STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZC 40

Cursus asymptotiek.

Eindhoven 1956/57.

N.G.de Bruijn.

1957

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

Cursus ASYMPTOTIEK te Eindhoven
door

Prof.Dr N.G. de Bruijn .

1956-'57

De syllabus dient tegelijkertijd als praemanuscript voor een onder de titel "Asymptotics" te verschijnen boek.

Inhoud

De meeste hoofdstukken staan min of meer op zichzelf; alleen de hoofdstukken 4,5 en 6 hangen sterk samen. De meeste hoofdstukken zijn eenvoudig van opzet, doch lopen ook uit op moeilijke uitgewerkte voorbeelden.

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 explanations 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 would be very happy to have an entirely different method for finding information about the number, at least giving some useful approximation to it. And usually this new method gives 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 called 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 asymptotical formula, is that it belongs to a part of analysis that already has got a name: the integral calculus. The safest, but not the vaguest definition is the following one: Asymptotics is the part of analysis that considers problems of the type of those dealt with in this book.

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

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

For each n, the number n! can be evaluated without any theoretical difficulty, and the larger n, 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, the smaller its relative error becomes.

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

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

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

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

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

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

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

When proving $f(n)\to 1$, one usually produces, hidden or not, information of the form (1.1.4) with explicit construction of a suitable function $N(\varepsilon)$. It is clear that the knowledge of $N(\varepsilon)$ does mean having numerical information about f. However, when using the notation $f(n)\to 1$, this information is suppressed. So 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.

To a certain extent, it is one of the reasons of the success of analysis, that a notation has been found that suppresses that much information and still remains useful. With quite simple theorems, for instance $\lim a_n b_n = \lim a_n . \lim b_n$, it is already 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. It does not suppress a function, but only a number. That is to say, it replaces the knowledge of a number with certain properties by the knowledge that such a number exists. The O -notation suppresses much less information than the limit notation, and yet it is easy enough to handle.

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

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

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

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

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

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

The logical connections are given by

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

Now (1.2.3) is the statement expressed by the symbolism

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

There are some minor differences between the various definitions of the O-symbol which occur in the literature, but these differences are unimportant. Usually, the O -symbol is meant to represent the words "something that is less than a constant number times". Instead, we shall use it in the sense of "something that is* $\leq$ a constant number times the absolute value of". So if S is any set, and if f and φ are real or complex functions defined on S, then the formula

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

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

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

---

* "in absolute value"

** means

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.

$$\begin{aligned} x^2 &= O(x) && (|x| < 2),\\ \sin x &= O(x) && (-\infty < x < \infty),\\ \sin x &= O(1) && (-\infty < x < \infty),\\ \sin x - x &= O(x^3) && (-\infty < x < \infty). \end{aligned}$$

Quite often we are interested in results of the type (1.2.6) only on a part of the set S, especially those parts of S where the information is non-trivial. For example, with the formula $\sin x - x = O(x^3)(-\infty < x < \infty)$ the only interest lies in small values of $|x|$. But those uninteresting values of the variable sometimes give some extra difficulties, though without being essential in any respect. An example is:

$$e^x - 1 = O(x^2) \qquad (-1 < x < 1).$$

We are obviously interested in small values of x, but it is the fault of the large values of x that the formula $e^x - 1 = O(x)(-\infty < x < \infty)$ 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 it takes some trouble to find a suitable interval. Now in order to eliminate these non-essential minor inconveniences one uses a modified O-notation, which again suppresses some information. We shall explain it for the case where the interest lies in large positive values of x $(x \to \infty)$, but by obvious modifications we get similar notations for cases like $x \to -\infty$, $|x| \to \infty$, $x \to c$, $x \uparrow c$ (i.e., x tends to c from the left), etc.

The formula

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

means that there exists a real number $\underline{a}$ such that

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

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

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

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

In many cases a formula of the type (1.2.7) can be replaced immediately by an O-formula of the type (1.2.5). For if (1.2.7) is given, and if f and φ are continuous in the interval $0 \leq x < \infty$, and if moreover

$\varphi(x)\neq 0$ throughout that interval, then we have $f(x)=O(\varphi(x))(0\leqslant x<\infty)$. This follows from the fact that f/φ is continous, and therefore bounded, over $0\leqslant x\leqslant a$.

The reader should notice that as far as yet, we did not define what $O(\varphi(s))$ means, we only defined the meaning of some complete formulas. It is obvious that it 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 would arise if someone, because the sign < fails on his typewriter, starts to write =L for the words "is less than", and so he writes $3=L(5)$. Now when 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 man's L-symbol. We give a few examples:

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

The last one, for example, has to be interpreted as follows: whenever the O(1) on the left is replaced by any function $f(x)$ satisfying $f(x)=O(1)$ $(0<x<\infty)$, then $x^{-1}f(x)$ can be written as $g(x)+h(x)$, where $g(x)=O(1)$ $(0<x<\infty)$ and $h(x)=O(x^{-2})$ $(0<x<\infty)$. Its proof is easy: take $g(x)=0$ when $0<x\leqslant 1$, and $h(x)=0$ when $x>1$.

We next take a general example, meant for discussing the matter of uniformity. 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

$$(1.2.9) \qquad (f(x)+g(x))^h = O((f(x)^k) + O((g(x))^k) \qquad (x\in S).$$

For, we have

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

Formula (1.2.9) means that A and B can be found such that

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

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

On the other hand, in

$$(1.2.10) \qquad \frac{k}{x^2+k^2} = O(\tfrac{1}{x}) \qquad (1 < x < \infty).$$

the constant involved in the O-symbol can be chosen independent of $k(-\infty < k < \infty)$, as $2|kx| \leqslant x^2+k^2$ for all real values of x and k. This part is expressed by saying that (1.2.10) holds uniformly in k.

We can also look upon (1.2.10) from a different point of view. The function $k(x^2+k^2)^{-1}$ 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 (1.2.10) expresses the same thing as

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

The set S referred to in (1.2.6) specializes to the half-plane $1 < x < \infty,\ -\infty < k < \infty$.

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

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

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

but this does not hold uniformly. If it did, we would have, by specializing $k=x^2$, that $x^4(1+x^4)^{-1}=O(x^{-1})$, which is obviously false. On the other hand, *one* of the two constants can be chosen independent of k: there is a function $a(k)$ such that for each k we have $|k^2(1+kx^2)^{-1}| < 1.x^{-1}$, if only $x > a(k)$. It suffices to take $a(k)=k$.

1.3. *The o-symbol*. The expression

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

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

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

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

In asymptotics, o's are less popular 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. Asymptotical equivalence. We say that $f(x)$ and $g(x)$ are asymptotical equivalent as $x \to \infty$, if the quotient $f(x)/g(x)$ tends to zero. The notation is

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

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

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

Examples: $x \approx x + 1 \qquad (x \to \infty)$,

$\sinh x \approx \frac{1}{2}e^x \qquad (x \to \infty)$,

$n! \approx e^{-n}\, n^n \sqrt{2\pi n} \quad (n \to \infty) \quad$ (cf.(1.1.1)),

$\pi(x) \approx x/(\log x) \quad (x \to \infty) \quad$ (cf.(1.1.2)).

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

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

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

(1.5.1) $\varphi_1(x)=o(\varphi_0(x)),\ \ \varphi_2(x)=o(\varphi_1(x)),\ \ \varphi_3(x)=o(\varphi_2(x)), \ldots \quad (x \to \infty),$

$$(15.2)\quad \begin{cases} f(x) = O(\varphi_0(x)) & (x \to \infty) \\ f(x) = c_0\varphi_0(x) + O(\varphi_1(x)) & (x \to \infty) \\ f(x) = c_0\,\varphi_1(x)+c_1\varphi_1(x) + O(\varphi_2(x)) & (x \to \infty) \\ \\ f(x) = c_0\varphi_0(x) + c_1\varphi_1(x)+\ldots+c_{n-1}\,\varphi_{n-1}(x)+O(\varphi_n(x)) & (x \to \infty) \end{cases}$$

Obviously, the second formula improves the first one, as

$$c_0\,\varphi_0(x) + O(\varphi_1(x)) = (c_0+o(1))\varphi_0(x) = O(\varphi_0(x)),$$

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

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

The right hand side is called an asymptotical series for $f(x)$, or an asymptotical expansion for $f(x)$. It is easy to see that the c's are uniquely determined when the φ's are given, assuming that such an asymptotic expansion exists.

The multiplication points between c_k and $\varphi_k(x)$ are used in order to make the notation reveal the sequence $\varphi_0(x), \varphi_1(x), \ldots$. It is evident, however, that $c_k .\ \varphi_k(x)$ may be replaced by $\frac{1}{2}c_k . 2\,\varphi_k(x)$, say, since $O(\varphi_k(x)) = O(2\varphi_k(x))$. But if the coefficient is zero, we are not allowed to replace $0.\varphi_k(x)$ by $1.(0.\varphi_k(x))$, as $O(\varphi_k(x))$ cannot be replaced by $O(0.\varphi_k(x))$. Also, the meaning of (1.5.3) would change slightly upon omitting the terms with coefficients 0.

The following example shows the importance of the multiplication points:

$$(1.5.4)\qquad e^{-x} \sim 0.1 + 0.x^{-1} + 0.x^{-2} +\ldots \qquad (x \to \infty)$$

is true, as it expresses the well known fact that $e^{-x}=O(x^{-n})$ for each n. On the other hand

$$e^{-x} \sim 0.e^{-x} + 0.e^{-2x} + 0.e^{-3x} +\ldots \qquad (x \to \infty)$$

is false (e^{-x} is not $0 + O(e^{-2x})$), and, finally, the line

$$e^{-x} \sim 0 + 0 + 0 + \ldots \qquad (x \to \infty)$$

has no meaning at all.

The series occurring in (1.5.3) need not be convergent. At first sight it seems strange that such a sequence, producing sharper and sharper approximations, does not automatically converge. The answer is, that convergence means something for some fixed x_0, whereas the O-formulas (1.5.2) are not concerned with $x=x_0$, but with $x \to \infty$. Convergence of the series, for all $x > 0$, say, means that for every individual x there is a statement about the case $n \rightarrow \infty$. On the other hand, the series being 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 $f(x)$, $\varphi_0(x)$, $\varphi_1(x), \ldots$, such that the series of (1.5.3) converges for all x, but such that the sum of the series does not have 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

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

(apart from an additional constant this is the so called exponential integral Ei t). By partial integration we obtain

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

The first term is $x^{-1}e^x - e$, but the second one if of smaller order:

$$\int_1^x \frac{e^t}{t^2}\,dt = \int_1^{\frac{1}{2}x} + \int_{\frac{1}{2}x}^x < \int_1^{\frac{1}{2}x} e^t dt + \int_{\frac{1}{2}x}^x \frac{4e^t}{x^2}dt < \tfrac{1}{2}xe^{\frac{1}{2}x} + e^x . \frac{4}{x^2} = O(\frac{e^x}{x^2}).$$

Consequently we have $f(x) = x^{-1}e^x + O(x^2e^x)$.

A next approximation can be obtained if we apply partial integration to the integral in (1.6.5). Repeating this procedure, we get

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

Dealing with the last integral in the same way as it was done above with its special case $n=1$, we find that it is $O(x^{-n-1}e^x)$. It follows that

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

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

A quite simple though rather trivial class of asymptotic series consists of the class of convergent power series. If R is a positive number, and if the function $f(z)$ is, when $|z| < R$, the sum of a convergent power series

$$(1.5.7) \qquad f(z) = a_0 + a_1z + a_2z^2 + \ldots \qquad (|z| < R),$$

then we also have asymptotically

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

The proof is easy. The series converges at $z = \frac{3}{4} R$, whence the terms $a_n . (\frac{3}{4} R)^n$ are bounded. It follows that at $z = \frac{1}{2} R$ the series converges absolutely. Put

$$\sum_{k=0}^{\infty} |a_k| (\tfrac{1}{2} R)^k = A.$$

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

$$\left| \sum_{k=n+1}^{\infty} a_k z^k \right| \leq (2z/R)^{n+1} \sum_{n+1}^{\infty} |a_k| R^k \leq |2z/R|^{n+1} . A,$$

and, therefore

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

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 discussions to asymptotic series of the form

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

though 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.

For these formal power series addition and multiplication can be defined in such a way that the set of all formal power series becomes a commutative ring, with $1 + 0.x + 0.x + \dots$ as the unit element (to be denoted by I). If the series $a_0 + a_1 x + \dots$ and $b_0 + b_1 x + \dots$ are represented by A and B, respectively, then we define

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

If $a_0 \neq 0$, then there is a uniquely determined C such that AC=I.

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). It is defined as follows: Let c_{kn} be the coefficient of x^k in the series $a_0 I + a_1 B + a_2 B^2 + \dots + a_n B^n$. Then it is easily seen that $c_{kk} = c_{k,k+1} = c_{k,k+2} = \dots$. Writing $c_{kk} = c_k$, it follows that

$$a_oI+a_1B+\ldots+a_nB^n=c_o+c_1x+\ldots+c_nx^n+c_{n+1,n}x^{n+1}+c_{n+2,n}x^{n+2}+\ldots$$

We now define

$$A(B) = c_o+c_1x+\ldots+c_nx^n+c_{n+1}x^{n+1}+c_{n+2}x^{n+2}+\ldots .$$

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

A further operation on formal power series is differentiation. The derivative of $A=a_o+a_1x+\ldots$ is defined by

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

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

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 the same 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 developments

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

Notice that A(x) stands for the function, and that A stands for the formal series $a_o + a_1x + \ldots$.

Now it is not difficult to show that

(1.6.2) $\quad A(x) + B(x) \sim A+B \quad (x\to 0),$

(1.6.3) $\quad A(x)\, B(x)\sim AB \ (x\to 0),$

and if $a_o\neq 0$,

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

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

(1.6.5) $\qquad A(B(x)) \sim A(B) \qquad (x \to 0).$

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

$$A(x) = a_o+\ldots+a_nx^n+ O(x^{n+1}),\ B(x)=b_o+\ldots+b_nx^n+O(x^{n+1}) \quad (x\to 0$$

and so

$$A(x)B(x)=(a_o+\ldots+b_nx^n)(b_o+\ldots+b_nx^n)+O(x^{n+1}) \quad (x\to 0).$$

Now

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

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

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

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

With the operation of differentiation the situation is somewhat different. If $A(x)$ has the asymptotical development $A(x)\sim A(x \longrightarrow 0)$, then $A'(x)$ does not necessarily exist. If it exists, it does not necessarily have an asymptotic expansion. But if it has an asymptotical expansion, in the form of a formal power series, it automatically coincides with the formal derivative A'.

For example, we have

$$e^{-\frac{1}{x}}\sin(e^{\frac{1}{x}}) \sim 0 + 0.x^{-1} + 0.x^{-2} +\dots \qquad (x\downarrow 0),$$

but the derivative

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

has no such asymptotical expansion.

The theorem that term-by-term differentiation of an asymptotical development is legitimate whenever the derivative of the function has an asymptotical expansion (in the form of a formal power series), is an immediate consequence of the following theorem on integration (at least if the derivative is continuous):

If $f(x)$ is continuous, and

$$f(x) \sim a_0 + a_1x + a_2x^2 + \dots \qquad (x \longrightarrow 0)$$

then we have

$$\int_0^x f(t)dt \sim a_0x + \tfrac{1}{2}a_1x^2 + \tfrac{1}{3}a_2x^3 + \dots \quad (x \longrightarrow 0).$$

This immediately follows from the fact that if $g(x)$ is continuous, then

$$g(x) = O(x^n) \quad (x \longrightarrow 0) \qquad \text{implies} \qquad \int_0^x g(t)dt = O(x^{n+1}) \ (x \longrightarrow 0)$$

1.7. <u>Asymptotics and Numerical Analysis</u>. 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 it 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 asymptotical formulas only stated the existence of such numbers or functions.

In most cases, the final estimates obtained in this way are rather weak, with constant 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 asymptotical result is presented in its best possible explicit form, it need not be satisfactory from the numerical point of view. The following dialogue between a Numerical Analyst and an Asymptotical Analyst is typical in several respects.

Numer.: I want to evaluate my function $f(x)$ for large values of x, with a relative error of at most 1%.

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

Numer.: ???

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

Numer.: But my value of x is only 100.

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

$$|f(x)-x^{-1}| < 47000\ x^{-2} \qquad (x \geqslant 100).$$

Numer.: This is no news to me. I know already that $0 < f(100) < \frac{1}{2}$.

Asympt.: I can gain a little on some of my estimates. Now I find that

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

Numer.: I asked for 1%, not for 20%.

Asympt.: It is almost the best thing I possibly can get. Why don't you take larger values of x?

Numer.: !!!

I think it's better to ask my electronic computing machine.

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

Asympt.: Haven't I told you so? My estimate of 20% was not very far from the 14% of the real error.

Numer.: !!!

Some days later, the Numerical Analyst wants to know the value of f(1000). He now asks his machine first, and notices that it will require a month, working at top speed. Therefore, the Numerical Analyst returns to his asymptotical colleague, and gets a fully satisfactory reply.

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 asymptotical behaviour of $\varphi(t)$ as $t \to \infty$.

We shall only discuss a few examples, since little can be said in general. In general, the question is rather vague, for what we really want is the asymptotic behaviour of $\varphi(t)$ expressed in terms of elementary functions, or at least in terms of explicit functions.

If no one had ever introduced logarithms, the question about the asymptotical behaviour of the positive solution of the equation $e^x-t=0$ (as $t \to \infty$) would have been a hopeless problem. But as soon as one considers logarithms as useful functions, the problem vanishes entirely.

In many cases occurring in practice it is possible to express the asymptotical 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

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

then we can easily verify that $x=e^{t\,\varphi(t)}$, where $\varphi(t)$ is the solution of $\varphi e^{\varphi} =t$. Now for φ we have 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 $t \to \infty$. This would require a formula for $\psi(t)$ with an error term of $o(t^{-1})$, and it is unlikely that such a formula could 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. 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.2.5).

2.2. The Lagrange inversion formula. Let the function f(z) be analytic in some neighbourhood of the point z=0 of the complex plane. Assuming that $f(0)\neq0$, we consider the equation

$$w=z/f(z), \tag{2.2.1}$$

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

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

where the coefficients c_k are given by

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

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

(2.2.4) $$g(z) = g(0) + \Sigma_0^{\infty} d_k w^k, \quad d_k=(k!)^{-1}(d/dz)^{k-1} \{g'(z)(f(z))^k\}_{z=0}.$$

Formula (2.2.2), usually quoted as the Bürmann-Lagrange formula, is a special case of a more general theorem on implicit functions: If f(z,w) is an analytic function of both z and w, in some region $|z| < a_1$, $|w| < b_1$, 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 $z= \Sigma_{k=0}^{\infty} c_k w^k$.

For proofs of these theorems we refer to standard textbooks on complex function theory.

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

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

when $t \to \infty$. As t^{-1} tends to zero, we apply the Lagrange formula (4.2.2) to the equation $ze^z=w$, so that $f(z)=e^{-z}$. It results that there are constants $a > 0$ and $b > 0$, such that for $|w| < a$ there is only one solution z satisfying $|z| < b$, viz.

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

(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 $|x| < b^{-1}$. But as xe^x increases from 0 to ∞ if x increases from 0 to ∞, the equation (2.3.1) has a positive solution, and this one cannot exceed b^{-1} if t is sufficiently large. So if t is large enough, the positive solution we are looking for, is given by

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

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

Our second example considers the positive solution of

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

when $t \to \infty$. The function x^t is increasing if $x > 0$, and e^{-x} is decreasing. We notice that x^t is small in the interval $0 \leqslant x \leqslant 1$ 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 root, close to 1, and tending to 1 as $t \to \infty$.

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

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

The function $f(z)$ is analytic at $z=0$: $f(z)=-1 + c_1 z + \dots$

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 one if $t \to \infty$, guarantees that the positive solution is represented by the power series, if t is sufficiently large.

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

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

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

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

where $f(z)$ is analytic at $z=0$. Therefore z is a power series in terms of powers of w, and we easily evaluate $z=w+c_2 w^2+\dots$. Therefore we have, if n is large enough,

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

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

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

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

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

$$x = \log t - \log x. \tag{2.4.2}$$

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. We must have something to start with. As t tends to infinity, we may assume $t > e$, and then we have

$$1 < x < \log t,$$

as $1.e^1 = e < t$, $\log t\, e^{\log t} = t \log t > t$. It follows that $\log x = $ $= O(\log \log t)$, and so, by (2.4.2),

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

Taking logarithms, we infer that

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

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

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

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

Again taking logarithms here, and inserting the result into (4.2), we get the third approximation

$$x = \log t - \log \left\{\log t - \log \log t + O(\log \log t/\log t)\right\} =$$

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

We shall carry out two further steps. Abbreviating

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

we obtain

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

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

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

$$- \frac{1}{3}\,(L_2L_1^{-1})^3 =$$

$$= L_1-L_2+L_2L_1^{-1}+(\tfrac{1}{2}L_2^2-L_2)L_1^{-2}+ \left\{\tfrac{1}{3} L_2^3- \tfrac{3}{2} L_2^2+ O(L_2)\right\} L_1^{-3}.$$

The next step can be verified to give

$$(2.4.4)\quad x=L_1-L_2+L_2L_1^{-1}+(\tfrac{1}{2}L_2^2-L_2)L_1^{-2}+(\tfrac{1}{3} L_2^3- \tfrac{3}{2} L_2^2+L_2)L_1^{-3} +$$

$$+ (\tfrac{1}{4} L_2^4- \tfrac{11}{6} L_2^3+3L_2^2+O(L_2)L_1^{-4}.$$

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

$$(2.4.5)\quad x \sim L_1-L_2 + L_2P_0(L_2)L_1^{-1}+L_2P_1(L_2)L_1^{-2}+L_2P_2(L_2)L_1^{-3}+\dots,$$

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: the series in (2.4.5) converges if t is sufficiently large, and its sum equals x.

The method is modelled after the usual proof of the Lagrange theorem. For abbreviation, we put

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

and we obtain from (2.4.2) that

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

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

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

$$|\sigma z-\tau| < \delta \qquad (|\sigma| < a,\ |\tau|<a,\ |z|= \pi).$$

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

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

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

For every z on the integration path we have $|\sigma z|+|\tau|<\frac{1}{2}|e^{-z}-1|$, 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 σ and τ). We notice that all terms not containing τ vanish. For, in (2.4.8) the terms with m=0 give rise to integrals

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

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

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

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

where the c_{km} are constants.

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

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

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 role of iteration in asymptotics. In the next sections we shall discuss some further aspects of asymptotical 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 asymptotical behaviour is required, as $t\to\infty$. 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 $\varphi_0(t), \varphi_1(t), \ldots$ be a sequence of functions and assume that, for each separate k, the asymptotical behaviour of $\varphi_k(t)$ is known. Assume that we have reasons to believe that the behaviour of $\varphi_0(t)$ is, in some sense, an approximation to the one of f(t). Moreover assume that there is a procedure that transforms φ_0 into φ_1, φ_1 into φ_2, etc., and that there are reasons to believe that this procedure turns any good approximation into a better approximation. What we hope for is this: it might happen that for some k φ_k is so close to f, that we may be able to prove this fact, in some specified sense. It may even happen that we are able to use the procedure itself for proving things. Namely, if we are able to show that (i) if φ_n is an approximation in some n-th sense, then automatically φ_{n+1} is an approximation in some (n+1)-th sense, and if moreover (ii) for some k it can be proved that φ_k is an approximation in the k-th sense. A simple example for this is the process which led to (2.4.4). 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 had to deal with (2.4.1) under consideration of complex values of t, the first step would already be more difficult. In order to be specific, we assume that Im t=1, and that we want to have a solution x with Re $x \longrightarrow \infty$, Im $x \longrightarrow 0$. 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 asymptotical behaviour, but just a number. We shall consider things of that type in sec.2.6, and compare them to asymptotical problems in sec.2.7.

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

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

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

A sufficient condition for the existence of J is, for instance, that f(x) has a positive second derivative throughout some neighbourhood of ξ. If the process converges at all, it does so very rapidly, as (2.6.1), together with some very light extra assumptions, guarantees that $x_{n+1}-\xi$ is of the order of the square of $x_n-\xi$.

Quite often very little is known about the function f(x), that is, for every special x the value of f(x) can be found, but in larger x-intervals there is not much information about lower and upper bounds of f(x), f'(x), etc. Usually such information can be obtained in very small intervals. In order to find a root of the equation f(x)=0, we then simply choose some number x_0, more or less at random, and we construct $x_1,x_2,\ldots$ by Newton's iteration process. If this sequence shows the tendency to converge, nothing as yet has been proved, as convergence can not 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, and that there are no other roots inside J. What we then have achieved is not the exact value of a root, but a small interval in which there is one; moreover we have a procedure to find smaller and smaller intervals to which it belongs. Therefore it is a perfectly happy situation from the point of view of the numerical analyst.

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

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

4.7. Asymptotical iteration. Now returning to asymptotical 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 out for asymptotical 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 asymptotical behaviour is precisely known.

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

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

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

$$\varphi_1= \varepsilon,$$

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

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

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.4.6.

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

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

and so

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

and so on. It is clear that this leads us nowhere. None of the φ_k's have any asymptotical resemblance to the true root x.

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

Next assume that we try $\varphi_0=\log t-\log\log t+a_0$ for some constant a_0. (admittedly, this example is not very natural, as no one would try this before trying $\varphi_0=\log t-\log\log t$). Then we easily calculate that $\varphi_n=\log t-\log\log t+a_n$, where $a_{n+1}=a_n+e^{-a_n}-1$. It can be shown (see ch.8) that a_n tends to 0 quite rapidly. However, not a single φ_k of this sequence gives an approximation essentially better than $\log t-\log\log t + O(1)$.

In some sense $\log t-\log\log t$ is the limit of this sequence $\varphi_0, \varphi_1, \varphi_2, \ldots$ If we now start the Newton method anew, with $\varphi_0^*= \log t - \log\log t$, we suddenly get much better approximations.

Actually it means that we consider the small-parameter case (2.4.6), starting with zero as a first approximation to **v**.

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

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 method 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 asymptotical 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.

In case d. 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 $s_n = \sum_{k=1}^{n} k^{-3}$. 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

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

and therefore

$$(3.2.2) \qquad s_n = s + O(n^{-2}).$$

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 some other source of information. The numerical analyst would prefer to evaluate explicitly $\sum_1^m k^{-3}$ for some suitably chosen value of m, and to estimate $\sum_{m+1}^{n} k^{-3}$.

Formula (3.2.2) can be improved by refinement of the argument that led us to (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 $\sum_1^{n-[\log n]} 2^k \log n \leq 2^{n-\log n+1} \log n =$ $= 2^{n+1} n^{-1} \log n$, and this is much smaller than the n-th term.

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

$$\log k = \log n - hn^{-1} + O(h^2 n^{-2}) \qquad (h=n-k),$$

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

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

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

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

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

The main error term is $O(2^n\, n^{-2})$; the terms involving $2^{\frac{1}{2}n}$ are much smaller than this one. Our result is

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

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

$$2^{-n} \sum_{k=1}^{n} 2^k \log k - 2 \log n \sim c_1 n^{-1} + c_2 n^{-2} + c_3 n^{-3} + \dots \qquad (n \to \infty),$$

with $c_k = (-1)^k\, k^{-1} \sum_{h=1}^{\infty} h^k\, 2^{-h}$.

3.3. *Case b.* We are often confronted with sums of positive terms, where each term is much larger than, or anyway not much smaller than, the sum of all previous terms. Our example is $\sum_{k=1}^{n} k!$ Dividing by the last term, we find that

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

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

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

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

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

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

(3.3.1) $$s_n/n! \sim c_0 + c_1 n^{-1} + c_2 n^{-2} + \dots \qquad (n \to \infty).$$

This series is not convergent, that is to say, the series $c_0+c_1x+c_2x^2+\dots$ does not converge unless x=0. For, the coefficients exceed those of the expansion of

$$1 + x + \frac{x^2}{1-x} + \frac{x^3}{(1-x)(1-2x)} + \dots + \frac{x^{k+1}}{(1-x)\dots(1-kx)},$$

in terms of powers of x. It follows that the infinite series in (3.3.1) diverges at $x=k^{-1}$, and this holds for any value of k.

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 $\exp(e^x-1) = \sum_0^\infty d_k x^k$. We leave the proof to the reader [Hint: first prove, e.g. by induction, that

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

The asumptotic behaviour of d_k, as $k \to \infty$, will be studied in sec.6.2.

3.4. Case c. A typical example is

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

We have $a_{k+1}(n)/a_k(n)=\left\{2\ (n-k)/(k+1)\right\}^2$. 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 $|k-2n/3| < An^{\frac{1}{2}}$. We shall not go into this matter now, as the easiest method consists of comparison with integrals, and the integrals which arise, are of the type of those studied in ch.4.

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 $\sum_{1}^{\infty} (a_k-u_k)$ to converge. The second step deals with $\sum_{k=1}^{n} (a_k-u_k)$. The first approximation to this sum is, like in sec.3.2, the infinite sum $S=\sum_{k=1}^{\infty} (a_k-u_k)$, and we have

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

In the last sum we try to approximate u_k-a_k by a sequence v_k, such that $\sum_{n+1}^{\infty} v_k$ is explicitly known, and such that $\sum_{n+1}^{\infty}(u_k-a_k-v_k)$ 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 $\int_0^n t^{\frac{1}{2}}dt = \frac{2}{3}n^{3/2}$. If we now try to take u_k such that $\sum_{1}^{n} u_k = \frac{2}{3}n^{3/2}$, we still fail. For

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

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 $\frac{1}{4} k^{-\frac{1}{2}} + O(k^{-3/2})$, 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

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

we easily obtain that

$$(3.5.3) \qquad u_k - a_k = k^{-3/2}/48 + O(k^{-5/2})$$

whence $\sum (a_k - u_k)$ converges.

In the second step we have to approximate to $u_k - a_k$ by a sequence v_k. We first take $v_k = V_{k-1} - V_k$, where $V_k = k^{-\frac{1}{2}}/24$, $\sum_{n+1}^{\infty} v_k = V_n$, suggested by the integral $\int_n^{\infty} (t^{-3/2}/48)dt = n^{-\frac{1}{2}}/24$. We so obtain

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

The term $O(n^{-3/2})$ can be replaced by an asymptotical 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). That is easily done, for $(u_k - a_k)k^{3/2}$ 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

$$(3.5.5) \qquad S = \sum_{k=1}^{\infty} \left\{ k^{\frac{1}{2}} - \tfrac{2}{3} k^{3/2} - \tfrac{1}{2} k^{\frac{1}{2}} + \tfrac{2}{3}(k-1)^{3/2} + \tfrac{1}{2}(k-1)^{\frac{1}{2}} \right\} ,$$

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,

$$(3.5.6) \qquad \sum_{k=1}^{n} k^{-z} = n^{1-z}(1-z)^{-1} + \tfrac{1}{2} n^{-z} + S(z) + O(n^{-z-1}) \qquad (n \to \infty),$$

if $\mathrm{Re}\, z > -1, z \neq 1$. 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 $\mathrm{Re}\, z > -1$, $z \neq 1$. If $\mathrm{Re}\, z > 1$, it represents the Riemann zeta function $\zeta(z)$, as can be seen from (3.5.6) by making $n \to \infty$. Therefore $S(z) = \zeta(z)$ in the whole region.

Especially, the value of (3.5.5) is

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

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

3.6. <u>The Euler-Maclaurin sum formula</u>. Our considerations of sec.3.5 were meant to demonstrate a method, rather than giving the shortest way to deal with $\sum_1^n k^{\frac{1}{2}}$. It seems that the shortest and most efficient way of dealing with such cases depends on the Euler-Maclaurin sum formula. We do not derive the formula here, as it is incorporated in most textbooks of advanced analysis.

The formula reads as follows. Let $B_0, B_1, \dots$ denote the Bernoulli numbers, defined by $z/(e^z-1)= \Sigma_0^\infty B_n z^n/n!$ (whence $B_0=1$, $B_1=-\frac{1}{2}$, $B_2= =\frac{1}{6}, \dots, B_3=B_5=B_7=\dots=0$, $B_4=-\frac{1}{30}$, $B_6=\frac{1}{42}$. Furthermore, $B_n(t)$ denote the Bernoulli polynomials, defined by $ze^{zt}/(e^z-1)=\Sigma_0^\infty B_n(t)z^n/n!$ Let $f(x)$ have, for $x \geqslant 1$, at least $2m$ continuous derivatives. Then we have

$$(3.6.1) \quad f(1)+\dots+f(n)= \Big\{\int_1^n f(x)dx + C + \tfrac{1}{2}f(n)+B_2f'(n)/2!+B_4f'''(n)/4!+ \\ +\dots+ B_{2m}f^{(2m-1)}(n)/(2m)! - \int_1^n f^{(2m)}(x)\ B_{2m}(x- [\ \ ])/2m)!dx$$

In this formula the number C does not depend on n (it can be determined by taking $n=1$), and $[x]$ denotes the largest integer $\leqslant x$. It is known that $|B_{2m}(x)| \leqslant |B_{2m}|$ $(0 \leqslant x \leqslant 1)$, and this gives a satisfactory estimate for the integral.

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

$$(3.6.2) \quad f(1)+\dots+f(n)=\int_1^n f(x)dx + S+\tfrac{1}{2}f(n)+ \Sigma_{k=1}^m B_{2k}f^{(2k-1)}(n)/(2k)!+ \\ + O(\int_n^\infty |f^{(2m)}(x)|\,dx),$$

where

$$(3.6.3) \quad S=\tfrac{1}{2}f(1)-B_2f'(1)/2!-\dots-B_{2m}f^{(2m-1)}(1)/(2m)! - \\ \int_1^\infty f^{(2m)}(x)\ B_2(x-[x])/(2m)!dx.$$

3.7. <u>A further example</u>. Let z be a complex number, and $f(x)=x^{-z}\log x$. Then (3.6.2) can be applied if $2m > 1-\mathrm{Re}\ z$:

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

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

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

The accents denote differentiation with respect to n.

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 $\zeta'(0)=-\frac{1}{2}\log 2\pi$.

3.8. <u>A remark</u>. Roughly speaking, the Euler-Maclaurin method does not work if the largest term is <u>not</u> small compared to the sum of all terms. 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)+...+f(n)=O(f(n)). One can illustrate this by the example $\sum_1^n k!$ of sec.3.3.

3.9. 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 (3.6.1) to (3.6.2) in that case, for then S will depend on n. An unspecified constant may often be tolerated in an asymptotical 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, the reason being that $\int_1^n |f^{(2m)}(x)|dx$ is relatively small.

As an example we take $s_n=\sum_{k=-n}^{n} e^{-k^2\alpha/n}$, where α is a positive constant. The Euler-Maclaurin formula gives, if $f(x)=e^{-x^2\alpha/n}$,

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

where

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

The first term $\int_{-n}^n f(x)dx$ equals $\int_{-\infty}^{\infty}=(\pi n/\alpha)^{\frac{1}{2}}$, apart from an error which is 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 what we can do about R_m.

It is not difficult to show that $|R_m|<C_m\ n^{1-m}$, where C_m is positive and independent of n, so that, for every m,

$$(3.9.3)\qquad s_m\ =\ (\pi n/\alpha)^{\frac{1}{2}}\ +\ O(n^{1-m})\ .$$

3.10. In the case just mentioned 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 $N\to\infty$, we obtain

$$(3.10.1)\quad S_n=\sum_{k=-\infty}^{\infty} e^{-k^2\alpha/n}=(\pi n/\alpha)^{\frac{1}{2}}-\int_{-\infty}^{\infty} f^{(2m)}(x)\ B_{2m}(x-[x])/(2m)!dx.$$

We denote the latter integral by R^*; it follows from (3.10.1) that R does not depend on m. What we shall call here the Euler-Maclaurin method consists of estimating

$$(3.10.2) \quad |R^*| \leqslant \frac{B_{2m}}{(2m)!} \int_{-\infty}^{\infty} |f^{(2m)}(x)| \, dx \, .$$

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

The theta transformation formula reads

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

and therefore

$$(3.10.3) \qquad -R^* = 2(\pi n/\alpha)^{\frac{1}{2}} \, e^{-\pi^2 n/\alpha} + O(n^{\frac{1}{2}} e^{-4\pi^2 n/\alpha}).$$

We shall now investigate whether (3.10.2) can give anything as strong as this. As we remarked in sec.3.9, (3.10.2) gives immediately that, for every m, we have $R^*=O(n^{1-m})$. For m fixed, this is very much weaker than (3.10.3), but by taking m to be a suitable function of n, we can obtain a better result. There is of course no hope of proving (3.10.3), but it is interesting to see that the Euler-Maclaurin method can still show that $R^*=O(n \, e^{-\pi^2 \, n/\alpha})$, restricting the losses to the unimportant factor $n^{\frac{1}{2}}$.

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

$$(3.10.4) \quad \int_{-\infty}^{\infty} |f^{(2m)}(x)| dx = (2\alpha/n)^{m-\frac{1}{2}} \int_{-\infty}^{\infty} \left|\left(\frac{d}{dy}\right)^{2m} e^{-\frac{1}{2}y^2}\right| dy,$$

and if we adopt the definition $H_x(y)=(-1)^k \, e^{\frac{1}{2}y^2} (d/dx)^k \, e^{-\frac{1}{2}y^2}$ for the Hermite polynomials, the integrand in the latter integral is $e^{-\frac{1}{2}y^2}|H_{2m}(y)|$.

Using the integral representation

$$H_{2m}(y) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} (y+iu)^{2m} \, e^{-\frac{1}{2}u^2} \, du$$

we infer that

$$\int_{-\infty}^{\infty} |f^{(2m)}(x)| dx \leqslant (2\alpha/n)^{m-\frac{1}{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y^2+u^2)} (y^2+u^2)^m \, du \, dy.$$

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 (3.10.2) is known to be equal to $2(2\pi)^{-2m} \zeta(2m)$, and therefore it is $< C(2\pi)^{-2m}$, where C is an absolute constant. It now follows from (3.10.2) that

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

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

$$(3.10.5) \qquad |R^*| < C_1 \quad (\alpha m/\pi^2 ne)^m \quad (nm/2\alpha)^{\frac{1}{2}}.$$

It is now the right moment for fixing the value of m. The minimum of $(\alpha t/\pi^2 ne)^t$ 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 $m=m_0$, (3.10.5) becomes

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

1 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

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

and we ask for the asymptotic behaviour of S(t) as $t\to\infty$. The function f(x) is decreasing, and tends to zero as $x\to\infty$. Therefore the sum 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 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

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

Using the Taylor series, we can express both in terms of f(2k), f'(2k),... . We have, if we stop the Taylor developments at the terms involving f",

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

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

On subtraction we find

$$(3.11.2) \qquad \left| f(2k)-f(2k+1)+\tfrac{1}{2}\int_{2k}^{2k+2} f'(x)dx \right| \leqslant \tfrac{1}{2}\int_{2k}^{2k+2} |f''(x)|\,dx \;.$$

In our case we have $f(x)\to\infty$, as $x\to\infty$, and therefore

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

It follows that

$$\left| S(t) - \tfrac{1}{2}f(0)\right| \leqslant \tfrac{1}{2}\int_0^{\infty} f''(x)\,dx.$$

We have $f''(x)=(2x^2-t^2)/(x^2+t^2)^{5/2}$. We transform the integral, substituting $x=yt$:

$$\int_0^{\infty}|f''(x)|\,dx = t^{-2}\int_0^{\infty}|1-2y^2|/(1+y^2)^{5/2}dy,$$

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

$$(3.11.4) \qquad S(t) = \tfrac{1}{2}t^{-1} + O(t^{-2}) \qquad (t\to\infty).$$

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

$$(3.11.5) \qquad \left| f(2k)-f(2k+1)+\tfrac{1}{2}\int_{2k}^{2k+2} f'(x)dx - \tfrac{1}{4}\int_{2k}^{2k+2} f''(x)dx \right| \leqslant C \;{}_{2k}^{2k+2}|f'''(x)|\,dx.$$

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

$$(3.11.6) \qquad S(t)= \tfrac{1}{2}t^{-1} + O(t^{-3}) \qquad (t\to\infty).$$

The fact that no term t^{-2} occurs is due to the circumstance that $f(x)$ is even, and for the same reason the coefficients of $t^{-3}, t^{-4}, \ldots$ will all vanish. In order to show this, it is easier to put the series in the following form:

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

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

$$S(t)-\tfrac{1}{2}t^{-1}=-\tfrac{1}{2}\int_{-\infty}^{\infty}f'(x)dx - \tfrac{1}{4}\int_{-\infty}^{\infty}f''(x)dx + O\left(\int_{-\infty}^{\infty} f'''(x)\,dx\right).$$

As $f(x)\to 0$, $f'(x)\to 0$, $f''(x)\to 0, \ldots$ as $x\to\infty$, we have

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

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

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

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

$$(3.11.6) \qquad S(t) \sim \tfrac{1}{2}t^{-1} + 0.t^{-2} + 0.t^{-3} + 0.t^{-4} + \ldots \qquad (t \to \infty).$$

It may be remarked that the general formula of which (3.11.2), (3.11.3) are special cases, is related in a trivial way to the Boole sum formula, which we shall not discuss here.

With (3.11.6) we have the same situation as in sec.3.10. 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.

In many textbooks the Euler-Maclaurin formula is derived from the Poisson sum formula, and actually a similar proof can be given for the Boole formula. Quite often, however, it happens that Poisson-formula itself gives better results than Euler-Maclaurin's or Boole's (unless, of course, one does not interprete these as inequalities, like (3.10.2), but as equalities, like (3.10.2); in the latter case one can return to Poisson by Fourier expansion of the Bernoulli polynomials.

3.12. Application of the Poisson sum formula. The formula reads

$$(3.12.1) \qquad \sum_{k=-\infty}^{\infty} f(k+a) = \sum_{\nu}^{*} e^{2\pi\nu i a} \int_{-\infty}^{\infty} e^{-2\pi\nu i y} f(y)dy.$$

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

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

The following set of conditions is easily seen to be sufficient.

(i) $\sum_{k=-\infty}^{\infty} f(k+x)$ converges uniformly for $0 \leqslant x \leqslant 1$.

(ii) The function $\phi(x)=\sum_{-\infty}^{\infty} f(k+x)$, which has the 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 $\phi(x)$:

$$\int_0^1 e^{-2\pi\nu iy}\,\phi(y)dy=$$

$$=\int_0^1 \Sigma_{k=-\infty}^{\infty} e^{-2\pi\nu iy}\, f(k+y)dy=\Sigma_{k=-\infty}^{\infty}\int_0^1=$$

$$=\Sigma_{k=-\infty}^{\infty}\int_k^{k+1} e^{-2\pi\nu iy} f(y)dy=\int_{-\infty}^{\infty} e^{-2\pi\nu iy}\, f(y)dy.$$

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

$$\int_k^{k+t} e^{-2\pi\nu iy}\, f(y)dy\to 0$$

as $k\to\infty$, uniformly in $0\leqslant t\leqslant 1$.

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

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

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

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

For, (iii)+(v) imply (i), (apply the mean value theorem to finite sums $\Sigma_{-N}^{M} f(k+x)$) 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 total variation over $-\infty< x<\infty$,

(vii) $\lim_{h\to 0}\{f(x+h)+f(x-h)\}=2f(x)$, 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), as well as the fact that $\phi(x)$ has bounded total variation over $0\leqslant x\leqslant 1$; (vi)+(vii) can be used to show that $\lim_{h\to 0}\{\phi(a+h)+\phi(a-h)\}=2\phi(a)$. This formula, in combination with the fact that $\phi(x)$ has bounded total 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 function sum

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

whence, by (3.11.1)

(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) has got the special value 0 here, and in applying (3.12.1) to (3.12.2), t is considered to be a fixed positive number.

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

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

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

$$\sum_{k=N}^{M} f'(x+k) = e^{\pi i x}\sum_{k=N}^{M}(-1)^k\left\{\pi i.((x+k)^2+t^2)^{-\frac{1}{2}}-(x+k)((x+k)^2+t^2)^{-3/2}\right\}$$

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

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

It follows that

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

As this holds uniformly in x ($0 \leq x \leq t$), we infer that, for t fixed, $\sum_1^{\infty} f'(x+h)$ 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

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

and so we have to study, for $b=\pm\pi, \pm3\pi, \pm5\pi, \dots,$

$$\phi(b,t) = \int_{-\infty}^{\infty} e^{biy}(y^2+t^2)^{-\frac{1}{2}}dy.$$

On substitution of $y=tx$ the integral becomes

$$(3.12.5) \qquad \phi(b,t) = \int_{-\infty}^{\infty} e^{btix}\,(x^2+1)^{-\frac{1}{2}}dx,$$

and on substitution of $x=-x$ we infer that ϕ is an even function of b. The integral is a Bessel function of zero order, of second kind and of imaginary argument, and in the standard notation (see Watson, Bessel functions, p.172)

$$\phi(b,t) = 2\,K_0(bt).$$

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

The integral (3.12.5) can be transformed in a well-known way, by Cauchy's theorem. If $b > 0$, we can deform the integration path $(-\infty,\infty)$ into a path leading from $i\infty$ to i along the imaginary axis, encircling i in the positive sense, and finally back from i to $i\infty$. It results that

$$\phi(b,t) = 2\int_1^\infty e^{-btz}(z^2-1)^{-\frac{1}{2}}dz = 2e^{-bt}\int_0^\infty e^{-btz}(z^2+2z)^{-\frac{1}{2}}dz$$

We have, if $t > 1$, $b \geqslant \pi$

$$\int_0^\infty e^{-btz}(z^2+2z)^{-\frac{1}{2}}dz < \int_0^\infty e^{-\pi z}\, z^{-\frac{1}{2}}dz = 1,$$

and so $\left|\phi(b,t)\right| < 2e^{-bt}$ $(b \geqslant \pi,\ t \geqslant 1)$. So by (3.12.4) we have

$$S_1(t) = 2\,\phi(\pi,t) + 2\,\phi(3\pi,t)+\dots ,$$

and therefore

$$(3.12.6) \qquad S_1(t) = 2\,\phi(\pi,t) + O(e^{-3\pi t})$$

It remains to find the asymptotic behaviour of $\phi(\pi,t)$, which is quite easy. If we bear in mind that if t is large, the integrand of $\int_0^\infty e^{-btz}(z^2+2z)^{-\frac{1}{2}}\,dz$ is very small for $z > 1$, say, it seems worth while to develop the factor $(z+2)^{-\frac{1}{2}}$ into a power series

$$(3.12.7) \qquad (2+z)^{-\frac{1}{2}} = 2^{-\frac{1}{2}}\sum_{n=0}^\infty c_n z^n \qquad (c_0=1)$$

valid for $0 \leqslant z \leqslant 2$, with $c_0=1$. We break off somewhere, that is, we choose an integer $M > 0$ and we deduce that

$$(3.12.8) \qquad \left|(2+z)^{-\frac{1}{2}} - 2^{-\frac{1}{2}}\sum_{n=0}^{M-1} c_n z^n\right| < Cz^M$$

where C depends on M only. This holds for all $z \geqslant 0$, if C is suitably chosen: in the interval $0 \leqslant 2 \leqslant 1$ it holds by virtue of (3.12.7), and in $1 \leqslant z < \infty$ it holds because each term on the left of (3.12.8) is $O(z^M)$, while the number of terms is fixed.

Now using (3.12.8), we obtain

$$\phi(\pi,t) = 2e^{-\pi t}.\ 2^{-\frac{1}{2}}\int_0^\infty e^{-\pi t}\left[\sum_{n=0}^{M-1} c_n z^n + O(z^M)\right] z^{-\frac{1}{2}}dz =$$
$$= 2e^{-\pi t}(2\pi t)^{-\frac{1}{2}}\left\{\sum_{n=0}^{M-1} c_n \Gamma(n+\tfrac{1}{2})\pi^{-n}t^{-n}. + O(t^{-M})\right\}.$$

This means that we have an asymptotic series for $\phi(\pi,t)$. 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 expansion. So our final result is, as $c_n=(-1)^n 2^{-3n}(2n)!(n!)^{-2}$

$$S(t)-\tfrac{1}{2}t^{-1} = \tfrac{1}{2}S_1(t) \sim e^{-\pi t}\sum_{n=0}^\infty (-1)^n t^{-n-\frac{1}{2}} 2^{\frac{1}{2}-5n}\pi^{-n}\left\{(2n)!\right\}^2\left\{n!\right\}^{-3} \qquad (n \to \infty).$$

3.13. Partial summation. We often meet the question of the asymptotical behaviour, as $n \to \infty$, of a sum $a_1b(1)+\dots+a_nb(n)$, 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 partial summation formula

$$(3.13.1)\quad a_1b(1)+\ldots+a_nb(n)=(a_1+\ldots+a_n)b(n)-$$
$$\left[a_1(b(2)-b(1))+(a_1+a_2)(b(3)-b(2))+\ldots.(a_1+\ldots+a_{n-1})(b(n)-b(n-1))\right].$$

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 $A(x)=\sum_{1\leq k\leq x} a_k$ (i.e. $A(x)=0$ if $x<1$, $A(x)=a_1+\ldots+a_{[x]}$ if $x\geqslant 1$). Then (3.13.1) becomes

$$(3.13.2)\quad a_1b(1)+\ldots+a_nb(n)=A(n)b(n)-\int_1^n A(x)\ b'(x)\ dx,$$

that is, a special case of the formula for integration by parts in the theory of Stieltjes integrals:

$$(3.13.3)\quad \int_0^n b(x)\ d\ A(x) = \left[A(x)b(x)\right]_0^n - \int_0^n A(x)d\ b(x),$$

but we need not discuss the conditions for (3.13.3) in general.
We shall discuss an example 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 ($m=1,2,3,\ldots$), we have

$$(3.13.4)\qquad \vartheta(x) = x + O(x(\log x)^{-m}) \qquad (x\to\infty)$$

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 approached by partial summation. We consider $\sum_{p\leq x} 1$, i.e. the number of primes not exceeding x, and usually denoted by $\pi(x)$. We have, by (3.12.2),

$$(3.13.5)\quad \pi(x) = \int_{3/2}^x (\log u)^{-1}\ d\,\vartheta(u) = x(\log x)^{-1}\,\vartheta(x) -$$
$$\int_{3/2}^x \vartheta(u)\ d(\log u)^{-1}.$$

(We have replaced the lower limit by 3/2, since $(\log u)^{-1}$ is singular at $u=1$; it makes no difference, as $\vartheta(u)=0$ for $u<2$). We compare this with

$$\int_{3/2}^x (\log u)^{-1}\ du = \left[(\log u)^{-1}u\right]_{3/2}^x - \int_{3/2}^x u\ d(\log u)^{-1}.$$

On subtraction we obtain, using (3.13.4),

$$\pi(x) - \int_{3/2}^x (\log u)^{-1}du = (\log x)^{-1}\ O(x(\log x)^{-m}) +$$
$$+\int_{3/2}^x\ O(u(\log u)^{-m})(\log u)^{-2}u^{-1}\ du\ .$$

The integral on the right can be written as

$$\int_{3/2}^{x^{\frac{1}{2}}} O(1)du + \int_{x^{\frac{1}{2}}}^x O\left\{(\log x^{\frac{1}{2}})^{m-2}\right\}du = O(x(\log x)^{-m-2}),$$

and therefore

$$\pi(x)-\int_{3/2}^{x}(\log u)^{-1}\,du = O\{x(\log x)^{-m-1}\} \qquad (x\to\infty).$$

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

$$(3.13.6)\quad \pi(x)\sim x\log^{-1}x + x\log^{-2}x + 2!\,x\log^{-3}x + 3!\,x\log^{-4}x + \dots \qquad (x\to\infty).$$

Meanwhile we notice that (3.13.4) is an example of the situation described in sec.3.9 and 3.11. Again there is an asymptotical expansion with zero coefficients:

$$e^{-y}\,\vartheta(y)-1\sim 0.y^{-1} + 0.y^{-2} + 0.y^{-3} + \dots \qquad (y\to\infty),$$

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

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 $t \to \infty$. 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

$$(4.1.1) \qquad F(t) = \int_{-\infty}^{\infty} f(x,t)\, dx \qquad (t \to \infty)$$

It often occurs that the graph of f(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 f 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.

It is by no means necessary 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 $x=ay+b$ in the integral, where a and b may depend on t.

A simple example of the method was already given at the end of sec.3,12. As a second example we roughly sketch how to deal with the integral

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

If t is large, the integrand is very small unless x is very close to 0. The function $\log(1+x+x^2)$ can be succesfully 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

$$\log(1+x+x^2)=\log\left\{(1-x^3)/(1-x)\right\} = x+\tfrac{1}{2}x^2-2x^3/3 + O(x^4) \qquad (-\tfrac{1}{2} < x < \tfrac{1}{2}).$$

So we are led to approximate F(t) by

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

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, 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 $(-\frac{1}{2},\frac{1}{2})$.

In the above argument the idea of comparing $(-\infty,\infty)$ with $(-\frac{1}{2},\frac{1}{2})$ 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

$$\log(1+x+x^2) - x-\tfrac{1}{2}x+2x^3/3 = g(x).$$

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

$$g(x) = O(x^4) \qquad (-\infty < x < \infty).$$

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

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

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

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

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

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

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

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:

$$(4.1.7) \qquad \int_{-\infty}^{\infty}\left|e^{-tx^2} \, x^k\right| dx = O\left\{(\mathrm{Re} \ t)^{-\frac{1}{2}(k+1)}\right\} \quad (\mathrm{Re} \ t > 0; \ k=0,1,2,\dots$$

This estimate is not uniform with respect to k.

The simpler formulas involving $e^{-tx} \, x^k$ will also be frequently applied:

$$(4.1.8) \qquad \int_0^{\infty} e^{-tx} \, x^k \, dx = t^{-k-1}k! \qquad (k=0,1,2,\dots; \ \mathrm{Re} \ t > 0),$$

$$(4.1.9) \qquad \int_0^{\infty}\left|e^{-tx} \, x^k\right| dx = O((\mathrm{Re} \ t)^{-k-1}) \qquad (k=0,1,2,\dots; \ \mathrm{Re} \ t > 0),$$

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

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. For, both width and locolization 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 $\int_{-\infty}^{\infty}-\int_J$ is small compared to $\int_J$, 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 a substitution $x=\varphi(y)$, the integral (4.1.1) becomes $\omega(t)\int_{-\infty}^{\infty}g(y,t)dy$, and that $g(y)=\lim_{t\to\infty}g(y,t)$ 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 have immediately $F(t)\sim\omega(t)\int_{-\infty}^{\infty}g(y)dy$, and further approximations can often be obtained from a closer study of the difference $g(y,t)-g(y)$.

4.2. <u>A general case</u>. We shall consider

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

assuming that $\psi(x)$ is real and continuous, that $\psi(x)$ has just one absolute maximum that $\int_{-\infty}^{\infty}e^{\,\psi(x)}dx$ converges, and that $\varphi(x)\to-\infty$ if $x\to\pm\infty$. Without loss of generality we may assume that the maximum is attained at $x=0$, and that $\psi(0)=0$. Furthermore we assume that $\psi'(x)$ exists in a neighbourhood of $x=0$, that $\psi''(0)$ exists, and finally that $\psi''(0)<0$.

It follows from our assumptions that there exists, to any $\delta>0$, a positive number $\eta(\delta)$ such that $\psi(x)\leqslant-\eta(\delta)$ both in $-\infty<x<-\delta$ and in $\delta<x<\infty$.

The contribution of these intervals satisfies

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

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

$$(4.2.2)\qquad |\psi(x)-\tfrac{1}{2}x^2\psi''(0)|\leqslant\varepsilon x^2 \qquad (-\delta\leqslant x\leqslant\delta).$$

For if a continuous function $\varphi(x)$ satisfies $\varphi(0)=\varphi'(0)=\varphi''(0)$, we have $\varphi(x)=x\,\varphi'(\theta x)=x\,o(x)$ $(x\to 0)$.

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

$$\int_{-\delta}^{\delta}e^{\frac{1}{2}tx^2(\psi''(0)-2\varepsilon)}dx<\int_{-\delta}^{\delta}e^{t\,\psi(x)}dx<\int_{-\delta}^{\delta}e^{\frac{1}{2}tx^2(\psi''(0)+2\varepsilon)}dx .$$

All three integrals $\int_{-\delta}^{\delta}$ differ from the corresponding integrals by an amount $O(e^{-t\alpha})$, where α is positive and independent of t. 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

$$(4.2.3)\quad \int_{-\infty}^{\infty} e^{-tf(x)}dx < (2\pi)^{\frac{1}{2}}(-\psi''(0)-2\varepsilon)^{-\frac{1}{2}}\; t^{-\frac{1}{2}}+O(e^{-t\alpha}),$$

and that there is a similar estimate below. The number ε being arbitrary, it follows that

$$(4.2.4)\qquad \int_{-\infty}^{\infty} e^{t\,\psi(x)}dx \sim (2\pi)^{\frac{1}{2}}(-t\,\psi''(0))^{-\frac{1}{2}} \qquad (t\to\infty).$$

If the restriction $\psi(0)=0$ is dropped, we of course get an extra factor $e^{t\,\psi(0)}$ 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 $\psi'=0$ at that point. If, however, the interval is finite, and if the minimum of ψ is attained at the left end-point, say, then ψ' will be usually > 0 at this point.

So assume that we have to deal with

$$\int_{0}^{\infty} e^{t\,\psi(x)}dx\ ,$$

where $\psi(x)$ is real and continuous in $0\leqslant x\leqslant 1$, attains its maximum at $x=0$, $\psi'(0)$ exists and $\psi'(0) < 0$. Moreover we assume that $\psi(x) < \psi(0)$ $(x>0)$, and that $\int_0^\infty e^{\psi(x)}dx$ 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

$$\left|\psi(x) - \psi(0) - x\psi'(0)\right| \leqslant \varepsilon x \qquad (0\leqslant x\leqslant \delta).$$

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

$$\int_{0}^{1} e^{t\,\psi(x)}dx \sim (-t\,\psi'(0))^{-1}\, e^{t\,\psi(0)} \qquad (t\to\infty).$$

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

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 $x\to 0$) about $\psi(x)$ leads to more asymptotic information about the integral (as $t\to\infty$). We shall restrict ourselves to the case of sec.4.2, the case of sec.4.3 being analogous.

For simplicity we shall assume now that $\psi(x)$ is, in some interval $-\delta\leqslant x\leqslant\delta$, the sum of a convergent power series $\psi(x)=a_2x^2+a_3x^3+\ldots$, and that $a_2<0$. We shall discuss the integral

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

where $g(x)$ is an integrable function, being in $-\delta \leqslant x \leqslant \delta$, the sum of a convergent power series $g(x)=b_0+b_1x+b_2x^2+\dots$.

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

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

not bothering about how such information can be obtained. And we shall assume the existence of two positive numbers η and C such that

(4.4.3) $$\psi(x) \leq -\eta x^2,\ |g(x)| < C\quad (-\delta \leq x \leqslant \delta).$$

In the integrand we shall consider $\exp(ta_2x^2)$ as the main factor. The remaining factor $g(x)\exp t\ x^3\{(a_3+a_4x+a_5x^2+\dots)\}$ 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 t.

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 $|x| \leqslant t^{-1/3}$. We may assume that t is so large that $t^{-1/3} \leqslant \delta$.

We first notice that the intervals $(t^{-1/3},\delta)$ and $(-\delta,-t^{-1/3})$ can be neglected. It is not difficult to show that, if $\tau=t^{-1/3}$, $\eta>0$, η fixed,

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

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 $N \geqslant 0$ we have

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

The factor x^N gives no difficulties. For, if $x \geqslant \tau$, $t>2/\eta$ we have $\eta tx \geqslant 2$, and so $x^N=O(e^x)=O\left\{\exp(\tfrac{1}{2}\eta tx^2)\right\}$. So (4.4.5) follows from (4.4.4), replacing η by $\tfrac{1}{2}\eta$.

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

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

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

$$(4.4.6) \qquad P - P_A = O((tx^3)^{A+1}) + O(x^{A+1}) \ ,$$

uniformly with respect to x and t. This step requires some explanation. If a double power series $\sum_{m \geqslant 0, n \geqslant 0} c_{mn} z^m w^n$ 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 $|z| < \frac{1}{3}R$, $|w| < \frac{1}{3}S$,

$$\sum_{m \geqslant 0,\ n \geqslant 0,\ m+n > A} c_{mn} z^m w^n = O(\sum |z/R|^m \ |w/S|^n) =$$

$$= O(\sum_{k=A+1}^{\infty} (|z/R| + |w/S|)^k) = O\left\{(|z/R| + |w/S|)^{A+1}\right\} =$$

$$= O\left\{(|z| + |w| \ )^{A+1}\right\} = O(|z|^{A+1}) + O(|w|^{A+1}).$$

For the last step, cf. (1.2.9). It should be stressed that the estimates are not uniform with respect to A.

As to (4.4.6) we remark that P is continuous if $-\delta \leqslant x \leqslant \delta$, $|tx^3| \leqslant 1$, so that the formula just proved for a smaller region can be extended to this larger range (cf. sec.1.2).

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

$$\left\{ \int_{\infty}^{\infty} - \int_{-\tau}^{\tau} \right\} P_A \exp\ (ta_2 x^2) dx = O\left\{ t^A \exp\ (-\tfrac{1}{2}\eta t^{1/3}) \right\} \qquad (t \to \infty),$$

a_2 being negative. And, if we combine (4.4.3) and (4.4.4),

$$\left\{ \int_{-\delta}^{\delta} - \int_{-\tau}^{\tau} \right\} g(x) e^{t\,\psi(x)} dx = O\left\{\exp(-\eta t^{1/3})\right\} \qquad (t \to \infty).$$

Hence, by (4.3.2) and (4.3.4), we have, for each positive integer M,

$$(4.4.7) \qquad \int_{-\infty}^{\infty} g(x) e^{t\,\psi(x)} dx - \int_{-\infty}^{\infty} P_A \exp\ (ta_2 x^2) dx = O(t^{-M}) +$$

$$+ O\left\{ \int_{-\infty}^{\infty} \exp(ta_2 x^2) \left\{ |tx^3|^{A+1} + |x|^{A+1} \right\} dx \right. \qquad (t \to \infty).$$

For these integrals we refer to (4.1.5) and (4.1.6). By virtue of (4.1.7), the last O-term in (4.4.7) is easily seen to be $O(t^{-\frac{1}{2}A-1})$. So we obtain

$$\int_{-\infty}^{\infty} g(x) e^{t\,\psi(x)} dx = \sum_{m \geqslant 0, n \geqslant \text{[illegible]},\ m+n \leqslant A} c_{mn} \varepsilon_{m+n} \ t^{-\frac{1}{2}(m+n+1)} \ \times$$

$$\times (-a_2)^{-\frac{1}{2}(3m+n+1)} \Gamma\left\{\tfrac{1}{2}(3m+n+1)\right\} + O(t^{-\frac{1}{2}A-1}) + O(t^{-1})$$

$$(t \to \infty),$$

where ε_{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

$$(4.4.8) \qquad \int_{-\infty}^{\infty} g(x) e^{t\,\psi(x)} dx \sim \sum_{\nu=0}^{\infty} d_\nu \ t^{-\frac{1}{2}-\nu} \qquad (t \to \infty),$$

where

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

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

$$\{-2\pi/t\,\psi''(0)\}^{\frac{1}{2}}\, g(0).$$

We assumed above that ψ 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

$$\psi(x) \sim a_2x^2+a_3x^3+\dots, \quad g(x)\sim b_0+b_1x+b_2x^2+\dots \qquad (x\to\infty)$$

in the sense of asymptotic series.

4.5. Asymptotical behaviour of the gamma function. If $t>-1$, the function $\Gamma(t+1)$ is defined by

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

We shall apply the result of sec.4.5 to the problem of the behaviour of $\Gamma(t+1)$ as $t\to\infty$.

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

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

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

$$\Gamma(t+1)\sim e^{-t}\, t^{t+\frac{1}{2}}\,(2\pi)^{\frac{1}{2}} \qquad (t\to\infty).$$

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. It is based upon the idea explained at the end of sec.4.1. We introduce a new integration variable z by $\frac{1}{2}z^2 = -\psi(x)$, or, more precisely,

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

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 δ and c, both independent of t, such that

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

and the method of sec.4.4 leads to

$$(4.5.2)\qquad (2\pi)^{\frac{1}{2}}\Gamma(t+1)e^{t}\, t^{-t-\frac{1}{2}} \sim 1+3c_3\cdot\frac{2}{t}\cdot\frac{2!}{1!2^2} + 5c_5\left(\frac{2}{t}\right)^2\frac{4!}{2!2^4} + \dots \qquad (t\to\infty).$$

Usually, e.g. by application of the Euler-Maclaurin method to

$$\Gamma(t+1) = \lim_{n\to\infty} n!n^t/(1+t)(2+t)\dots(n+t),$$

(cf. sec.3.7) one derives the asymptotic series for the logarithm of the left-hand-side of (4.5.2):

$$(4.5.3)\qquad \log\left\{(2\pi)^{-\frac{1}{2}}\Gamma(t+1)e^{t}\, t^{-t-\frac{1}{2}}\right\} \sim \frac{B_2}{1.2t} + \frac{B_4}{3.4t^3} + \frac{B_6}{5.6t^5} + \dots \qquad (t\to\infty).$$

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, then 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

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

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

$$(4.6.1)\qquad \psi(x_1,\dots,x_n)=-\tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j + \\ +o(x_1^2+\dots+x_n^2)(x_1^2+\dots+x_n^2\to 0),$$

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

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

$$(4.6.2)\qquad F(t)\sim A\ t^{-\frac{1}{2}n}(t\to\infty),\ A=\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}\exp(-\tfrac{1}{2}\sum\sum a_{ij}x_ix_j)dx_1\dots dx_n.$$

It is well-known that

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

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 f admits a development into powers of $x_1,\dots,x_n$, we can obtain an asymptotical 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 described at the end of sec.4.5. If $y\geq 0$, let $\phi(y)$ denote the volume of that part of the cube $-1\leq x_1\leq 1,\dots,-1\leq x_n\leq 1$ which satisfies $\psi(x_1,\dots,x_n)\geq -\frac{1}{2}y^2$. Then we have

$$F(t)=\int_0^{\infty} e^{-\frac{1}{2}ty^2}\,d\phi(y),$$

and so the problem has been reduced to a question about a single integral. Usually $\phi(y)$ will be differentiable, and $\phi'(y)\sim ny^{n-1}\ D^{-\frac{1}{2}}V_n(y\to 0)$, where V_n is the volume of the n-dimensional unit sphere. For the main term we now obtain

$$F(t)\sim n\ D^{-\frac{1}{2}}\ V_n\int_0^{\infty}e^{-\frac{1}{2}ty^2}\ y^{n-1}\ dy = n\,D^{-\frac{1}{2}}\,V_n.2^{-1+\frac{1}{2}n}\Gamma(\tfrac{1}{2}n)\ t^{-\frac{1}{2}n}.$$

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. <u>An application</u>. We shall discuss an instructive example of the multidimensional Laplace method. We consider the sum

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

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 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 upon the existence of such a formula is to determine the asymptotic behaviour of S(s;n) as $n \to \infty$ (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,... . 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 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 $2^{2n}(\pi n)^{-\frac{1}{2}}$). This means, roughly, that the Euler-Maclaurin method (in the version of sec.3.11, because of the alternating signs) give a result of the type $S/t_n \sim 0+0.n^{-1}+0.n^{-2}+\ldots$, 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 S(s,n) equals the coefficient of $z_1{}^0 z_2{}^0 \ldots z_r{}^0$ in the product

$$(-1)^n(1+z_1)^{2n}(1+z_2)^{2n}\ldots(1+z_r)^{2n}\left\{1-(z_1\ldots z_r)^{-1}\right\}^{2n}, \text{ where } r=s-1.$$

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

By Cauchy's formula we have

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

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

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

$$(4.7.2) \qquad S(r+1,n) = 2^{2rn+2n}\pi^{-r}\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi}\ldots\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi}\cos\varphi_1\ldots\cos\varphi_r\sin(\varphi_1+\ldots+\varphi_r)^{2n} d\varphi_1\ldots d\varphi_r,$$

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 extremal points of G. As G=0 on the boundary of the cube $-\frac{1}{2}\pi \leqslant \varphi_1 \leqslant \frac{1}{2}\pi, \dots, -\frac{1}{2}\pi \leqslant \varphi_n \leqslant \frac{1}{2}\pi$, 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

$$(4.7.3) \qquad \partial G/\partial\varphi_j = -\{\tan\varphi_j + \cot(\varphi_1+\dots+\varphi_r)\}\, G \quad (j=1,\dots,r).$$

Hence our conditions implies that $\tan\varphi_1 = \dots = \tan\varphi_r$. The φ_j's being restricted to the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ it follows that all φ_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 ν is an odd integer, $|\nu| < s$. The values 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 Ω of $(\beta,\dots,\beta)$

$$G(\varphi_1,\dots,\varphi_r) = G(\beta,\dots,\beta)\exp(2n\psi(\beta+x_1,\dots,\beta+x_r)),$$

and we have to deal with

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

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

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

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

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

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 δ_{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 substraction 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

$$S(s,n)\sim 2^{2rn+2n}\pi^{-r}\, 2.(2\pi)^{\frac{1}{2}r}\, D^{-\frac{1}{2}}.(2n)^{-\frac{1}{2}r}.\{G(\beta,\dots,\beta)\}^{2n} \quad (n\to\infty),$$

$$D = s\cos^{-2r}(\pi/2s), \quad G(\beta,\dots,\beta) = \cos^{s}(\pi/2s),$$

and finally

$$S(s,n) \sim \{2\cos(\pi/2s)\}^{2ns+s-1}\, 2^{2-s}\, (\pi n)^{\frac{1}{2}(1-s)} s^{-\frac{1}{2}} \quad (n\to\infty). \tag{4.7.4}$$

As a verification we take s=3, then we find

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

in accordance with Dixon's formula.

5. The saddle point method.

5.1. The method. The saddle point method is one of the most important and powerful methods in asymptotics. Its object is to obtain useful approximations to integrals in the complex plane.

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

where P is the integration path, and $\varphi(z)$ is analytic along P. We assume that both the integration path and the function φ depend on a parameter t, and we want to study the asymptotic behaviour of F(t) as $t \to \infty$.

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, fit for application of the second stage.

(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 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,

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

where $\max_P |\varphi(z)|$ is the maximum of $|\varphi(z)|$ along P, and l_P is the length of 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 φ. We shall call these paths C admissible.

We now wish to determine C such that the value of

$$l_C \max_C |\varphi(z)| \tag{5.1.3}$$

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, any way 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 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

$$\max_C \quad |\varphi(z)| \tag{5.1.4}$$

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

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.4.4).

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 endpoints of the path, or they turn out to be saddlepoints, 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. Consider the surface in three dimensional x-y-w-space, whose equation is $w = |\varphi(x+iy)|$. 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

$$\frac{\partial|\varphi|}{\partial x} - i\,\frac{\partial|\varphi|}{\partial y} = \frac{|\varphi|}{\varphi}\,\frac{d\varphi}{dz}\,.$$

(for simplicity, we shall forget a moment about points where $\varphi(x+iy)=0$, and where there is not always a tangent plane). 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 small 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 road. 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 $w = \varphi(x,y)$. 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 guaranteed. 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 altitude is $\leq h$. 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 k is this smallest value of h, and if P is a path from A to B entirely in R_k, 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 mountainer would do would be to look up the altitudes of the various passes 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. It is not difficult to show that in our case ($w = |\varphi(x+iy)|$) a sufficient condition for a path to cross the saddle point is having a tangent at that point.

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, then this saddle point is not the highest point of the path. (We take it that speaking about the highest point includes that the path has no other points of the same altitude). In order to show this it is sufficient to restrict to closed paths without double points, for otherwise we can easily find a closed part of the path, without double points, and containing the saddle point (mathematically we are still on the level of the mountainer).

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 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 leading from A to B crosses a saddle point, and if this saddle point is the highest point of the path, 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 (provided that it has just one highest point). So 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 inspect the collection of 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, 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 mountainer 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. The situation can be compared with the auxiliary line in problems of elementary geometry: the correct proof of the theorem uses the auxiliary line, but need not discuss why just that line 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 ζ, is the highest point of the path, we have $\varphi(\zeta)\neq 0$. As we are further only interested in a neighbourhood of ζ, we can select for $\psi(z)$ any branch of the logarithm, and we have no difficulties due to the multivaluedness.

If follows from $\varphi(\zeta)\neq 0$ that the conditions $\varphi'(\zeta)=0$ and $\varphi'(\zeta)=0$ are equivalent.

As far as yet we took $|\varphi(z)|$ to be the altitude of the landscape. If we replace it by 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, 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 derivatives $\psi'(\zeta), \psi''(\zeta), \psi'''(\zeta), \ldots$ 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 ζ is mainly determined by the value of $\psi'(\zeta)$, for

$$\psi(z) = \psi(\zeta) + \tfrac{1}{2}\psi''(\zeta).(z-\zeta)^2 + \ldots .$$

We shall define the axis of the saddle point ζ as the straight line in the complex plane defined by

$$\psi''(\zeta)(z-\zeta)^2 \text{ real and } \leq 0.$$

This is a line passing through ζ ; the line where $\psi''(\zeta)(z-\zeta)^2\geq 0$ is clearly perpendicular to this one. 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 α, then we say that $\arg\alpha$ is the argument of l. The argument of l is obviously uniquely determined apart from additional multiply of π).

The curves where Re $\psi(z)$=Re $\psi(\zeta)$ obviously intersect the axis at ζ , making angles of $\pm\frac{1}{4}\pi$ with the axis. These curves are drawn in Fig.1 as heavy lines. The curves where Im $\psi(z)=$ Im $\psi(\zeta)$ are drawn as dotted lines. Their tangents at ζ are the axis of the saddle point and the line perpendicular to it, respectively. The axis itself is denoted by a.

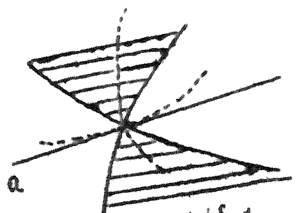

Fig.1

The curves Re $\psi(z)=$Re $\psi(\zeta)$ divide the neighbourhood of ζ into 4 parts. The two of these which do not contain the axis, are hatched. In these regions we have Re $\psi(z)>$ Re $\psi(\zeta)$, whence it follows that our integration path, of which ζ 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 isexactly one of the lines with Re $\psi(z)=$Re $\psi(\zeta)$, and where ζ 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.4.8), and we shall not further discuss it.

In the general case, where the integration path uses, apart from ζ itself, only inner points of the non-hatched regions, we can always deform it such that its tangent at ζ coincides with the axis a, and such that, in the cases of violent behaviour of Re $\psi(z)$ we are interested in, the value of Re $\psi(z)$ is in all points of the path very much smaller than Re $\psi(\zeta)$, apart from a small segment of the path around ζ. If we start at the saddle point, and go in one of the two directions of the axis, the function Re $\psi(z)$ decreases. It is easily verified that this decrease is stronger than the decrease of 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 might take any other curve connecting the non-hatched regions, provided that the angle it makes with the axis is less than $\pi/4$, and does not tend to $\pi/4$ 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 discuss 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 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 Re $\psi(z)$=Re $\psi(A)$. At one side of this curve, the side of the axis, we have Re $\psi(z) <$ Re $\psi(A)$, at the other side we have Re $\psi(z) >$ 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. <u>The second stage</u>. Suppose that we have found a curve that minimizes $\max_C$ 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 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 Re $\psi(z)$ 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 Re $\psi(z)$ 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 Lagrange 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. <u>A general simple case</u>. 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 asymptotical behaviour of

$$f(t) = \int_A^B g(z)\, e^{t\, h(z)} dz$$

as $t > 0$, $t \to \infty$. Assume that there is a point $\zeta \in D$ where $h'(\zeta)=0$, $h''(\zeta) \neq 0$. Properly speaking, ζ is not a saddle point of ge^{th}, but of e^{th}; nevertheless it will turn out that ζ can be used for the problem about ge^{th} as well.

The landscape of $|e^{th}|$ around ζ is of the type illustrated in fig.1, p.59. If δ 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 ζ and radius ρ, 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

$$\text{(i)} \quad 0 < |z-\zeta| < \rho, \quad |\arg(z-\zeta) + \tfrac{1}{2}\pi - \tfrac{1}{2} \arg h''(\zeta)| < \tfrac{1}{4}\pi - \delta$$

$$\text{(ii)} \quad 0 < |z-\zeta| < \rho, \quad |\arg(z-\zeta) - \tfrac{1}{2}\pi + \tfrac{1}{2} \arg h''(\zeta)| < \tfrac{1}{4}\pi - \delta .$$

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 $\operatorname{Re} h(z) =$

$$\operatorname{Re} h(\zeta) + \tfrac{1}{2} \operatorname{Re}\{h''(\zeta)(z-\zeta)^2\} + O(|z-\zeta|^3) < \operatorname{Re} h(\zeta) - |z-\zeta|^2 \sin 2\delta + O(|z-\zeta|^3),$$

and this is < 0 if ρ is sufficiently small.

Obviously, both δ and ρ 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, $\operatorname{Re}(\psi(z) - \psi(\zeta))$ has a negative upper bound $-c$, and therefore the contribution of these parts to the integral is $O(e^{-ct}.e^{t \operatorname{Re} h(\zeta)})$. 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 $z = \zeta + \alpha x$, $a \leq x \leq b$ ($-\rho < a < 0 < b < \rho$, $\alpha = \exp(\frac{1}{2}\pi i - \frac{1}{2} i \arg h''(\zeta))$), and the integral from A_2 to B_2 becomes

$$\alpha \int_a^b g(\zeta + \alpha x)\, e^{t\, h(\zeta + \alpha x)}\, dx.$$

We have thus obtained an integral that was discussed extensively in sec.4.4, as $h(\zeta + \alpha x) = h(\zeta) + \frac{1}{2} h''(\zeta)\, \alpha^2 x^2 + \ldots$, and $\frac{1}{2}h''(\zeta)\alpha^2 < 0$. The results of that section can be applied immediately It results that we have an asymptotic series of the form

$$(5.7.1)\quad \int_{A_2}^{B_2} g(z)\, e^{t\, h(z)} dz \sim e^{t\, h(\zeta)} t^{-\frac{1}{2}} \sum_0^\infty d_n t^{-n} \qquad (t \to \infty).$$

If $g(\zeta)\neq 0$, the main term is easily evaluated:

$$(5.7.2)\quad \int_{A_2}^{B_2} g(z)\, e^{t\, h(z)} dz \sim (2\pi)^{\frac{1}{2}} \alpha\, t^{-\frac{1}{2}} |h''(\zeta)|^{-\frac{1}{2}} g(\zeta) e^{t\, h(\zeta)} \quad (t \to \infty).$$

We note that α 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 ζ. 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 directions, the contribution is of course multiplied by -1, as $\int_{B_2}^{A_2} = -\int_{A_2}^{B_2}$.

As the integrals from A_1 to A_2 and from B_2 to B_1 are exponentially small compared to $e^{t\, h(\zeta)}$, 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 ζ. 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 Re $h(z)$ is less than 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.

A similar, but simpler, discussion can be given for the contribution of an end-point to our integral $\int g(z)e^{t\, h(z)}dz$. 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 Re $h(A)$ decreases, then the first term of the contribution of a neighbourhood of A equals (cf. sec.4.3)

$$g(A)\, e^{t\, h(A)}\, (t\, h'(A))^{-1}.$$

5.8. Path of constant altitude. We again consider integrals $_A\int^B \varphi(z)dz=\int_A^B e^{\psi(z)}dz$.

If the points A and B are connected by a path, all points of which have the same altitude in the landscape $w=\text{Re}\,\psi(z)$, 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 devided 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, $f'(x)\neq 0$ $(0 \leq s \leq 1)$ 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

$$d\psi/ds = \psi'(z) \,.\, f'(s) \neq 0 \qquad (0 < s < 1)$$

As Re ψ 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

$$i d\psi/ds \text{ real and } < 0 \qquad (0 < s < 1).$$

It now follows from the Cauchy-Riemann equations that the derivative of 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 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 the 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 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 side we came from).

The asymptotical behaviour of the integral from A to B is now given by the sum of the contributions of the points $A=A_0,A_1,\ldots,A_n=B$.

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. The minimum problem should be interpreted, in this case, as the problem to find a closed curve C, encircling z=0 just once, such that (5.1.4) is minimal.

5.10. Range of a saddle point. It is often quite difficult to determine a saddle point exactly. However, for asymptotical purposes it is not necessary to take a path exactly through the saddle point. (In sec.5.7 we actually did not use the exact saddle point of $g(z)e^{t\,h(z)}$, but an approximation to it, viz. the saddle point of $e^{t\,h(z)}$).

If ζ is a saddle point of the function ψ , then the range of ζ is a circular neighbourhood of ζ , 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, how 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 role 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 ζ may depend on t. But even if the saddle point is fixed, the range may still depend on t. E.g., if $\psi(z)=-tz^2$ $\zeta=0$ 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 ζ is a saddle point, then it is very important to know whether in the formula

(5.10.1) $\psi(z) = \psi(\zeta) + \frac{1}{2}\psi''(\zeta)(z-\zeta)^2 + \frac{1}{6}\psi'''(\zeta)(z-\zeta)^3 + \dots$

the sum of the terms

(5.10.2) $\frac{1}{6}\psi'''(\zeta)(z-\zeta)^3 + \dots$

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 o-formula as $t \to \infty$. 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 ζ is concerned, the integral can be successfully compared to

(5.10.3) $\int_L \exp\{\psi(\zeta) + \frac{1}{2}\psi''(\zeta)(z-\zeta)^2\}\,dz = (2\pi)^{\frac{1}{2}}\alpha|\psi''(\zeta)|^{-\frac{1}{2}}e^{\psi(\zeta)}$.

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 α has absolute value 1, and its argument indicates the direction on L (cf.5.7.2).

In the special case that $\psi(z)=t\,h(z)$, 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 3rd 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 only 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 $O((z-\zeta)^3)$ if $|z-\zeta|<\frac{1}{2}R$. It follows that the sum of (5.10.2) is $O\{t^{-\frac{1}{2}}.(\psi''(\zeta)(z-\zeta)^2)\}$ $(t>4R^2,\ |z-\zeta|<t^{-\frac{1}{2}})$. This means that inside the range the second order term dominates all other terms.

If, on the other hand, (5.10.2) is not small compared to $\frac{1}{2}\psi''(\zeta)(z-\zeta)^2$ throughout the range of the saddle point ζ, it is difficult to say anything in general. Usually it means that there are other significant saddle points in the range of ζ, or that there are even singularities of ψ in that range. We shall come across a few examples in sec.5.14.

5.11. Examples. In the next sections we shall give some simple examples. They 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, all possible difficulties occur in one and the same problem. We shall give some of these more complicated problems in the next chapter.

5.12. Our first example is

$$f(t) = \int_0^\infty \exp((z+iz-z^3)t)dz,$$

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, ζ and $-\zeta$, where $\zeta=2^{\frac{1}{4}}.3^{-\frac{1}{2}}.e^{\pi i/8}$. 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 ζ 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. Furtunately 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 $z=e^{\pi i/8}x$, $0 \leqslant x < \infty$, the function h(z) becomes

$$h(z) = h(e^{\pi i/8}x) = e^{3\pi i/8}(2^{\frac{1}{2}}x-x^3)t \ ,$$

and $2^{\frac{1}{2}}x-x^3$ is maximal at $x=x_0= 2^{\frac{1}{4}}3^{-\frac{1}{2}}$. Therefore, the same thing applies to Re h(z).

On the line l we have Re $h(z) < Cx$ $(C<0)$ for x sufficiently large, and therefore, the contribution of the part from x_0+1, say, to ∞ is exponentially small compared to the contribution of the saddle point.

In order to evaluate the contribution of the saddle point, we evaluate $h(\zeta)=(1+i)\zeta-\zeta^3=\frac{2}{3}(1+i)\zeta=2^{7/4}3^{-3/2}e^{3\pi i/8}$,

$$h''(\zeta) = -6\zeta = -2^{5/4}\ 3^{\frac{1}{2}}\ e^{\pi i/8}.$$

As we cross the saddle point from left to right, the number α (see sec.5.7) equals $\alpha=e^{-\pi i/16}$. So by (5.7.2) the first term of the contribution equals

$$(2\pi)^{\frac{1}{2}}\ e^{-\pi i/16}\ t^{-\frac{1}{2}}(2^{5/4}3^{\frac{1}{2}})^{-\frac{1}{2}}\ e^{t\ h(\zeta)}=$$

$$=e^{-\pi i/16}\ 2^{-1/8}3^{-1/4}\pi^{\frac{1}{2}}\ t^{-\frac{1}{2}}\ e^{t\ h(\zeta)}.$$

This gives at the same time the asymptotic behaviour of f, and so

$$f(t) \sim e^{-\pi i/16}\ 2^{-1/8}3^{-1/4}\pi^{\frac{1}{2}}t^{-\frac{1}{2}}\ \exp\left\{2^{7/4}3^{-3/2}e^{3\pi i/8}t\right\}\quad (t\to\infty).$$

5.13. Our next example is

$$F(t) = \int_{-\infty}^{\infty}\varphi(z,t)dz, \qquad \varphi(z,t)=e^{it(3z-z^3)},$$

and it is meant to illustrate sec.5.8. There is an extra difficulty, the path having infinite length.

Although φ 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 $a\to\infty$ (t fixed), but it is not true that $\int_a^\infty$ is exponentially small if a is fixed and $t\to\infty$. Therefore it is not advisable to apply the method of sec.5.8 to $F_a=\int_{-a}^{a}$ and to make $a\to\infty$ 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 $t\to\infty$.

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 $\mathrm{Re}(it(3z-z^3))=t(3y(x^2-1)-y^3)$. 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 ε and δ 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

$$|\varphi(x-i\delta)| = |\varphi(1+\varepsilon-i\delta)|\ .\exp(-3\delta t(x^2-(1+\varepsilon)^2)) \leq$$

$$\leq |\varphi(1+\varepsilon-i\delta)|\ \exp(-6\delta t(x-1-\varepsilon)) \qquad (x>1+\varepsilon),$$

and a similar estimate holds on the part of P extending to $-\infty$. Moreover we have

$$\left|\int_{1+\varepsilon-i\delta}^{\infty-i\delta}\varphi dz\right| \leq \left|\varphi(1+\varepsilon-i\delta)\right| /6\delta t,$$

and $|\varphi(1+\varepsilon-i\delta)|$ is exponentially small, as $\mathrm{Re}(i(3z-z^3))<0$ at $z=1+\varepsilon-i\delta$. It follows that the asymptotical behaviour of the integral along P just consists of the contributions of the saddle points 1 and -1.

We still have the question whether $\int_P = \int_{-\infty}^{\infty}$. When investigating this, we can consider t as a positive constant. If b is a large positive number we have, by Cauchy's theorem,

$$\int_{-b}^{b} = \int_P - \int_{b-i\delta}^{\infty-i\delta} - \int_{-\infty-i\delta}^{-b-i\delta} + \int_{b-i\delta}^{b} - \int_{-b-i\delta}^{-b},$$

where the integration paths, apart from P, are straight horizontal or vertical lines. It follows from the convergence of $\int_P$ that the second and the third integral on the right-hand-side tend to 0 as $b \to \infty$. The same thing is true for the fourth and for the fifth, as

$$\left| \int_{b-i\delta}^{b} \right| \leq \int_0^{\delta} \exp(-t(3u(b^2-1)-u^3))du <$$

$$< C(\delta,t) \int_0^{\delta} \exp(-3tub^2)du < C(\delta,t)/3tb^2,$$

and this tends to 0 as $b \to \infty$ ($C(\delta,t)$ is independent of b). This proves that $\int_{-b}^{b}$ tends to $\int_P$ as $b \to \infty$. 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-east to north-west, and the number α (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.

$$(2\pi)^{\frac{1}{2}} e^{\pi i/4} t^{-\frac{1}{2}}.6^{-\frac{1}{2}} e^{-2it}.$$

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

$$F(t) = 2(\pi/3t)^{\frac{1}{2}} \cos(2t-\pi/4) + O(t^{-3/2}).$$

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. It is not true that $F(t) \sim 2(\pi/3t)^{\frac{1}{2}}\cos(2t-\pi/4)$. The latter formula would imply that for large values of t, F(t) vanished at exactly the same places where $\cos(2t-\pi/4)$ vanishes, and this is not necessarily so.

5.14. We shall now discuss some examples illustrating the notion of the range of a saddle point (sec.5.10). We start from the integral

$$(5.14.1) \qquad \int_{-\infty}^{\infty} e^{-tz^2} dz,$$

whose saddle point is $z=0$. The circle $|z| \leq t^{-\frac{1}{2}}$ can be considered as its range. The notion of range has some importance when one wants to discuss integrals obtained from (5.14.1) by small perturbations. For example, in the integral

(5.14.2) $$\int_{-\infty}^{\infty}(1+iz)^{\frac{1}{2}}\, e^{-tz^2}\, dz$$

the factor $(1+iz)^{\frac{1}{2}}$ behaves quite smoothly within the range of the saddle point $z=0$ of (5.14.1). 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 (4.13.2), the integration path $-\infty < x < \infty$ of (4.13.1) can still successfully be used for the calculation of the asymptotic behaviour of (4.13.2). 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

(5.14.3) $$F_\alpha(t) = \int_{-\infty}^{\infty} e^{-tz^2}\, \frac{e^z}{1+t^\alpha z^2}\, dz,$$

where α is a positive parameter. For each fixed value of α the asymptotic behaviour (as $t \to \infty$) is required. We shall investigate in what respect this can still be considered as a "minor modification" of (5.14.1).

The factor e^z is harmless: it behaves smoothly within the range. It gets large, however, 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 ring R where $t^\alpha z^2$ is neither large nor small, and actually there are two poles in that ring, viz. $z=\underline{+}it^{-\frac{1}{2}\alpha}$. Now there are these possibilities:

(i). $0 < \alpha < 1$. Then the ring R lies far outside the range.

(ii). $\alpha = 1$. The ring R covers a considerable part of the range.

(iii). $\alpha > 1$. The ring 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 (5.14.2). 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 ∞ and from -T to $-\infty$ are easily seen to be $O(\exp(Tct^{1-2\beta}))$, with some positive constant c. 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 $-T \leq z \leq T$,

$$e^z(1+t^\alpha z^2)^{-1} = \left\{1+z+\tfrac{1}{2}z^2 + O(z^3)\right\} \left\{1-t^\alpha z^2+O(t^{2\alpha}z^4)\right\} =$$

$$= 1+z+(\tfrac{1}{2}-t^\alpha)z^2 - t^\alpha z^3 + O(t^{2\alpha}z^4) + O(z^3).$$

We now easily find (cf sec.4.1)

$$f\alpha(t) = \int_{-\infty}^{\infty}\left\{1+z+(\tfrac{1}{2}-t^\alpha)z^2-t^\alpha z^3\right\}e^{-tz^2}dt + t^{-\frac{1}{2}}\, O(t^{2\alpha-2}+t^{-3/2}),$$

and therefore

$$F_\alpha(t) = \pi^{\frac{1}{2}}t^{-\frac{1}{2}}\left\{1-\tfrac{1}{2}t^{\alpha-1} + O(t^{2\alpha-2})\right\} \qquad (t \to \infty).$$

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:

$$F_1(t) = t^{-\frac{1}{2}} \int_{-\infty}^{\infty}\exp(-w^2+t^{-\frac{1}{2}}w)(1+w^2)^{-1}\, dw.$$

The role which thus far was played by integrals of the type $\int_{-\infty}^{\infty}\exp(-w^2)w^k dw$, is now taken over by integrals $\int_{-\infty}^{\infty}\exp(-w^2)(1+w^2)^{-1}dw$. Otherwise the technique is the same as in sec.4.4. We obtain the asymptotic series

$$f_1(t)\sim\sum_{\nu=0}^{\infty} c_\nu\ t^{-\nu-\frac{1}{2}} \qquad (t \to \infty),$$

where

$$c_\nu = \left\{(2\nu)!\right\}^{-1} \int_{-\infty}^{\infty} e^{-w^2} w^{2\nu}\ (1+w^2)^{-1}dw.$$

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 part 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

$$r(t) = \tfrac{1}{2} i t^{-\frac{1}{2}\alpha} \exp(t^{1-\alpha} - i t^{-\frac{1}{2}\alpha}),$$

and we obtain

$$(5.14.4) \qquad F_\alpha(t) = -2\pi i r(t) + \int_P \exp(-tz^2+z).(1+t^\alpha z^2)^{-1}\, dz,$$

the integration path P being the straight line from $-2i.t^{-\frac{1}{2}\alpha}-i\infty$ to $-2i.t^{-\frac{1}{2}\alpha}+i\infty$. 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 $|t^\alpha z^2| \geqslant 2$ 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 that its sum equals $F_\alpha(t)+2\pi i\, r(t)$. To this end we write

$$F_\alpha(t)+2\pi i\, r(t) = \int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \exp(-tz^2)(z^k/k!)(-1)^n t^{-(n+1)\alpha} z^{-2(n}$$

where $z = x-2it^{-\frac{1}{2}\alpha}$, and x is the new integration variable.

We now apply a theorem on integration under the sum sign. It can be stated in a general form for Lebesgue integrals, but the following more elementary form will do. Let $f_{kn}(x)$ $(k,n=0,1,2,\dots)$ be continuous functions $(-\infty < x < \infty)$. Assume that $\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} |f_{kn}(x)|$ converges for all x $(-\infty < x < \infty)$, and that the double sum $\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} f_{kn}(x)$ represents a continuous function of x. Finally assume that the funct $g(x) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} |f_{kn}(x)|$ is such that $\int_{-\infty}^{\infty} g(x)dx$ converges. Then we have

$$\int_{-\infty}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} f_{kn}(x)dx = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} f_{kn}(x)dx,$$

and both the integral on the left and the double series on the right converge absolutely. It is quite easy to prove the theorem. The fact that we have a double series instead of the usual single series, does not cause any difficulty as it converges absolutely.

The assumptions of this theorem are satisfied in our case. We have

$$g(x) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \left|\exp(-tz^2)\right| (|z|^k/k!).t^{-(n+1)\alpha}|z|^{-2(n+1)} =$$

$$= \left|\exp(-tz^2)\right| . e^{|z|} . (t^\alpha |z|^2 - 1)^{-1}.$$

The function g(x) is continuous $(-\infty < x < \infty)$, as $t^\alpha |z|^2 \geqslant 4$ for all z on the path P. Moreover $\int_{-\infty}^{\infty} g(x)dx$ converges, by virtue of the overwhelming power of the factor $\exp(-tz^2)$. (Notice that t is a fixed positive number through [illegible] Therefore the theorem may be applied, and we infer 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 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

$$\int_P \exp(-tz^2)z^{k-2n-2}dz=-\tfrac{1}{2}t^{-1}\int_P \exp(-tz^2)z^{k-2n-3}\ d(-tz^2) =$$

$$= -\tfrac{1}{2}t^{-1}\int_C e^w.(i)^{-k+2n+3}\ t^{+\frac{1}{2}(2n+3-k)}\ w^{\frac{1}{2}(k-2n-3)}dw =$$

$$= \pi t^{n-\frac{1}{2}(k-1)}\ i^{2n-k+2}/\Gamma(n-\tfrac{1}{2}(k-3)),$$

according to Hankel's formula 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 $\alpha>1$, onl 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

$$F_\alpha(t)=\pi t^{-\frac{1}{2}\alpha}\sum_{h=0}^{\infty}\ \sum_{m=0}^{\infty}\frac{(-1)^{h+m}\ t^{-\frac{1}{2}m(\alpha-1)-h}}{(2h)!\ \Gamma(\frac{1}{2}m+1-h)}\qquad (t>0).$$

5.15. <u>Exercises</u>. 1. Show that

$$\int_{-\infty}^{\infty} e^{itx}(1+x^2)^{-t}\ dx \sim (\pi(1-c)/t)^{\frac{1}{2}}\ e^{-ct}(2c)^{-t}\qquad (t\to\infty).$$

where $c=-1+2^{\frac{1}{2}}$.

2. Show that

$$\int_{-1}^{\infty}(x^3+3x-2i)^{-n}\ e^{ix}\ dx \sim 2e(i/4)^n(\pi/3n)^{\frac{1}{2}}\qquad (n\to\infty)$$

where n runs through the integers.

(Hoofdstuk 6 (Applications of the saddle point method) wordt voorlopig overgeslagen).

7. Indirect Apymptotics.

7.1. Direct and indirect asymptotics. The first six chapters of this book (with the exception of some arguments in ch.I) have been devoted to asymptotical methods which we shall call direct methods. The common features are (i) the function f(t) whose asymptotical behaviour (as $t \longrightarrow \infty$) 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).

The remaining chapters of this book will show 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 role. 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 estimations.

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 α such that $a_n=\alpha+o(1)\,(n\longrightarrow\infty)$, 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 $f(t)=t^2+O(t)$ $(t\longrightarrow\infty)$ we show, for instance, that $|f(t)| < 10$ $(2<t\le 3)$ and that, for all $T\geq 3$, the assumption

$$|f(t)-t^2| < 80t \qquad (T-1<t\le T)$$

leads to

$$|f(t)-t^2| < 80t \qquad (T<t\le T+1).$$

Now the principle of induction shows that $|f(t)-t^2| < 80t$ $(t>2)$.

The difficulty with such arguments is the same 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 inequalities of the type $|f(t)-t^2| < At^{\frac{1}{2}}$. 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 asymptotical 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 explicit expressions of the function $f(t)$ (whose asymptotical behaviour is required), not in terms of known expressions, but 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 $t\longrightarrow\infty$; 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

$$f(t) = \cos t^{-1} + \int_t^\infty \left\{\tau^2+(f(\tau))^2\right\}^{-1} d\tau \qquad (t>1),$$

then it is easily seen that the integral is $O(t^{-1})$. If follows that $f(t)=\cos t^{-1}+O(t^{-1})=1+O(t^{-1})\,(t>1)$. Inserting this into the integral, we get $\int_t^\infty\{\tau^2+1+O(\tau^{-1})\}^{-1}d\tau = \int_t^\infty\{\tau^{-2}-\tau^{-4}+O(\tau^{-5})\}d\tau = t^{-1}-\frac{1}{3}t^{-3}+O(t^{-4})$, whence $f(t)=1+t^{-1}-\frac{1}{6}t^{-2}-\frac{1}{3}t^{-3}+O(t^{-4})$.

The procedure can be carried on, and an asymptotic series for $f(t)$ is easily obtained.

(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 asymptotical behaviour itself. An example: Suppose that we are dealing with a bounded function $f(t)$ (in $0 \leq t < \infty$) which satisfies a differential equation. Suppose it turns out that this equation has just one bounded solution. Now the problem is again 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.

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.

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 $\lim_{x\to 1, 0\leq x<1} f(x)=f(1)$. Subtracting the constant value $f(1)$, we can put this theorem into the form: If

$$a_0 + a_1 + \ldots + a_n = o(1) \qquad (n\to\infty), \tag{7.2.1}$$

then we have

$$\sum_{k=0}^{\infty} a_k x^k = o(1) \qquad (0\leq x<1,\ x\to 1). \tag{7.2.2}$$

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))

$$\sum_{k=0}^{n} a_k x^k = A(n)x^n - \log x \int_0^n A(y)\, x^y\, dy. \tag{7.2.3}$$

Taking a fixed value of x in the interval $0<x<1$, we observe that $A(n)x^n=O(x^n)\to 0$ as $n\to\infty$, and that $\int_0^{\infty} A(y)x^y dy$ converges. It therefore follows from (7.2.3) that

(7.2.4) $$f(x) = -\log x \int_0^\infty A(y)x^y dy.$$

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 $\int_0^\infty = \int_0^b + \int_b^\infty$,

$$|f(x)| \leq |\log x| \int_0^b |A(y)|\, dy + \tfrac{1}{2}\varepsilon|\log x| \int_b^\infty x^y dy.$$

The first term is $< \frac{1}{2}\varepsilon$ if x is sufficiently close to 1. The second term is $\leq \frac{1}{2}\varepsilon |\log x| \int_0^\infty x^y dy = \frac{1}{2}\varepsilon$. 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. Abels' 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

$$f(x) = \tfrac{1}{2} - x + x^2 - x^3 + \ldots \qquad = \tfrac{1}{2}(1-x)/(1+x) \qquad (0 \leq x < 1),$$

then $f(x) \to 0$ as $x \to 1$, i.e. (7.2.2) holds. However, (7.2.1) is false in this case:

$$a_0 + a_1 + \ldots + a_n = \tfrac{1}{2}.(-1)^n \neq o(1).$$

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) $$a_n = o(n^{-1}) \qquad (n \to \infty)$$

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 conditions that there exists a positive constant C such that

(7.2.6) $$a_n > Cn^{-1} \qquad (n=1,2,3\ldots).$$

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, asymptotical information about some kind of average of a function from asymptotical 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 sec. 7.3 we shall deal with a quite simple case, and a more difficult Tauberian theorem will be proved in sec.7.4. A quite general theory about Tauberian theorems was developed by N. Wiener. For this and for further details we refer to: G.A. Hardy, Divergent Series (Oxford, 1949).

7.3. Differentiation of an asymptotic formula. Let f(x) be integrable over any finite interval, and put

(7.3.1) $$F(t) = \int_0^t f(x)dx.$$

Assuming some asymptotical behaviour of f(t), say

(7.3.2) $$f(t) \sim t^{\alpha} \qquad (t \to \infty),$$

where α is a constant ≥ 0, it is easy to derive an Abelian result about F(t), in this case

(7.3.3) $$F(t) \sim (\alpha+1)^{-1}\, t^{\alpha+1}. \qquad (t \to \infty).$$

It was already pointed out in sec.1.6 that the converse is not always true, i.e. formal differentiation of (7.3.3) is not always legitimate. We need a Tauberian condition, and as such we take

(7.3.4) f(t) is real and non-decreasing $(0 \leq t < \infty)$.

We shall prove the Tauberian theorem that (7.3.3) and (7.3.4) together imply (7.3.2).

Let ε be a positive number. By (7.3.3) we can take T such that

$$\left|F(t) - (\alpha+1)^{-1}\, t^{\alpha+1}\right| < \varepsilon\, t^{\alpha+1} \qquad (t \geq T).$$

Let p and t be numbers such that $t \geq T$, $t+p \geq T$; p may be negative. Then we have

$$\left|F(t+p) - (\alpha+1)^{-1}(t+p)^{\alpha+1}\right| < \varepsilon \ (t+p)^{\alpha+1} \qquad \text{and}$$

taking differences we find that

(7.3.5) $$\left|\int_t^{t+p} f(x)dx - \int_t^{t+p} x^{\alpha}\, dx\right| < \varepsilon t^{\alpha+1} + \varepsilon(t+p)^{\alpha+1}.$$

First assume $p > 0$. It 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

$$f(t) \leq t^{\alpha} \left\{ (1+q)^{\alpha} + 2\varepsilon\, q^{-1}(1+q)^{\alpha+1} \right\} .$$

It is not necessary to find the exact minimum. We only remark that q should be small enough to keep $(1+q)^{\alpha}$ within reasonable bounds, and that q should be large with respect to ε in order to keep εq^{-1} small. So we take $q = \varepsilon^{\frac{1}{2}}$, and we get

$$(7.3.6) \qquad f(t) \leq t^{\alpha} \left\{ (1+\varepsilon^{\frac{1}{2}})^{\alpha} + 2\varepsilon^{\frac{1}{2}}(1+\varepsilon^{\frac{1}{2}})^{\alpha+1} \right. \qquad (t \geq T) .$$

A lower bound is obtained by taking, in (7.3.5), $p < 0$. We immediately take $p = -\varepsilon^{\frac{1}{2}}t$. Assuming $\varepsilon < 1$, and $t \geq 2T$, we have $p+t \geq T$. Moreover we have $|p|\ f(t) \geq \int_{t+p}^{t} f(x)dx$, $\int_{t+p}^{t} x^{\alpha}\, dx \geq |p|\,(t+p)^{\alpha}$, and it follows from (7.3.5) that

$$(7.3.7) \qquad f(t) \geq t^{\alpha} \left\{ (1-\varepsilon^{\frac{1}{2}})^{\alpha} - 2\varepsilon^{\frac{1}{2}} \right\} \qquad (t \geq 2T).$$

From (7.3.6) and (7.3.7) we can deduce $f(t) \sim t^{\alpha}$ $(t \to \infty)$. For, if $\varepsilon' > 0$ is given, then ε 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 ε 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. <u>A similar problem.</u> We again consider $F(t) = \int_0^t f(x)dx$, as in the previous section, but instead of (7.3.2) we consider

$$(7.4.1) \qquad f(t) = 2t + O(1) \qquad (t \to \infty).$$

Then we can derive by an Abelian argument

$$(7.4.2) \qquad F(t) = t^2 + O(t) \qquad (t \to \infty).$$

We again ask whether the supplementary condition

$$(7.4.3) \qquad f(t) \text{ is non-decreasing} \qquad (0 \leq t < \infty)$$

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):

$$p\ f(t) \leq \int_t^{t+p} f(x)dx = F(t+p) - F(t) = 2pt + p^2 + O(t) + O(p),$$

$$f(t) \leq 2t + p + O(tp^{-1}) + O(1).$$

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

$$(7.4.4) \qquad f(t) = 2t + O(t^{\frac{1}{2}}).$$

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 $\lim_{t\to\infty}(f(t)-2t)t^{-\frac{1}{2}}=0$.

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 functions $y=F(t)$ 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 parabole $y=t^2+t$ by π_1, and the parabole $y=t^2-t$ by π_2. The condition (7.4.3) means that the graph of $F(t)$ is convex. We now want to draw a convex curve between π_1 and π_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 paraboles is occasionally large.

We therefore choose the graph of F as follows. We take a sequence of points $P_0=(t_0,y_0)$, $P_1=(t_1,y_1),\ldots$ on π_2, such that, for each value of k, the line connecting P_k and P_{k+1} touches π_1 in a point somewhere between these two. Now the graph of F is the broken line $P_0P_1P_2P_3\ldots$.

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 π_2 is, by elementary analytic geometry, easily translated into the relation $(t_k+t_{k+1}-2)^2=4t_kt_{k+1}$, whence

$$t_{k+1} = t_k + 2 + (8t_k)^{\frac{1}{2}}.$$

If x_0 is chosen arbitrarily, then $t_1,t_2,\ldots$ 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 $k\to\infty$, and that can be done by methods indicated in ch.8). If we take $t_0=0$, then $t_1=0+2+0=2$, $t_2=2+2+4=8$, $t_3=8+2+8=18$, $t_4=32$, $t_5=50,\ldots$ 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

$$f(t) = (2k+1)^2 \qquad (2k^2\leq t<2(k+1)^2,\ k=0,1,2,\ldots).$$

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 obviously 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

$$F(t) = t^2+t - (t-2k^2-2k)^2 =$$
$$= t^2-t + (2(k+1)^2-t)(t-2k^2) \qquad (2k^2\leq t\leq 2(k+1)^2).$$

From these formulas it is evident that $t^2-t \leq F(t) \leq t^2+t$ $(0 \leq t < \infty)$.

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 assume a much stronger condition: we assume that f has a non-negative second derivative:

$$f''(t) \geq 0 \qquad (t \geq 0). \tag{7.4.5}$$

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

$$F(t) - t^2 - F(t_1) + t_1{}^2 = \int_{t_1}^{t} (f(x)-2x)dx \leq (t-t_1)(f(t_1)-2t_1).$$

If A is such that $|F(t)-t^2| < At$ $(t > 1)$, 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 $t \to \infty$. 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 $a_0+a_1x+a_2x^2+\ldots$ be a power series, convergent if $|x| < 1$, and let some asymptotical information be given about the partial sums $a_1+\ldots+a_n$, as $n \to \infty$. If this behaviour is not too irregular, we can deduce, by Abelian arguments, the behaviour of the sum function

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots \qquad (|x| < 1)$$

as $0 < x < 1$, $x \to 1$. (A special case of this problem and of its Tauberian counterpart was discussed in sec.7.2).

Next assume that the asymptotical behaviour of $f(x)$ (as $x \to 1$) 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

$$A(n) = \sum_{k=0}^{\infty} a_k\, g(e^{-k/n}) \qquad (n > 0), \tag{7.5.1}$$

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 $P(x)$. If

$$P(x) = \sum_{j=1}^{m} p_j x^j,$$

then the sum corresponding with (7.5.1) is

$$(7.5.2) \qquad \sum_{k=0}^{\infty} a_k\, P(e^{-k/n}) = \sum_{j=1}^{m} p_j\, f(e^{-j/n}) \; .$$

If P is fixed, the asymptotical behaviour (as $n \to \infty$) 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 γ and assuming

$$(7.5.3) \qquad f(x) = \sum_{k=0}^{\infty} a_k\, x^k = o((1-x)^{-\gamma}) \qquad (0<x<1,\ x \to 1).$$

Using the Tauberian condition

$$(7.5.4) \qquad a_k > -C\ (k+1)^{\gamma-1} \qquad (k=0,1,2,\ldots),$$

where C is a positive constant, we shall prove that

$$(7.5.5) \qquad A(n) = \sum_{k=0}^{n} a_k = o(n^{\gamma}) \qquad (n \to \infty).$$

The special case $\gamma=0$ already has been announced 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).

As to the approximations of $g(x)$ by polynomials we shall stipulate the following conditions. Let h be an integer, $h \geq 0$, $h > -\gamma$. We may take for h the smallest integer satisfying these inequalities. Let be a positive number, $0 < \varepsilon < \frac{1}{2}$. Then we want to have polynomials $P_1(x)$, $P_2(x)$ such that

$$P_1(x) \leq g(x) \leq P_2(x) \qquad (0 \leq x \leq 1),$$

$$|P_2(x) - P_1(x)| \leq 2 \qquad (0 \leq x \leq 1),$$

$$|P_2(x)-P_1(x)| \leq \varepsilon\, x \qquad (0 \leq x \leq e^{-1-\varepsilon}),$$

$$|P_2(x)-P_1(x)| \leq \varepsilon\,(1-x)^h \quad (e^{-1+\varepsilon} \leq x \leq 1).$$

We do not require anything regarding the degree or the coefficient 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 $g_1(x) \leq g(x) \leq g_2(x)$

$(0 \leq x \leq 1)$, $g_2(x) = g_1(x) = g(x)$ if $x \leq e^{-1-\varepsilon}$ or $x \geq e^{-1+\varepsilon}$, $g_2(x)-g_1(x) \leq 1$ if $e^{-1-\varepsilon} \leq x \leq e^{-1+\varepsilon}$. Next we determine a polynomial $Q(x)$ such that $Q(0)=g(0)$, $Q(1)=g(1)$, $Q'(1)=g'(1), \ldots, Q^{(h-1)}(1)=g^{(h-1)}(1)$. Actually $Q(x)=1-(1-x)^h$ is already suitable, but the explicit form of Q does not matter. Now $(g_2(x)-Q(x))x^{-1}(1-x)^{-h}=\varphi_2(x)$ is a continuous function in $0 \leq x \leq 1$. By the Weierstrass approximation theorem we can find a polynomial $R_2(x)$ such that $|\varphi_2(x)+\frac{1}{4}\varepsilon - R_2(x)| < \frac{1}{4}\varepsilon$ $(0 \leq x \leq 1)$. Putting $Q + x(1-x)^h R_2=P_2$ we observe that $g_2(x) \leq P_2(x) \leq g_2(x)+\frac{1}{2}\varepsilon x(1-x)^{-h}$ $(0 \leq x \leq 1)$. Similarly we construct $P_1(x)$ such that $g_1(x)-\frac{1}{2}\varepsilon x(1-x)^{-h} \leq P_1(x) \leq g_1(x)$ $(0 \leq x \leq 1)$. 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,

$$(7.5.6) \qquad -A_n + \sum_{k=0}^{\infty} a_k P_1(e^{-k/n}) \leq \sum_{k=0}^{\infty} C(k+1)^{\gamma-1} (g(e^{-k/n})-P_1(e^{-k/n}))$$

$$(7.5.7) \qquad A_n - \sum_{k=0}^{\infty} a_k P_2(e^{-k/n}) \leq \sum_{k=0}^{\infty} C(k+1)^{\gamma-1}(P_2(e^{-k/n})-g(e^{-k/n})),$$

in 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

$$\sum_{0 \leq k \leq n(1-\varepsilon)} \leq \sum_{0 \leq k \leq n(1-\varepsilon)} C(k+1)^{\gamma-1} (1-e^{-k/n})^h \leq$$

$$\leq \sum_{0 \leq k \leq n} \varepsilon C(k+1)^{\gamma-1} k^h n^{-h} \leq \varepsilon C D_1 n^{\gamma},$$

where D_1 depends on γ, but not on ε and n. For, $(k+1)^{\gamma-1} k^h \leq (k+1)^{\gamma+h-1}$, and $\gamma+h$ is positive, and if β is a positive number, we have $\sum_0^n (k+1)^{\beta-1} = O(n^{\beta})$ $(n > 0)$.

The second part of the sum has at most $2\varepsilon n+1$ terms, and so

$$\sum_{n(1-\varepsilon) \leq k \leq n(1+\varepsilon)} \leq 2C \sum_{n(1-\varepsilon) \leq k \leq n(1+\varepsilon)} (k+1)^{\gamma-1} \leq 2C.(2\varepsilon n+1).D_2 n^{\gamma-1},$$

where D_1 depends on γ only (we can take $D_2=3^{|\gamma-1|}$, since $0 < \varepsilon < \frac{1}{2}$ guarantees that $n/3 < k+1 < 3n$ if $n(1-\varepsilon) \leq k \leq n(1+\varepsilon)$).

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 γ 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) are at most $\varepsilon C n^{\gamma} (D_1+4D_2+D_3+2n^{-1}\varepsilon^{-1}D_2)$, where the D's depend on γ 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 may be noticed that $P_1(0)=0$). As $(1-e^{-j/n})^{-\gamma} \sim j^{-\gamma} n^{\gamma}$ $(n \to \infty$, j fixed), the sum is $o(n^{\gamma})$. More precisely, we can determine a number n_0, depending on γ, ε and on the polynomial P_1, such that the sum is, in absolute value, less than εn^{γ}. 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 ε and n. As ε is arbitrary, this proves (7.5.5).

If $\gamma > 0$, we have

$$\sum_{k=0}^{\infty} (k+1)^{\gamma-1} x^k \sim \Gamma(\gamma)(1-x)^{-\gamma} \quad (0 < x < 1,\ x \to 1),$$

$$\sum_{k=0}^{n} (k+1)^{\gamma-1} \sim \gamma^{-1} n^{\gamma} \quad (n \to \infty).$$

So for $\gamma > 0$ our result can be put into the following form. If

$$\sum_{k=0}^{\infty} a_k x^k \sim (1-x)^{-\gamma} \quad (0 < x < 1,\ x \to 1),$$

$$a_k > -C(k+1)^{\gamma-1} \quad (k=0,1,2,\ldots),$$

then we have

$$\sum_{k=0}^{n} a_k \sim (\Gamma(\gamma+1))^{-1} n^{\gamma} \quad (n \to \infty).$$

8. Iterated functions.

8.1. Introduction. Many problems in asymptotics can be stated in the following terms: Let a sequence of functions $F_1, F_2, \ldots$ be given, and let x_0 be a number. Now we put $x_1=F_1(x_0)$, $x_2=F_2(x_1)$, $x_3=F_3(x_2), \ldots$, assuming that F_1 is defined at x_0, F_2 is defined at x_1, etc. The problem is to find the asymptotical behaviour of x_n as n tends to infinity.

In the present chapter we shall discuss only a very special case of the problem, taking all functions $F_1, F_2, \ldots$ 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 (for instance, see sec.8.5). This remark holds for still 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 $x_n=F_n(x_0, x_1, \ldots, x_{n-1})$. A further generalization is obtained if we replace the x's by functions and the F's by operators. Under this heading fall many asymptotical problems about the solutions of differential or integral equations.

8.2. Iterates of a function. From now on we shall take all functions F_k to be equal to a fixed function f, and the problem becomes what is usually called an iteration problem. We have

$$x_1 = f(x_0),\ x_2 = f(x_1),\ x_3 = f(x_2), \ldots$$

We shall denote by f_n the n-th iterate of f, which is defined by

$$f_1(x_0) = f(x_0), \quad f_{n+1}(x_0) = f\{f_n(x_0)\} = f_n\{f(x_0)\} \qquad (n=0,1,2,\ldots).$$

Therefore we have $x_n=f_n(x_0)$ $(n=1,2,\ldots)$.

For a moment we assume that f is defined everywhere, so that there is no question about the f_n's being defined or not.

It may be possible that the sequence $x_0, x_1, x_2, \ldots$ tends to a limit c. If the function f is continuous at c, the relation $x_{n+1}=f(x_n)$ 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 $x_{n+1}, x_{n+2}, \ldots$ converges to c. In such cases it is likely that the asymptotical behaviour of x_n (as $n \longrightarrow \infty$) can be studied in detail, especially if f(x) is analytical 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, 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 $\lim_{n\to\infty} f_n(x)=c$. For, our assumptions imply that f maps J into itself, and therefore the same thing can be said about $f_2, f_3, \ldots$. Further $f(x) < x$ $(x \in J)$ guarantees that $x \geq f_1(x) \geq f_2(x) \geq \ldots$. As all $f_n(x)$ are in J, the sequence is bounded below. So lim $f_n(x)$ exists, and so 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 $\lim_{n\to\infty} f_n(x)$ need not be a continuous function of x. If, for example, we apply our previous results to the function f(x)=x+sin x it follows that the function $\lim_{n\to\infty} f_n(x)= \varphi(x)$ exists for all x, and is 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 just discussed. 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 $c < x < d$, $n \to \infty$) 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 can 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 $f_n^*(x)=f_n(x+c)-c$), and on behalf 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:

$$(8.2.1) \qquad f(x) = a_1x + a_2x^2 + a_3x^3 + \ldots \qquad (|x| < p)$$

where p is some positive number. The coefficients $a_1, a_2, \ldots$ 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 $x_0, x_1, x_2, \ldots$ converges to 0 indeed, provided that the starting point x_0 is sufficiently close to 0. Moreover the asymptotical 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 consta t. If $a_1=0$ the convergence is even much faster (see sec.8.4). 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 $x_n=f_n(x_o)$) we can derive a new formula, from which the asymptotical behaviour of x_n (as $n\longrightarrow\infty$) can be obtained, provided that $|x_o|$ is not too large.

We assume here that $a_1\neq 0$; the case $a_1=0$ will be discussed in sec. 8.3a.

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

$$(8.3.1)\qquad 0<|f_n(x)|<b^n|x| \qquad (n=1,2,\ldots;\ 0<|x|<p_1).$$

For, the power series for $x^{-1}f(x)$ has this 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_o satisfies $0<|x_o|<p_1$ we have $0<|x_1|<b|x_o|$, 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 $0<|x_n|<b^n|x_o|$, and (8.3.1) follows.

We next prove that, if x_o is fixed $(0<|x_o|<p_1)$, then $x_na_1^{-n}$ tends to a limit which we denote by $\omega(x_o)$. We have

$$(8.3.2)\qquad \frac{x_{n+1}}{a_1x_n}=\frac{f(x_n)}{a_1x_n}=1+\frac{a_2}{a_1}x_n+\frac{a_3}{a_1}x_n^2+\ldots,$$

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 $\prod_{k=0}^{n-1}$ equals $x_na_1^{-n}x_o^{-1}$, we infer that $x_na_1^{-n}$ tends to a limit $\omega(x_o)$, where

$$(8.3.3)\qquad \omega(x_o)=x_o\prod_{n=0}^{\infty}(1+r_n).$$

As $f_n(x_o)=f_{n-1}(f(x_o))=f_{n-1}(x_1)$, we also have $x_n\sim a_1^{n-1}\omega(x_1)$, and therefore $\omega(x_1)=a_1\omega(x_o)$. That is, the function ω satisfies the so-called Schroeder equation

$$(8.3.4)\qquad \omega(f(x))=a_1\omega(x) \qquad (|x|<p_1).$$

It should be remarked that $\omega(x)$ is analytical 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 analytical function of x_o, 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

$$(8.3.5) \qquad \omega(x) = x + d_2x^2 + d_3x^3 + \ldots \qquad (|x| < p_1).$$

The coefficients $d_2, d_3, \ldots$ can be determined recursively from the identity (8.3.4). Once we have determined $d_2, d_3, \ldots, d_{n-1}$, we can evaluate d_n by equating the coefficients of x^n in (8.3.4). This gives a linear equation for d_n, which does not contain $d_{n+1}, d_{n+2}, \ldots$. 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 get an explicit formula. The Lagrange inversion formula gives the inverse function Ω of ω , satisfying $\Omega(\omega(x))=x$ in a suitable circle $|x| < p_2$; and

$$\Omega(x) = x + e_2x^2 + e_3x^3 + \ldots \qquad (|x| < p_2).$$

We remark that the coefficients of Ω can be evaluated recursively without using the coefficients of ω . To this end we can use the identity $f(\Omega(y)) = \Omega(a_1y)$, which follows from (8.3.4) by putting $f(x)=y$.

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

$$(8.3.6) \qquad x_n = \Omega(a_1^n \omega(x_0)) = a_1^n \omega(x_0) + e_2 a_1^{2n} \omega(x_0)^2 + $$
$$+ e_3 a_1^{3n} \omega(x_0)^{3n} + \ldots \qquad (|x_0| < p_3).$$

This formula gives very satisfactory information about the behaviour of x_n when $n \longrightarrow \infty$.

Although it has not 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 $f_\lambda(f_\mu(x)) = f_{\lambda+\mu}(x)$ $(\lambda \geq 0, \mu \geq 0)$, $f_1 = f$, and f_0 is the identity $(f_0(x)=x)$. If λ is a positive integer f_λ is the λ-th iterate of f. The functions f_λ can be defined by

$$f_\lambda(x) = \Omega\{a_1^\lambda \omega(x)\} \qquad (|x| < p_3).$$

8.3a. Very rapid convergence. In sec. 8.3 we assumed that $a_1 \neq 0$. Here we 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 take $a_k = 1$; otherwise consider $f^*(x) = \alpha^{-1} f(\alpha x)$ where α 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:

$$\lim_{n \to \infty} \{f_n(x)\}^{k^{-n}} = \omega(x).$$

$$\omega(0) = 0, \quad \omega'(0) = 1.$$

$$\omega(f(x)) = (\omega(x))^k.$$

$$f_n(x) = \Omega \left\{ (\omega(x))^{k^n} \right\} \qquad (\Omega = \text{inverse of } \omega).$$

This is only the formal side of the matter, but it is not difficult to finish the details in the way it was done in sec.8.3.

8.4 . Slow convergence. Our next case is the iteration problem for a function f, in the form (8.2.1), with $|a_1|=1$. We shall treat a typical example, viz. $f(x)=\sin x = x - x^3/3! + x^5/5! - \dots$. 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 $\lim x_n = c$ exists, and that $c \geqslant 0$. It was remarked in sec.8.2 that c has to satisfy the relation $c=f(c)$, therefore we have $c=0$. We now raise the question of the asymptotic behaviour.

There is a difference between this case and the case of sec.8.3. In sec.8.3 (and also in sec.8.3a) 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, $x_0 = i\, t_0$ $(t_0 > 0)$, then we have $x_n = i\, t_n$, where $t_n = \sinh_n t_0$. And it is easily seen that $0 < t_0 < t_1 < \dots$, 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 $0 < x_0 < \pi$ ($x_1 = \sin x_0$ satisfies $-\pi < x_1 < \pi$ anyway, 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 solutions for the problem of the asymptotical 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 solution (secs.8.7 and 8.8) is more effective, but essentially restricted to iteration problems of the type we are presently dealing with.

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

$$(8.5.1) \qquad u_{n+1} = u_n - u_n^2 + O(u_n^3). \qquad (n=1,2,3,\dots).$$

What can be said about the asymptotical behaviour of u_n as $n \longrightarrow \infty$?

In the first place it is clear that nothing can be said if we do not assume something like $u_n \to 0$. 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

$$(8.5.2) \qquad |u_{n+1} - u_n + u_n^2| \leq A\, u_n^3 \qquad (n=1,2,3,\dots).$$

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 $u_k \to 0$. 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 $|c - c + c^2| < Ac^3$. As $Ap \leq 1$, $0 \leq c < p$, it follows that $c=0$.

We just learned that either all u_n are $\geq p$ or u_k tends monotonically to 0. For the latter case we can prove a much sharper result:

$$(8.5.3) \qquad \text{If } u_n \to 0, \text{ then } \; u_n = n^{-1} + O(n^{-2}\log n).$$

This can be proved as follows. By some simple 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

$$(8,5.4) \qquad 0 < x < n^{-1} + K\, n^{-2} \log n$$

we have

$$(8.5.5) \qquad 0 < x - x^2 + Ax^3 < (n+1)^{-1} + K(n+1)^{-2} \log(n+1).$$

Now let k be chosen such that $0 < u_k < N^{-1} + K\, N^{-2} \log N$; this is possible by virtue of the assumption $u_n \to 0$. Then it is easy to prove by induction that

$$0 < u_{k+m} < (N+m)^{-1} + K(N+m)^{-2} \log(N+m) \qquad (m=0,1,2,\dots).$$

Therefore $u_n < n^{-1} + O(n^{-2}\log n)$. A similar (but slightly more complicated) 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 $K\, n^{-2}\log n$ creates the possibility of passing 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 (7.5.1) 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)

$$v_{n+1} = v_n(1-u_n + O(u_n^2))^{-1} = v_n(1+u_n+O(u_n^2)),$$

and therefore

$$(8.5.6) \qquad v_{n+1} - v_n = 1 + O(v_n^{-1}), \qquad v_n \to \infty.$$

As $v_n > 2$ for all large n, we infer that $v_{n+1}-v_n > \frac{1}{2}$ and therefore $v_n > n/4$ for all large n, and it follows that $v_n^{-1}=O(n^{-1})$. Consequently $v_{n+1}-v_n=1+O(n^{-1})$, and this leads to $v_n=n+O(\log n)$. As $\{n+O(\log n)\}^{-1}= =n^{-1}+O(n^{-2}\log n)$, 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 is, however, not restricted to real sequences.

(i) If the sequence $\{a_n\}$ satisfies, for all n,

$$(8.5.7) \qquad a_{n+1}-a_n = O(n^{-1}) + O(n^{-2}a_n^2), \qquad \lim n^{-1}a_n=0,$$

then we have $a_n=O(\log n)$.

(ii) If the sequence $\{b_n\}$ satisfies, for all n

$$b_{n+1}-b_n = O(n^{-2}b_n^2), \qquad \lim n^{-1}b_n=0,$$

then we have $b_n=O(1)$.

Theorem (ii) can be used for the proof of (i). Theorem (i) is related to (8.5.1): assuming that we have already established $nu_n \to 1$

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 showed already that $x_n \to 0$. We have

$$x_{n+1} = \sin x_n = x_n - x_n^3/6 + x_n^5/120 + \dots$$

As the series contains only odd powers of x_n, the formula can be simplified by putting $x_n^2 = y_n$

$$y_{n+1} = y_n(1 - y_n/6 + y_n^2/120 \dots)^2 .$$

Writing $y_n = 3z_n$ we obtain something of the form of (8.5.1):

$$(8.6.1) \qquad z_{n+1} = z_n(1 - z_n/2 + 3z_n^2/40 - \dots)^2 = z_n - z_n^2 + 2z_n^3/5 + \dots$$

$$(z_n = x_n^2/3).$$

As z_n 0, (8.5.3) gives an asymptotic formula for z_n, viz. $z_n = n^{-1} + O(n^{-2} \log n)$.

Further results can be found by inserting this result into the equation (8.6.1). This leads to a better approximation of $z_{n+1} - z_n$, from which a new asymptotic formula for z_n can be obtained. This procedure of step-by-step improvement of an asymptotic formula was already described at the end of ch.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 $w_n \to \infty$. Further, w_n satisfies the recurrence relation

$$(8.6.2) \qquad w_{n+1} = w_n + 1 + 3/5w_n + O(w_n^{-2}).$$

It is not difficult to obtain the full development. By differentiation of the well-known series for cotg x it follows that

$$(8.6.3) \qquad (\sin x)^{-2} = \sum_{k=0}^{\infty} (-4)^k (1-2k) B_{2k}\, x^{2k-2}/(2k)! ,$$

where the B's are the Bernoulli numbers. If follows from $x_{n+1} = \sin x_n$ that

$$w_{n+1} = w_n \sum_{k=0}^{\infty} w_n^{-k} (1-2k)(-12)^k\, B_{2k}/(2k)! =$$

$$(8.6.4) \qquad = w_n + 1 + \frac{3}{5} w_n^{-1} + \frac{2}{7} w_n^{-2} + \frac{3}{25} w_n^{-3} + \frac{18}{385} w_n^{-4} + \dots .$$

Just as in our conclusions about (8.5.6), we have $w_n^{-1}=O(n^{-1})$, 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

$$t_n=C+\sum_{k=n}^{\infty} (t_k-t_{k+1}) - C+O(\sum_n^{\infty} k^{-2}\log k) = C+O(n^{-1}\log n).$$

Substituting $w_n=n+ \frac{3}{5} \log n + C + r_n$, $r_n=O(n^{-1}\log n)$ into (8.6.4), we obtain

$$r_{n+1}-r_n = - \frac{3}{25} \frac{\log n}{n^2} + \frac{41-42C}{70n^2} + O(\frac{\log^2 n}{n^3}),$$

and we easily infer by summation that

$$r_n = \frac{9}{25} \frac{\log n}{n} + \frac{-79+210C}{350n} + O(\frac{\log^2 n}{n^2}).$$

This procedure can be continued, and it is not difficult to show that there exists an asymptotic series for r_n, of the form

$$\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} c_{kl} (n^{-1}\log n)^k n^{-l}$$

In other words, w_n is of the form $n\, Q(n^{-1}\log n, n^{-1})$, where Q is an asymptotic series of the same type. Finally we get for $x_n=3^{\frac{1}{2}}w_n^{-\frac{1}{2}}$ an asymptotic series, of which we produce a few terms here:

(8.6.5) $$x_n=\sin_n x_o=3^{\frac{1}{2}}n^{-\frac{1}{2}}\left\{1- \frac{3}{10} \frac{\log n}{n} - \frac{C}{2n} + n^{-2}(\alpha \log^2 n+ \beta\log n+ \gamma)+ + O(n^{-3}\log^3 n)\right\},$$

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_o. It is remarkable that the first two terms of the asymptotic series are independent of x_o.

In order to find out something about how C depends on x_o, we replaced C by $\psi(x_o)$ and we consider the following abbreviation of (8.6.5):

(8.6.6) $$x_n=3^{\frac{1}{2}}n^{-\frac{1}{2}} \left\{1- \frac{3}{10} \frac{\log n}{n} - \frac{\psi(x_o)}{2n} + O(n^{-2}\log^2 n)\right\}.$$

As $x_{n+1}=\sin x_n=\sin_n(\sin x_o)$, we have

$$x_{n+1}=3^{\frac{1}{2}}n^{-\frac{1}{2}}\left\{1-\frac{3}{10}\,\frac{\log n}{n}-\frac{\psi(x_o)}{2n}+O(n^{-2}\log^2 n)\right\},$$

As $x_{n+1}=\sin x_n=\sin_n(\sin x_o)$, we have

$$x_{n+1}=3^{\frac{1}{2}}n^{-\frac{1}{2}}\left\{1-\frac{3}{10}\,\frac{\log n}{n}-\frac{\psi(\sin x_o)}{2n}+O(n^{-2}\log^2 n)\right\},$$

and so

$$x_{n+1}-x_n=\tfrac{1}{2}.3^{\frac{1}{2}}n^{-3/2}\left\{\psi(x_o)-\psi(\sin x_o)\right\}+O(n^{-5/2}\log^2 n).$$

On the other hand, we have $x_{n+1}-x_n=\sin x_n-x_n=-x_n{}^3/6+O(x_n{}^5)=-3^{3/2}n^{-3/2}/6+$ $+O(n^{-5/2}\log n)$. Therefore, it results that ψ satisfies the equation

$$(8.6.7) \qquad \psi(\sin x_o)-\psi(x_o)=1.$$

The function ψ is uniquely determined by (8.6.6):

$$\psi(x_o)=\lim_{n\to\infty}\left\{2n-\frac{3}{5}\log n-2x_n\,n^{3/2}\,3^{-\frac{1}{2}}\right\}.$$

It can be observed from this formula that $\psi(x)=\psi(\pi-x)$ $(0<x<\pi)$, and that $\psi(x)$ is non-increasing in the interval $0<x<\pi/2$. (For, if x_o decreases, and n is fixed, then x_n decreases). Furthermore we have $\psi(x)\to\infty$ if $x\to 0$ $(x>0)$.

8.7. An alternative method. Our second method for dealing with the asymptotic behaviour of $\sin_n x_o$ has some analogy to the contents of sec.8.3. It starts from the Schroeder equation (8.3.4), or, what is slightly simpler in our case, the socalled Abel equation $\psi\{f(x)\}-\psi(x)=1$ (the connection between the two is expressed by the relation $\omega(x)=a_1{}^{\psi(x)}$). It is a functional equation for the unknown function ψ, whereas f is a given function. In the present case it becomes

$$(8.7.1) \qquad \psi(\sin x)-\psi(x)=1.$$

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 $0<x\leq\pi/2$, it is quite easy to describe the general solution of (8.7.1): Choose an arbitrary function ψ^* 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 ψ such that (8.7.1) is satisfied. Next we define it in the interval $\sin_2 1\leq x\leq\sin 1$, such that (8.7.1) 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 asumptotic behaviour (as $x\to 0$) 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 (8.7.1), we start with a simpler problem: we ask for elementary functions ψ for which $\psi(\sin x)-\psi(x)$ tends rapidly to 1 as $x \to 0$. One of the first functions to try is a monomial $\psi_1(x)=ax^{-b}$ $(b>0)$. It gives

$$\psi_1(\sin x) - \psi_1(x) = a(\sin x)^{-b}-ax^{-b} =$$
$$= ax^{-b}\left\{(1-x^2/6+\dots)^{-b}-1\right\} = ab\ x^{-b+2}/6 + \dots$$

We want it to approximate 1, and this is achieved by taking b=2, a=3. So $\psi_1(x)=3x^{-2}$, and

$$\psi_1(\sin x) - \psi_1(x) = 1 + \frac{x^2}{5} + \frac{2x^4}{63} + \frac{x^6}{225} + \dots$$

(the series on the right equals $12\sum_0^\infty (-4)^k(2k+1)\ B_{2k+2}\ x^{2k}/(2k+2)!$, cf.(8.6.3)).

We next want to modify ψ_1 in order to compensate the term $x^2/5$. Therefore we try to find a function χ_1 such that $\chi_1(\sin x)-\chi_1(x)$ is approximately $-x^2/5$. We notice that a monomial does not work now. The following argument shows that it is worth while to try log x: 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

$$\chi_1(\sin x)-\chi_1(x) = \frac{6}{5}\log(\sin x/x) = -\frac{x^2}{5} - \frac{x^4}{150} + \dots .$$

If we take $\psi_2=\psi_1+\chi_1$, we have

$$\psi_2(\sin x) - \psi_2(x) = 1 + p(x),$$

$$p(x) = (+\frac{2}{63} - \frac{1}{150})\ x^4+\dots = \frac{3}{\sin^2 x} - \frac{3}{x^2} + \frac{6}{5}\log\frac{\sin x}{x} .$$

We can go on this way, by choosing χ_2 such that $\chi_2(\sin x)-\chi_2(x)$ equals approximately $-p(x)$, and so on. In this way we would still be constructing approximate solutions instead of approximations to a solution. Fortunately we are able to indicate an exact solution of the equation $\chi(\sin x)-\chi(x)=-p(x)$. This equation is satisfied by the function

$$(8.7.2) \qquad \chi_2(x) = p(x) + p(\sin x) + p(\sin_2 x) + p(\sin_3 x) + \dots .$$

We only need to show that the series is convergent for all real values of x. This immediately follows from the formulas $p(y)=O(y^{-4})$ $(y \to 0)$ and $\sin_n x = O(n^{-\frac{1}{2}})$ $(n \to \infty)$, whence $p(\sin_n x) = O(n^{-2})(n \to \infty)$.

The formula $\sin_n x = O(n^{-\frac{1}{2}})$ can be proved in the way we proved (8.5.3). We shall, however, indicate a short independent proof: If ε is a positive constant, then the iterates of the function $f(x)=x(1+\varepsilon x^2)^{-\frac{1}{2}}$ can be evaluated explicitly. We easily find $f_n(x)=x(1+n\varepsilon x^2)^{-\frac{1}{2}}$. Furthermore it is not difficult to show that ε can be chosen such that $\sin x < f(x) < x$ $(0 < x < 1)$. Now if x is such that $\sin x > 0$ we 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 $\sin_n x < f_n(x)$, and so $\sin_n x = O(n^{-\frac{1}{2}})$. The assumption $\sin x > 0$ is, of course, no essential restriction.

We just established the convergence of (8.7.2), but we shall need something more, viz. an estimate for $\chi_2(x)$ as $x \to 0$. We shall prove that

$$(8.7.3) \qquad \chi_2(x) = O(x^2) \qquad (0 < x \leq \tfrac{1}{2}\pi) \quad .$$

We established before that $p(x)=O(x^4)$, $p(\sin_n x)=O(n^{-2})$. We can easily be more specific: There are constants C_1 and C_2 such that $|p(x)| < C_1 x^4$, $|p(\sin_n x)| < C_2 n^{-2}$ $(0 < x \leq \frac{1}{2}\pi$, $n=1,2,\dots)$. 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

$$|\chi_2(x)| < C_1 n^{-1} + C_2 n x^4.$$

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

$$|\chi_2(x)| < x^2(\tfrac{1}{3} C_1 + 6C_2),$$

which proves (8.7.3).

As χ_2 exactly satisfies $\chi_2(\sin x) + \chi_2(x) = -p(x)$, the function

$$(8.7.4) \qquad \psi(x) = 3x^{-2} + \frac{6}{5} \log x + \chi_2(x),$$

χ_2 given by (7.7.2), is an exact solution of (8.7.1). We have the estimate (8.7.3) for χ_2, but it is not difficult to see that there is an asymptotic series

$$\chi_2(x) \sim c_2x^2 + c_4x^4 + c_6x^6 + \dots \; . \tag{8.7.5}$$

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 $\chi_2(\sin x) - \chi_2(x) = -p(x) = \alpha x^4+\dots$, with $\alpha=-79/3150$. Furthermore $(\sin x)^k-x^k= -kx^{k+2}/6+\dots$. $(k=1,2,\dots)$; here we use the case $k=2$. Now put $\chi_2(x)=-3\alpha x^2+\chi_3(x)$, then $\chi_3(\sin x)-\chi_3(x)=-q(x)$ is a power series starting with a term βx^6. And from $\chi_3(x)= =q(x)+q(\sin x)+ q(\sin_2 x)-\dots$ we argue, in the way we proved (8.7.3), that $q(x)=0(x^4)$. So $\chi_2(x)=-3\alpha x^2+0(x^4)$. 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 $\chi_3(x),\dots$).

8.8. Final discussion about the iterated sine. The asymptotic information about ψ, obtained in sec.8.7, will now be applied to the problem about $\sin_n x$. To this end we start from the formula

$$\psi(\sin_n x) = n + \psi(x), \tag{8.8.1}$$

which is a direct consequence of (8.7.1). We shall consider x as a fixed number, and n as a large integer. As ψ 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 asymptotical behaviour of the solution $\sin_n x$ in terms of the parameter n is a problem of the type discussed in ch.1.

We replace the unknown $\sin_n x$ by the single letter u. The equation can be written as

$$3u^{-2} + \frac{6}{5}\log u + \frac{79}{1050}u^2 +\dots \sim n + \psi(x) \qquad (n\to\infty). \tag{8.8.2}$$

In some respects the question is more difficult than the problems considered in ch.1. Firstly, the left-hand-side represents the function $\psi(u)$ asymptotically, but probably not exactly, and $\psi(u)$ is probably not analytic at $u=0$. Secondly, we are not yet sure, that ψ 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.1 is that the asymptotic behaviour of u is unknown.

To start with, 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 left-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 right-hand-side. At the same time we substract on both sides (3/5) log 3 in order to simplify further calculations. So we have

$$(8.8.3) \qquad 3u^{-2} + \frac{6}{5}\log(3^{-\frac{1}{2}}u) = n + \psi(x) - \rho - \frac{3}{5}\log 3.$$

Here ρ is equal to $\psi(u)$ minus the left-hand-side of (8.8.3). So by (8.7.4) and (8.7.5) it is asymptotically

$$(8.8.4) \qquad \rho \sim c_2u^2 + c_4u^4 + \dots \; (u \to 0), \text{ and therefore } \rho = O(n^{-1})(n \to \infty).$$

Replacing the right-hand-side of (8.8.3) by y, we have

$$3u^{-2} + \frac{6}{5}\log(3^{-\frac{1}{2}}u) = y,$$

and $y=n+O(1)$, for x is a fixed number. Putting $u=(3/y)^{\frac{1}{2}}v$ we get

$$(8.8.5) \qquad v^{-2} = 1 + \frac{3}{5}\frac{\log y}{y} - \frac{6}{5y}\log v.$$

We know that $v \to 1$, and so $\log v = O(1)$. Raising both sides of (8.8.5) to the power $-\frac{1}{2}$, we get

$$(8.8.6) \qquad v = 1 - \tfrac{1}{2}\left\{\frac{3}{5}\frac{\log y}{y} - \frac{6}{5y}\log v\right\} + O(y^{-2}\log^2 y)$$

and so $v=1 + O(y^{-1}\log y)$, $\log v = O(y^{-1}\log y)$.

Inserting this into (8.8.6) we get

$$v = 1 - \frac{3}{10}\frac{\log y}{y} + O(y^{-2}\log^2 y),$$

and so $\log v = -\frac{3}{10}y^{-1}\log y + O(y^{-2}\log^2 y)$. Now we again raise (8.8.5) to the power $-\frac{1}{2}$, and we develop a little further; according to the formula

$$(1+z)^{-\frac{1}{2}} = 1 - \tfrac{1}{2}z + \tfrac{3}{8}z^2 + O(z^3).$$

Then we obtain

$$v = 1 - \tfrac{1}{2}\left\{\frac{3}{5}\frac{\log y}{y} - \frac{6}{5y}\log v\right\} + \frac{3}{8}\left(\frac{3}{5}\frac{\log y}{y}\right)^2 + O(y^{-3}\log^3 y).$$

Contenting ourselves with a third order error, we thus have

$$(8.8.7) \qquad v = 1 - \frac{3}{10}\frac{\log y}{y} + \frac{27}{200}\frac{\log^2 y}{y^2} - \frac{9}{50}\frac{\log y}{y^2} + O(y^{-3}\log^3 y).$$

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 $v=e^{\frac{1}{2}w}$), and it follows that the asymptotic series is even convergent.

We next replace, in (8.8.7), v by $3^{-\frac{1}{2}}y^{\frac{1}{2}}u$, and y by $n+\psi(x)+$ $+\rho-\frac{3}{5}\log 3$, $\rho=O(n^{-1})$, and then we get a $\sin_n x$ expressed in terms of n, with an error term $O(n^{-5/2}\log^2 n)$:

$$u=\sin_n x = (3/n)^{-\frac{1}{2}}\left\{1-\frac{3}{10}n^{-1}\log n-\frac{1}{2}(\psi(x)-\frac{3}{5}\log 3)n^{-1}+\right. \tag{8.8.8}$$
$$\left.+O(n^{-2}\log^2 n)\right\} .$$

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 (8.7.1). Incidentally this shows that $\psi(x)$ is decreasing $(0 < x < \pi/2)$.

Refinements of (8.8.8) can be found as follows. On behalf of (8.8.8) we can improve our formula $\rho=O(n^{-1})$. Formula (8.8.4) gives

$$\rho = 3\, c_2 n^{-1} + O(n^{-2}\log n).$$

Inserting this into (8.8.7), with $y=n+\Psi(x)-\rho$, we get 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 one cannot be expected to be convergent in the ordinary sense.

Ch.9. Differential Equations

9.1. Introduction. Many problems in pure and applied mathematics are concerned with the behaviour of the solutions of a differential equation at a singular point. Such problems are obviously of asymptotical nature, for by a transformation of the independent variable it is always possible to transform the singularity to infinity, and the question takes the following form. Let $F(t,y,y',\ldots,y^{(m)})=0$ be a given differential equation for the unknown function $y=y(t)$. How do the solutions behave as $t\longrightarrow\infty$?

Such questions arise, for instance, in stability problems, in problems about linear or non-linear oscillations and in quantum mechanics. Another type of application lies in the fact that one can study the asymptotical 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 a vast literature about it. It is out of 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 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 on asymptotic formula, or even anasymptotic series, and more often than not it is much 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 asymptotical 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 $\varphi'(t)<F(t,\varphi(t))$ $(a\leq t\leq b)$, $\varphi(a)\leq y(a)$. 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 φ satisfying $\varphi'<F(t,\varphi)$ are of course quite easy to find).

Proof: 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 again follows. 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 $\leq$, depends on more delicate considerations about the uniqueness of solutions. For, if φ and y were different solutions of the differential equation, both having the same value at $t=a$, then $\varphi(t) \leq y(t)$ $(a < t \leq b)$ 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 ≥ 0. We consider the differential equation

$$t^{-k} \rho'(t) = \alpha(t) + \beta(t)\rho(t) + \gamma(t)\rho^2(t) \tag{9.2.1}$$

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).

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 $t > 0$, the following existence theorem holds. If $t_0 > 0$ and ρ_0 are given real numbers, then either there exists a solution ρ, with $\rho(t_0)=\rho_0$, in the whole interval $t_0 \leq t < \infty$, or there is a number $t_1 > t_0$ and a solution $\rho(t)$ in the interval $t_0 \leq t < t_1$, such that $\rho(t_0)=\rho_0$ and $\rho(t)$ tends to $+\infty$ or to $-\infty$ if t tends to t_1 from the left.

This fact will be repeatedly used in the following way. If we have numbers t_0, ρ_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 $\rho(t_0)=t_0$ (and satisfying $|\rho(t)| \leq A$ $(t_0 \leq t < \infty)$).

As, in our case, $F(t, \rho)$ has a continuous derivative with respect to ρ, there is no difficulty about the uniqueness: any solution is uniquely determined by its initial conditions.

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 ρ has a singularity at $t=t_1$: it tends to $\pm\infty$ if t tends to t_1. The existence and location of roots t_1, however, obviously depends on the choice of the ratio A:B of the integration constants in the general solution $y(t)=Ay_1(t)+By_2(t)$.

Returning to the equation (9.2.1), we shall make the further assumption that the asymptotical behaviour of each of the coefficients α, β, γ (as $t\to\infty$) is given by an asymptotic series:

$$\begin{aligned} \alpha(t) &\sim a_0 + a_1t^{-1} + a_2t^{-2} + \dots && (t\to\infty), \\ \beta(t) &\sim b_0 + b_1t^{-1} + b_2t^{-2} + \dots && (t\to\infty), \\ \gamma(t) &\sim c_0 + c_1t^{-1} + c_2t^{-2} + \dots && (t\to\infty), \end{aligned} \tag{9.2.2}$$

and we assume that

$$b_0 < 0, \quad c_0 = 0. \tag{9.2.3}$$

It is our aim to prove that there is a class of solutions, whose behaviour is also given by an asymptotic series

$$\rho(t) \sim r_0 + r_1t^{-1} + r_2t^{-2} + \dots . \qquad (t\to\infty). \tag{9.2.4}$$

More precisely, we shall show the existence of numbers $t_0 > 0$ and $A > 0$ such that for every number ρ_0 in the interval $-A \le \rho_0 \le A$ the solution with $\rho(t_0)=\rho_0$ can be continued indefinitely to the right, and has an asymptotic series of the form (9.2.4). Moreover, the coefficients $r_0, r_1, r_2, \dots$ are independent of ρ_0.

It is not generally true that all solutions have 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_1 can be determined such that

$$\beta(t) < 0, \qquad 2|\alpha(t)| < A|\beta(t)| \qquad (t > t_1), \tag{9.2.5}$$

and we can determine $t_0 > t_1$ such that

$$2A|\gamma(t)| < |\beta(t)| \qquad (t \ge t_0). \tag{9.2.6}$$

Let φ_1 denote the constant function $\varphi_1(t)=-A$ $(t \ge t_0)$, and φ_2 the constant function $\varphi_2(t)=A$ $(t \ge t_0)$. It follows from (9.2.5) and (9.2.6) that

$$t^{-k}\varphi_1'(t) < \alpha(t) + \beta(t)\varphi_1(t) + \gamma(t)\varphi_1^2(t) \qquad (t \ge t_0),$$

$$t^{-k}\varphi_2'(t) > \alpha(t) + \beta(t)\varphi_2(t) + \gamma(t)\varphi_2^2(t) \qquad (t \ge t_0).$$

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 indefinitly 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: $\rho(t)=O(1)$ $(t \to \infty)$. A solution will be called bounded if it is bounded in some interval (t_3, ∞), although it is possible that this solution can be continued over some interval (t_4, t_3) or even over $(-\infty, t_4)$ without being bounded over the extended interval.

In our next step we show that $\rho(t)=O(1)$ $(t \to \infty)$ implies that

$$\rho(t) = r_0 + O(t^{-1}) \quad (t \to \infty), \text{ where } r_0=-a_0/b_0. \tag{9.2.7}$$

We consider a special solution $\rho(t)$, bounded in some interval $t_0 \leq t < \infty$. Again we introduce two auxiliary functions

$$\varphi_3(t) = r_0 - A\,t^{-1}, \qquad \varphi_4(t) = r_0 + A\,t^{-1},$$

and we try to determine $A > 0$ and $t_1 > t_0$ such that

$$t^{-k}\varphi_3'(t) < \alpha(t) + \beta(t)\varphi_3(t) + \gamma(t)\varphi_3^2(t) \qquad (t \geq t_1), \tag{9.2.8}$$

$$t^{-k}\varphi_4(t) > \alpha(t) + \beta(t)\varphi_4(t) + \gamma(t)\varphi_4^2(t) \qquad (t \geq t_1), \tag{9.2.9}$$

$$\varphi_3(t_1) \leq \rho(t_1) \leq \varphi_4(t_1). \tag{9.2.10}$$

To this end we remark that both sides of (9.2.8) have asymptotic series. In both sides the constant term of the series vanishes. The coefficient of the term t^{-1} vanishes on the left; on the right it equals (notice that $c_0=0$)

$$a_1 - b_0 A + b_1 r_0 + c_1 r_0^2 .$$

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 $\varphi_3(t)=r_0+O(At^{-1})$, $\varphi_3^2(t)=O(1)+O(A^2t^{-2})$, $\alpha(t)=a_0+O(t^{-1})$, $\beta(t)=b_0+O(t^{-1})$, $\gamma(t)=O(t^{-1})$, $t^{-k}\varphi_3(t)=O(At^{-k-2})$, where all O-symbols refer to $t \geq t_0$, $A \geq 1$, say. It follows that $\beta(t)\varphi_3(t)=b_0r_0-b_0At^{-1}+O(t^{-1})+O(At^{-2})$

$\gamma(t)\varphi_3^2(t)=O(t^{-1})+O(A^2t^{-3})$,

and the right-hand side of (9.2.8) exceeds the left-hand-side by the amount

$$-b_oAt^{-1}+O(t^{-1})+O(At^{-2})+O(A^2t^{-3})=At^{-1}\{-b_o+O(A^{-1})+O(t^{-1})+O(At^{-2})\}.$$

It follows that (9.2.8) is true when A,t_1 and $t_1^2A^{-1}$ are sufficiently large, and the same thing applies, of course, to (9.2.9). Next we consider (9.2.10). As $\rho(t)-r_o=O(1)$, (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

$$A>C,\ t_1>C,\ t_1^2A^{-1}>C,\ At_1^{-1}>C.$$

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

$$\varphi_3(t)\le\rho(t)\le\varphi_4(t)\qquad(t\ge t_1),$$

and thus we have proved (9.2.7).

Next we write $\rho(t)=r_o+t^{-1}\rho_1(t)$, 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

$$(9.2.11)\qquad t^{-k}\rho_1'(t)=\alpha_1(t)+\beta_1(t)\rho_1(t)+\gamma_2(t)\rho_1^2(t).$$

with

$$\alpha_1=t(\alpha+r_o\beta)+r_o^2t\gamma,$$
$$\beta_1=\beta+2r_o\gamma+t^{-k-1},$$
$$\gamma_1=t^{-1}\gamma.$$

The new coefficients α_1, β_1, γ_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 $r_o,r_1,r_2,\dots$ follow successively from the above procedure. It is easier, however, to proceed by the method of undetermined coefficients. Just substitute the formal series $r_o+r_1t^{-1}+r_2t^{-2}+\dots$ 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,...), the coefficients of t^{-n} on both sides are equal. This produces a set of equations from which, in virtue of $b_o \neq 0$, $c_o = 0$, the numbers $r_o, r_1, \ldots$ 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, and if t is given, then there exists a positive number ε such that any solution whose value at t_o satisfies $|\rho_2(t_o) - \rho_1(t_o)| < \varepsilon$, also satisfies $\rho_2(t) - \rho_1(t) \to 0$ $(t \to \infty)$. For a special value of t_o it is contained in our previous results, and it is not difficult to show it for arbitrary values of t_1. It is quite a strong type of stability, for ρ_1 and ρ_2 have the same asymptotic series, and therefore $\rho_2 - \rho_1 = O(t^{-n})$ $(n = 1, 2, 3, \ldots)$. We shall even show that $\rho_2 - \rho_1 = O(e^{-ct})$ with some positive constant c.

Let ρ_1 and ρ_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, in 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 $\beta \sim b_o + \ldots$, $b_o < 0$, and $\eta(t) > 0$ for all t, we have, for some $t_1 > 0$,

$$\eta'/\eta < \tfrac{1}{2}b_o t^k \qquad (t > t_1).$$

It follows by integration that $\eta = O\left\{\exp(b_o t^{k+1}/2(k+1))\right\}$.

9.3. An unstable case. We again consider the equation (9.2.1)

$$(9.3.1) \qquad \rho'(t) = \alpha(t) + \beta(t)\rho(t) + \gamma(t)\rho^2(t),$$

and again we assume that α, β and γ have asymptotic developments (9.2.2), but instead of (9.2.3) we assume

$$(9.3.2) \qquad b_o > 0, \quad c_o = 0.$$

Formally nothing has been changed, and therefore we can again find a series $r_o + r_1 t^{-1} + r_2 t^{-2} + \ldots$ which formally satisfies the equation (in this formal procedure the sign of b_o is irrelevant; it is only the condition $b_o \neq 0$ that matters). The difference with the case $b_o < 0$ is that in the present case there is only one solution having this series as its development, and that one is the only bounded solution (i.e. the only solution which is $O(1)$ when $t \to \infty$).

We start by defining the functions $\varphi_3(t) = r_o - At^{-1}$, $\varphi_4(t) = r_o + At^{-1}$

$(r_0=-a_0/b_0)$, just as in sec.9.2. And again we can fix $A>0$ and $t_1>0$ such that

$$(9.3.3) \qquad t^{-k}\varphi_3'(t) > \alpha(t) + \beta(t)\,\varphi_3(t) + \gamma(t)\,\varphi_3^2(t) \qquad (t \geq t_1)$$

$$(9.3.4) \qquad t^{-k}\varphi_4'(t) < \alpha(t) + \beta(t)\,\varphi_4(t) + \gamma(t)\,\varphi_4^2(t) \qquad (t \geq t_1).$$

The inequality signs are different from those in (9.2.8) and (9.2.9). In the present case the conclusion is, that if a solution $\rho(t)$ exists in an interval (t_2,t_3), where $t_2 \geq t_1$, 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 $\varphi_3(t)$ at $t=t_2$, where $t_2 \geq t_1$, then $\rho(t)>\varphi_3(t)$ 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.

We just made a statement about solutions $<\varphi_3$ or $>\varphi_4$, but we would rather know something about solutions between φ_3 and φ_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 is as follows: Assume $t_1<t_2$, and let the number ρ_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 $\varphi_3(t)<\rho(t)<\varphi_4(t)$ $(t_1 \leq t \leq t_2)$. (The solution determined by $\rho(t_2)=\rho_2$ cannot tend to $\pm\infty$ when t tends to some t_3 $(t_1 \leq t_3 < t_2)$ from the right, for then there would be a number t_4 $(t_3<t_4<t_2)$ where $\rho(t_4) = \varphi_3(t_4)$ or $\varphi_4(t_4)$, and we would have a contradiction).

If ρ_2 ranges through the closed interval $\varphi_3(t_2) \leq \rho_3 \leq \varphi_4(t_2)$, then the value $\rho_1=\rho(t_1)$ ranges through a sub-set of the interval $\varphi_3(t_1)<\rho<\varphi_4(t_1)$.

If follows from the general theory of differential equations that ρ_1 is a continuous and increasing function of ρ_3, and therefore this sub-set 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 ρ_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. It follows that $i(t_2) \supset i(t_2+1)$. Now applying this with $t_2=t_1+n$ $(n=1,2,3,\ldots)$ we get a sequence of closed intervals

$$i(t_1+1) \supset i(t_1+2) \supset i(t_1+3) \supset \ldots ,$$

and therefore these intervals have a number ρ_1^* in common. Denoting the solution with the initial value ρ_1^* at t_1 by $\rho^*(t)$, we know that this one can be continued up to t_1+n, and stays between $\varphi_3(t)$ and $\varphi_4(t)$ $(t_1 \leq t < t_1+n)$. This holds for any value of n, and therefore

$\rho^*(t)$ can be continued to infinity, and

$$\varphi_3(t) \leq \rho^*(t) \leq \varphi_4(t) \qquad (t > t_1).$$

Thus we have proved that our equation has a solution of the form $r_o+O(t^{-1})$. We next show that there is only one such solution. Let $\rho_1(t)$ and $\rho_2(t)$ denote solutions which are both bounded in an interval $t_o \leq t < \infty$. We suppose them to be different: $\rho_2(t) > \rho_1(t)$ $(t > t_o)$, say. Putting $\rho_2(t)-\rho_1(t) = \eta(t)$, we would have (cf. the end of sec. 9.2.), for some t_1,

$$\eta'/\eta > \tfrac{1}{2}b_o t^k \qquad (t > t_1)$$

and it would follow that

$$\eta > c \exp\left\{b_o t^{k+1}/2(k+1)\right\} \qquad (t > t_1) \tag{9.3.5}$$

with some positive constant c. So η would tend to infinity, whereas ρ_1 and ρ_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, as in sec.9.2,

$\rho(t)=r_o+t^{-1}\rho_1(t)$, and for ρ_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 $r_o+t^{-1}\rho_1^*(t)$ is a solution of (9.3.1), and it has the form $r_o+r_1t^{-1}+O(t^{-2})$. As (9.3.1) has only one bounded solution, we have identically

$$r_o + t^{-1}\rho_1^*(t) = \rho^*(t),$$

and therefore

$$\rho^*(t) = r_o + r_1t^{-1} + O(t^{-2}).$$

This procedure can be continued, and it follows that $\rho^*(t)$ has an asymptotic development (9.2.4). The numbers $r_o, r_1, \ldots$ can again be obtained by formal substitution of (9.2.4) into the differential equation, and 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 Riccatti 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 Riccatti type (i.e. a linear relation between w', w^2, w and 1). Now one can try to obtain a Riccatti equation of one of the types discussed in the previous sections, or anyway an equation to which the techniques of those sections can be applied.

As an example we take the equation

$$y''(t) - t^4y(t) = 0, \tag{9.4.1}$$

for which the substitution $y'/y=v$ leads to

$$v' + v^2 - t^4 = 0. \tag{9.4.2}$$

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 solutions $v=(t-t_0)^{-1}$ with 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 is 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, $v=t^2+\rho(t)$, transforms (9.4.2) into

$$t^{-2}\rho'=-2t^{-1}-2\rho -t^{-2}\rho^2, \tag{9.4.3}$$

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 $a_0=0$, $b_0<0$). We infer that there is a solution $\rho(t)$ with an asymptotic series

$$\rho(t)\sim r_0 + r_1\ t^{-1} + r_2\ t^{-2} + \dots , \qquad (t\to\infty),$$

(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 substitute $\rho(t)=t^{-1}\sigma(t^3)$, $t^3=\tau$, 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

$$(9.4.4) \qquad y \sim C\,t^{-1}e^{\frac{1}{3}t^3}(1+a_1t^{-3}+a_2t^{-6}+\dots) \qquad (t\to\infty),$$

where $a_0+a_1x+a_2x^2+\dots$ is the formal development of $\exp(-\frac{1}{3}r_4x-\frac{1}{6}r_7x^2-\dots)$, and $C\neq0$.

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 we are in the unstable case of sec.9.3. We now infer that there exists just <u>one</u> solution of the form

$$\rho(t) \sim s_0 + s_1t^{-1} + s_2t^{-2} + \dots \qquad (t\to\infty).$$

In terms of y it means that there is, apart form the arbitrary constant $C\neq0$ just one solution of the form

$$(9.4.5) \qquad y \sim C\,t^{-1}\,e^{-\frac{1}{3}t^3}(1+b_1t^{-3}+b_2t^{-6}+\dots) \qquad (t\to\infty).$$

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 $y_2/y_1\to0$). Now the general solution of (9.4.1) is $A\,y_1(t) + B\,y_2(t)$ (A and B constants). This illustrates the instability of the solution y_2: Every solution with $A\neq0$ is easily seen to have (9.4.4) as its asymptotical behaviour (with some value of C), and only if $A=0$, $B\neq0$ we have something of the type (9.4.5). Moreover, it is easily seen that when adding $B\,y_2(t)$ to $A\,y_1(t)$, the asymptotical series for $A\,y_1(t)$ 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

$$(9.4.6) \qquad b_n = (-1)^n a_n \qquad (n=0,1,2,\dots).$$

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) is P(-t). This is easily deduced from the state of affairs with the Riccatti equation (9.4.2). If we substitute for v the formal series $Q_1(t)=t^2+r_0+r_1t^{-1}+r_2t^{-2}+\dots$, 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 $-t^2+s_0+s_1t^{-1}+s_2t^{-2}+\dots$. It follows that $s_n=-(-1)^nr_n$ $(n=0,1,2,\dots)$. Now (9.4.6) is an easy consequence.

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 $y''-t^4y=0$ 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 $y''-t^4y=0$. It follows from the asymptotical behaviour that $y_2(-t)$ is positive and increasing if t is negative and large. It follows from the equation $y''=t^4y$ that the solutions are convex whenever they are positive, and therefore $y_2(-t)$ keeps increasing as t tends to $+\infty$. As $y_2(t)\longrightarrow 0$ $(t\rightarrow\infty)$, 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

$$y_2(t)\sim e^{-\frac{1}{3}t^3}\sum_{n=0}^{\infty}\frac{(3n)!}{18^n(n!)^2}(-1)^n\ t^{-3n-1}\qquad(t\rightarrow\infty).$$

9.5. <u>Oscillatory cases</u>. 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 $y''(t)+y(t)f(t)=0$ (again with $f(t)>0$). Under very general conditions it can be proved that all solutions are oscillating, i.e. they have infinitely many zeros in the interval $(0,\infty)$. We shall consider the special case

$$y''(t)+(1+t^{-1})\ y(t)=0. \tag{9.5.1}$$

We obtain a Riccati equation

$$v'+v^2+(1+t^{-1})=0$$

by the substitution $y'/y=v$. Applying the same heuristic argument we used in the case of (9.4.2), we are led to a substitution $v=i+t^{-1}\rho(t)$, and $\rho(t)$ satisfies

$$\rho'(t)=-1+(-2i+t^{-1})\ \rho(t)-t^{-1}\ \rho^2(t). \tag{9.5.2}$$

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 can not 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 Re $b_0 < 0$, and $b_0 > 0$ in (9.3.2) by 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 coeffcients. The first few terms are

$$\tfrac{1}{2}i + (2-i)/(8t) - (4+3i)/(16t^2) + \dots .$$

However, at the present stage we cannot say whether this formal series has any significance.

We shall attempt an entirely different method. We consider an equation of the type

$$y''(t) + \{1 + g(t)\} y(t) = 0, \tag{9.5.3}$$

where the given function g is continuous and satisfies $\int_0^{\infty} g(t)\,dt < \infty$. This means that the results can not be applied directly 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

$$y(t) = A(t) \cos t + B(t) \sin t,$$
$$y'(t) = -A(t) \sin t + B(t) \cos t, \quad A'(t)\cos t + B'(t) \sin t = 0,$$
$$y''(t) = -y(t) - A'(t) \sin t + B'(t) \cos t, \quad -A'(t)\sin t + B'(t)\cos t = +h(t).$$

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

$$y(t) = C_1 \cos t + C_2 \sin t - \int_a^t g(\tau)\, y(\tau) \sin (t-\tau)\,d\tau , \tag{9.5.4}$$

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 such that $\int_a^{\infty} |g(t)|\,dt < \tfrac{1}{2}$, which is possible by virtue 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

$$M \leq |C_1| + |C_2| + M\int_a^b |g(\tau)|\,d\tau \leq |C_1| + |C_2| + \tfrac{1}{2}M,$$

and so $M \leq 2|C_1| + 2|C_2|$, irrespective of the value of b. This shows the boundedness.

The boundedness implies the convergence of $\int_0^\infty g(\tau)y(\tau)\sin(t-\tau)\,d\tau$, and therefore we can rewrite (9.5.4) into the form

$$(9.5.5)\qquad y(t) = C_3\cos t + C_4\sin t - \int_t^\infty g(\tau)y(\tau)\sin(\tau-t)\,d\tau\ ,$$

with new constants C_3 and C_4:

$$C_3=C_1+\int_0^\infty g(\tau)y(\tau)\sin\tau\ d\tau\ ,\quad C_4=C_2-\int_0^\infty g(\tau)y(\tau)\cos\tau\ d\tau\ .$$

If C_3 and C_4 are given, at most one solution of (9.5.3) satisfies (9.5.5). In order to show this, it suffices to consider the case $C_3=C_4=0$ (otherwise consider the difference of two solutions). Then we infer from (9.5.5) that

$$\sup_{a\leq t<\infty}|y(t)| \leq \sup_{a\leq t<\infty}|y(t)|\ .\ \int_a^\infty g(\tau)\,d\tau\ .$$

By taking a so large that the integral is less than 1, we obtain that $y(t)$ vanishes identically if $t \geq a$.

On the other hand the general solution of (9.5.3) involves two constants, and we now infer that C_3 and C_4 can be arbitrarily prescribed: to any choice of C_3 and C_4 there corresponds just one solution of the differential equation. If y_1 is the solution corresponding with $C_3=1$, $C_4=0$, and y_2 the one with $C_3=0$, $C_4=1$, then Ay_1+By_2 (with constants A and B) gives the general solution. 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. We shall not do this, as our only aim is to obtain asymptotic information about $y(t)$. This is achieved by iteration, in a way similar to the derivation of the Neumann series. We choose the solution y_1, say, with $C_3=1$, $C_4=0$. We know already that $y(t)$ is bounded, and so we infer from (9.5.5) that

$$(9.5.6)\qquad y_1(t) = \cos t + O\Big\{\int_t^\infty |g(\tau)|\ d\tau\Big\} = \cos t + o(1).$$

Next we insert this result into the integral on the right-hand-side of (9.5.5):

$$y_1(t)=\cos t - \int_t^\infty g(\tau)\cos\tau\ \sin(\tau-t)\,d\tau + O\Big(\int_t^\infty |g(s)|\,ds\int_s^\infty g(\tau)|\,d\tau\Big),$$

and so on.

For the calculations it may be easier to deal with the complex combinations $y_1(t)+i\,y_2(t) = e^{it}+o(1)$, $y_1(t)-i\,y_2(t) = e^{-it}+o(1)$.

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

$$(9.5.7)\qquad y''(t)+(1+\lambda t^{-2})\,y(t)=0,$$

which is indeed of the type discussed above (λ is a constant). So we know that there is a solution of the type $e^{it}+o(1)$ $(t\to\infty)$. 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) $(C_3=1,\ C_4=i)$

$$(9.5.8)\qquad y(t) = e^{i} - \lambda\int_t^\infty \tau^{-2}e^{i\tau}\sin(\tau-t)d\tau + O(\int_t^\infty \tau^{-3}d\tau).$$

We first consider the integrals $\int_t^\infty \tau^{-k}e^{it}\sin(\tau-t)d\tau = f_k(t)$, with k fixed, and $k\geqslant 2$. We have, as $2i\sin x = e^{ix}-e^{-ix}$,

$$e^{-it}f_k(t) = \int_0^\infty (x+t)^{-k}\,e^{ix}\sin x\,dx = \tfrac{1}{2}i(k-1)^{-1}t^{-k+1}-\tfrac{1}{2}i\int_0^\infty (x+t)^{-k}e^{2ix}dx.$$

This last integral can be expanded asymptotically by means of repeated partial integration, or by steepest descent. It results that $e^{-it}f_k(t)$ has an asymptotical expansion of the form $t^{-k+1}(c_0+c_1t^{-1}+c_2t^{-2}+\dots)$ $(t\to\infty)$.

Now (9.5.8) gives $y(t)=e^{it}(1+a_1t^{-1}+O(t^{-2}))$ (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 $y(t)=e^{it}(1+a_1t^{-1}+a_2t^{-2}+O(t^{-3}))$. Continuing this procedure, we get an asymptotic series

$$(9.5.9)\qquad y(t)\sim e^{it}(1+a_1t^{-1}+a_2t^{-2}+a_3t^{-3}+\dots)\qquad (t\to\infty).$$

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. $\varphi=y(t)e^{-it}$ satisfies $\varphi''+2i\varphi'+\lambda t^{-2}\varphi=0$. Formally substituting $\varphi(t)\sim 1+a_1t^{-1}+a_2t^{-2}+\dots$, and equating coefficients, we can evaluate the a_i. In order to justify this procedure it is sufficient to show that $\phi'(t)$ and $\phi''(t)$ also have asymptotic series (for then those series automatically are the derived series of the one for $\varphi(t)$). It follows from (9.5.7) and (9.5.9) that

$$y''(t)\sim -e^{it}\,(1+\lambda t^{-2})(1+a_1t^{-1}+\dots)\qquad (t\to\infty).$$

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 $e^{it}(b_0+b_1t^{-1}+b_2t^{-2}+\dots)$. 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 ' and " have asymptotic series of the required type.

So in order to determine the coefficients in $\varphi(t) \sim 1+a_1 t^{-1}+a_2t^{-2}+\dots$, we have the right of formal substitution into $\varphi''+2i\varphi'+\lambda t^{-2}\varphi=0$. Equating coefficients, we get the relations

$$a_{k+1} = a_k(k^2+k+\lambda)/2(k+1)i \qquad (k=0,1,2,\dots),$$

where $a_0=1$.

The 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 becomes $e^{-it}(1-a_1t^{-1}+a_2t^{-2}-a_3t^{-3}+\dots)$. So every solution of (9.5.7) has an expansion of the type

$$A\ e^{it}(1+a_1t^{-1}+a_2t^{-2}+\dots) + B\ e^{-it}(1-a_1t^{-1}+a_2t^{-2}-\dots).$$

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 $t\to 0$, and there our knowledge about $t\to\infty$ 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 $e^{-it}+o(1)$ $(t\to\infty)$, provided that $\int_0^\infty |g(t)|\,dt$ converges. We shall now try to reduce the more general equation

$$y''(t) + (p(t))^2\, y(t) = 0 \tag{9.6.1}$$

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

$$x = \varphi(t),\ y = \psi(t)\ z\ .$$

Here φ and $\psi(t)$ are functions of t, to be determined later. We assume that $\varphi(t)$ tends monotonically to $+\infty$ when $t\to +\infty$. We obtain (accents denoting differentiation with respect to t):

$$y' = \psi\varphi'\frac{dz}{dx} + z\psi'\ ,$$

$$y''(t) = \psi(\varphi')^2\frac{d^2z}{dx^2} + (\psi\varphi''+2\psi'\varphi')\frac{dz}{dx} + \psi''z.$$

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 chosing φ and ψ such that

$$(9.6.2) \qquad \frac{\varphi''}{\varphi'} + 2\,\frac{\psi'}{\psi} = 0, \qquad \int_0^{\infty}\left|\left(\frac{p}{\varphi'}\right)^2 + \frac{\psi''}{\psi\varphi'^2} - 1\right| dx < \infty ,$$

then we know from what we proved about (9.5.3) that (9.6.1) has solutions

$$\psi(t)\left\{e^{i\varphi(t)} + o(1)\right\} \quad \text{and} \quad \psi(t)\left\{e^{-i\varphi(t)} + o(1)\right\} \qquad (t\to\infty).$$

Bearing in mind that $\psi''/\psi = (\psi'/\psi)' + (\psi'/\psi)^2$, and $dx=\varphi' dt$, we can replace (9.6.2) by

$$(9.6.3) \qquad \int_0^{\infty}\left|\frac{p^2}{\varphi'^2} - \frac{1}{2}\,\frac{\varphi'''}{\varphi'^3} + \frac{3}{4}\left(\frac{\varphi''}{\varphi'^2}\right)^2 - 1\right|\varphi'\, dt < \infty$$

If, for example

$$(9.6.4) \qquad \int_0^{\infty}\left|3\, p'^2\, p^{-3} - 2\, p''\, p^{-2}\right| dt < \infty ,$$

it is possible to choose φ such that $\varphi'=p$, i.e. $\varphi(t)=\int_0^t p(\tau)d\tau$, and then (9.6.3) is satisfied (see A. Wintner, Phys.Rev. 72, 516-517 (1947)).

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 asymptotical behaviour is given by

$$(9.6.5) \qquad y_1 \sim p(t)^{-\frac{1}{2}} \exp\left\{i\int_0^t p(\tau)d\tau\right\}, \quad y_2 \sim p(t)^{-\frac{1}{2}} \exp\left\{-i\int_0^t p(\tau)d\tau\right\}. \qquad (t\to\infty).$$

The substitutions $x=\varphi(t)$, $y=p(t)^{-\frac{1}{2}}z$ which have to be carried out in this case, transform the equations into the form $d^2z/dx^2+(1+g(x))z=0$, 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.

In a quite wide range of cases the substitution $\varphi'=p$, can be used. The condition (9.6.4) is satisfied, for example, if $p(t)=t^{\lambda}$ $(\lambda>-1)$. It fails,however, if $p(t)=t^{-1}$. In that case we can still satisfy (9.6.3), but then with $\varphi'(t)=3^{\frac{1}{2}}/(2t)$. 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 $\leq 1/4$, 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), that is the special case of (9.6.1) with $p(t)=(1+t^{-1})^{\frac{1}{2}}$. 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 $\varphi(t)=\int^{t}(1+\tau^{-1})^{\frac{1}{2}}d\tau$, $\psi(t)=(1+t^{-1})^{-\frac{1}{2}}$. This means that $\varphi(t)=t+\log t + C + O(t^{1}-1)$, $\psi(t)=1+O(t^{-1})$, where C is a constant. It follows that (9.5.1) has solutions $y_1(t)$ and $y_2(t)$ with the asymptotical behaviour (when $t\rightarrow\infty$)

$$y_1(t) = e^{it+\frac{1}{2}i \log t}\{1+o(1)\}, \quad y_2(t)=e^{-it-\frac{1}{2}i \log t}\{1+o(1)\} .$$

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

$$(9.6.6) \qquad (1+t^{-1})S^{-2} = \tfrac{1}{2} S''S^{-3} - \tfrac{3}{4}(S'S^{-2})^{2} + 1.$$

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 $y_1(t)=\psi(t)e^{i\varphi(t)}(1+O(t^{-n}))$, 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

$$y_1(t)\sim e^{it+\frac{1}{2}i \log t}\, S(t)^{-\frac{1}{2}} \exp(-is_2t^{-1}-\tfrac{1}{2}is_3t^{-2}-\dots) \qquad (t\rightarrow\infty),$$

$$y_2(t)\sim e^{-it-\frac{1}{2}i \log t}\, S(t)^{-\frac{1}{2}} \exp(is_2t^{-1}+\tfrac{1}{2}is_3t^{-2}+\dots) \qquad (t\rightarrow\infty).$$

These lines have to be interpreted in the following sense. The function $y_1(t)e^{-it-\frac{1}{2}i \log t}$ has an asymptotic series which is obtained by formal multiplication of the formal series $S(t)^{-\frac{1}{2}}$ 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.9.6).

6. Applications of the saddle point method

In ch.5 we gave several quite simple applications of the saddle point method, intended to illustrate the principles of the method with easy examples. As 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 be extensively studied in ch.2.

The second problem (sec.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 problem.

In the third example 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 collection 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 5 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 $n \longrightarrow \infty$.

There is a recurrence relation, expressing d_{n+1} in terms of $d_1,\ldots,d_n$:

$$(6.1.1) \qquad d_{n+1} = \binom{n}{0} d_0 + \binom{n}{1} d_1 + \ldots + \binom{n}{n} d_n \qquad (n=0,1,2,\ldots),$$

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 summing 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 $D'(z) = e^z.D(z)$, and, as $D(o)=1$, it follows that $D(z)=\exp(e^z-1)$. Therefore, the $d_n/n!$ are the coefficients in the expansion

$$(6.1.2) \qquad \exp(e^z-1) = \sum_{n=0}^{\infty} d_n z^n/n!$$

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 sub-sets; let s_j denote the number of subsets having j elements. So $s_j \geq 0$ ($j=1,2,3,\dots$), and $n=s_1+2s_2+3s_3+\dots$. We now fix the 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

$$n!\left\{(1!)^{s_1}(2!)^{s_2}(3!)^{s_3}\dots s_1!\,s_2!\,s_3!\,\dots\right\}^{-1} .$$

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

$$\exp(z) . \exp(z^2/2!) \exp(z^3/3!)\dots = \exp(z+z^2/2!+\dots)=\exp(e^z-1).$$

6.2. Asymptotical behaviour. We shall study the asymptotical behaviour of the coefficients in (6.1.2) by expressing them via Cauchy's formula for the coefficients of a power series:

$$(6.2.1) \qquad 2\pi i\, e\, d_n/n! = \int_C \exp(e^z)\, z^{-n-1}\, dz,$$

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 $(2k-1)\pi < \mathrm{Im}\ z < (2k+1)\pi$ $(k=0,\pm 1,\pm 2,\dots)$, 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 u we have $|\exp(e^z)| \leq \exp(e^u)$, as, on that line, we have $\mathrm{Re}\ e^z \leq |e^z| = e^{\mathrm{Re}\ z} = e^u$. 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 contributions 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 $\mathrm{Re}\ z \leq u$. 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

$$(6.2.2) \qquad 2\pi e\, d_n/n! = \exp(e^u-ue^u\log u)\int_{-\infty}^{\infty}\exp(\phi(y))dy,$$

where

$$\phi(y) = e^u\left[(e^{iy}-1)-u\log(1+iyu^{-1})\right].$$

As $|\exp\phi(y)| = \exp\mathrm{Re}\,\phi(y)$, we have to study

$$\mathrm{Re}\,\phi(y) = e^u\left[-1+\cos y - u\log(1+y^2u^{-2})^{\frac{1}{2}}\right].$$

As long as y is not too large, the terms $-1+\cos y$ are predominant. Therefore, there is a maximum at $y=0$, 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 $\log(1+y^2/u^2)^{\frac{1}{2}} > \frac{1}{2}y^2/u^2$, and therefore

$$\left|\int_{\pi}^{u}\exp\phi(y)dy\right| < u\ .\ \exp\left\{-e^u.\tfrac{1}{2}\,\pi^2u^{-1}\right\}.$$

If $y > u$, then we use $1+y^2/u^2 > 2y/u$, and so we have, putting $y=ux$,

$$\left|\int_{u}^{\infty}\exp(\phi(y))dy\right| \leq u\int_{1}^{\infty}\exp\left\{-e^u u\log(2x)\right\}dx.$$

It is easily seen that $\int_1^{\infty}(2x)^{-p}\,dx = O(e^{-\frac{1}{2}p})$, and therefore

$$\int_u^{\infty}\exp(\phi(y)dy = O(u.\exp\left[-\tfrac{1}{2}ue^u\right]).$$

It follows that

$$(6.3.3)\qquad \int_{-\infty}^{\infty}\exp(\phi(y))dy - \int_{-\pi}^{\pi}\exp(\phi(y))dy = O\{\exp(-4e^u/u)\},$$

and we can now direct our attention to the interval $(-\pi,\pi)$, where the saddle point at $y=0$ gives the main contribution.

As

$$\phi(y) = e^u\left[-\frac{y^2}{2!} + \frac{(iy)^3}{3!} + \ldots - \frac{y^2}{2u} + \frac{(iy)^3}{3u^2} - \ldots\right],$$

we find by the Laplace method (ch.4),

$$(6.2.4)\qquad \int_{-\pi}^{\pi}\exp\,\phi(y)dy = (2\pi e^{-u})^{\frac{1}{2}}\,(1+O(u^{-1})).$$

In order to get an asymptotic expansion, it seems to be inconvenient to follow the method of sec.4.4, as ϕ 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 $\exp(\phi(y))$ is transformed into $\exp(-f(u).w^2)$ ($f(u)$ depending on u only).

Thus far y was real; we shall now treat it as a complex variable. If $|y|$ is small, we define w by

$$e^{iy}-1-u\log(1+iyu^{-1}) = -\tfrac{1}{2}w^2(1+u^{-1}),$$

choosing the root w which satisfies $dw/dy=+1$ at $y=0$. Now w can be written as

$$(6.2.5)\qquad w = y + y^2\,P(y,u^{-1})\ ,$$

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)):

$$y=y(w) = \frac{1}{2\pi i}\int_C\left\{\eta+\eta^2P(\eta,u^{-1})-w\right\}^{-1}\eta\,\frac{d}{d\eta}\left\{\eta+\eta^2P(\eta,u^{-1})\right\}d\eta\ ,$$

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 need its derivative

$$dy/dw = 1 + w\,\psi_1(u^{-1}) + w^2\psi_2(u^{-1}) + \ldots\ ,$$

where $\psi_1, \psi_2, \ldots$ are convergent for large values of u.

It is not quite sure that the integral (6.2.4) can be transformed succesfully 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 paths becomes a curve which crosses the saddle point at $w=0$. The errors made by braking 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 $\exp(-Ce^u)$.

Our final result is

$$(2\pi)^{-\frac{1}{2}}U^{\frac{1}{2}} \int_{-\infty}^{\infty} \exp(\phi(y))dy \sim 1+a_1\psi_2(u)U^{-1} + a_2\psi_4(u)U^{-2}+\dots \quad (u\to\infty),$$

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:

$$(6.2.6) \qquad d_n = n!\, e^{-1}(2\pi)^{-\frac{1}{2}} (1+u^{-1})^{-\frac{1}{2}} \left\{1+O(e^{-u})\right\} \exp(e^u-ue^u\log u-\tfrac{1}{2}u).$$

Here u is related to n, by the formula

$$ue^u = n+1, \qquad u>0.$$

The asymptotic behaviour of u, as $n\to\infty$, 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, $\sigma=(\log t)^{-1}$, $\tau=\log\log t/\log t$, and $t=n+1$. 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

$$e^u-ue^u\log u -\tfrac{1}{2}u = t(u^{-1}-\log u) + O(\log t) =$$

$$= \left\{-\log\log t + \frac{1}{\log t} + \frac{\log\log t}{\log t} + \frac{1}{2}\left(\frac{\log\log t}{\log t}\right)^2 + O\left(\frac{\log\log t}{(\log t)^2}\right)\right\},$$

where t is still n+1, but it is easy to see that on replacing n+1 by n, we make an error which is much smaller than the error already involved. Further, we use a rough estimate for n! :

$$\log n! = n\log n - n + O(\log n),$$

and we find that

$$(6.2.7) \qquad \frac{\log d_n}{n} = \log n-\log\log n-1+ \frac{\log\log n}{\log n} + \frac{1}{\log n} + \frac{1}{2}\left(\frac{\log\log n}{\log n}\right)^2 + O\left(\frac{\log\log n}{(\log n)^2}\right).$$

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. <u>Alternative method</u>. We indicate a 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 $\exp(e^x)$ as $\sum_{k=0}^{\infty} e^{kx}/k!$, 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

$$(6.3.1) \qquad d_n = e^{-1} \sum_{k=0}^{\infty} k^n/k!$$

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 solution of the equation $ue^u=n+1$. For, $k^n/k!$ equals, roughly, $(2\pi)^{\frac{1}{2}}k^{-\frac{1}{2}}$. 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 $|k-k_{max}| < n^{-\frac{1}{2}}$ that gives a substantious 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. <u>The sum S(s,n).</u> In sec.4.7 we obtained the asymptotical behaviour of S(s,n), defined by

$$(6.4.1) \qquad S(s,n) = \sum_{k=0}^{2n} (-1)^{n+k} \binom{2n}{k}^s ,$$

if s is a fixed integer >1, and $n \to \infty$. The method was definitely restricted to this case, as s-1 occurred as the number of dimensions of an euclidean space. In the present section we shall study this sum for general real values of s (s fixed, $n \to \infty$). 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 quite difficult problem with various interesting aspects.

If s=2,3,... we have, by (4.7.4),

$$(6.4.2) \qquad S(s,n) \sim \{2\cos(\pi/2s)\}^{2ns+s-1}\, 2^{2-s}(\pi n)^{\frac{1}{2}(1-s)} s^{-\frac{1}{2}} \qquad (n \to \infty).$$

This formula is definitely false if s=0, for we have $S(0,n)=(-1)^n$. If it is negative, it does not hold either. In that case, we are in the situation that the first few terms, and the last few, are prevailing. We evidently have the asymptotic series

$$S(s,n) \sim 2.(-1)^n \left\{1 - \binom{2n}{1}^s + \binom{2n}{2}^s - \dots\right\} \qquad (s < 0,\ n \to \infty),$$

or, more explicitly, for h and s fixed,

$$(6.4.3) \qquad S(s,n) = 2.(-1)^n \sum_{k=0}^{h} (-1)^k \binom{2n}{k}^s + O(n^{(h+1)s})$$
$$(s < 0,\ n \to \infty).$$

(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 is at most 2n times the first neglected term).

If $s > 0$, there is not such a trivial way. We shall then replace the sum by an integral, by application of the residue theorem:

$$(6.4.4) \qquad S(s,n) = \int_C \left(\frac{\Pi(2n)}{\Pi(n+z)\Pi(n-z)}\right)^s \frac{dz}{2i \sin \pi z} .$$

Here the integration path C is a curve which encircles the points $-n,-n+1,\ldots,-1, 0,1,2,\ldots,n$ just once, but does not encircle the points $\pm(n+1), \pm(n+2),\ldots$. We may 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 Π by the formula $\Pi(z)=\Gamma(z+1)$.

The function $\Pi(n+z)\Pi(n-z)$ has poles at the points $\pm(n+1)$, $\pm(n.2),\ldots$, 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 lies entirely inside the domain D. Now it is easy to prove (6,4,4): the residue at z=h ($h=0,\pm1,\ldots,\pm n$) is $\{(2n)!/(n+h)!(n-h)!\}^s$. $(-1)^h.(2\pi i)^{-1}$, 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 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 $|(n-z)/(n+z+1)|^s$. 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. It is well-known that the Stirling formula holds in the complex plane provided that we remove a sector containing the negative real axis. Precisely, if δ is a positive number ($0 < \delta < \pi$), and if R_δ denotes the sector $|\arg z| < \pi - \delta$, then we have

$$(6.4.5)\qquad \Pi(z) = (2\pi)^{\frac{1}{2}}\, z^{z+1}\, e^{-z}\left\{1+O(|z^{-1}|)\right\} \qquad (z\in R_{\delta}).$$

This means that the integrand 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,∞). 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 second 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 p be a positive number. We define P_N and Q_N by

$$P_N = \int_{-N+\frac{1}{2}+ip}^{N+\frac{1}{2}+ip} \Omega\, dz, \qquad Q_N = \int_{(N+\frac{1}{2})} \Omega\, dz.$$

Here Ω 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 p is a positive constant. The integration path of Q_N starts at $N+\frac{1}{2}$, proceeds 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 sifficiently large, that

$$\lim_{N\to\infty} P_N = P \quad , \quad \lim_{N\to\infty} Q_N = Q$$

both exist, and that

$$(6.4.6)\qquad S(s,n) = -2P + 2Q.$$

These statements easily follow from the fact that the integrand Ω is $O(|z|^{-2})$ in the domain described by $\mathrm{Re}\, z > 1$, $|\mathrm{Im}\, z| \leq p$, $|z-2| > \frac{1}{4}$, $|z-3| > \frac{1}{4}, \ldots$, provided that n is fixed but sufficiently large (it suffices that $(2n+1)s > 2$). For, by the functional equation $\Pi(z)\Pi(-z) = \pi z/\sin \pi z$, we have

$$\Omega = \left\{(2n)!\right\}^{s}\left\{\Pi(z-n)/\Pi(z+n)\right\}^{s}(2i)^{-1}\pi^{-s}z^{-s}(\sin\pi z)^{s-1} = \\ =0(|z|^{-2ns-s}).$$

As to formula (6.4.6) we refer to remark (i), which shows that $S(s,n)$ equals $2P_N+2Q_N$ 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 $N\longrightarrow\infty$, since $\Omega=0(|z|^{-2})$(n fixed).

There are now two separate problems, viz. the asymptotic behaviour of P and Q respectively, as $n\longrightarrow\infty$.

6.5. We shall first deal with $P=\int_{ip-\infty}^{ip+\infty}\Omega\,dz$, where p is a fixed positive number. Clearly P is independent of p (cf.6.4.6). We shall approximate Ω with the aid of Stirling's formula. However, a slight adjustment of (6.4.5) is necessary, as our values $n+z$ and $n-z$ do not always belong to the sector R_{δ}. If we use the relation $\Pi(z)\Pi(-z)= =\pi z/\sin\pi z$, the behaviour of $\Pi(z)$ in the quadrant $\mathrm{Im}\, z>0$, $\mathrm{Re}\, z<0$ can be deduced from its behaviour in the opposite quadrant. By a careful discussion of the arguments of $z^{z+\frac{1}{2}}$ and $(-z)^{-z+\frac{1}{2}}$ we find that

$$\Pi(z)\left\{1-e^{2\pi iz}\right\}\quad(2\pi)^{\frac{1}{2}}\,z^{z+\frac{1}{2}}\,e^{-z}\left\{1+0|z^{-1}|\right\}\qquad(\mathrm{Re}\, z<0,\ \mathrm{Im}\, z>0).$$

It now easily follows that we have, if p_0 is positive and fixed,

$$\Pi(z)=(2\pi)^{\frac{1}{2}}\,z^{z+\frac{1}{2}}\,e^{-z}\left\{1+0(|z^{-1}|)+0(e^{-2\pi|\mathrm{Im}\, z|})\right\},\qquad(\mathrm{Im}\, z>p_0).$$

From this formula we can immediately deduce an estimate for Ω (i.e. the integrand of (6.4.4)) by some trivial calculations.

$$(6.5.1)\quad \Omega = -2^{2ns}(\pi n)^{-\frac{1}{2}s}(1-\zeta^2)^{-\frac{1}{2}s}\exp\left[n\left\{-s\log(1-\zeta^2)-s\zeta\log\frac{1+\zeta}{1-\zeta}+\pi i\zeta\right\}\right]\times$$

$$\times\left\{1+0(n^{-1}+n^{-1}|1+\zeta|^{-1}+n^{-1}|1-\zeta|^{-1}+e^{-2\pi n\,\mathrm{Im}\,\zeta})\right\}\quad(n>1,\ \mathrm{Im}\, n\zeta>p_0).$$

Here p_0 is positive and fixed, and ζ stands for z/n.

The integrand is, roughly speaking, of the type considered in sec.5.7, if we put

$$(6.5.2)\qquad \varphi(\zeta) = -s\log(1-\zeta^2) - s\zeta\log\frac{1+\zeta}{1-\zeta}+\pi i\zeta .$$

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 in the upper half-plane. Disregarding the O-terms for the moment, we start looking for saddle points. We have

$$(6.5.3)\qquad \phi'(\zeta)=\pi i - s\log(1+\zeta)/(1-\zeta),$$

and therefore $\phi'(\eta)=0$ if $\eta=i\tan(\pi/2s)$. So we observe that there is a saddle point in the upper half-plane if and only if $s>1$. Assuming $s>1$, we take as the integration path, the infinite straight horizontal line through the saddle point η. As we have $\phi''(\eta)=-2s(1-\eta^2)^{-1}=$ $=-2s(1+\tan^2(\pi/2s))^{-1}=-2s\cos^2(\pi/2s)<0$, 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 η is "the highest point" of the path. Re $\phi(\zeta)$ decreases if x increases from 0 to ∞, and also if x runs from 0 to $-\infty$. For, by (6.5.3) we have

$$\text{Re}\,\frac{d\phi}{dx}=-s\log\left|\frac{1+\zeta}{1-\zeta}\right|<0 \qquad (x>0)$$

as $|1+\zeta|>|1-\zeta|$ if $x>0$. If $x<0$, we have a similar thing.

The fact that Re $\phi(\zeta)$ decreases (if $x\longrightarrow\infty$), 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 $\phi(\zeta)$ as $x\longrightarrow\infty$. Actually we infer that

$$\left|\int_{\eta+1}^{\infty}\Omega\,d\zeta\right|<C2^{2ns}(\pi n)^{-\frac{1}{2}s}.\ \left|\exp(n\phi(\eta+1))\right|\int_{\eta+1}^{\infty}\left|(1-\zeta^2)^{-\frac{1}{2}s}\right|d\zeta$$

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 $\int_{-\infty}^{\eta-1}$ can be estimated in the same way.

By sec.5.7 we have

$$(6.5.4)\quad \int_{\eta-1}^{\eta+1}(1-\zeta^2)^{-\frac{1}{2}s}\exp\{n\phi(\zeta)\}d\zeta=(2\pi)^{\frac{1}{2}}n^{-\frac{1}{2}}|\varphi''(\eta)|^{-\frac{1}{2}}.$$
$$\exp\{n\varphi(\eta)\}(1-\eta^2)^{-\frac{1}{2}s}\{1+0(n^{-1})\}.$$

We have $\varphi''(\eta)=-2s(\cos\pi/2s)^2$; $\exp\{n\phi(\eta)\}=(\cos(\pi/2s))^{2ns}$; $(1-\eta^2)^{-\frac{1}{2}s}=(\cos\pi/2s)^s$ and therefore the value of (6.5.4) reduces to

$$(6.5.5)\qquad (\pi/ns)^{\frac{1}{2}}(\cos\pi/2s)^{2ns+s-1}\{1+0(n^{-1})\}.$$

The 0-terms on (6.5.1) can be reduced to one term $0(n^{-1})$, by virtue of the fact that Im ζ is positive and fixed on our path. It is not difficult to show that

$$\int_{\eta\div1}^{\eta+1}\left|\exp(n\phi(\zeta)).0(n^{-1})\right|d\zeta=0\{n^{-3/2}\exp(n\phi(\eta))\},$$

and therefore these 0-terms result in a correction of the same order as the 0-term in (6.5.5).

Collecting the various results, we obtain our final formula for P:

$$(6.5.6)\quad P=\int_{\eta-\infty}^{\eta+\infty}\Omega\,dz=n\int_{\eta-\infty}^{\eta+\infty}\Omega\,d\zeta=-\{2\cos(\pi/2s)\}^{2ns+s-1}\,2^{1-s}(\pi n)^{\frac{1}{2}}(1-s)_s{}^{-\frac{1}{2}}.$$
$$\{1+0(n^{-1})\},$$

valid if s is fixed and >1.

It may be remarked that the factor $\{1+0(n^{-1})\}$ 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 η 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 to infinity in the vertical direction. In fact we can show that, if $0 < s \leq 1$, $n > \frac{1}{2}(s^{-1}-1)$, the integral P vanishes. For, if n is fixed and Im $z > p_0$ (p_0 positive and fixed), $|z| \to \infty$, we have, using for Ω the expression in the integrand of (6.4.4)

$$\Omega = O((\Pi(n+z)\Pi(n-z))^{-s} |\sin \pi z|^{-1}) =$$

$$= O(|z|^{-2ns-s}|\sin \pi z|^{s-1}) = O(|z|^{-(2n+1)s}) .$$

So if $(2n+1)s > 1$, we have $\Omega = O(|z|^{-\lambda})$, with a constant $\lambda > 1$. Therefore the integral P vanishes.

6.6. We next turn our attention to Q, defined in sec.6.4. We of course 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 of 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.

For a first orientation, we remark that the only factor depending "heavily" on n is, apart form the trivial multiplier $\{(2n)!\}^s$, the factor $\Pi(z+n)$. By a well-known formula (actually a consequence of the Stirling formula) we have, $\Pi(z+n) \sim \Pi(n).n^z$ (z fixed, $n \to \infty$). Therefore it seems to be sensible to pull the integration path to the right, as far as possible. We can of course not pull it over the branch point at z=n+1.

We shall now deform the integration path of the integral Q by making a path following the real axis from ∞ 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 ∞ with the values from the upper half-plane. This path has to be provided with small semi-circles around n+2, n+3,... 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 not defined originally as an integral from ∞ to ∞ , 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 is no essential restriction.

The factor $\{\Pi(n-z)\}^{-s}$ seems to be awkward. We can replace it by $(\Pi(z-n-1))^s \pi^{-s}(\sin \pi(z-n))^s$, using the functional equation $\Pi(w)\Pi(-w)=\pi w/\sin \pi w$. By this procedure, the singularities are shifted from the non-elementary function Π to the elementary function sin, where they are easier to handle. So for Ω we write

$$\Omega=(-1)^n \pi^{-s}(2i)^{-1}((2n)!)^s\{\Pi(z-n-1)/\Pi(z+n)\}^s(\sin \pi(z-n))^{s-1}.$$

The Π-factor here is single-valued in the half-plane 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 with $\pi(s-1)$ if we pass the branch point $n+k(k=1,2,3,\dots)$ by a semi-circle 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, writing $z=n+x$,

$$(6.6.1)\quad Q=(-1)^{n+1}\pi^{-s}\int_1^\infty\{(2n)!\,\Pi(x-1)/\Pi(x+2n)\}^s.|\sin \pi x|^{s-1}.\\ .\sin\{\pi(s-1)[x]\}dx.$$

(Here $[x]$ 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 the formula

$$(6.6.2)\qquad (2n)!/\Pi(x+2n) = (2n)^{-x}\{1+O(n^{-1})\},$$

where the constant implied in the O-symbol does not depend on x. (This uniformity in x is lost if we take the interval $1\leq x<\infty$). Formula (6.6.2) is a well-known consequence of the Stirling formula.

Therefore we obtain, with the integrand of (6.6.1),

$$\int_1^2 = \sin \pi(s-1).\int_1^2 (2n)^{-xs}\{\Pi(x-1)\}^s|\sin \pi x|^{s-1}\{1+O(n^{-1})\}dx.$$

This integral is of the type of those discussed in sec.4.3, log 2n playing the rôle of the parameter t occurring in that section. Writing $x=1+y$, and comparing the integral with

$$\int_0^\infty e^{-y\, s\log 2n}\, y^{s-1}\,dy = \Gamma(s)\ (s\log 2n)^{-s},$$

we easily obtain that

$$(6.6.3)\qquad \int_1^2 = \pi^{s-1}\Gamma(s)\ (2ns\log 2n)^{-s}\ \{\sin \pi(s-1)+O((\log n)^{-1})\}.$$

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 $\int_1^\infty$ requires some careful attention as (6.6.2) does not hold uniformly with respect to **x**. On the other hand a rough estimate will do: we shall show that it 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) $$\Pi(x+2n)/\{(2n)!\ \Pi(z-1)\} > C_1\ x^K n^2. \quad (x \geq 2,\ n \geq K).$$

The left hand side can be written as

$$x(1+\frac{x}{1})(1+\frac{x}{2})\dots(1+\frac{x}{2n}) ,$$

and this is

$$> x^K\{(K-1)!\}^{-1} \prod_{h=K}^{2n} (1+2h^{-1}) \geq C_2 x^K \prod_{h=1}^{2n} (1+2h^{-1}),$$

with $C_2^{-1}=(K-1)!\prod_{h=1}^{K-1} (1+2h^{-1})$. The product $\prod_1^{2n}$ can be compared to $\Pi(1+2h^{-1}+h^{-2})$:

$$\prod_1^{2n}(1+2h^{-1})=\prod_1^{2n} (1+h^{-1})^2.\prod_1^{2n}\{1-(h+1)^{-2}\} \geq (2n+1)^2\prod_1^{\infty}\{1-(h+1)^{2}\}.$$

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 $\int_2^\infty$ (with the integrand of (6.6.1)): $\left|\int_2^\infty\right| < n^{-2s}\int_2^\infty C_1^{-s}x^{-Ks}|\sin \pi x|^{s-1}dx$, and as the integral on the right is convergent (as $K > s^{-1}$), and independent of n, we have

$$\left|\int_2^\infty\right| < C_3 n^{-2s} \qquad\qquad (C_3 \text{ independent of } n).$$

Comparing this to (6.6.3) we observe that n^{-2s} can be absorbed into the O-term $n^{-s}(s \log n)^{-s}\ O((\log n)^{-1})$. Therefore, our final result for Q (cf.(6.6.1)) is $(s > 0)$.

(6.6.5) $$Q=(-1)^n\ \pi^{-1}\ \Gamma(s)\ (2ns \log 2n)^{-s}\{\sin \pi s + O((\log n)^{-1})\} .$$

6.7. We have thus, for all real values of s, obtained the asymptotic behaviour of $S(s,n)$. 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 $1 < s < 3/2$ P is much smaller than Q as $2 \cos(\pi/2s) < 1$ in that case, (cf.(6.5.6) and (6.6.5)); if $s > 3/2$ P is much larger than Q (if $s=3/2$ 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:

$s < 0$: $\quad S(s,n) =2.(-1)^n + O(n^s)$. $\quad s=0$: $\quad S(0,n)=(-1)^n$.
$0 < s < 3/2$: $\quad S(s,n) = 2(-1)^n.\ \pi^{-1}\ \Gamma(s)(2ns \log(2n))^{-s}\{\sin \pi s+O((\log n)^{-1})\}$.
$s=1$: $\quad S(1,n)=0$
$s \geq 3/2$: $\quad S(s\ ,n) = 2^{2-s}\{2 \cos(\pi/2s)\}^{2ns+s-1}(\pi n)^{\frac{1}{2}(1-s)}s^{-\frac{1}{2}}\{1+O(n^{-1})\}$.

In the cases $s < 0$, and $s > 3/2$ the O-term can be replaced by an asymptotic series in terms of powers of n^{-1}; the same thing can be done if $0 < s < 3/2$, but then in terms of powers of $(\log n)^{-1}$. If $s=3/2$ the asymptotic series is more complicated, as both P and Q give their contributions:

$$S(3/2,n) \sim 2 3^{-\frac{1}{2}} . \pi^{-\frac{1}{4}} \{n^{-1/4} + c_1 n^{-5/4} + c_2 n^{-3/2} (\log n)^{-3/2} + c_3 n^{-3/2} (\log n)^{-5/2} + \dots\};$$

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 ascertain the asymptotical behaviour of the integral for all complex values of s, when $|s| \to \infty$. 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

$$G(s) = \int_0^\infty e^{-P(u)} u^{s-1} \, du, \tag{6.8.1}$$

if 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 Re $s > 0$, because of the functional equation $\Gamma(z) = \Gamma(1-z) = \pi/\sin(\pi z)$. Moreover, it can be done by other methods as well, e.g. by application of the Euler-Maclaurin technique to the infinite product for $\Gamma(z)$.

As in the case of the Γ-integral, it is easily seen that the integral (6.8.1) converges only if 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 α, 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. If k is any integer ≥ 0, we have

$$e^{-P(u)} = a_0 + a_1 u + a_2 u^2 + \ldots + a_k u^k + R(u),$$

$$R(u) = O(u^{k+1}) \qquad (|u| < 1),$$

and therefore

$$G(s) = \int_1^{\infty} e^{-P(u)}\, u^{s-1}\, du + \int_0^1 R(u) u^{s-1} du + \sum_{j=0}^{k} a_j (s+j)^{-1}$$
$$(\mathrm{Re}\ s > 0).$$

The first integral is analytic for all complex values of s, and the second one is analytic in the half-plane $\mathrm{Re}\ s > -k-1$. 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

$$s\ G(s) = NG(s+N) + N-1\ \alpha_{N-1}\ G(s+N-1) + \ldots + \alpha_1\ G(s+1),$$

which is, if $\mathrm{Re}\ s > 0$, 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

$$(6.8.2) \qquad G(s) = \int_{-\infty}^{\infty} \exp\left\{-P(e^z) + sz\right\} dz \qquad (\mathrm{Re}\ s > 0).$$

The integral converges if $\mathrm{Re}\ s > 0$, 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 δ, and we define the path C_δ, consisting of two half lines

(i) the half line described by $z = ixe^{i\delta}$, $\infty > x \geq 0$.

(ii) the positive real axis $z=x$, $0 \leq x < \infty$.

It is not difficult to show that the integral along C_δ is equal to $G(s)$ if $\mathrm{Re}\ s > 0$, $\mathrm{Im}\ s \geq 0$, and that the integral is an analytic function of s in the half-plane defined by $-\delta < \arg s < \pi - \delta$. As δ 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_δ can be tackled by saddle point analysis. The saddle points are the solutions of

$$(6.8.3) \qquad e^z\ P'(e^z) = s\ .$$

If $|s|$ is large, the solutions of (6.8.3) are easily localised. We write $s = |s| e^{i\theta}$, with $-\frac{1}{2}\pi < \theta < \frac{3}{2}\pi$. Now in every horizontal strip $S_k (k=0, \underline{+}1, \underline{+}2, \ldots)$, defined by

(6.8.4) 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 on 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. (In order to have a bounded domain it is necessary, of course, 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 holds for $e^z P'(e^z)-s$, by virtue of the Rouché theorem.

Moreover, if $|s|$ is large enough, the roots of (6.8.3) can be expanded into powers of $s^{-1/N}$

(6.8.5) $$\zeta_k = z_k + c_{k1}\, s^{-1/N} + c_{k2}\, s^{-2/N} + \ldots ,$$

where ζ_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), and refer to a similar case in sec.2 (see (2.4.7)).

Our integration path C_δ leads to $+\infty$ through the strip S_0 . We have to move the path such that it leads over the saddle point ζ_0. However, we cannot keep the path entirely inside the strip S_0; as C_δ starts at $ie^{i\delta}.\infty$, it has to cross the strips $S_1, S_2, \ldots$. And, it has to be feared that (in order to avoid values of the integrand greater than its value at ζ_0) the crossings have to be made quite close to the saddle points $\zeta_1, \zeta_2, \zeta_3, \ldots$. Actually this makes our problem awkward to deal with. It could easily be done 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 mucn 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 would 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, in virtue of the fact that the exponents in $P(e^z)=e^{Nz}+\alpha_1 e^{(N-1)z}+\ldots$ are integers, which implies that $P(e^z)$ is periodic mod $2\pi i$. We shall first define new paths $L_k (k=0,\pm1,\pm2,\ldots)$.

L_k consists of three parts: (i) The half-line $z=2k\pi i/N+x$ $(\infty > x \geqslant 0)$, (ii) The segment $z=ix$ $(2k\pi/N \leq x \leq 2(k+1)\pi/N$, (iii) The half-line $z=(2k+2)\pi/N+x$ $(0 \leq x < \infty)$.

Let the function $G_k(s)$ be defined by

(6.8.6) $$G_k(s) = \int_{L_k} \exp\left\{-P(e^z)+sz\right\} dz.$$

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 2π, we have

(6.8.7) $$G_{k+N}(s) = e^{2\pi is} \, G_k(s) \qquad (k=0,\pm1,\pm2,\dots).$$

The function $G(s)$ can be expressed in terms of $G_0,\dots,G_{N-1}$. Assume, for a moment, that Re $s>0$, so that $G(s)$ is represented by (6.8.2). Obviously we have

$$\int_{2\pi i-\infty}^{2\pi i+\infty} = e^{2\pi is}\int_{-\infty}^{\infty},$$

in analogy to (6.8.7). Furthermore, it can be shown that

$$\int_{2\pi i-\infty}^{2\pi i+\infty} - \int_{-\infty}^{\infty} = G_0 + G_1 + \dots + G_{N-1},$$

shifting the paths $L_0,\dots,L_{N-1}$ indefinitely to the left. Therefore, if Re $s>0$,

(6.8.8) $$G(s) = -(1-e^{2\pi is})^{-1}\left\{G_0(s)+\dots+G_{N-1}(s)\right\}.$$

The right-hand side is analytic for all s, except for possible poles at $s=0, \pm1,\pm2,\dots$. 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 by (6.8.8) we get 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. <u>The entire function $G_0(s)$.</u> 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 δ . Then the saddle point ζ_0 of $\exp(-P(e^z)+sz)$ lies inside the path L_0 if $|s|$ is large, for ζ_0 is close to z_0 (see (6.8.5)), and $z_0=N^{-1}\left\{\log(|s|/N)+\theta i\right\}$, 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 ζ_0 coincides with z_0. In that special case we write G_0^* instead of G_0.

By the substitution $z=z_o+w$, the saddle point is shifted to the origin. The path is also shifted; it becomes $(L_o|-z_o)$, by which we denote the path described by $z-z_o$, if z describes the path L_o. It follows from Cauchy's theorem that a horizontal shift of the path $(L_o,-z_o)$ has no influence upon the value of the integral, and therefore we may replace it by $(L_o|-i\theta)$. This path passes through the saddle point $w=0$.

The integrand becomes

$$\exp(-e^{Nz}+sz) = \exp(s(z_o-N^{-1})).\exp\left\{-sN^{-1}(e^{Nw}-Nw-1)\right\},$$

where the splitting has been made so as to make the second factor of the form $\exp(-\frac{1}{2}sN(w^2+\ldots))$, for small values of w. Now $G_o^*(s)$ becomes

$$(6.9.1)\qquad G_o^*(s)=\exp(s(z_o-N^{-1}))\int_{(L_o|-i\theta)}\exp\left\{-sN^{-1}(e^{Nw}-Nw-1)\right\}dw.$$

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 $|\mathrm{Im}\ w|\leq 2\pi/N$. The path $(L_o|-i\theta)$ lies, for all θ 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 there. Instead, we consider the function $\zeta(w)=\left\{2(e^{Nw}-Nw-1)\right\}^{\frac{1}{2}}$, where in a neighbourhood of $w=0$ the sign of the square root has been chosen such that $\zeta(w)=Nw+\ldots$. By analysing what happens to ζ if w runs through the boundary of a long horizontal rectangle $|\mathrm{Im}\ w|\leq 2\pi/N$, $|\mathrm{Re}\ w|\leq M$, and making $M\longrightarrow\infty$, we find that the strip is mapped one-to-one onto a set S which is obtained from the complete ζ-plane by deleting two hyperbolic arcs. These arcs can be described by

$$(\mathrm{Re}\ \zeta)\ .\ (\mathrm{Im}\ \zeta) = \pm 2\pi,\ \mathrm{Re}\ \zeta\leq -|\mathrm{Im}\ \zeta|\ .$$

We want to have ζ as a new integration variable, and therefore we have to investigate $dw/d\zeta$. Needless to say, this is an analytic function of ζ throughout the set S. As $\zeta^2=2(e^{Nw}-Nw-1)$, we have $\zeta d\zeta = N(e^{Nw}-1)dw$.

Now $e^{Nw}-1$ is, as far as our strip is concerned, close to 0 only if w is close to either 0, or $2\pi i/N$, or $-2\pi i/N$. Therefore we have

$$(6.9.2)\qquad dw/d\zeta = O(\zeta)\qquad (|\arg\zeta|\leq\frac{3}{4}\ ,\ |\zeta|>3).$$

It is not difficult to show that $dw/d\zeta$ is even $O(\zeta^{-1})$ in this region, yet $O(\zeta)$ is sufficient for our purpose.

Our integral becomes

$$(6.9.3)\qquad G_o^*(s) = \exp(s(z_o-N^{-1}))\int_C \exp(-\tfrac{1}{2}sN^{-1}\zeta^2)\ .\ \frac{dw}{d\zeta}\ .\ d\zeta\ ,$$

and the next problem is what C is. Analysing the image of $(L_o|-i\theta)$

under the conformal mapping we find that C is a curve starting at $e^{-\frac{1}{2}i\theta}.\infty$ and tending to $e^{\pi i-\frac{1}{2}i\theta}.\infty$, avoiding the hyperbolic arcs. We have, however, considerable freedom in modifying this path. Along the line through the origin, with arguments $-\frac{1}{2}\theta$ and $\pi-\frac{1}{2}\theta$ (a straight line through the origin has <u>two</u> arguments!) the expression $\frac{1}{2}sN^{-1}\zeta^2$ is positive. It is easily seen that this line may be replaced by other lines, 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 $\frac{1}{2}sN^{-1}\zeta^2$ is still larger than a constant positive multiple of $|\zeta|^2$, 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 lines are, at the same time, satisfactory from the point of view of the saddle-point method: on these lines the absolute value of $\exp(-\frac{1}{2}sN^{-1}\zeta^2)$ is maximal at the saddle point $\zeta=0$. 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 $-e^{i\eta}.\infty$ to $+e^{i\eta}.\infty$, where $\pi/4<\eta<3\pi/4$. This is no objection: as $\eta-(\pi-\frac{1}{2}\theta)$ is allowed to lie between $-\pi/4$ and $\pi/4$, we can find a satisfactory value for η to any θ in the interval $\delta<\theta<2\pi-\delta$. Actually we could take $\eta=(3\pi-\theta)/4$.

It is, however, preferable to have a fixed integration path, not depending on η. This can only be achieved under restriction of θ, and therefore we shall consider two different values of η:

(i) $\eta=(3\pi-\delta)/4$. This can be used as long as $\frac{1}{2}\delta<\theta<\pi+\frac{1}{2}\delta$.

(ii) $\eta=(\pi+\delta)/4$. 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 $\delta<\theta<\pi+\delta/4$. So we replace, in (6.9.3), the path C by the straight line through the origin, from $-e^{i(3\pi-\delta)/4}.\infty$ to $+e^{i(3\pi-\delta)/4}$. This path will be denoted by D. Along D we have

$$|\exp(-\tfrac{1}{2}sN^{-1}\zeta^2)| \leq \exp\left\{-\tfrac{1}{2}|s|.N^{-1}.|\zeta|^2.\sin(\delta/4)\right\}.$$

Now the stage has been set for application of the method of sec.4.4. The value of $dw/d\zeta$ at $\zeta=0$ equals N^{-1} (for, $\zeta=Nw+\ldots$). Furthermore, $dw/d\zeta$ is an analytic function of ζ along the line D, and $dw/d\zeta=O(\zeta)$ if $|\zeta|>3$.

The integral can be compared with the formula

$$\int_D \exp(-\tfrac{1}{2}sN^{-1}\zeta^2)d\zeta=-(2\pi N)^{\frac{1}{2}}s^{-\frac{1}{2}},$$

where $s^{-\frac{1}{2}}$ is to be interpreted as $|s|^{-\frac{1}{2}}e^{-\frac{1}{2}i\theta}$. So finally we obtain

$$(6.9.4)\quad G_0^*(s) = -\exp(s(z_0-N^{-1}))(2\pi N)^{\frac{1}{2}}s^{-\frac{1}{2}}N^{-1}\left\{1+O(s^{-1})\right\},$$

and the O-term can be replaced by an asymptotic series $c_1s^{-1}+c_2s^{-2}+\ldots$.

It should be noted that (6.9.4) holds uniformly in the region $|s| > 1$, $\delta < \arg s < \pi + \delta/4$.

Formula (6.9.4) is not new, of course. We have (cf.(6.8.8); notice that $G_k^*(s) = e^{2\pi iks/N}\, G_o^*(s))$, $G_o^*(s) = -N^{-1}(1-e^{2\pi is/N})\,\Gamma(s/N)$, and therefore, if we replace s/N by w, (6.9.4) reduces to

$$\Gamma(w) = (2\pi)^{\frac{1}{2}}(1-e^{2\pi iw})^{-1}\, e^{(w-\frac{1}{2})\log w - w}\left\{1+0(w^{-1})\right\} \;(\,\delta < \arg w < \pi+\delta/4),$$

and this is an easy consequence of the Stirling formula ($\log w$ is given the value with $\delta < \operatorname{Im} \log w < \pi + \delta/4$). It was not, however, our purpose to deduce well-known results on the gamma function for which easier methods exist, but rather to develop a technique for $G_o^*(s)$ which can be modified to a technique for $G_o(s)$.

As a preparation to the problem of the asymptotical behaviour of $G_o(z)$ we shall first investigate the following integral

$$(6.9.5) \qquad f(t,\omega) = \int_{-g(x)}^{g(x)} \exp(-tx^2 - t\, Q(x,\omega))\, dx.$$

Here $g(x)$ is a fixed positive number, $Q(x,\omega)$ is a double power series in x and ω:

$$Q(x,\omega) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} c_{nm}\, x^n \omega^m ,$$

absolutely convergent if $|x| < 2p$ and $|\omega| < b$, where b is a fixed positive number. The function $g(x)$ is assumed to be analytic in the circle $|x| < 2p$, and it is assumed that $g(o) \neq 0$ We want to determine the behaviour of $f(t,\omega)$ when $|\omega|$ is small and $|t|$ is large, where t is restricted to a sector $|\arg t| < \frac{1}{2}\pi - \delta$.

Searching for saddle points, we investigate the equation

$$(6.9.6) \qquad -2tx - t\, Q'(x,\omega) = 0,$$

the dash indicating differentiation with respect to x. By the Rouché theorem (see sec.2.4), this equation has exactly one solution x_o inside the circle $|x| < p$ 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_o can be written as a sum of a power series:

$$x_o = \sum_{k=1}^{\infty} d_k \omega^k,$$

convergent if $|\omega|$ is sufficiently small.

The straight line from $-p$ to $+p$ is a satisfactory path for applying the saddle point method to the simplified integral $\int \exp(-tx^2)dx$. This path makes an angle $< \frac{1}{4} - \delta$ with the axis of the saddle point in the origin. In the modified integral (6.9.5) the saddle point x_o 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 $< \frac{1}{4}\pi - \frac{1}{2}\delta$ with the new axis, and therefore it can be used, anyway in a neighbourhood of the saddle point. These remarks are, of course, only given as an orientation; the work remains to be done.

We replace (6.9.5) by

$$(6.9.7) \qquad \int_{-p}^{-p+x_0} + \int_{-p+x_0}^{p+x_0} + \int_{p+x_0}^{p} \quad ,$$

where the integration is 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

$$(6.9.8) \quad f(t,\omega)=\exp(-tx_0^2-tQ(x_0,\omega)) \cdot \int_{-p}^{p} g(x_0+y) \exp(-ty^2-t\,Q_1(y,\omega))dy,$$

where $Q_1(y,\omega) = Q(x_0+y,\omega)-Q(x_0,\omega)-y\,Q'(x_0,\omega)$ 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 $|y| < 3p/2$ and $|\omega|$ sufficiently small (if ω is small enough, this condition on y implies that $|y+x_0| < 2p$).

We can now proceed along various methods. For example, we can expand $\exp(-t\,Q_1(y,\omega))$ in terms of powers of y. Another method is to apply conformal transformation again, putting $y^2+Q(y,\omega)=z^2$. Then z can be solved as a double power series in y and ω, and we get an integral

$$\int Q_3(z,\omega) \exp(-tz^2)dz .$$

Omitting the details, which are all implied in the usual saddle point routine, we state the result:

$$(6.9.9) \quad f(t,\omega) = (\pi)^{\frac{1}{2}} t^{-\frac{1}{2}} \left\{1+\tfrac{1}{2}Q''(x_0,\omega)\right\}^{-\frac{1}{2}} \left\{g(x_0)+O(t^{-1})\right\} \exp\left\{-tx_0^2-tQ(x_0,\omega)\right\}$$

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 ω, 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 ω.

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 ω is small enough, there is just one saddle point near $x=0$, and the contribution of this saddle point gives an asymptotic series for $f(t,\omega)$.

In the integral (6.8.6) for $G_0(s)$, we carry out the same substitutions as we did in the case of $G_o^*(s)$. That is, $z=z_o+w$, where $Nz_o=\log(|s|/N)+i\theta$, $\theta=\arg s$, $2(e^{Nw}-Nw-1)=\zeta^2$, and we discuss the integral in the ζ-plane:

$$(6.9.10)\qquad G_o(s)=\exp(s(z_o-N^{-1}))\int_D \exp\left\{-\tfrac{1}{2}sN^{-1}\zeta^2+(e^{Nz}-P(e^z))\right\}\frac{dw}{d}\,d\zeta\,.$$

D is the straight line through the origin, from $-e^{i(3\pi-\delta)/4}.\infty$ to $+e^{i(3\pi-\delta/4}.\infty$. We shall again restrict θ by $\delta<\theta<\pi+\delta/4$.

We first investigate the term $e^{Nz}-P(e^z)$, which embodies the deviation of G_o from G_o^*. It is equal to

$$(6.9.11)\qquad e^{Nz_o+Nw}\left\{-\alpha_{N-1}e^{-z_o-w}-\alpha_{N-2}e^{-2z_o-2w}-\dots\right\}.$$

It is not difficult to show from the properties of the conformal mapping that Re w tends to $+\infty$ if $|\zeta|$ tends to infinity, provided that ζ runs along D. It easily follows, for any positive number p, that $e^w=O(\zeta^{2/N})$ (ζ on D, $|\zeta|>p$). Therefore

$$|e^{Nz}-P(e^z)|\le c_1|s\,\zeta^2|^{(N-1)/N}\qquad(\zeta \text{ on D}, |\zeta|>p, |s|>c_2),$$

where the c's are suitable positive numbers independent of s and ζ. 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 $|s|\to\infty$.

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 $-(3\pi-\delta)/4$, by the substitution $\zeta=xe^{i(3\pi-\delta)/4}$, and to put $s=e^{-i(3\pi-\delta)/2}t$, so that t is restricted by the condition $-\pi/2+\delta/2<\arg t<\pi/2-\delta/4$, and $s\zeta^2$ becomes tx^2. Furthermore, we have, in some circle $|\zeta|<2p$,

$$e^{Nz}-P(e^z)=N^{-1}s\,Q(\zeta,\omega),$$

where (cf.(6.9.11)) $\omega=s^{-1/N}$, and Q is a double power series with $Q(\zeta,0)=0$. (For $w=N^{-1}\zeta+\dots$ and $e^{-w},\dots,e^{-(N-1)w}$ are convergent power series in powers of ζ for small values of ζ, and $e^{-z_o}=N^{1/N}\omega$). It follows that for $G_o(s)$ we have an asymptotic series, which is entirely given by the contribution of the saddle point. The series equals a certain function multiplied by an asymptotic series $c_o(\omega)+c_1(\omega)s^{-1}+\dots$, and the $c(\omega)$'s are convergent series in powers of $\omega=s^{-1/N}$. Therefore, the series $c_o(\omega)+\dots$ can be rewritten as an asymptotic series of the form $c_o+c_1s^{-1/N}+c_2s^{-2/N}+\dots$.

We do not evaluate explicitly the contribution of the saddle point in the integral in the ζ-plane, as it is easier to do it in the original z-plane (see (6.8.6), with the saddle point ζ_o, 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 ζ and ζ to x. So our final result is (cf.sec.4.4) an asymptotic series

$$(6.9.12) \quad G_o(s) \sim - \exp\{-P(e^{\zeta_o})+s\zeta_o \,.\, (-2\pi/\psi''(\zeta_o))^{\frac{1}{2}} \sum_0^\infty c_k\, s^{-k/N},$$

where $c_o=1$, 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

$$-\psi''(\zeta_o) = N^2e^{Nz}+(N-1)^2\alpha_{N-1}e^{N-1}z+\ldots = Ns\left\{1+a_1s^{-1/N}+a_2s^{-2/N}+\ldots\right\},$$

so that (6.9.12) can be slightly simplified, introducing new coefficients d_k (with $d_o=1$):

$$(6.9.13) \quad G_o(s) \sim - \exp\left\{-P(e^{\zeta_o}) + s\,\zeta_o\right\} .\, (2\pi/Ns)^{\frac{1}{2}} \sum_0^\infty d_k s^{-k/N},$$

and the formula has been proved if $\delta < \arg s < \pi + \delta/4$, $|s| \to \infty$. The argument of $s^{\frac{1}{2}}$ is understood to lie between $\frac{1}{2}\delta$ and $\frac{1}{2}\pi + \delta/8$. 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 asymptotical behaviour of $G_o, G_1, \ldots, G_{N-1}$. The problem about G_k $(k=1,\ldots,N-1)$ is easily reduced to the problem for G_o. It follows from (6.8.6) that

$$(6.10.1) \quad G_k(s) = \exp\,(2\pi iks/N) \int_{L_o} \exp\left\{-P(e^{2\pi ik/N}\, e^z)+sz\right\} dz,$$

and this integral equals the function $G_o(s)$, constructed for the polynomial $P(e^{2\pi ik/N}u)$ instead of P(u). The leading term of the new polynomial is again u^N.

Replacing P(u) by $P(e^{2\pi ik/N})$ 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 (6.10.1) there occurs, a term $s\zeta_o$ in the exponent, the influence upon the asymptotical behaviour can be considerable.

It is quite easy to state a simpler but weaker result;

$$(6.10.2) \quad G_k(s) = e^{2\pi iks/N}\,(2\pi/Ns)^{\frac{1}{2}} \exp\left\{-\frac{s}{N} + \frac{s}{N}\log\frac{s}{N} + O(s^{(N-1)/N})\right\},$$

where $\delta < \arg s < \pi + \delta/4$. From this we infer that G_k $(k>0)$ is negligible compared to G_o as soon as the factor $\exp(2\pi iks/N)$ beats the $\exp\left\{O(s^{(N-1)/N})\right\}$. This is certainly the case if s is restricted to $\delta < \arg s < \pi - \delta$. And, under that assumption, the factor $(1-e^{2\pi is})^{-1}$ in (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 have, from (6.9.13)

$$(6.10.3) \qquad G(s) \sim (2\pi/Ns)^{\frac{1}{2}} \exp(-P(e^{\zeta_0})+s\zeta_0) \sum_0^{\infty} d_k\, s^{-k/N} \qquad (d_0=1)$$

in the sector $\delta < \arg s < \pi - \delta$.

For the sector $-\pi+\delta < \arg s < \delta$ we can obtain the same result, provided that ζ_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 that 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 $|\mathrm{Im}\, s| < 1$, $\mathrm{Re}\, s < 1$, then in the remaining region we have

$$|G(s)/\Gamma(s)| < C_1 \exp(C_2|s|^{(N-1)/N}).$$