u_2, u_3$  if and only if  $x \geq c_2$ . Again,  $u_3(x)$  is continuous and increasing if  $x \geq c_2$ , and  $u_3(c_2) = 3^{-\frac{1}{2}}c_2 < 1$  (remember that  $u_n \geq (n+1)^{-\frac{1}{2}}$  if  $x \geq 2^{\frac{1}{2}}$ , whence  $c_2 < 2^{\frac{1}{2}}$ ). We thus find a sequence  $c_2, c_3, \dots$ , with

$$1 < c_2 < c_3 < \dots < 2^{\frac{1}{2}},$$

and the infinite system has a positive solution if and only if  $x \geq c_k$  for all  $k$ . That means, if and only if  $x \geq \gamma$ , where  $\gamma = \lim c_k$ . Moreover, we find  $1 < \gamma \leq 2^{\frac{1}{2}}$ .

Proof of (ii): If  $x \geq \gamma$ , we deduce from (8.9.3) that

$$u_n(x) \leq n^{-\frac{1}{2}}x + u_{n-1}(x) \quad (n = 2, 3, \dots),$$

and therefore that  $u_n(x) \leq x(1 + 2^{-\frac{1}{2}} + \dots + n^{-\frac{1}{2}})$ . It follows that

$$(8.10.1) \quad \limsup_{n \rightarrow \infty} n^{-\frac{1}{2}}u_n \leq 2x \quad (x \geq \gamma).$$

In the other direction we can prove

$$(8.10.2) \quad \liminf_{n \rightarrow \infty} n^{-\frac{1}{2}}u_n \geq (2x)^{-1} \quad (x \geq \gamma).$$

Assume that this is false for a certain  $x \geq \gamma$ . Then we can find a number  $c$  ( $0 < c < (2x)^{-1}$ ) such that  $u_n(x) < cn^{\frac{1}{2}}$  for infinitely many values of  $n$ . So there is a number  $m$  such that

$$(8.10.3) \quad u_n < cn^{\frac{1}{2}}, \quad n > 2x^2(1 - 2cx)^{-1}$$

holds for  $n = m$ . We can now show by induction that (8.10.3) holds for all  $n \geq m$ . The induction step is:

$$\begin{aligned} u_{n+1}^2 &= u_n^2 - 1 + (n+1)^{-1}x^2 + 2x(n+1)^{-\frac{1}{2}}(u_n^2 - 1)^{\frac{1}{2}} < \\ &< u_n^2 - 1 + (n+1)^{-1}x^2 + 2xc < u_n^2 - \frac{1}{2}(1 - 2cx) < c(n+1)^{\frac{1}{2}}. \end{aligned}$$

At the same time we have proved that  $u_{n+1}^2 - u_n^2$  is less than the negative constant  $\frac{1}{2}(2cx - 1)$  if  $n \geq m$ . It follows that  $u_n^2 \rightarrow -\infty$ . As this is contradictory, we have proved (8.10.2).

For convenience we put  $u_n = n^{\frac{1}{2}}s_n$ . So if  $x$  is fixed ( $x \geq \gamma$ ), we

have  $s_n = O(1)$ ,  $s_n^{-1} = O(1)$ . The relation between  $s_n$  and  $s_{n+1}$  can be written as

$$s_{n+1} = (n+1)^{-1}x + s_n \left(1 - \frac{s_n^{-2} + 1}{n+1}\right)^{\frac{1}{2}} \quad (n = 1, 2, \dots).$$

As  $s_n$  and  $s_n^{-1}$  are bounded, we infer that

$$(8.10.4) \quad s_{n+1} - s_n = n^{-1}(x - \frac{1}{2}s_n - \frac{1}{2}s_n^{-1}) + O(n^{-2}) \quad (n \rightarrow \infty).$$

The sign of  $x - \frac{1}{2}s_n - \frac{1}{2}s_n^{-1}$  will be important. Let  $\lambda_1$  and  $\lambda_2$  denote the roots of the equation  $\lambda^2 - 2\lambda x + 1 = 0$ :

$$\lambda_1 = x - (x^2 - 1)^{\frac{1}{2}}, \quad \lambda_2 = x + (x^2 - 1)^{\frac{1}{2}},$$

whence  $(2x)^{-1} < \lambda_1 < 1 < \lambda_2 < 2x$ . Now we have  $x - \frac{1}{2}s - \frac{1}{2}s^{-1} = -\frac{1}{2}s^{-1}(s - \lambda_1)(s - \lambda_2)$ . And (8.10.4) expresses, roughly speaking, that if  $s_n$  lies to the right of  $\lambda_1$ , then it will move (when  $n$  increases) into the direction of  $\lambda_2$ , and if  $s_n$  lies to the left of  $\lambda_1$ , then it will move farther away to the left. Therefore we expect that either  $s_n \rightarrow \lambda_1$  or  $s_n \rightarrow \lambda_2$ . Indeed we shall prove that the assumption

$$(8.10.5) \quad \limsup s_n > \lambda_1$$

leads to  $s_n \rightarrow \lambda_2$ , whereas the assumption

$$(8.10.6) \quad \limsup s_n \leq \lambda_1$$

leads to  $s_n \rightarrow \lambda_1$ . It will depend on  $x$ , which one of these two alternatives actually occurs.

First assume (8.10.5). We write (8.10.4) in the form

$$(8.10.7) \quad |(s_{n+1} - \lambda_2) - (s_n - \lambda_2)(1 - n^{-1}\sigma_n)| < An^{-2},$$

where  $A$  is a constant, and

$$\sigma_n = \frac{1}{2}s_n^{-1}(s_n - \lambda_1).$$

It follows from (8.10.5) that there is a number  $\beta$  with  $\lambda_1 < \beta < \lambda_2$ , such that  $s_n > \beta$  infinitely often. We shall deduce that  $s_n > \beta$  for all large values of  $n$ . To this end we shall show the existence of a number  $n_0$  with the property that  $n > n_0$ ,  $s_n > \beta$  imply  $s_{n+1} > \beta$ .

If  $s_n \geq \lambda_2$ , we derive from (8.10.7) that  $s_{n+1} \geq \lambda_2 - An^{-2}$ , and if  $\beta < s_n < \lambda_2$  we infer that  $\sigma_n > 0$ , and

$$s_{n+1} > \beta + (\lambda_2 - \beta)n^{-1}\sigma_n - An^{-2}.$$

We have  $s_n^{-1} = O(1)$ , so  $s_n^{-1} \leq B$ , where  $B$  is a constant. It follows that  $\sigma_n > \frac{1}{2}B^{-1}(\beta - \lambda_1)$ , whence

$$s_{n+1} > \beta + \frac{1}{2}(\lambda_2 - \beta)(\beta - \lambda_1)B^{-1}n^{-1} - An^{-2}.$$

So in order to meet all requirements we take  $n_0$  such that

$$An_0^{-2} < \lambda_2 - \beta, \text{ and } An_0^{-1} < \frac{1}{2}(\lambda_2 - \beta)(\beta - \lambda_1)B^{-1}.$$

Then  $n > n_0$ ,  $s_n > \beta$  imply  $s_{n+1} > \beta$ . It follows that  $s_{n+1} > \beta$ ,  $s_{n+2} > \beta$ , ... .

So from (8.10.5) we deduced that  $s_n > \beta$  for all large  $n$  (for some  $\beta$ ,  $\lambda_1 < \beta < \lambda_2$ ), and from this we infer that there is a positive constant  $c$  such that  $\sigma_n > c$  for all  $n$ . Putting  $s_n - \lambda_2 = w_n$ , we obtain from (8.10.7) that

$$(8.10.8) \quad |w_{n+1}| \leq (1 - cn^{-1})|w_n| + An^{-2} \quad (n = 1, 2, \dots).$$

We may of course assume  $0 < c < 1$ , whence  $0 < 1 - cn^{-1} < 1$ . By repeated application of (8.10.8) we get

$$\begin{aligned} |w_{n+2}| &\leq (1 - c(n+1)^{-1})(1 - cn^{-1})|w_n| + An^{-2} + A(n+1)^{-2}, \\ |w_{n+3}| &\leq (1 - c(n+2)^{-1})(1 - c(n+1)^{-1})(1 - cn^{-1})|w_n| + \\ &\quad + An^{-2} + A(n+1)^{-2} + A(n+2)^{-2}, \end{aligned}$$

and so on. As the infinite product  $\prod_n^\infty (1 - ck)^{-1}$  diverges to zero, we obtain that

$$\limsup_{p \rightarrow \infty} |w_{n+p}| \leq A \sum_{k=n}^{\infty} k^{-2} < A(n-1)^{-1}.$$

Now making  $n \rightarrow \infty$ , we infer that the left-hand-side is zero. It follows that  $w_n \rightarrow 0$ , so on the assumption (8.10.5) we have deduced that  $s_n \rightarrow \lambda_2$ .

Next we assume (8.10.6). Moreover, we assume that  $s_n \rightarrow \lambda_1$  is false. Then we can find a number  $\mu$  ( $0 < \mu < \lambda_1$ ) such that  $s_n < \mu$  infinitely often. As in the previous case, but interchanging the rôles of  $\lambda_1$  and  $\lambda_2$ , we can show that  $s_n < \mu$  for all large values of  $n$ . Putting  $\lambda_1 - s_n = z_n$ , we get an inequality of the type

$$z_{n+1} \geq (1 + cn^{-1})z_n - An^{-2} \quad (n = 1, 2, \dots),$$

with a positive constant  $c$ . Instead of following the method applied

to (8.10.8), we now show a simple trick. We derive that

$$z_{n+1} - c^{-1}A(n+1)^{-\frac{1}{2}} \geq (1 + cn^{-\frac{1}{2}})(z_n - c^{-1}An^{-\frac{1}{2}}),$$

whence  $z_n - c^{-1}An^{-\frac{1}{2}} \rightarrow \infty$ . This is contradictory, as  $z_n = \lambda_1 - s_n < \lambda_1$ .

So far we have proved that  $s_n = n^{-\frac{1}{2}}u_n = n^{-\frac{1}{2}}u_n(x)$  converges if  $x \geq \gamma$  and  $n \rightarrow \infty$ , either to  $x - (x^2 - 1)^{\frac{1}{2}}$  or to  $x + (x^2 - 1)^{\frac{1}{2}}$ . It is easy to see that the limit is  $x + (x^2 - 1)^{\frac{1}{2}}$  if  $x > \gamma$ . For, every  $u_n(x)$  is an increasing function of  $x$ , whence  $\lim n^{-\frac{1}{2}}u_n(x)$  is a non-decreasing function of  $x$ . If the limit were  $x_1 - (x_1^2 - 1)^{\frac{1}{2}}$  at a point  $x_1 > \gamma$ , it would be less than  $x_1 - (x_1^2 - 1)^{\frac{1}{2}}$  at  $x = \gamma$ . But this cannot be true, both  $\gamma - (\gamma^2 - 1)^{\frac{1}{2}}$  and  $\gamma + (\gamma^2 - 1)^{\frac{1}{2}}$  being larger.

The remaining difficulty is to decide the value of the limit at  $x = \gamma$ . We, of course, have to use the minimum property of  $\gamma$ . We shall show that if the limit would be  $\gamma + (\gamma^2 - 1)^{\frac{1}{2}}$  at  $x = \gamma$ , we should be able to find a number  $\gamma' < \gamma$  such that the system (8.9.3) still has a positive solution at  $x = \gamma'$ .

So we assume that  $u_n(\gamma) \sim n^{\frac{1}{2}}(\gamma + (\gamma^2 - 1)^{\frac{1}{2}})$ . It follows that  $u_n(\gamma) > n^{\frac{1}{2}}$  for all large values of  $n$ . Put  $\gamma'' = \frac{1}{2}(1 + \gamma)$  (so  $1 < \gamma'' < \gamma$ ), and take  $N$  so large that

$$u_N(\gamma) > N^{\frac{1}{2}}, \quad (n+1)^{\frac{1}{2}} - (n-1)^{\frac{1}{2}} < \gamma''n^{-\frac{1}{2}} \quad (n \geq N)$$

(notice that  $(n+1)^{\frac{1}{2}} - (n-1)^{\frac{1}{2}} \sim n^{-\frac{1}{2}}$ ). The functions  $u_1(x)$ ,  $u_2(x)$ , ...,  $u_N(x)$  are continuous for  $x \geq c_{N-1}$  (See page 170). Hence we can fix  $\gamma'$  such that

$$c_{N-1} < \gamma' < \gamma, \quad \gamma'' < \gamma' < \gamma, \quad u_N(\gamma') > N^{\frac{1}{2}}.$$

Then we can prove by induction that  $u_n(\gamma') > n^{\frac{1}{2}}$  ( $n = N, N+1, N+2, \dots$ ). The induction step is produced by the formula

$$u_{n+1}(\gamma') > (n-1)^{\frac{1}{2}} + n^{-\frac{1}{2}}\gamma' > (n-1)^{\frac{1}{2}} + n^{-\frac{1}{2}}\gamma'' > (n+1)^{\frac{1}{2}}.$$

Thus we have proved that (8.9.3) still has a solution at  $x = \gamma'$ , in spite of the minimum property of  $\gamma$ . So the case  $\gamma + (\gamma^2 - 1)^{\frac{1}{2}}$  is ruled out, and it follows that  $n^{-\frac{1}{2}}u_n(\gamma) \rightarrow \gamma - (\gamma^2 - 1)^{\frac{1}{2}}$ .

This completes the proof of (ii).

We finally return to the questions left open at the end of sec. 8.9.

In the first place it has to be shown that  $v_1 \dots v_{k-1} \rightarrow 0$  ( $x \geq \gamma$ ,  $n \rightarrow \infty$ ).

We have  $v_n = 1 - u_n^{-2}$ , and  $u_n^{-2} > cn^{-1}$  ( $c$  not depending on  $n$ , see (8.10.2)). Therefore, the infinite product  $\prod(1 - u_n^{-2})$  diverges to zero, whence  $v_1 \dots v_{k-1} \rightarrow 0$ .

Secondly, we have to show the convergence of  $\sum_1^\infty u_k^{-1}(v_1 \dots v_{k-1})^{\frac{1}{k}}$  and of  $\sum_1^\infty k^{-\frac{1}{k}}(v_1 \dots v_{k-1})^{\frac{1}{k}}$ , in the case that  $x = \gamma$ . As  $u_n^{-1} = O(n^{-\frac{1}{2}})$ , it suffices to consider the latter series. We have  $u_n^{-2} \sim \rho n^{-1}$ , with  $\rho = (\gamma - (\gamma^2 - 1)^{\frac{1}{2}})^{-2}$ . Since  $\rho > 1$ , we can find a constant  $\eta > 1$  such that  $u_n^{-2} > \eta n^{-1}$  for all large  $n$ .

It follows that

$$v_1 \dots v_{k-1} = O(\prod_1^{k-1}(1 - \eta n^{-1})) = O(k^{-\eta}).$$

As  $\frac{1}{2}(1 + \eta) > 1$ , we find that  $\sum k^{-\frac{1}{k}}(v_1 \dots v_{k-1})^{\frac{1}{k}}$  converges.

From a discussion of the system (8.9.3) it was shown before<sup>1)</sup> that the best possible constant  $\gamma$  for Copson's inequality lies between 1.08 and 1.13334.

Much better numerical results can be obtained, if we study the asymptotic behaviour of  $u_n(\gamma)$  in some detail. It can be shown that we have, for  $n = 1, 2, 3, \dots$ ,

$$n^{-\frac{1}{k}}u_n(\gamma) = \lambda_1 + \frac{1}{2}\gamma n^{-1} - \frac{1}{8}\gamma^2(\lambda_1 + \gamma)^{-1}n^{-2} + \theta_n n^{-4},$$

with  $0 < \theta_n < 2^{-4}(\lambda_1 + \gamma)^{-2}\gamma^3$ . This can be proved by showing that (i) if some  $\theta_n$  lies outside this interval, then all further  $\theta_n$  lie outside, and (ii)  $\theta_n$  tends to the limit

$$2^{-7}\gamma^3(5\gamma + 8\lambda_1)(\gamma + \lambda_1)^{-2}(\gamma + 3\lambda_1)^{-1},$$

which lies inside the interval. This formula was taken as a basis for numerical computations, made by the electronic computer ARMAC at the Mathematical Centre, Amsterdam. Using values of  $n$  up to  $n = 50$ , the ARMAC proved that

$$4752363148 < 2^{32}\gamma < 4752363155.$$

Therefore  $\gamma = 1.1064957714$ , with an error of at most 9 units at the last decimal place.

<sup>1)</sup> Wiskundige Opgaven, vol. 20 (3) (1957), solution of problem nr. 83 (problem by R. P. BOAS, solution by N. G. DE BRUIJN).

For this numerical work on the ARMAC the author is indebted to Prof. A. van Wijngaarden and his staff.

8.11. EXERCISES. 1. Let  $k$  be an integer  $> 1$ , and let, in some interval  $0 < x < \delta$ , the real function  $f$  be the sum of a convergent power series

$$f(x) = x + a_k x^k + a_{k+1} x^{k+1} + \dots,$$

with  $a_k < 0$ . Show that there is an interval  $0 < x < \delta_1$  such that for any  $x$  of this interval the  $n$ -th iterate  $f_n(x)$  is asymptotically equivalent to  $((1 - k)a_k n)^{-1/(k-1)}$ . Furthermore, show that there is, in some interval  $0 < x < \delta_2$ , a solution  $\psi$  of the equation  $\psi(f(x)) - \psi(x) = 1$ , with

$$\psi(x) \approx c_{1-k} x^{1-k} + \dots + c_{-1} x^{-1} + c_0 \log x + c_1 x + c_2 x^2 + \dots \quad (x > 0, x \rightarrow 0).$$

The coefficients  $c_{1-k}, \dots, c_0$  are such that

$$\psi'(x) - \frac{1}{2}(1 + f'(x))(f(x) - x)^{-1} = O(1) \quad (x > 0, x \rightarrow 0).$$

More precise results about  $f_n(x)$  can be obtained from the relation

$$\psi(f_n(x)) = n + \psi(x).$$

2. Let  $P(x, y)$  be a double power series

$$P(x, y) = \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} c_{ij} x^i y^j,$$

convergent for sufficiently small values of  $|x|$  and  $|y|$ , with real coefficients, and assume that  $c_{10} = 0$ ,  $c_{11} > 0$ . Let  $\{w_n\}$  be a sequence satisfying  $w_n = o(1)$  ( $n \rightarrow \infty$ ) and  $w_{n+1} - w_n = P(n^{-1}, w_n)$  for all large values of  $n$ . Show that there is an asymptotic development

$$w_n \approx a_1 \cdot n^{-1} + a_2 \cdot n^{-2} + \dots \quad (n \rightarrow \infty).$$

3. Let  $x$  be the minimal number such that the system (8.9.3) has a solution in positive numbers  $u_1, u_2, u_3, \dots$ . Show that there is an asymptotic expansion

$$u_n \approx n^{\frac{1}{k}}(c_0 + c_1 n^{-1} + c_2 n^{-2} + \dots) \quad (n \rightarrow \infty).$$

Show also that  $c_3 = c_5 = c_7 = \dots = 0$ .

## CHAPTER 9

### DIFFERENTIAL EQUATIONS

#### 9.1. Introduction

Many problems in pure and applied mathematics are dealing with the behaviour of the solutions of a differential equation at a singular point. Such problems are obviously of asymptotic nature, for by a transformation of the independent variable it is always possible to transform the singularity to infinity, and after that, the question takes the following form: Let  $F(t, y, y', \dots, y^{(m)}) = 0$  be a given differential equation for the unknown function  $y = y(t)$ . How do the solutions behave as  $t \rightarrow \infty$ ?

Such questions arise, for instance, in stability problems, in problems about linear or non-linear oscillations and in quantum mechanics. Another type of applications is given by the fact that one can study the asymptotic behaviour of a given function, a Bessel function, say, from the differential equation which it satisfies, instead of from one of the explicit expressions for that function.

There is a wide variety of problems in this field, and there is an extensive literature about it<sup>1)</sup>. It is out of the question that an adequate survey could be given within the scope of this book. Nevertheless, in the few problems we shall discuss here, several ideas appear which can be applied in many other cases.

Problems on differential equations are usually very flexible owing to the possibility of transforming both the dependent and the independent variable. After such a transformation has been carried out, the problem usually looks different.

Another general trend in the asymptotics of differential equations is the following one. Usually it is quite easy to guess an asymptotic

---

<sup>1)</sup> For references, see A. ERDÉLYI, Asymptotic Expansions, Dover Publ. 1956.

formula, or even an asymptotic series, and more often than not it is less easy to prove that it is an asymptotic formula indeed.

If we have to prove a certain asymptotic formula for a certain solution of a differential equation, then the obvious thing to try is to enclose this solution between functions whose asymptotic behaviour is known. In many cases such inequalities can be derived by simple theorems of the following type:

*Let  $y(t)$  be a solution of the first-order differential equation  $y' = F(t, y)$  ( $a \leq t \leq b$ ), and let  $\varphi(t)$  be a function satisfying*

$$\varphi'(t) < F(t, \varphi(t)) \quad (a \leq t \leq b), \quad \varphi(a) \leq y(a).$$

*Then we have  $\varphi(t) < y(t)$  ( $a < t \leq b$ ). (So in order to prove an inequality for  $y(t)$  it is not necessary to solve the differential equation explicitly, as functions  $\varphi$  satisfying  $\varphi' < F(t, \varphi)$  are of course quite easy to find).*

**Ptoof:** We first show that there is an interval  $(a, a + \varepsilon)$  where  $\varphi(t) < y(t)$ . This is trivial if  $\varphi(a) < y(a)$ . If  $\varphi(a) = y(a)$ , we have

$$\varphi'(a) < F(t, \varphi(a)) = F(t, y(a)) = y'(a),$$

and the existence of such an interval follows again. Suppose that the inequality  $\varphi(t) < y(t)$  can not be continued over the whole interval  $a < t \leq b$ . Then there exists a number  $c$  ( $a < c \leq b$ ) such that  $\varphi(c) = y(c)$ , and  $\varphi(t) < y(t)$  ( $a < t < c$ ). This implies  $\varphi'(c) \geq y'(c)$ , but now  $\varphi'(c) < F(c, \varphi(c)) = F(c, y(c)) = y'(c)$  leads to a contradiction.

The question whether in the above theorem the signs  $<$  can be replaced by  $\leq$ , is not very important for our purpose. It depends on more delicate considerations as to the uniqueness of solutions. For, if  $\varphi$  and  $y$  were different solutions of the differential equation, both having the same value at  $t = a$ , then  $\varphi(t) \leq F(t, \varphi(t))$  would still be valid, whereas  $\varphi(t) \leq y(t)$  ( $a < t \leq b$ ) would not be necessarily true.

## 9.2. A Riccati equation

Let  $\alpha(t)$ ,  $\beta(t)$ ,  $\gamma(t)$  be continuous real functions for  $t > 0$ , and let  $k$  be an integer  $\geq 0$ . We consider the differential equation

$$(9.2.1) \quad t^{-k} \rho'(t) = \alpha(t) + \beta(t)\rho(t) + \gamma(t)\rho^2(t)$$

for the unknown function  $\rho(t)$ . (We choose Greek letters for the functions  $\alpha, \beta, \gamma, \rho$  in order to be able to use the corresponding Latin letters for the coefficients of their asymptotic expansions).

Everything we shall do with (9.2.1) can be extended to equations of the type

$$t^{-k} \rho'(t) = \alpha_0(t) + \alpha_1(t)\rho(t) + \dots + \alpha_m(t)\rho^m(t),$$

but as the case  $m > 2$  hardly ever occurs in practical problems, we shall restrict ourselves to (9.2.1).

As to the existence of solutions, the state of affairs is only slightly less simple than with linear equations. Since the equation can be written as  $\rho'(t) = F(t, \rho)$ , where  $F$  is a continuous function in both variables as long as  $t > 0$ , with a continuous derivative  $\partial F / \partial \rho$ , the following theorem holds.

*Let a positive number  $t_0 > 0$  and a real number  $\rho_0$  be given. Then these numbers uniquely define an interval  $J$ , which is either the infinite interval  $t_0 \leq t < \infty$  or a finite interval  $t_0 \leq t < t_1$ . There is a solution  $\rho(t)$ , defined on  $J$ , and satisfying the differential equation, with  $\rho(t_0) = \rho_0$ . If  $J$  is the finite interval  $t_0 \leq t < t_1$ , then  $\rho(t)$  tends either to  $+\infty$  or to  $-\infty$  if  $t$  tends to  $t_1$  from the left.*

*The solution  $\rho(t)$  is uniquely determined: If  $J^*$  is any other interval with left end-point  $t_0$ , and if  $\rho^*$  is a solution throughout  $J^*$ , with  $\rho^*(t_0) = \rho_0$ , then  $J^*$  is contained in  $J$ , and  $\rho^* = \rho$  on  $J^*$ .*

**Proof:** Throughout this proof the word “solution” will be restricted to solutions of the differential equation  $\rho'(t) = F(t, \rho)$  which satisfy the initial condition  $\rho(t_0) = \rho_0$ .

It follows from the fundamental existence and uniqueness theorem<sup>1)</sup> that there is a maximal interval  $J$  (with left end-point  $t_0$ ) carrying a solution  $\rho(t)$ . For, in the first place, the fundamental theorem asserts that there is a number  $b > t_0$  such that there is a solution for  $t_0 \leq t < b$ . Secondly, if  $b_1 < b_2$ , and if both  $t_0 \leq t < b_1$  and  $t_0 \leq t < b_2$  carry a solution, then these solutions coincide in the common interval  $t_0 \leq t < b_1$ . If not, we could obtain a contradiction by application of the uniqueness theorem at a point  $c$ , where  $c$

<sup>1)</sup> See E. L. INCE, Ordinary Differential Equations, 1st ed., New York, § 3.1.

is the maximal number such that both solutions coincide in the interval  $t_0 \leq t \leq c$ .

Now the set of numbers  $b$  which have the property that the interval  $t_0 \leq t < b$  carries a solution, is either unbounded or bounded. In the first case we denote the interval  $t_0 \leq t < \infty$  by  $J$ . In the second case we denote the least upper bound of the set of all possible  $b$ 's by  $t_1$ , and then  $J$  stands for the interval  $t_0 \leq t < t_1$ . From the arguments above it follows that in both cases  $J$  carries a uniquely determined solution, and in the case of finite  $J$  we can assert that there is no solution in the closed interval  $t_0 \leq t \leq t_1$ .

It remains to be shown that, in the case of finite  $J$ , the solution  $\rho(t)$  tends either to  $+\infty$  or to  $-\infty$  if  $t$  tends to  $t_1$  from the left.

If  $\rho$  is bounded over  $t_0 \leq t < t_1$ , then  $F(t, \rho(t))$  is bounded, so  $\rho'(t)$  is bounded. It follows that  $\int_{t_0}^{t_1} \rho'(t) dt$  converges, so  $\rho(t)$  tends to a limit  $\rho_1$  if  $t$  tends to  $t_1$  from the left. As  $\rho'(t) = F(t, \rho(t))$ , the function  $\rho'(t)$  tends to  $F(t_1, \rho(t_1))$ , and it follows that the closed interval  $t_0 \leq t \leq t_1$  carries a solution. This being contradictory,  $\rho$  is unbounded over  $J$ .

Now assume that  $\rho(t)$  is unbounded over  $J$ , and neither tends to  $+\infty$ , nor to  $-\infty$ . Then it is not difficult to show that there are numbers  $A, B, x_1, x_2, \dots, y_1, y_2, \dots$  satisfying

$$A < B, \quad t_0 < x_1 < y_1 < x_2 < y_2 < \dots < t_1,$$

such that for  $n = 1, 2, 3, \dots$  we have

$$A \leq \rho(t) \leq B \quad (x_n < t < y_n), \quad \rho(x_n) = A, \quad \rho(y_n) = B.$$

The function  $F(t, \rho)$  is continuous in the closed rectangle  $t_0 \leq t \leq t_1$ ,  $A \leq \rho \leq B$  of the  $t$ - $\rho$ -plane. Therefore  $F(t, \rho)$  is bounded above by some constant  $c$ , as far as that rectangle is concerned. It follows that

$$B - A = \int_{x_n}^{y_n} \rho'(t) dt \leq c(y_n - x_n).$$

This contradicts the fact that  $\sum_{n=1}^{\infty} (y_n - x_n)$  converges (its sum is less than  $t_1 - t_0$ ), and so the proof of the theorem is complete.

We especially emphasize the fact that the length of the interval  $J$  may depend on the initial value  $\rho_0$ .

The above theorem will be repeatedly used in the following way. If we have numbers  $t_0, \rho_0$  and  $A$  ( $A > 0$ ), and if we have proved

that for no value of  $t_1$  ( $t_1 > t_0$ ) there exists a solution  $\rho(t)$  ( $t_0 \leq t \leq t_1$ ) with  $\rho(t_0) = \rho_0$ ,  $|\rho(t_1)| > A$ , then we know that there is a solution  $\rho(t)$  ( $t_0 \leq t < \infty$ ) with the initial condition  $\rho(t_0) = \rho_0$  (and satisfying  $|\rho(t)| \leq A$  ( $t_0 \leq t < \infty$ )).

A clear idea about the existence of solutions of a Riccati equation, and of their singularities, can be obtained from their relation to linear second-order equations. For example, the equation  $\rho' + \rho^2 = \alpha(t) + \rho\beta(t)$  is related to  $y'' - \beta(t)y' - \alpha(t)y = 0$  by the substitution  $y'/y = \rho$ . If  $y(t_1) = 0$  for some solution  $y$ , then the corresponding function  $\rho$  has a singularity at  $t = t_1$ : it behaves like  $(t - t_1)^{-1}$  if  $t$  tends to  $t_1$ . The existence and location of roots  $t_1$ , however, obviously depend on the choice of the ratio  $A:B$  of the integration constants in the general solution  $y(t) = A y_1(t) + B y_2(t)$ .

Returning to the equation (9.2.1), we shall introduce the further assumption that the asymptotic behaviour of each of the coefficients  $\alpha$ ,  $\beta$ ,  $\gamma$  is given by an asymptotic series:

$$(9.2.2) \quad \begin{cases} \alpha(t) \approx a_0 + a_1 t^{-1} + a_2 t^{-2} + \dots & (t \rightarrow \infty), \\ \beta(t) \approx b_0 + b_1 t^{-1} + b_2 t^{-2} + \dots & (t \rightarrow \infty), \\ \gamma(t) \approx c_0 + c_1 t^{-1} + c_2 t^{-2} + \dots & (t \rightarrow \infty), \end{cases}$$

and we assume that

$$(9.2.3) \quad b_0 < 0, \quad c_0 = 0.$$

It is our aim to prove that there is a class of solutions, whose behaviour is also given by an asymptotic series

$$(9.2.4) \quad \rho(t) \approx r_0 + r_1 t^{-1} + r_2 t^{-2} + \dots \quad (t \rightarrow \infty).$$

More precisely, we shall show the existence of numbers  $t_0 > 0$  and  $A > 0$  such that for every number  $\rho_0$  in the interval  $-A \leq \rho_0 \leq A$  the solution with  $\rho(t_0) = \rho_0$  can be continued indefinitely to the right, with an asymptotic series of the form (9.2.4). Moreover, the coefficients  $r_0, r_1, r_2, \dots$  are independent of  $\rho_0$ .

It is not generally true that *all* solutions show this behaviour. For example, the equation

$$\rho'(t) = -1 - \rho(t) - t^{-1}\rho^2(t)$$

has the solution  $\rho(t) = -t$ , and moreover it has solutions escaping to  $-\infty$  if  $t$  tends to some finite number  $t_1$ . (This can be seen from

the equation  $y'' + (1 + t^{-1})y' + t^{-1}y = 0$ , and the substitution  $y'/y = t^{-1}\rho$ .

Let  $A$  be a number exceeding  $2|a_0/b_0|$ . By virtue of (9.2.3),  $t_0$  can be determined such that

$$(9.2.5) \quad \beta(t) < 0, \quad 2|\alpha(t)| < A|\beta(t)| \quad (t \geq t_0),$$

and such that

$$(9.2.6) \quad 2A|\gamma(t)| < |\beta(t)| \quad (t \geq t_0).$$

Let  $\varphi_1$  denote the constant function  $\varphi_1(t) = -A$  ( $t \geq t_0$ ), and  $\varphi_2$  the constant function  $\varphi_2(t) = A$  ( $t \geq t_0$ ). It follows from (9.2.5) and (9.2.6) that

$$\begin{aligned} t^{-k}\varphi_1'(t) &< \alpha(t) + \beta(t)\varphi_1(t) + \gamma(t)\varphi_1^2(t) \quad (t \geq t_0), \\ t^{-k}\varphi_2'(t) &> \alpha(t) + \beta(t)\varphi_2(t) + \gamma(t)\varphi_2^2(t) \quad (t \geq t_0). \end{aligned}$$

By virtue of the theorem of sec. 9.1, we infer that any solution  $\rho(t)$  of our equation  $t^{-k}\rho' = \alpha + \beta\rho + \gamma\rho^2$  with  $|\rho(t_0)| \leq A$  automatically satisfies  $|\rho(t)| \leq A$  for all  $t > t_0$  as long as  $\rho(t)$  exists. So these solutions can not escape to  $\pm\infty$ , and it follows that they can be continued indefinitely to the right and that they satisfy  $|\rho(t)| \leq A$  ( $t \geq t_0$ ). We shall only use the fact that they are bounded:  $\rho(t) = O(1)$  ( $t \rightarrow \infty$ ). A solution will be called bounded if it is bounded in some interval  $(t_3, \infty)$ , although it is possible that this solution can be continued over some interval  $(t_4, t_3)$  or even over  $(-\infty, t_3)$  without being bounded over the extended interval.

In our next step we shall show that  $\rho(t) = O(1)$  ( $t \rightarrow \infty$ ) implies that

$$(9.2.7) \quad \rho(t) = r_0 + O(t^{-1}) \quad (t \rightarrow \infty), \text{ where } r_0 = -a_0/b_0.$$

We consider a special solution  $\rho(t)$ , bounded in some interval  $t_0 \leq t < \infty$ . Again we introduce two auxiliary functions

$$\varphi_3(t) = r_0 - At^{-1}, \quad \varphi_4(t) = r_0 + At^{-1},$$

and we try to determine  $A > 0$  and  $t_1 > t_0$  such that

$$(9.2.8) \quad t^{-k}\varphi_3'(t) < \alpha(t) + \beta(t)\varphi_3(t) + \gamma(t)\varphi_3^2(t) \quad (t \geq t_1),$$

$$(9.2.9) \quad t^{-k}\varphi_4'(t) > \alpha(t) + \beta(t)\varphi_4(t) + \gamma(t)\varphi_4^2(t) \quad (t \geq t_1),$$

$$(9.2.10) \quad \varphi_3(t_1) \leq \rho(t_1) \leq \varphi_4(t_1).$$

To this end we remark that both sides of (9.2.8) have asymptotic series. On both sides the constant term of the series vanishes. The coefficient of the term  $t^{-1}$  vanishes on the left; on the right (notice that  $c_0 = 0$ ) it equals

$$a_1 - b_0 A + b_1 r_0 + c_1 r_0^2.$$

Since  $b_0 < 0$ , we can determine  $A > 0$  such that this is  $> 1$ , say. Therefore it is easy to find  $t_1$  such that (9.2.8) holds. Moreover, we can argue similarly about (9.2.9), and it results that  $A$  and  $t_1$  can be chosen such that both (9.2.8) and (9.2.9) are satisfied. However, (9.2.10) gives a difficulty: it is easily satisfied for a special  $t_1$  by making  $A$  sufficiently large, but in the previous argument the choice of  $t_1$  was depending on  $A$ . We therefore restate things more carefully, considering both  $t$  and  $A$  as variables.

We have

$$\varphi_3(t) = r_0 + O(At^{-1}), \quad \varphi_3^2(t) = O(1) + O(A^2t^{-2}), \quad \alpha(t) = a_0 + O(t^{-1}), \\ \beta(t) = b_0 + O(t^{-1}), \quad \gamma(t) = O(t^{-1}), \quad t^{-k}\varphi_3'(t) = O(At^{-k-2}),$$

where all  $O$ -symbols refer to  $t \geq t_0$ ,  $A \geq 1$ , say. It follows that

$$\beta(t)\varphi_3(t) = b_0 r_0 - b_0 At^{-1} + O(t^{-1}) + O(At^{-2}), \\ \gamma(t)\varphi_3^2(t) = O(t^{-1}) + O(A^2t^{-3}),$$

and the right-hand-side of (9.2.8) exceeds the left-hand-side by the amount

$$-b_0 At^{-1} + O(t^{-1}) + O(At^{-2}) + O(A^2t^{-3}) = \\ = At^{-1}\{-b_0 + O(A^{-1}) + O(t^{-1}) + O(At^{-2})\}.$$

It follows that (9.2.8) is true when  $A$ ,  $t_1$  and  $t_1^2 A^{-1}$  are sufficiently large, and of course the same thing applies to (9.2.9). Next we consider (9.2.10). As  $\rho(t) - r_0 = O(1)$ , (9.2.10) holds if  $At_1^{-1}$  is sufficiently large. So the question remains, if  $C$  is any large number, whether  $A$  and  $t_1^{-1}$  can be found such that

$$A > C, \quad t_1 > C, \quad t_1^2 A^{-1} > C, \quad At_1^{-1} > C.$$

This can be achieved by making  $A = t_1^{3/2}$  and taking  $t_1$  sufficiently large. Thus we have found  $A$  and  $t_1$  such that (9.2.8), (9.2.9), (9.2.10) hold simultaneously.

By the theorem of sec. 9.1 we now infer that

$$\varphi_3(t) \leq \rho(t) \leq \varphi_4(t) \quad (t \geq t_1),$$

and we have proved (9.2.7). (For a second proof, see sec. 9.7, ex. 2).

Next we write  $\rho(t) = r_0 + t^{-k}\rho_1(t)$ , so that the result we just proved, is: if  $\rho(t) = O(1)$ , then  $\rho_1(t) = O(1)$ . We easily derive the differential equation

$$(9.2.11) \quad t^{-k}\rho_1'(t) = \alpha_1(t) + \beta_1(t)\rho_1(t) + \gamma_1(t)\rho_1^2(t),$$

with

$$\begin{aligned} \alpha_1 &= t(\alpha + r_0\beta) + r_0^2t\gamma, \\ \beta_1 &= \beta + 2r_0\gamma + t^{-k-1}, \\ \gamma_1 &= t^{-1}\gamma. \end{aligned}$$

The new coefficients  $\alpha_1, \beta_1, \gamma_1$  turn out to have asymptotic series again, and the analogue of (9.2.3) holds. Applying the above result to our new equation, we infer that there exists a constant  $r_1$  such that, if  $\rho_1(t) = r_1 + t^{-1}\rho_2(t)$ , then  $\rho_1(t) = O(1)$  implies  $\rho_2(t) = O(1)$ . As this procedure can be continued, it is clear that  $\rho(t)$  has an asymptotic development of the type (9.2.4).

The coefficients  $r_0, r_1, r_2, \dots$  follow successively from the above procedure. It is easier, however, to proceed by the method of undetermined coefficients. Just substitute the formal series  $r_0 + r_1t^{-1} + r_2t^{-2} + \dots$  into the equation (9.2.1), taking as its derivative the formal derivative  $-r_1t^{-1} - 2r_2t^{-2} - \dots$ . Then require that, for each value of  $n$  ( $n = 0, 1, 2, \dots$ ), the coefficients of  $t^{-n}$  on both sides are equal. This produces a set of equations from which, by virtue of  $b_0 \neq 0, c_0 = 0$ , the numbers  $r_0, r_1, \dots$  can be solved successively. The validity of this procedure is easily proved from the fact that there exists asymptotic series both for  $\rho(t)$  and for  $t^{-k}\rho'(t)$ , but it can also be shown by comparing the two procedures from an algebraic point of view.

With our equation we have a typical case of *stability*. If  $\rho_1(t)$  is one of the bounded solutions, bounded in  $t_0 \leq t < \infty$ , say, then there exists a positive number  $\varepsilon$  such that any solution whose value at  $t_0$  satisfies  $|\rho_2(t_0) - \rho_1(t_0)| < \varepsilon$ , also satisfies  $\rho_2(t) - \rho_1(t) \rightarrow 0$  ( $t \rightarrow \infty$ ). It is quite a strong type of stability, for  $\rho_1$  and  $\rho_2$  have the same asymptotic series, and therefore  $\rho_2 - \rho_1 = O(t^{-n})$ , for

every fixed value of  $n$  ( $n = 1, 2, 3, \dots$ ). We shall even show that  $\rho_2 - \rho_1 = O(e^{-ct})$  with some positive constant  $c$ .

Let  $\rho_1$  and  $\rho_2$  be two bounded solutions, and put  $\rho_2 - \rho_1 = \eta$ . If  $\rho_1(t) = \rho_2(t)$  for some  $t$ , it holds identically, by virtue of the uniqueness. So we may assume that  $\rho_2(t) > \rho_1(t)$  for all values of  $t$ . We evidently have

$$t^{-k}\eta' = \beta\eta + \gamma\eta(\rho_1 + \rho_2),$$

and consequently

$$t^{-k}\eta' = \eta(\beta + O(t^{-1})).$$

As  $\beta \approx b_0 + \dots$ ,  $b_0 < 0$ , and  $\eta(t) > 0$  for all  $t$ , we have, for some  $t_1 > 0$ ,

$$\eta'/\eta < \frac{1}{2}b_0 t^k \quad (t > t_1).$$

It follows by integration that

$$\eta = O\{\exp(b_0 t^{k+1}/2(k+1))\} \quad (t \rightarrow \infty).$$

### 9.3. An unstable case

We again consider the equation (9.2.1)

$$(9.3.1) \quad t^{-k}\rho'(t) = \alpha(t) + \beta(t)\rho(t) + \gamma(t)\rho^2(t),$$

and again we assume that  $\alpha$ ,  $\beta$  and  $\gamma$  have asymptotic developments (9.2.2), but instead of (9.2.3) we assume

$$(9.3.2) \quad b_0 > 0, \quad c_0 = 0.$$

Formally nothing has been changed, and therefore we can again find a series  $r_0 + r_1 t^{-1} + r_2 t^{-2} + \dots$  which formally satisfies the equation (in this formal procedure the sign of  $b_0$  is irrelevant; it is only the condition  $b_0 \neq 0$  that matters). The difference with the case  $b_0 < 0$  is that in the present case there is only one solution having this series as its development, and this one is the only bounded solution (i.e. the only solution which is  $O(1)$  when  $t \rightarrow \infty$ ).

In order to show this, we start by defining the functions

$$\varphi_3(t) = r_0 - At^{-1}, \quad \varphi_4(t) = r_0 + At^{-1} \quad (r_0 = -a_0/b_0),$$

just as in sec. 9.2. And we can fix  $A > 0$  and  $t_1 > 0$  such that

$$(9.3.3) \quad t^{-k}\varphi_3'(t) > \alpha(t) + \beta(t)\varphi_3(t) + \gamma(t)\varphi_3^2(t) \quad (t > t_1),$$

$$(9.3.4) \quad t^{-k}\varphi_4'(t) < \alpha(t) + \beta(t)\varphi_4(t) + \gamma(t)\varphi_4^2(t) \quad (t > t_1).$$

The inequality signs are different from those in (9.2.8) and

(9.2.9), as  $b_0 > 0$  in the present case. The conclusion is, that if a solution  $\rho(t)$  exists in an interval  $(t_2, t_3)$ , where  $t_3 > t_2 \geq t_1$ , and if  $\rho(t_2) < \varphi_3(t_2)$ , then we have  $\rho(t) < \varphi_3(t)$  ( $t_2 \leq t \leq t_3$ ). Similarly, if a solution  $\rho(t)$  exceeds  $\varphi_4(t)$  at  $t = t_2$ , where  $t_2 \geq t_1$ , then  $\rho(t) > \varphi_4(t)$  for all  $t > t_2$ , as long as  $\rho(t)$  exists. So the situation is just the opposite of the one in sec. 9.2.

Just now we made a statement about solutions  $< \varphi_3$  or  $> \varphi_4$ , but we would rather know something about solutions between  $\varphi_3$  and  $\varphi_4$ . This can still be achieved with the aid of the theorem of sec. 9.1, by taking  $\tau = -t$  as the independent variable, so that  $d\rho/dt = -d\rho/d\tau$ . That is, we observe what the solutions do if  $t$  is decreasing instead of increasing. Our conclusion reads as follows: Assume  $t_1 < t_2$ , and let the number  $\rho_2$  satisfy  $\varphi_3(t_2) \leq \rho_2 \leq \varphi_4(t_2)$ . Then the given equation has a solution  $\rho(t)$  in the interval  $t_1 \leq t \leq t_2$ , with  $\rho(t_2) = \rho_2$  and  $\varphi_3(t) < \rho(t) < \varphi_4(t)$  ( $t_1 \leq t < t_2$ ).

If  $\rho_2$  ranges through the closed interval  $\varphi_3(t_2) \leq \rho_2 \leq \varphi_4(t_2)$ , then the value  $\rho_1 = \rho(t_1)$  ranges through a subset of the interval  $\varphi_3(t_1) < \rho < \varphi_4(t_1)$ .

If follows from the general theory of differential equations<sup>1)</sup> that  $\rho_1$  is a continuous and increasing function of  $\rho_2$ , and therefore this subset is again a closed interval. We shall denote it by  $i(t_2)$  ( $t_1$  is considered as fixed, and  $t_2$  will be varied).

The set  $i(t_2)$  can be interpreted as the set of all numbers  $\rho_1$  with the property that there exists a solution  $\rho(t)$  in the interval  $t_1 \leq t \leq t_2$  satisfying  $\varphi_3(t) \leq \rho(t) \leq \varphi_4(t)$  throughout that interval, with  $\rho(t_1) = \rho_1$ . It follows that  $i(t_2) \supset i(t_2 + 1)$ . Now applying this with  $t_2 = t_1 + n$  ( $n = 1, 2, 3, \dots$ ) we get a sequence of closed intervals

$$i(t_1 + 1) \supset i(t_1 + 2) \supset i(t_1 + 3) \supset \dots,$$

and therefore these intervals have a number  $\rho_1^*$  in common. Denoting the solution with the initial value  $\rho_1^*$  at  $t_1$  by  $\rho^*(t)$ , we know that this one can be continued up to  $t_1 + n$ , and remains between  $\varphi_3(t)$  and  $\varphi_4(t)$  ( $t_1 \leq t \leq t_1 + n$ ). This holds for any value of  $n$ , and therefore  $\rho^*(t)$  can be continued to infinity, and

$$\varphi_3(t) \leq \rho^*(t) \leq \varphi_4(t) \quad (t > t_1).$$

<sup>1)</sup> See E. L. INCE, Ordinary Differential Equations, 1st ed., New York, § 3.22.

Thus we have proved that our equation has a solution of the form  $r_0 + O(t^{-1})$ . We next show that there is only one such solution, and even that there is only one bounded solution. Let  $\rho_1(t)$  and  $\rho_2(t)$  denote solutions which are both bounded in an interval  $t_0 < t < \infty$ . We suppose them to be different:  $\rho_2(t) > \rho_1(t)$  ( $t > t_0$ ), say. Putting  $\rho_2(t) - \rho_1(t) = \eta(t)$ , we should have (cf. the end of sec. 9.2), for some  $t_1$ ,

$$\eta'/\eta > \frac{1}{2}b_0 t^k \quad (t > t_1)$$

and it would follow that

$$(9.3.5) \quad \eta > c \exp\{b_0 t^{k+1}/2(k+1)\} \quad (t > t_1)$$

with some positive constant  $c$ . So  $\eta$  would tend to infinity, whereas  $\rho_1$  and  $\rho_2$  are bounded. This is contradictory; hence there is just one bounded solution.

In order to get the full asymptotic expansion for  $\rho^*(t)$ , we write  $\rho(t) = r_0 + t^{-1}\rho_1(t)$ , as in sec. 9.2, and for  $\rho_1$  we get a differential equation (9.2.11) of the same type as the equation for  $\rho(t)$ . So we infer that there is just one solution  $\rho_1^*(t)$  of the form  $r_1 + O(t^{-1})$ . Now  $r_0 + t^{-1}\rho_1^*(t)$  is a solution of (9.3.1), having the form  $r_0 + r_1 t^{-1} + O(t^{-2})$ . As (9.3.1) possesses only one bounded solution, we have identically

$$r_0 + t^{-1}\rho_1^*(t) = \rho^*(t),$$

and therefore

$$\rho^*(t) = r_0 + r_1 t^{-1} + O(t^{-2}).$$

This procedure can be continued, and it follows that  $\rho^*(t)$  has an asymptotic development (9.2.4). The numbers  $r_0, r_1, \dots$  can again be obtained by formal substitution of (9.2.4) into the differential equation, and by equating coefficients of corresponding powers of  $t$ .

#### 9.4. Application to a linear second-order equation

If  $y(t)$  is the unknown function in a linear homogeneous second-order equation

$$y''(t) + P(t)y'(t) + Q(t)y(t) = 0,$$

then the substitution  $y'/y = v$  leads to a first-order equation, of the Riccati type, for the function  $v$ :

$$v'(t) + v^2(t) + P(t)v(t) + Q(t) = 0.$$

By linear substitution  $v(t) = a(t) + b(t)w(t)$ , where  $w(t)$  is the new unknown function, we get for  $w$  an equation, again of the Riccati type (i.e. a linear relation between  $w'$ ,  $w^2$ ,  $w$  and 1). Now one can try to obtain a Riccati equation of one of the types discussed in the previous sections, or at least an equation to which the techniques of those sections can be applied.

As an example we take the equation

$$(9.4.1) \quad y''(t) - t^4y(t) = 0,$$

for which the substitution  $y'/y = v$  leads to

$$(9.4.2) \quad v' + v^2 - t^4 = 0.$$

This one is not yet of the right type. In order to get a rough idea about the behaviour of the solutions, we argue as follows. There are three terms in the equation, and so at least two have to be of the same order of magnitude. So we first try to neglect one of the terms, and we investigate the remaining equation.

First neglect the term  $t^4$ . The remaining equation  $v' + v^2 = 0$  has the solution  $v = (t - t_0)^{-1}$ , with an arbitrary constant  $t_0$ . Now for these functions  $v$  the neglected term is much larger than  $v'$  and  $v^2$ , and so we are left nowhere. Next neglect the term  $v^2$ . Then there remains  $v' = t^4$ , and therefore  $v = \frac{1}{5}t^5 + C$ . And again the neglected term is much larger than the others.

So our last attempt will be to neglect  $v'$ . The fact that the remaining equation is no longer a differential equation does not disturb us in the least. We obtain  $v = \pm t^2$ , and now the term  $v'$  is small indeed. Nothing has been proved yet, but we have now sufficient reason to try the substitutions  $v = t^2 + \rho(t)$  and  $v = -t^2 + \rho(t)$ . The first one,  $v = t^2 + \rho(t)$ , transform (9.4.2) into

$$(9.4.3) \quad t^{-2}\rho' = -2t^{-1} - 2\rho - t^{-2}\rho^2,$$

and this one is of the stable type discussed in sec. 9.2 (with  $k = 2$ ,  $\alpha(t) = -2t^{-1}$ ,  $\beta(t) = -2$ ,  $\gamma(t) = -t^{-2}$ , so indeed  $b_0 < 0$ ,  $c_0 = 0$ ). We infer that there is a solution  $\rho(t)$  with an asymptotic series

$$\rho(t) \approx r_0 + r_1t^{-1} + r_2t^{-2} + \dots, \quad (t \rightarrow \infty),$$

(and even that there are infinitely many such solutions). Upon formal substitution of the series into (9.4.3) we obtain that  $r_0 = 0$ ,

$r_1 = -1$ ,  $r_2 = r_3 = 0$ , and generally, that  $r_n = 0$  unless  $n$  is of the form  $3k + 1$ . So it may have some advantage to put  $\rho(t) = t^{-1}\sigma(t^3)$ ,  $t^3 = \tau$ , which transforms the equation into

$$3 \frac{d\sigma}{d\tau} = -2 + (-2 + \tau^{-1})\sigma(\tau) - \tau^{-1}\sigma^2(\tau).$$

As  $y'/y = v$ , we obtain  $\log y$  by integration of  $t^2 + t^{-1}\sigma(t^3)$ . It follows that the equation (9.4.1) has solutions of the form

$$(9.4.4) \quad y \approx Ct^{-1}e^{\frac{1}{3}t^3}(1 + a_1t^{-3} + a_2t^{-6} + \dots) \quad (t \rightarrow \infty),$$

where  $C \neq 0$ , and  $1 + a_1x + a_2x^2 + \dots$  is the formal development of

$$\exp(-\frac{1}{6}r_4x - \frac{1}{12}r_7x^2 - \dots).$$

We can also try the second substitution, viz  $v = -t^2 + \rho(t)$ . We then get

$$t^{-2}\rho'(t) = 2t^{-1} + 2\rho - t^{-2}\rho^2,$$

and so we are in the unstable case of sec. 9.3. We now infer that there exists just *one* solution of the form

$$\rho(t) \approx s_0 + s_1t^{-1} + s_2t^{-2} + \dots \quad (t \rightarrow \infty).$$

In terms of  $y$  it means that there is, apart from the arbitrary constant  $C \neq 0$ , just one solution of the form

$$(9.4.5) \quad y \approx Ct^{-1}e^{-\frac{1}{3}t^3}(1 + b_1t^{-3} + b_2t^{-6} + \dots) \quad (t \rightarrow \infty).$$

If we select a solution  $y_1(t)$  of the form (9.4.4), and a solution  $y_2(t)$  of the form (9.4.5), then  $y_1$  and  $y_2$  are obviously linearly independent (as  $y_2/y_1 \rightarrow 0$ ). Now the general solution of (9.4.1) is  $Ay_1(t) + By_2(t)$  ( $A$  and  $B$  are constants). This illustrates the instability of the solution  $y_2$ : Every solution with  $A \neq 0$  is easily seen to have (9.4.4) as its asymptotic behaviour (with some value of  $C$ ), and only if  $A = 0$ ,  $B \neq 0$  we have something of the type (9.4.5). Moreover, it is easily seen that when adding  $By_2(t)$  to  $Ay_1(t)$ , the asymptotic series for  $Ay_1(t)$  is not altered.

There is a quite simple relation between the  $a$ 's of (9.4.4) and the  $b$ 's of (9.4.5), due to the fact that the coefficients in (9.4.1) are even functions of  $t$ . Its effect is that

$$(9.4.6) \quad b_n = (-1)^n a_n \quad (n = 0, 1, 2, \dots).$$

In other words, if the right-hand-side of (9.4.4) is denoted formally

by  $P(t)$ , then the right-hand-side of (9.4.5) equals  $P(-t)$ . This is easily deduced by studying the Riccati equation (9.4.2). If we substitute for  $v$  the formal series

$$Q_1(t) = t^2 + r_0 + r_1 t^{-1} + r_2 t^{-2} + \dots,$$

then (9.4.2) is formally satisfied. As  $t^4$  is an even function, it follows that the formal series  $Q_2(t) = -Q_1(-t)$  also satisfies (9.4.2). On the other hand we observe (cf. secs. 9.2 and 9.3) that the Riccati equation has only one formal solution of the type  $-t^2 + s_0 + s_1 t^{-1} + s_2 t^{-2} + \dots$ . It follows that  $s_n = -(-1)^n r_n$  ( $n = 0, 1, 2, \dots$ ). Now (9.4.6) is easily shown.

Let  $y_2(t)$  denote the unstable solution of the form (9.4.5) (with  $C = 1$ ). We shall show, as a consequence of (9.4.6), that the general solution of  $y'' - t^4 y = 0$  can be written in the form

$$y = A y_2(-t) + B y_2(t).$$

First we remark that  $y_2(t)$  can be continued over  $(-\infty, \infty)$ , the equation being linear. As  $t^4$  is even, the function  $y_2(-t)$  also satisfies  $y'' - t^4 y = 0$ . It follows from the asymptotic behaviour that  $y_2(-t)$  is positive and increasing if  $t$  is negative and large. It follows from the equation  $y'' = t^4 y$  that the solutions are convex whenever they are positive, and therefore  $y_2(-t)$  keeps on increasing as  $t$  tends to  $+\infty$ . As  $y_2(t) \rightarrow 0$  ( $t \rightarrow \infty$ ), it follows that  $y_2(t)$  and  $y_2(-t)$  are linearly independent solutions.

It is not difficult to evaluate the coefficients of the asymptotic series for  $y_2(t)$ . The reader may verify that

$$y_2(t) \approx e^{-\frac{1}{4}t^4} \sum_{n=0}^{\infty} \frac{(3n)!}{18^n (n!)^2} (-1)^n t^{-3n-1} \quad (t \rightarrow \infty).$$

## 9.5. Oscillatory cases

The analysis of sec. 9.4 applies to many equations of the type  $y''(t) - y(t)f(t) = 0$ , where  $f(t) > 0$ . The situation is entirely different, however, with equations of the form  $y''(t) + y(t)f(t) = 0$  (again with  $f(t) > 0$ ). Under very general conditions it can be proved that all solutions are oscillating, i.e. that they have infinitely many zeros in the interval  $(0, \infty)$ . We shall consider the special case

$$(9.5.1) \quad y''(t) + (1 + t^{-1})y(t) = 0.$$

We obtain a Riccati equation

$$v' + v^2 + (1 + t^{-1}) = 0$$

by the substitution  $y'/y = v$ . Guided by the same heuristic argument we used in the case of (9.4.2), we introduce a new function  $\rho(t)$  by the substitution  $v = i + t^{-1}\rho(t)$ . This function  $\rho(t)$  satisfies

$$(9.5.2) \quad \rho'(t) = -1 + (-2i + t^{-1})\rho(t) - t^{-1}\rho^2(t).$$

With the notation of (9.2.2), we thus have  $b_0 = -2i$ ,  $c_0 = 0$ . As both in sec. 9.2 and sec. 9.3 the reality of the coefficients was postulated, these sections cannot be applied to (9.5.2). Admittedly, we may be able to show that the main results of sec. 9.2 and sec. 9.3 remain valid for complex equations, provided that we replace  $b_0 < 0$  in (9.2.3) by  $\operatorname{Re} b_0 < 0$ , and  $b_0 > 0$  in (9.3.2) by  $\operatorname{Re} b_0 > 0$ . But in the present case  $b_0$  is purely imaginary, and therefore we neither have the strong type of stability of sec. 9.2, nor the strong type of instability of sec. 9.3.

We remark that it is not difficult to find an asymptotic series, formally satisfying (9.5.2), just by substituting the series and equating coefficients. The first few terms are

$$\frac{1}{2}i + (2 - i)/(8t) - (4 + 3i)/(16t^2) + \dots$$

However, at the present stage we cannot say whether this formal series has any significance.

We shall attempt an entirely different method, based upon the use of an integral equation. It should be noted that it is not difficult to modify this method so that it will work equally well in the exponential case of sec. 9.4.

We consider an equation of the type

$$(9.5.3) \quad y''(t) + \{1 + g(t)\}y(t) = 0,$$

where the given function  $g$  is continuous and moreover satisfies  $\int_0^\infty |g(t)|dt < \infty$ . This means that the results will not be directly applicable to (9.5.1), although (9.5.1) can easily be transformed into an equation of the present type (see the end of sec. 9.6).

We shall first transform (9.5.3) into an integral equation. It

can be obtained as follows. We write the equation in the form

$$y''(t) + y(t) = -g(t)y(t),$$

and we treat this equation as if the right-hand-side were a given function  $h(t)$ . Using the method of variation of constants, we put

$$\begin{aligned} y(t) &= A(t) \cos t + B(t) \sin t, \\ y'(t) &= -A(t) \sin t + B(t) \cos t, \quad A'(t) \cos t + B'(t) \sin t = 0, \\ y''(t) &= -y(t) - A'(t) \sin t + B'(t) \cos t, \\ &\quad -A'(t) \sin t + B'(t) \cos t = h(t). \end{aligned}$$

Thus we have  $A' = -h(t) \sin t$ ,  $B' = +h(t) \cos t$ . So if  $a$  is a real number, every solution of  $y''(t) + y(t) = h(t)$  can be written in the form

$$y = C_1 \cos t + C_2 \sin t + \int_a^t h(\tau) (\sin t \cos \tau - \cos t \sin \tau) d\tau,$$

with suitable constants  $C_1$  and  $C_2$ . So if  $y(t)$  is a solution of (9.5.3), it also satisfies

$$(9.5.4) \quad y(t) = C_1 \cos t + C_2 \sin t - \int_a^t g(\tau) y(\tau) \sin(t-\tau) d\tau,$$

with suitable constants  $C_1$  and  $C_2$ .

We can now show that every solution of (9.5.3) is bounded in the interval  $0 \leq t < \infty$ . To this end we choose a positive number  $a$  such that  $\int_a^\infty |g(t)| dt < \frac{1}{2}$ , which is possible on the strength of the convergence of  $\int_0^\infty |g(t)| dt$ . Let  $y(t)$  be any solution, and let  $b$  be any number  $> a$ . Put  $M = \max_{a \leq t \leq b} |y(t)|$ . From (9.5.4) we infer that

$$M \leq |C_1| + |C_2| + M \int_a^b |g(\tau)| d\tau \leq |C_1| + |C_2| + \frac{1}{2}M,$$

and so  $M \leq 2|C_1| + 2|C_2|$ , irrespective of the value of  $b$ . This shows the boundedness.

The boundedness implies that  $\int_0^\infty g(\tau) y(\tau) \sin(t-\tau) d\tau$  is convergent, and therefore we can rewrite (9.5.4) into the form

$$(9.5.5) \quad y(t) = C_3 \cos t + C_4 \sin t - \int_t^\infty g(\tau) y(\tau) \sin(\tau-t) d\tau,$$

with new constants  $C_3$  and  $C_4$ :

$$C_3 = C_1 + \int_a^\infty g(\tau) y(\tau) \sin \tau d\tau, \quad C_4 = C_2 - \int_a^\infty g(\tau) y(\tau) \cos \tau d\tau.$$

We have proved that for every solution of (9.5.3) there exist constants  $C_3, C_4$  such that (9.5.5) holds. We shall next show that to every choice of the pair of numbers  $C_3, C_4$  there corresponds exactly one solution.

Let  $y_1(t)$  denote the solution determined by the initial conditions  $y_1(a) = 1, y_1'(a) = 0$ , and let  $y_2(t)$  be determined by  $y_2(a) = 0, y_2'(a) = 1$ . Then every solution of (9.5.3) is a linear combination of  $y_1(t)$  and  $y_2(t)$ :

$$y(t) = y(a)y_1(t) + y'(a)y_2(t),$$

and  $y(a), y'(a)$  can be prescribed arbitrarily.

The constants  $C_3, C_4$ , corresponding to the solution  $y(t)$  in (9.5.5), are easily determined by differentiation of that relation at  $t = a$ . For simplicity we assume that  $a$  is an integral multiple of  $2\pi$ . Then we obtain that

$$\begin{aligned} C_3 &= y(a) + y(a) \int_a^\infty g(\tau) y_1(\tau) \sin \tau d\tau + y'(a) \int_a^\infty g(\tau) y_2(\tau) \sin \tau d\tau, \\ C_4 &= y'(a) - y(a) \int_a^\infty g(\tau) y_1(\tau) \cos \tau d\tau - y'(a) \int_a^\infty g(\tau) y_2(\tau) \cos \tau d\tau. \end{aligned}$$

We observe that  $C_3$  and  $C_4$  are linear combinations of  $y(a)$  and  $y'(a)$ . We have to show that these combinations are linearly independent. In other words, we have to prove that  $C_3 = C_4 = 0$  occurs only if  $y(a) = y'(a) = 0$ , that is, if  $y(t)$  vanishes identically.

Let  $y(t)$  be a solution with the property that  $C_3$  and  $C_4$  both vanish. As  $y(t)$  is bounded over  $(a, \infty)$ , we deduce from (9.5.5) that

$$\sup_{a \leq t \leq \infty} |y(t)| \leq \sup_{a \leq t \leq \infty} |y(t)| \cdot \int_a^\infty |g(\tau)| d\tau.$$

As the integral on the right can be assumed to be less than  $\frac{1}{2}$ , we infer that  $y(t)$  vanishes identically.

The integral equation (9.5.5) can be used in order to solve the differential equation explicitly, in the form of the so-called Neumann series<sup>1)</sup>. We shall not do this, as our only aim is to obtain asymptotic

<sup>1)</sup> See E. T. WHITTAKER and G. N. WATSON, *Modern Analysis*, 4th ed., Cambridge 1946, § 11.4.

information about  $y(t)$ . This is achieved by the iteration method which is used in the derivation of the Neumann series. We choose the solution  $y_3$ , say, determined by  $C_3 = 1$ ,  $C_4 = 0$ . We know already that  $y(t)$  is bounded, and so we infer from (9.5.5) that

$$(9.5.6) \quad y_3(t) = \cos t + O(\int |g(\tau)| d\tau) = \cos t + o(1).$$

Next we insert this result into the integral on the right-hand-side of (9.5.5):

$$y_3(t) = \cos t - \int g(\tau) \cos \tau \sin(\tau - t) d\tau + O(\int |g(s)| ds \int |g(\tau)| d\tau),$$

and so on.

For the calculations it may be easier to deal with the complex combinations

$$y_3(t) + iy_4(t) = e^{it} + o(1),$$

$$y_3(t) - iy_4(t) = e^{-it} + o(1),$$

where  $y_4$  is the solution determined by  $C_3 = 0$ ,  $C_4 = 1$ .

We take a specific example. If in the differential equation for the  $n$ -th Bessel function we write  $y = t^{\frac{1}{2}} J_n(t)$ ,  $\frac{1}{4} - n^2 = \lambda$ , we obtain the equation <sup>1)</sup>

$$(9.5.7) \quad y''(t) + (1 + \lambda t^{-2})y(t) = 0,$$

which is indeed of the type discussed above ( $\lambda$  is a constant). So we know that there is a solution of the type  $e^{it} + o(1)$  ( $t \rightarrow \infty$ ). Denoting this one by  $y(t)$ , we have

$$y(t) = e^{it} + O(\int_t^\infty \tau^{-2} d\tau) = e^{it} + O(t^{-1}).$$

In the next step we get from (9.5.5) ( $C_3 = 1$ ,  $C_4 = i$ )

$$(9.5.8) \quad y(t) = e^{it} - \lambda \int_t^\infty \tau^{-2} e^{i\tau} \sin(\tau - t) d\tau + O(\int_t^\infty \tau^{-3} d\tau).$$

We first consider the integrals  $\int_t^\infty \tau^{-k} e^{i\tau} \sin(\tau - t) d\tau = f_k(t)$ , with  $k$

<sup>1)</sup> See E. L. INCE, Ordinary Differential Equations, 1st ed., New York, § 7.32.

fixed, and  $k \geq 2$ . We have, as  $2i \sin x = e^{ix} - e^{-ix}$ ,

$$\begin{aligned} e^{-it} f_k(t) &= \int_0^\infty (x+t)^{-k} e^{ix} \sin x \, dx = \\ &= \frac{1}{2} i (k-1)^{-1} t^{-k+1} - \frac{1}{2} i \int_0^\infty (x+t)^{-k} e^{2ix} \, dx. \end{aligned}$$

This last integral can be expanded asymptotically by means of repeated partial integration, or by steepest descent. It results that we have an asymptotic expansion of the form

$$e^{-it} f_k(t) \approx t^{-k+1} (c_0 + c_1 t^{-1} + c_2 t^{-2} + \dots) \quad (t \rightarrow \infty).$$

Now (9.5.8) gives  $y(t) = e^{it}(1 + a_1 t^{-1} + O(t^{-2}))$  (with  $a_1 = -\frac{1}{2}i\lambda$ ). Again inserting this into the right-hand-side of (9.5.5), we get a formula of the type  $y(t) = e^{it}(1 + a_1 t^{-1} + a_2 t^{-2} + O(t^{-3}))$ . Continuing this procedure, we obtain an asymptotic series

$$(9.5.9) \quad y(t) \approx e^{it}(1 + a_1 t^{-1} + a_2 t^{-2} + a_3 t^{-3} + \dots) \quad (t \rightarrow \infty).$$

Once we know that there exists an asymptotic series of this type, it is quite easy to determine the coefficients directly from the differential equation. If we put  $y(t)e^{-it} = \varphi$ , then  $\varphi$  satisfies  $\varphi'' + 2i\varphi' + \lambda t^{-2}\varphi = 0$ . Formally substituting

$$(9.5.10) \quad \varphi(t) \approx 1 + a_1 t^{-1} + a_2 t^{-2} + \dots,$$

and equating coefficients, we can evaluate the  $a_i$ . In order to justify this procedure it is sufficient to show that  $\varphi'(t)$  and  $\varphi''(t)$  also have asymptotic series (for then those series automatically represent the formal derivatives of the asymptotic series of  $\varphi(t)$ ). It follows from (9.5.7) and (9.5.9) that

$$y''(t) \approx -e^{it}(1 + \lambda t^{-2})(1 + a_1 t^{-1} + \dots) \quad (t \rightarrow \infty).$$

From the asymptotic formulas of the functions  $\int_t^\infty e^{i\tau} \tau^{-k} d\tau$  ( $k = 1, 2, 3, \dots$ ) it can now be shown that  $\int_t^\infty (y''(\tau) + e^{i\tau}) d\tau$  converges, and that it has an asymptotic series of the form  $e^{it}(b_0 + b_1 t^{-1} + b_2 t^{-2} + \dots)$ . On the other hand, this integral equals  $-y'(t) + ie^{it} + C$ , where  $C$  is a constant. By a second integration we infer, since  $y(t)$  is bounded, that  $C$  vanishes. It now follows that  $\varphi'$  and  $\varphi''$  have asymptotic series of the required type.

So in order to determine the coefficients in (9.5.10), we have the right of formal substitution into  $\varphi'' + 2i\varphi' + \lambda t^{-2}\varphi = 0$ . Equating coefficients, we get the relations

$$a_{k+1} = a_k(k^2 + k + \lambda)/2(k + 1)i \quad (k = 0, 1, 2, \dots),$$

where  $a_0 = 1$ .

Our differential equation (9.5.7) has also a solution of the form  $e^{-it} + O(1)$ . For this one we get similar results, and it is easily seen that its asymptotic series is  $e^{-it}(1 - a_1t^{-1} + a_2t^{-2} - a_3t^{-3} + \dots)$ . So every solution of (9.5.7) has an expansion of the type

$$Ae^{it}(1 + a_1t^{-1} + a_2t^{-2} + \dots) + Be^{-it}(1 - a_1t^{-1} + a_2t^{-2} - \dots).$$

The contents of this section do not give us a method to determine the values of  $A$  and  $B$  belonging to the special solution  $t^{\frac{1}{2}}J_n(t)$ . The choice of this special solution from the set of all solutions depends on its behaviour as  $t \rightarrow 0$ , and here our knowledge about  $t \rightarrow \infty$  has no direct value. On the other hand, a quite rough estimate for the behaviour of  $t^{\frac{1}{2}}J_n(t)$ , obtained by any other method, will be sufficient in order to evaluate the numbers  $A$  and  $B$ .

## 9.6. More general oscillatory cases

In sec. 9.5 we learned that  $y'' + (1 + g(t))y = 0$  has solutions  $e^{it} + o(1)$  and  $e^{-it} + o(1)$  ( $t \rightarrow \infty$ ), provided that  $\int_0^\infty |g(t)|dt$  converges. We shall now try to reduce the more general equation

$$(9.6.1) \quad y''(t) + (p(t))^2y(t) = 0$$

to this special case. It is assumed that  $p(t)$  is a positive continuous function.

We shall replace the variables  $t$  and  $y$  by new variables  $x$  and  $z$ . We put

$$x = \varphi(t), \quad y = \psi(t)z.$$

Here  $\varphi(t)$  and  $\psi(t)$  are functions of  $t$ , to be determined later. We assume that  $\varphi(t)$  tends monotonically to  $+\infty$  when  $t \rightarrow +\infty$ . We obtain (accents denoting differentiation with respect to  $t$ ):

$$y' = \psi\varphi' \frac{dz}{dx} + z\psi',$$

$$y'' = \psi(\varphi')^2 \frac{d^2z}{dx^2} + (\psi\varphi'' + 2\psi'\varphi') \frac{dz}{dx} + \psi''z.$$

Our differential equation becomes

$$\frac{d^2z}{dx^2} + \frac{1}{\varphi'} \left( \frac{\varphi''}{\varphi'} + 2 \frac{\psi'}{\psi} \right) \frac{dz}{dx} + \left\{ \left( \frac{p}{\varphi'} \right)^2 + \frac{\psi''}{\psi \varphi'^2} \right\} z = 0.$$

If we now succeed in choosing  $\varphi$  and  $\psi$  such that

$$(9.6.2) \quad \frac{\varphi''}{\varphi'} + 2 \frac{\psi'}{\psi} = 0, \quad \int_0^\infty \left| \left( \frac{p}{\varphi'} \right)^2 + \frac{\psi''}{\psi \varphi'^2} - 1 \right| dx < \infty,$$

then we know from what we proved about (9.5.3) that (9.6.1) has solutions

$$\varphi(t)\{e^{t\varphi(t)} + o(1)\} \text{ and } \psi(t)\{e^{-t\varphi(t)} + o(1)\} \quad (t \rightarrow \infty).$$

Bearing in mind that  $\psi''/\psi = (\psi'/\psi)' + (\psi'/\psi)^2$ , and  $dx = \varphi' dt$  we can replace (9.6.2) by

$$(9.6.3) \quad \int_0^\infty \left| \frac{p^2}{\varphi'^2} - \frac{1}{2} \frac{\varphi'''}{\varphi'^3} + \frac{3}{4} \left( \frac{\varphi''}{\varphi'^2} \right)^2 - 1 \right| \varphi' dt < \infty$$

Moreover, we have our condition that  $\varphi(t) \rightarrow \infty$  when  $t \rightarrow \infty$ , which means that  $\int_0^\infty \varphi' dt = \infty$ . If, for example

$$(9.6.4) \quad \int_0^\infty |3p'^2 p^{-3} - 2p'' p^{-2}| dt < \infty, \quad \int_0^\infty p dt = \infty,$$

it is possible to choose  $\varphi$  such that  $\varphi' = p$ , i.e.  $\varphi(t) = \int_0^t p(\tau) d\tau$ , and then (9.6.3) is satisfied<sup>1)</sup>. Moreover, we take  $\psi(t) = (p(t))^{-\frac{1}{2}}$ .

If (9.6.4) holds, we know by virtue of sec. 9.5 that (9.6.1) has solutions  $y_1$  and  $y_2$  whose asymptotic behaviour is given by

$$(9.6.5) \quad \begin{aligned} y_1 &\sim (p(t))^{-\frac{1}{2}} \exp\left\{i \int_0^t p(\tau) d\tau\right\}, \\ y_2 &\sim (p(t))^{-\frac{1}{2}} \exp\left\{-i \int_0^t p(\tau) d\tau\right\} \quad (t \rightarrow \infty). \end{aligned}$$

The substitutions  $x = \varphi(t)$ ,  $y = z(p(t))^{-\frac{1}{2}}$  which in this case have to be carried out, transform the equation into the form  $d^2y/dx^2 + (1 + g(x))y = 0$ , with a relatively small function  $g(x)$ . In many cases it will be possible to apply the method of sec. 9.5 in order to obtain asymptotic series for the functions  $y_1$  and  $y_2$ .

<sup>1)</sup> A. WINTNER, Phys. Rev. 72 (1947), pp. 516–517.

In a fairly wide range of cases the substitution  $\varphi' = \dot{\rho}$  can be used. The condition (9.6.4) is satisfied, for example, if  $\rho(t) = t^\lambda$  ( $\lambda > -1$ ). It fails, however, if  $\rho(t) = t^{-1}$ . In this case we can still satisfy (9.6.3), but then with  $\varphi'(t) = 3t/(2t)$ . Then the integrand of (9.6.3) vanishes identically, corresponding to the fact that  $y'' + t^{-2}y = 0$  has the simple solutions  $y = t^\lambda$ ,  $\lambda = \frac{1}{2} \pm \frac{1}{2}i\sqrt{3}$ .

If  $c$  is a positive constant  $\leq \frac{1}{2}$ , then the solutions of  $y'' + ct^{-2}y = 0$  are no longer oscillatory, and the same thing can be said about the equations  $y'' + t^{2\lambda}y = 0$  if  $\lambda < -1$ .

The contents of this section can easily be applied to (9.5.1), i.e. the special case of (9.6.1) with  $\rho(t) = (1 + t^{-1})^t$ . It is easily seen that (9.6.4) is satisfied in this case, if we replace the integration interval  $(0, \infty)$  by  $(1, \infty)$ , which obviously does not matter. So we can take

$$\varphi(t) = \int_1^t (1 + \tau^{-1})^t d\tau, \quad \psi(t) = (1 + t^{-1})^{-t}.$$

This means that

$$\varphi(t) = t + \frac{1}{2} \log t + C + O(t^{-1}), \quad \psi(t) = 1 + O(t^{-1}),$$

where  $C$  is a constant. It follows that (9.5.1) has solutions  $y_1(t)$  and  $y_2(t)$  with the asymptotic behaviour (when  $t \rightarrow \infty$ )

$$y_1(t) \sim e^{it + \frac{1}{2}t \log t}, \quad y_2(t) \sim e^{-it - \frac{1}{2}t \log t}.$$

If we want to have the asymptotic series for  $y_1$  and  $y_2$ , we can argue as follows. Let

$$S(t) = 1 + \frac{1}{2}t^{-1} + s_2t^{-2} + \dots$$

be a formal series satisfying

$$(9.6.6) \quad (1 + t^{-1})S^{-2} = \frac{1}{2}S''S^{-3} - \frac{3}{4}(S'S^{-2})^2 + 1.$$

Then the integrand of (9.6.3) vanishes upon formal substitution of  $\varphi' = S$ . The existence of  $S$  is easily established. If

$$S_n = 1 + \frac{1}{2}t^{-1} + s_2t^{-2} + \dots + s_nt^{-n}$$

is inserted into the right-hand-side of (9.6.6), then  $S_{n+1}$  can be evaluated.

Next take, for some  $n$ ,  $\varphi(t)$  such that  $\varphi'(t) = S_n(t)$ . Then it is easily seen that the integrand of (9.6.3) becomes  $O(t^{-n-1})$ . It follows

(cf. (9.5.6)) that we have a solution

$$y_1(t) = \psi(t)e^{i\varphi(t)}(1 + O(t^{-n})),$$

where  $\psi(t)$  is related to  $\varphi(t)$  according to (9.6.2). It is now quite easy to prove that we have asymptotic expansions

$$y_1(t) \approx e^{it+\frac{1}{4}\log t} S(t)^{-\frac{1}{4}} \exp(-is_2t^{-1} - \frac{1}{2}is_3t^{-2} - \dots),$$

$$y_2(t) \approx e^{-it-\frac{1}{4}\log t} S(t)^{-\frac{1}{4}} \exp(is_2t^{-1} + \frac{1}{2}is_3t^{-2} + \dots),$$

if  $t \rightarrow \infty$ .

These lines have to be interpreted in the following sense. The function  $y_1(t)e^{-it-\frac{1}{4}\log t}$  has an asymptotic series which is obtained by formal multiplication of the formal series  $S(t)^{-\frac{1}{4}}$  by the formal series  $\exp(-is_2t^{-1} - \frac{1}{2}is_3t^{-2} - \dots)$ , and similarly for  $y_2(t)$ .

Once we know the existence of these expansions, it is of course possible to derive them directly from the Riccati equation (see the beginning of sec. 9.5).

**9.7. EXERCISES.** 1. Let  $\rho(t)$  be real and differentiable ( $1 < t < \infty$ ) and assume that

$$t\rho'(t) + \rho(t) - t^{-1}\rho^2(t) = O(1) \quad (t \rightarrow \infty),$$

$$\limsup t^{-1}\rho(t) < 1 \quad (t \rightarrow \infty).$$

Show that  $\rho(t) = O(1)$  ( $t \rightarrow \infty$ ).

2. Give an alternative proof of (9.2.7) (assuming (9.2.1), (9.2.2), (9.2.3),  $k \geq 0$  and  $\rho(t) = O(1)$ ), based upon integration of

$$\{\psi(t) (b_0\rho(t) + a_0)\}' = \psi(t) O(t^{k-1}),$$

where  $\psi(t) = \exp(-b_0 t^{k+1}/(k+1))$ .

3. Show that there are constants  $a_0, a_1, a_2, \dots$  with the following property: If  $\rho(t)$  is any real solution of

$$\rho'(t) = \rho(t) + \rho^2(t) + t^{-1}\rho^4(t)$$

in any interval  $(t_0, \infty)$ , then we have

$$\rho(t) \approx a_0 + a_1t^{-1} + a_2t^{-2} + \dots \quad (t \rightarrow \infty),$$

unless  $\rho(t)$  vanishes identically.

## INDEX

- Abel, *Abelian theorems* 139; *A.'s equation* 160  
asymptotic, behaviour 10; equivalence 10; expansion 11; series 11  
asymptotics 1  
alternating sums 49  
argument of the axis of a saddle point 84  
ARMAC 174  
axis of a saddle point 84
- Bachmann 3  
Bernoulli, *numbers* 41, 158; *polynomials* 41  
Bessel function 54, 193  
Boas 174  
Boole sum formula 51  
bounded variation 53
- class partitions 102  
closed path theorem 82  
conformal mapping 124  
continuous iteration 153, 160  
contribution of a saddle point 88  
Copson 167  
crossing of a saddle point 81
- Debye 77  
differentiation of an asymptotic formula 15, 17, 139  
direct asymptotics 134  
Dixon 72
- Euler 47, *E.'s constant* 58; *E.-MacLaurin sum formula* 40, 109  
exponential integral 13  
exponentially small 44
- Fourier series 52  
functional equations of gamma function 48  
Gauss' factorial 110  
Gamma function 46, 69, 119
- Hankel 101  
Hardy 138  
Hermite polynomials 45
- Implicit functions 21  
indirect asymptotics 134  
integration of an asymptotic formula 17  
iteration 13, 28, 31, 148, 193
- Karamata 143  
Koenigs 153  
Kronecker symbol 75
- Lagrange inversion formula 22  
Landau 3  
Laplace 1, 60, 77  
Lebesgue integrals 100  
Littlewood 138
- mountaineering 80  
multiple integrals 71
- Neumann series 192  
Newton 30, 31  
numerical, *analysis* 18, 174; *equations* 30
- oscillatory 189

- perturbations 96  
 Poincaré 11  
 Poisson's sum formula 52  
 power series 14  
 Prime numbers 2, 57, 59  
 range of a saddle point 91  
 refinement 7  
 Riccati 177  
 Riemann 77, *R.'s zeta function* 40, 42  
 roots of equations 30  
 Rouché 26, 122  
 saddle point 80  
 Schröder 152, 160  
 stability 183  
 stationary phase 84  
 steepest descent 85, 86  
 Stirling's formula 1, 42, 46, 70, 127  
 summation by parts 56
- Tauber 138, *Tauberian theorems* 139  
 theta functions 45  
 unstable 184  
 uniformity with respect to a parameter 7  
 volume of unit sphere 72
- Weierstrass 145  
 Wielandt 144  
 van Wijngaarden 175  
 Wintner 196
- Zeta function** 40, 42
- |                 |      |
|-----------------|------|
| <b>Symbols:</b> | O 3  |
|                 | o 10 |
|                 | ~ 10 |
|                 | ≈ 11 |