

THE THEORY OF METRIC SPACES APPLIED TO INFINITELY DIFFERENTIABLE FUNCTIONS

THØGER BANG

1. An infinitely differentiable function $f(x)$ of a real variable has at each point a an element (in the sense of Weierstrass), viz. the sequence $\{f^{(n)}(a)\}$ consisting of the value of the function and the values of its derivatives at that point. This element varies with a . Its variation depends on the derivatives of the function, and if these are restricted in some way, the variation of the element will be restricted. Thus the problem of quasi-analytic functions, as posed by Hadamard, consists in finding conditions for the bounds of the derivatives in order that the function be uniquely determined when one of its elements is given. This is the case when in Taylor's formula the remainder term corresponding to a certain derivative $f^{(n)}(x)$ tends to zero as $n \rightarrow \infty$, so that the function is analytic. The difficulties in the general case arise from the fact that a remainder term, in order to be small enough, cannot be expressed in terms of one derivative only, but must contain several derivatives.

In order to study the variation of the elements, we organize the set of elements as a metric space. Restrictions on the derivatives can be expressed in a simple way in this metric. We shall apply this to prove results on infinitely differentiable functions. An outline of the present article was given in [3].

In § 2 we define a norm in the vectorspace of numerical sequences and state its main properties. The norm makes the space a metric space in the sense of Fréchet; it would be desirable to make it a Banach space, but it seems difficult to perform this in a way suitable for our purpose. In § 3 we apply this norm to the elements. We suppose that the derivatives are bounded uniformly in x , $|f^{(n)}(x)| \leq m_n$, where $\{m_n\}$ is a sequence of positive numbers which also enter in the definition of the norm, and then we prove an inequality that is fundamental for our applications. In § 4 we prove the necessity of a condition on the sequence $\{m_n\}$ in order that $f(x)$ might have an element in common with the zero-function with-

out being identically zero. Thus, we prove the more difficult half of the Denjoy-Carleman theorem, which gives the answer to the above-mentioned problem of quasi-analytic functions. In § 5 we suppose this condition to be satisfied and investigate how fast $f(x)$ can deviate from the zero-function.

In the rest of the article we suppose that the above condition is not satisfied so that the functions are quasi-analytic. We also suppose that the sequence $\{m_n\}$ is logarithmically convex; this is not essential for the method, but it simplifies the notations (in § 6 it is mentioned how a theorem would read without this assumption). In § 6 we give a procedure which generates the (uniquely determined) function from one of its elements. In particular, for an analytic function the procedure may produce the values outside the circle of convergence, but it does not consist in a summation of the Taylor series. In § 7 we consider a sequence of zeros of the successive derivatives, and the result is applied in § 8 to get theorems on the zeros of quasi-analytic functions and on the zeros of their successive derivatives. These theorems contain results on analytic functions, which are thus proved by real variable methods.

2. The class of sequences $U = \{u_n\} = u_0, u_1, \dots, u_n, \dots$ (of real or complex numbers) forms a vectorspace in an obvious way. In this space we shall introduce a metric.

Let $\{\varepsilon_n\}$ be a monotone decreasing sequence tending to 0,

$$\varepsilon_0 > \varepsilon_1 > \varepsilon_2 > \dots \rightarrow 0,$$

and let P be an infinite set of non-negative integers.

We define the norm $\|U\|$ in the following way:

$$(1) \quad \|U\| = \inf_{p \in P} [\max\{\varepsilon_p, \max_{0 \leq n \leq p} |u_n|\}].$$

It is obvious that $\|U\| \geq 0$. Moreover, $\|U\|$ is smaller than or equal to ε_p for a $p \in P$ if, and only if, $|u_n| \leq \varepsilon_p$ for all $n \leq p$. Thus, roughly speaking, a small norm means that a large section of the sequence consists of numbers near 0. The only sequence whose norm is 0 is the sequence consisting of zeros.

Moreover, if $\varepsilon_{p_1} \leq \|U\| \leq \varepsilon_{p_2}$, where p_1 and p_2 lie in P , it is possible to find a $p \in P$, such that

$$(2) \quad \|U\| = \max\{\varepsilon_p, \max_{0 \leq n \leq p} |u_n|\} \quad \text{where } p_2 \leq p \leq p_1.$$

Indeed, a p smaller than p_2 cannot give the "inf" in (1), as $\|U\| \leq \varepsilon_{p_2}$; and from $\|U\| \geq \varepsilon_{p_1}$ it easily follows that the quantity in square brackets

in (1) cannot decrease for $p > p_1$ and $p \in P$, so that there exists a $p \leq p_1$ for which (2) is valid.

It is obvious that $\| -U \| = \| U \|$, and we have also $\| tU \| \leq t \| U \|$ for $t > 1$ (but the norm is not linear). Moreover, the triangle inequality

$$(3) \quad \| U + V \| \leq \| U \| + \| V \|$$

is satisfied. This is immediate if $\| U \|$ or $\| V \|$ is 0. In the general case, (2) shows that there exist numbers p and q in P such that

$$\| U \| = \max_{0 \leq n \leq p} \{ \varepsilon_p, \max |u_n| \}$$

and

$$\| V \| = \max_{0 \leq n \leq q} \{ \varepsilon_q, \max |u_n| \} .$$

Supposing for example $p \leq q$, we infer that

$$\max_{0 \leq n \leq p} \{ \varepsilon_p, \max |u_n + v_n| \} \leq \| U \| + \| V \| ,$$

which shows that (3) is valid.

If we define the distance between U and V as $\| U - V \|$, we have organized the vectorspace of sequences U as a metric space. The topology induced by this metric is that in which convergence means convergence in each coordinate without any uniformity. This implies that if the elements u_n are continuous functions $u_n(x)$ of a parameter x , the norm $\| U \|$ will also be a continuous function of x .

3. Let $\{ m_n \}$ be a sequence of positive numbers, $n = 0, 1, 2, \dots$. We suppose that the sequence increases so rapidly that $m_n^{1/n} \rightarrow \infty$; then there exists (see for instance [8], p. 73) a unique largest minorant sequence $\{ m_n^c \}$ which is logarithmically convex, that is a sequence satisfying

$$(4) \quad m_n^c / m_{n-1}^c \leq m_{n+1}^c / m_n^c .$$

The equation $m_n^c = m_n$ is satisfied for infinitely many n , in particular for $n = 0$. In the following we shall use the set of these integers as the set P in the definition (1). We remark that when n is not contained in P the sign of equality prevails in (4).

Let $f(x)$ be a function defined in an interval I (finite or infinite) of the real axis and possessing derivatives of every order. We now put $u_n(x) = f^{(n)}(x) / (e^n m_n^c)$ and $\varepsilon_p = e^{-p}$. The norm defined in this way by (1) shall be denoted by $\| f \|_x$,

$$(5) \quad \| f \|_x = \inf_{p \in P} \left[\max \left\{ e^{-p}, \max_{0 \leq n \leq p} \frac{|f^{(n)}(x)|}{e^n m_n^c} \right\} \right] .$$

It is a continuous function of x since each $u_n(x)$ is continuous.

For this norm we shall prove a theorem that will be fundamental throughout the following. First we need a lemma.

LEMMA. *If $\|f\|_x \geq e^{-q}$ for some $q \in P$, and $|f^{(p)}(\xi)| \leq m_p$ for all ξ , then*

$$\|f\|_{x+h} \leq \|f\|_x \cdot \exp\left(e|h| \frac{m_p^c}{m_{p-1}^c}\right),$$

$p \in P$ being an integer for which the infimum on the right-hand side of equation (5) is attained.

The existence of an integer p for which the infimum in (5) is attained and for which even $p \leq q$ follows from (2) with q instead of p_1 .

PROOF OF THE LEMMA. We have the two inequalities

$$(6) \quad e^{-p} \leq \|f\|_x$$

and

$$(7) \quad |f^{(n)}(x)| \leq e^n m_n^c \|f\|_x \quad \text{for } 0 \leq n \leq p,$$

and with the same p we get from (5)

$$\|f\|_{x+h} \leq \max \left\{ e^{-p}, \max_{0 \leq n \leq p} \frac{|f^n(x+h)|}{e^n m_n^c} \right\}.$$

If here e^{-p} is the larger of the two terms between the braces on the right-hand side, it follows from (6) that the statement of the lemma holds true. Otherwise we shall estimate $|f^{(n)}(x+h)|/(e^n m_n^c)$. Using Taylor's formula with remainder term we find that the term in question is majorized by

$$\sum_{j=0}^{p-n-1} \frac{|f^{n+j}(x)| \cdot |h|^j}{j! e^n m_n^c} + \frac{|f^{(p)}(\xi)| \cdot |h|^{p-n}}{(p-n)! e^n m_n^c}.$$

In the first sum we use (7). In the last term we use that $|f^{(p)}(\xi)| \leq m_p = m_p^c$ (as $p \in P$) and multiply the term by $e^p \|f\|_x$, which is greater than or equal to 1 according to (6). The last term may then be placed under the summation sign, and we get the upper bound

$$\sum_{j=0}^{p-n} \frac{\|f\|_x e^{n+j} m_{n+j}^c |h|^j}{j! e^n m_n^c} = \|f\|_x \cdot \sum_{j=0}^{p-n} \frac{e^j |h|^j}{j!} \cdot \frac{m_{n+j}^c}{m_n^c}.$$

In view of (4) this is less than or equal to

$$\|f\|_x \cdot \sum_{j=0}^{p-n} \frac{e^j |h|^j}{j!} \left(\frac{m_p^c}{m_{p-1}^c} \right)^j.$$

Here the series is a partial sum of the expansion of $\exp(e|h|m_p^c/m_{p-1}^c)$. Hence the statement of the lemma is verified in all cases.

THEOREM I. *If $f(x)$ possesses derivatives of every order, and if these are bounded for all n ,*

$$(8) \quad |f^{(n)}(\xi)| \leq m_n \quad \text{uniformly in } \xi,$$

then from the inequality $\|f\|_x \geq e^{-q}$ it follows that

$$(9) \quad \|f\|_{x+h} \leq \|f\|_x \cdot \exp\left(e|h| \frac{m_q^c}{m_{q-1}^c}\right).$$

If $q \in P$, Theorem I is an immediate consequence of the lemma. In fact, we remarked above that we can choose the integer p smaller than or equal to q . Hence it follows from (4) that the right-hand side of the inequality in the lemma is smaller than or equal to the right-hand side of (9). If q is not contained in P , we may replace q by $q+1$ in the theorem, since the assumption $\|f\|_x \geq e^{-q-1}$ is satisfied a fortiori, while the right-hand side of (9) is unchanged as the sign of equality is valid in (4) ($n = q$). We repeat this process until q is replaced by a number contained in P , and the theorem is proved.

The statement of the theorem may be expressed in the following manner. If $f(x)$ satisfies (8), then $L(x) = \log \|f\|_x$ satisfies a Lipschitz condition at every point where it is defined, and at a point x where $L(x) > -q$ and $L'(x)$ exists we have $|L'(x)| \leq em_q^c/m_{q-1}^c$.

In Theorem I we have considered only functions $f(x)$ for which the condition (8) is satisfied, the sequence $\{m_n\}$ being subjected to the restriction $m_n^{1/n} \rightarrow \infty$ imposed at the beginning of this section. If, however, $\liminf m_n^{1/n} < \infty$, i.e., $m_n < K^n$ for a suitable K and arbitrarily large n , the corresponding remainder terms $f^{(n)}(\xi)h^n/n!$ in Taylor's formula tend to 0 as $n \rightarrow \infty$. Hence $f(x)$ is the limit of the corresponding partial sums, which implies that $f(x)$ is quasi-analytic. In view of this, the above restriction concerning $\{m_n\}$ means no actual loss of generality, and we shall therefore maintain it in the sequel.

4. The theorem of Denjoy and Carleman on quasi-analytic functions states (see e.g. [9], p. 78) that an infinitely differentiable function $f(x)$ which satisfies (8) is uniquely determined by the sequence $\{f^{(n)}(a)\}$, $n = 0, 1, 2, \dots$, if and only if

$$(10) \quad \sum_{n=0}^{\infty} \frac{m_n^c}{m_{n+1}^c} = \infty.$$

Mandelbrojt and Bang found independently that the proof of the necessity of the condition (i.e., the construction of a counterexample when the series in (10) converges) is both simple and elementary ([9], pp. 80-84, [2], pp. 55-56). Here we shall prove the sufficiency of the condition.

If the sequences $\{f_1^{(n)}(a)\}$ and $\{f_2^{(n)}(a)\}$ are identical, and $f_1(x)$ and $f_2(x)$ both satisfy (8) in an interval I , then $f = (f_1 - f_2)/2$ also satisfies (8) in I , and $\|f\|_a = 0$. Suppose now that f_1 and f_2 are not identical. Then $\|f\|_x \neq 0$, and there exists a number x_1 satisfying $\|f\|_{x_1} = e^{-p_1}$ where $p_1 \in P$. Between x_1 and a there exists a monotone sequence x_1, x_2, x_3, \dots such that

$$(11) \quad \|f\|_{x_j} = e^{-p_j}$$

where $p_1 < p_2 < p_3 < \dots$ are precisely all numbers $\geq p_1$ in P .

On setting $x = x_j$ and $x+h = x_{j-1}$ in Theorem I, we have by (9)

$$e^{-p_{j-1}} \leq e^{-p_j} \cdot \exp\left(e|x_j - x_{j-1}| \frac{m_{p_j}^c}{m_{p_{j-1}}^c}\right),$$

or

$$e|x_j - x_{j-1}| \geq (p_j - p_{j-1}) \frac{m_{p_{j-1}}^c}{m_{p_j}^c}.$$

Here the right-hand side *can* be written

$$(12) \quad \sum_{n=p_{j-1}}^{p_j-1} \frac{m_n^c}{m_{n+1}^c}$$

since, as remarked above, the terms of this sum are all equal.

As x_1, x_2, \dots form a monotone sequence in the interval between x_1 and a , we have

$$\sum |x_j - x_{j-1}| \leq |a - x_1|.$$

By adding the expressions (12) for $j = 2, 3, \dots$, we thus get

$$(13) \quad \sum_{n=p_1}^{\infty} \frac{m_n^c}{m_{n+1}^c} \leq e|a - x_1|,$$

which implies the convergence of the series in (10). This completes the proof of the deeper half of the Denjoy-Carleman theorem.

5. Let us consider a non-quasi-analytic function $f(x)$, i.e., a function for which the series in (10) converges, m_n denoting the upper bound of $|f^{(n)}(x)|$. If $f(a) = f'(a) = f''(a) = \dots = 0$, formula (13) restricts the rate of increase of $f(x)$. In fact, if $e|a - x|$ is less than the left-hand side of inequality (13), we have $\|f\|_x < e^{-p_1}$, which by (5) implies $|f(x)| < e^{-p_1} m_0^c = e^{-p_1} m_0$, and hence we infer that

$$(14) \quad |x - a| < \frac{1}{e} \sum_{n=p_1}^{\infty} \frac{m_n^c}{m_{n+1}^c} \quad \text{implies} \quad |f(x)| < e^{-p_1} m_0.$$

If the infinite series converges slowly, the estimate (14) is substantially

better than the result obtained by identifying $f(x)$ with the remainder term in Taylor's formula from the point a , viz.

$$(15) \quad |f(x)| \leq m_n \frac{|x-a|^n}{n!},$$

where, of course, n should be chosen such that the right-hand side is minimal. (The inequality (15) has been applied for instance by Gontcharoff [6], p. 49, and by Bang [2], p. 57.)

As a typical example let

$$m_n = n!(\log n)^{n(1+\beta)}$$

with $\beta > 0$. The series in (10) is then convergent (when $\beta = 0$ it diverges). It follows from (14) that, if $x \rightarrow a$, $f(x)$ tends to 0 so rapidly that

$$\log |\log |f(x)|| \geq A \cdot |x-a|^{-1/\beta},$$

while (15) only yields an inequality of the type

$$\log |\log |f(x)|| \geq B \cdot |x-a|^{-1/(1+\beta)},$$

A and B being positive constants. In the last inequality the essential difference between the cases $\beta > 0$ and $\beta = 0$ has disappeared.

6. Consider now again a quasi-analytic function, i.e., a function for which the series in (10) diverges, m_n denoting the upper bound of $|f^{(n)}(x)|$. The theorem of Denjoy and Carleman states that $f(x)$ is uniquely determined by an element $\{f^{(n)}(a)\}$. It is therefore natural to ask how to calculate the function when the element is given. In general the corresponding Taylor series diverges. Carleman proved, however, that a certain summation method (in the sense of Toeplitz), depending on the sequence $\{m_n\}$, produces $f(x)$ when applied to the Taylor series ([4], chapt. VII; a recent proof is due to L. Carleson [5]). This result can also be deduced from a general theorem on linear functionals in the vectorspace of numerical sequences ([1], p. 66).

Here we shall give a generating method for $f(x)$, a method that in a rather direct way yields the function and all its derivatives when the element $\{f^{(n)}(a)\}$ is given. The idea is, roughly speaking, that a large partial sum of the Taylor series from $x = a$ gives a good approximation to $f(x)$ in a small neighbourhood of a . In the expansion of this partial sum from a neighbouring point x_{N-1} we cancel the term of highest degree whereby a good approximation is obtained in a neighbourhood of x_{N-1} . Next we expand this new polynomial from a certain point x_{N-2} of this second neighbourhood, etc. Eventually every point x of the interval of definition is reached.

To simplify the notations we suppose in the sequel that the sequence $\{m_n\}$ itself is logarithmically convex. Thus $m_n = m_n^c$ for all n , and P consists of all non-negative integers.

The precise formulation of the above method of approximation is as follows.

THEOREM II. *For any given (positive) integer N we introduce the following points:*

$$(16) \quad \begin{aligned} x_N &= a \\ x_{N-1} &= x_N + m_{N-1}/(em_N) \\ x_{N-2} &= x_{N-1} + m_{N-2}/(em_{N-1}) \quad \text{etc.} \end{aligned}$$

On the interval (x_N, x_{N-1}) we take the polynomial $S_N(x)$ of degree $< N$ for which

$$(17) \quad S_N^{(n)}(x_N) = f^{(n)}(x_N), \quad n = 0, 1, \dots, N-1;$$

on the interval (x_{N-1}, x_{N-2}) we take the polynomial $S_{N-1}(x)$ of degree $< N-1$ for which

$$S_{N-1}^{(n)}(x_{N-1}) = S_N^{(n)}(x_{N-1}), \quad n = 0, 1, \dots, N-2;$$

and so on. The function $S(x)$ composed of these polynomials converges then to $f(x)$ as $N \rightarrow \infty$, and each of its derivatives converges to the corresponding derivative of $f(x)$, the convergence being uniform with respect to x in every finite interval to the right of $x = a$ and contained in the interval of definition.

If we do not suppose that $\{m_n\}$ is logarithmically convex, the theorem has to be modified as follows. The numbers m_n should be replaced everywhere by m_n^c . Further, we take $N \in P$ and use $S_N(x)$ on the interval (x_N, x_{N_1}) , where N_1 is the largest number in P smaller than N . Similarly, on an interval (x_{N_1}, x_{N_2}) we use the polynomial $S_{N_1}(x)$ of degree $< N_1$, for which the element at x_{N_1} coincides as far as possible with the element of $S_N(x)$ at x_{N_1} . In this way we continue.

Let us remark that if $\|g\|_x \leq e^{-q}$, $g(\xi)$ being a function for which $|g^{(q)}(\xi)| \leq m_q$ for all ξ (but without any assumption that $g^{(n)}(\xi)$ be bounded for $n \neq q$), we have

$$(18) \quad \|g\|_{x+h} \leq e^{-q+1} \text{ for } 0 < h \leq m_{q-1}/(em_q).$$

If $\|g\|_x = e^{-q}$, this follows immediately from the lemma in § 3, putting $p = q$, which is possible in view of (2). If $\|g\|_x < e^{-q}$, the statement (18) is a fortiori true since $\|g\|_\xi$ is continuous.

We begin the proof of the theorem by considering the function $f - S_N$. In view of (17) all of its first derivatives vanish at $a = x_N$. Hence the definition of the norm (5) with $p = N$ implies

$$\|f - S_N\|_{x_N} \leq e^{-N},$$

and since $|f^{(N)}(\xi) - S_N^{(N)}(\xi)| = |f^{(N)}(\xi)| \leq m_N$, we infer from (18) that

$$\|f - S_N\|_{x_{N-1}} \leq e^{-N+1}.$$

From this it follows that also

$$(19) \quad \|f - S_{N-1}\|_{x_{N-1}} \leq e^{-N+1},$$

the only change being that on the left-hand side of (19) we have to consider

$$|f^{(N-1)}(x_{N-1}) - S_{N-1}^{(N-1)}(x_{N-1})| = |f^{(N-1)}(x_{N-1})| \leq m_{N-1},$$

which means that in the maximum occurring in the definition of the norm (5) a quantity enters which is $\leq e^{-N+1}$, and this can never conflict with (19).

In the same way we get

$$\|f - S_{N-2}\|_{x_{N-2}} \leq e^{-N+2}$$

and generally

$$\|f - S_{R+1}\|_{x_{R+1}} \leq e^{-R-1}.$$

Hence, by (18),

$$\|f - S_{R+1}\|_x \leq e^{-R} \text{ for } x_{R+1} \leq x \leq x_R,$$

so that altogether

$$(20) \quad \|f - S\|_x \leq e^{-R}$$

uniformly in the interval $x_N \leq x \leq x_R$. In the case of a finite interval of definition the uniform convergence is restricted to the part of the interval (x_N, x_R) which is contained in the interval of definition.

From equations (16) we have

$$x_R - a = \sum_{j=R+1}^N (x_{j-1} - x_j) = \frac{1}{e} \sum_{j=R+1}^N \frac{m_{j-1}}{m_j}.$$

In view of the assumption

$$(21) \quad \sum \frac{m_{n-1}}{m_n} = \infty$$

it is, therefore, possible to tie R to N in such a manner that $N \rightarrow \infty$ implies $R \rightarrow \infty$ and, at the same time, $x_R - a \rightarrow \infty$. Hence (20) shows that $\|f - S\|_x$ tends to 0 for $N \rightarrow \infty$, uniformly in every finite interval to the right of $x = a$ and contained in the interval of definition. Thus, the theorem is proved.

If we consider the class of all functions $f(x)$ which in a given interval satisfy

$$(22) \quad |f^{(n)}(x)| \leq Cm_n, \quad n = 0, 1, \dots,$$

where C is a constant depending on $f(x)$, and $\{m_n\}$ is a given sequence, we observe that the points $x_q = x_{q,N}$, as defined by (16) where m_n is replaced by Cm_n , are independent of C . Thus the above method of approximation is applicable to the whole class with the same points $x_q = x_{q,N}$. Even for quasi-analytic classes defined, as is usual in the literature, by inequalities

$$(23) \quad |f^{(n)}(x)| \leq CK^n m_n, \quad n = 0, 1, \dots,$$

where C and K are constants depending on $f(x)$, there exists an approximation method that is common to the whole class. This follows easily from the fact that a class of the type (23) is contained in some class of the type (22), where the sequence $\{m_n\}$ is replaced by a more rapidly increasing sequence $\{m_n^*\}$ for which the condition (21) is maintained. (See [4], p. 68; from the convergence of an approximation method, common to a whole class of type (23), at a fixed point $x_0 \neq a$ follows also the existence of a Toeplitz summation which produces $f(x)$ for all x when applied to the Taylor series from the point a , see [4], pp. 71–72).

7. In §§ 5 and 6 we have considered functions for which all derivatives vanish at the same point. To investigate the situation when the derivatives have different zeros, we shall now apply the definition of the norm in a special way. As usual we suppose $|f^{(n)}(x)| \leq m_n$, the sequence $\{m_n\}$ being logarithmically convex.

We put $\|f\|_x = \|f\|_x^{(0)}$ and generally

$$(24) \quad \|f\|_x^{(N)} = \inf_{p \geq N} \left[\max \left\{ e^{-p}, \max_{N \leq n \leq p} \frac{|f^{(n)}(x)|}{e^n m_n} \right\} \right],$$

simply omitting the derivatives of order less than N in the definition (5). Another way of connecting $\|f\|_x^{(N)}$ with the norm as defined by (5) would be to put

$$(25) \quad \|f\|_x^{(N)} = e^{-N} \|f^{(N)}\|_x^*,$$

where $\| \cdot \|^*$ is given by (5) with the sequence m_0, m_1, \dots replaced by m_N, m_{N+1}, \dots

Let us first note that if we put $p = N$ in (24) we get

$$(26) \quad \|f\|_x^{(N)} \leq e^{-N}.$$

A theorem corresponding to Theorem I is also valid for $\|f\|_x^{(N)}$, i.e., $\|f\|_x^{(N)} \geq e^{-q}$ implies

$$(27) \quad \|f\|_{x+h}^{(N)} \leq \|f\|_x^{(N)} \cdot \exp\left(e|h| \frac{m_q}{m_{q-1}}\right).$$

In fact, (26) shows that the assumption can only be satisfied if $q \geq N$,

and hence, by (25), the result follows from Theorem I, q being replaced by $q-N$.

Finally we have the inequality

$$(28) \quad \|f\|_x^{(N)} \geq \|f\|_x^{(N+1)}.$$

It follows from (26) if $\|f\|_x^{(N)} > e^{-N-1}$, whereas, if $\|f\|_x^{(N)} \leq e^{-N-1}$, the infimum in (24) is attained for some $p \geq N+1$, and then the right-hand side in (24) can never increase when we omit $n = N$. For our applications it will be essential that the sign of equality holds in (28) when $f^{(N)}(x) = 0$, which is obvious in view of (24).

Suppose now that $f(x)$ is not identically 0 and that each of its derivatives $f^{(n)}(x)$, $n = 0, 1, \dots$, possesses at least one zero x_n .

Consider $\|f\|_x^{(N)}$ on the closed interval from x_{N-1} to x_N . This norm is a continuous function $B_N(t)$ of the parameter t defined by

$$(29) \quad t = |x_0 - x_1| + |x_1 - x_2| + \dots + |x_{N-2} - x_{N-1}| + |x_{N-1} - x|.$$

Thus t denotes the length of the shortest path from x_0 passing through the points x_1, \dots, x_{q-1} , successively, and ending up at the point x . The norm $B_N(t)$ is defined in the interval $\tau_{N-1} \leq t \leq \tau_N$, where

$$(30) \quad \tau_q = |x_0 - x_1| + \dots + |x_{q-1} - x_q|.$$

Since $f^{(N)}(x_N) = 0$ we have $B_N(\tau_N) = B_{N+1}(\tau_N)$. Hence, there exists a continuous function $B(t)$ with the property that $B_N(t) = B(t)$ when $\tau_{N-1} \leq t \leq \tau_N$. The function $B(t)$ is defined for $0 \leq t < \tau$, where $\tau = \lim \tau_N \leq \infty$.

In (26) we give N the values $q+1, q+2, \dots$ and infer that

$$(31) \quad B(t) < e^{-q} \text{ for } t \geq \tau_q.$$

In particular, $B(t) \rightarrow 0$ as $t \rightarrow \tau$.

Since $f(x)$ is not identically 0, it is easily shown that $B(t)$ is not identically 0. It follows that, for a suitable integer r , the range of values of $B(t)$ contains all the numbers e^{-q} , $q \geq r$. We may now select an increasing sequence $\{t_q\}$, $q = r, r+1, \dots$, such that t_q is the smallest value of t exceeding t_{q-1} for which $B(t) = e^{-q}$. Thus $B(t_q) = e^{-q}$ and $B(t) > e^{-q}$ for $t_{q-1} < t < t_q$.

The values τ_N (if any) for which $t_{q-1} < \tau_N < t_q$ give rise to a subdivision of the interval (t_{q-1}, t_q) into smaller intervals for each of which the inequality (27) implies

$$B(\alpha) \leq B(\beta) \cdot \exp\left(e(\beta - \alpha) \frac{m_q}{m_{q-1}}\right),$$

α and β denoting the left and the right end points respectively. When

multiplying together the inequalities obtained in this way, we infer that

$$(32) \quad B(t_{q-1}) \leq B(t_q) \cdot \exp\left(e(t_q - t_{q-1}) \frac{m_q}{m_{q-1}}\right),$$

or, since $B(t_q) = e^{-q}$ for every $q \geq r$,

$$t_q - t_{q-1} \geq \frac{1}{e} \frac{m_{q-1}}{m_q}.$$

Hence

$$(33) \quad t_q \geq t_r + \frac{1}{e} \sum_{j=r+1}^q \frac{m_{j-1}}{m_j}.$$

It follows from (31) that $t_q < \tau_q$. Substituting the expression (30) for τ_q , we conclude that

$$(34) \quad \sum_{j=1}^q |x_{j-1} - x_j| > \frac{1}{e} \sum_{j=1}^q \frac{m_{j-1}}{m_j} - \text{const.}$$

8. Let us apply the last result to quasi-analytic functions. We therefore suppose that the series (21) is divergent.

Hence, the right-hand side in (34) tends to infinity as $q \rightarrow \infty$, and we have the following theorem:

THEOREM III. *If a quasi-analytic function $f(x)$ is not identically 0, and $f(x_0) = f'(x_1) = f''(x_2) = \dots = 0$, then the series $\sum |x_{j-1} - x_j|$ is divergent.*

A corollary (suggested by Professor R. P. Boas) that gives the answer to a question put forward by Borel (see [4], p. 75) is the following:

COROLLARY: *If $f(x)$ is quasi-analytic in an interval, and all its derivatives are positive at the left end point a , then all the derivatives are positive in the whole interval, i.e., $f(x)$ is absolutely monotone (and consequently analytic).*

For, suppose there exists a zero of one of the derivatives, $f^{(j)}(x_j) = 0$, then by Rolle's theorem there would exist a monotone sequence of zeros of the subsequent derivatives $x_j > x_{j+1} > \dots > a$, and from Theorem III it would follow that $f^{(j)}(x)$ (which is also quasi-analytic) is identically 0, which contradicts the assumption.

The following theorem, which is also an immediate consequence of (34), is more general than Theorem III:

THEOREM IV. *If $f(x)$ belongs to the quasi-analytic class defined by the inequalities (23), then*

$$(35) \quad \liminf_{q \rightarrow \infty} \left(\sum_{j=1}^q |x_{j-1} - x_j| \bigg/ \sum_{n=1}^q \frac{m_{n-1}}{m_n} \right) > 0.$$

We remark that the inequality (35) is valid also in the complex domain. Let $f(x)$ be analytic in a domain D in the complex plane and continuous on the boundary Γ of D . We assume that $\sup_{x \in D} |f^{(n)}(x)| \leq C \cdot K^n \cdot m_n$ for suitable constants C and K , $\{m_n\}$ being logarithmically convex. As above we assume that there exists a sequence of numbers $x_j \in D + \Gamma$ for which $f^{(j)}(x_j) = 0$. Finally we assume the existence of a constant L such that any two consecutive points x_{j-1} and x_j can be connected by a polygonal line Π_j in $D + \Gamma$ whose length does not exceed $L \cdot |x_{j-1} - x_j|$. The last assumption is satisfied, for instance, if D is convex ($L = 1$), or if the set $\{x_j\}$ is bounded and has a positive distance from Γ . In the case where $\sum(m_{n-1}/m_n)$ diverges, the proof may then be accomplished along the same lines as that of § 7 (if the series converges, the inequality (35) is immediate with an obvious exception when $x_0 = x_1 = x_2 = \dots$). The polygonal lines Π_1, Π_2, \dots form an infinite polygonal line Π in $D + \Gamma$. As a parameter t we take the distance along Π from x_0 to the arbitrary point x of Π , and as the numbers τ_q we now choose the values of t corresponding to the points x_q . Moreover, we define $B(t) = \|f\|_x^{(N)}$ when $x \in \Pi_N$, i.e., $\tau_{N-1} \leq t \leq \tau_N$. As in § 7, $B(t)$ is a continuous function defined for $0 \leq t < \tau = \lim \tau_N$ and $B(t) \rightarrow 0$ as $t \rightarrow \tau$. The sequence $\{t_q\}$, $q \geq r$, is defined as above, and the only change in the proof of (32) is that the subdivision of the interval (t_{q-1}, t_q) is performed by means of *all* values t in the interval which correspond to vertices in Π . Equation (32) implies (33) without alterations and the inequality $t_q < \tau_q$ remains valid. Since, by assumption, $\tau_q \leq L \cdot \sum |x_{j-1} - x_j|$, $1 \leq j \leq q$, the inequality (34) subsists except for a constant factor. This completes the proof.

Special values of m_n give special classes of analytic functions. For instance, the values $m_n = n!^\alpha$ ($\alpha < 1$) give entire functions of order $1/(1-\alpha)$, and thus we get theorems due to Gontcharoff and others (see e.g. Whittaker [11], chapt. III).

Let us return to quasi-analytic functions (with $|f^{(n)}(x)| \leq m_n$) in the interval $0 \leq x < \infty$. If such a function has an infinity of zeros $z_0 < z_1 < z_2 < \dots$, then by Rolle's theorem each of its derivatives has an infinity of zeros, and it is possible to choose $x_0 = z_0 < x_1 < x_2 < \dots$ such that $x_1 < z_1$, $x_2 < z_2, \dots$ and $f^{(j)}(x_j) = 0$. Hence the left-hand side of (34) is equal to $x_q - x_0 < z_q - z_0$, and we obtain the following theorem concerning the distribution of the zeros z_q of the function:

THEOREM V. *If $f(x)$ is not identically 0 and has the zeros $z_0 < z_1 < z_2 < \dots$, then*

$$z_q > \frac{1}{e} \sum_{j=1}^q \frac{m_{j-1}}{m_j} - C$$

where C is independent of q .

Theorems of essentially this type have been proved by I. I. Hirschman [7]. He assumes the sequence $\{m_n\}$ to be somewhat more regular (while in Theorem V it is only supposed to be logarithmically convex), so that his results cannot be applied to subclasses of the class of analytic functions, while Theorem V gives a unified approach to quasi-analytic functions and special analytic functions (e.g. entire functions of order $\alpha > 1$). On the other hand, instead of the constant $1/e$ in the theorem above he has the better constant $2/\pi$, and this seems to be the best possible value.

A quasi-analytic function $f(x)$, not identically 0, can only possess a finite number of zeros in a finite interval, and the same is valid for its derivatives (which are also quasi-analytic). Let N_n denote the number of zeros of $f^{(n)}(x)$ in the given closed interval. Then N_n cannot increase too rapidly as $n \rightarrow \infty$.

THEOREM VI. *If $f(x)$, defined in an interval of length l , is not identically 0 and $|f^{(n)}(x)| \leq m_n$ for all x and n , then*

$$(36) \quad \liminf_{n \rightarrow \infty} \sum_{j=n+1}^{n+N_n} \frac{m_{j-1}}{m_j} \leq el.$$

If $f(x)$ is not quasi-analytic, the statement of the theorem is void. We shall therefore assume that $\sum(m_{j-1}/m_j) = \infty$. On the other hand, if m_{j-1}/m_j is bounded away from 0, then $f(x)$ is an entire function of order 1 and (36) follows from certain well-known results on the zeros of such functions. Hence we may suppose that $m_{j-1}/m_j \rightarrow 0$ as $j \rightarrow \infty$. In view of this, the inequality (36) is obvious if $\liminf N_n < \infty$. We may, therefore, assume that $N_n > 1$ for sufficiently large n , say $n \geq q_0$.

We begin by defining the increasing sequence q_0, q_1, \dots by induction as follows: If $q_s = n$ we put $q_{s+1} = n + N_n - 1$.

Denoting by x_n the smallest (largest) zero of $f^{(n)}(x)$ we may apply Rolle's theorem in the above way and find an increasing (decreasing) set of zeros $x_n, x_{n+1}, \dots, x_{n+N_n-1}$ of the subsequent derivatives ($f^{(j)}(x_j) = 0$), where x_{n+N_n-1} is chosen as the largest (smallest) zero of $f^{(n+N_n-1)}(x)$.

Choosing x_{q_0} as the smallest zero of $f^{(q_0)}(x)$, the above procedure determines a sequence $\{x_j\}$ ($j \geq q_0$) for which $f^{(j)}(x_j) = 0$ and

$$(37) \quad \sum_{j=q_s+1}^{q_{s+1}} |x_{j-1} - x_j| = \left| \sum_{j=q_s+1}^{q_{s+1}} (x_{j-1} - x_j) \right| \leq l.$$

To this sequence we wish to apply (34). Since, however, x_j is defined only when $j \geq q_0$, it is necessary to modify (34) by restricting the summations on both sides so as to begin with $j = q_0 + 1$ instead of $j = 1$.

This modification is easily justified by applying (34) to $f^{(q_0)}(x)$ instead of $f(x)$.

We may now complete the proof of (36) indirectly. In fact, if

$$\sum_{j=n+1}^{n+N_n} \frac{m_{j-1}}{m_j} > (1 + \varepsilon) \cdot el$$

for all sufficiently large values of n , then

$$\sum_{j=n+1}^{n+N_{n-1}} \frac{m_{j-1}}{m_j} > \left(1 + \frac{\varepsilon}{2}\right) \cdot el$$

for all sufficiently large n since $m_{j-1}/m_j \rightarrow 0$ as $j \rightarrow \infty$. We apply this inequality only when n has the form $n = q_s$ and $n + N_n - 1 = q_{s+1}$. In view of (37) this gives

$$(38) \quad \sum_{j=q_{s+1}}^{q_{s+1}} \frac{m_{j-1}}{m_j} > \left(1 + \frac{\varepsilon}{2}\right) \cdot e \cdot \sum_{j=q_{s+1}}^{q_{s+1}} |x_{j-1} - x_j|$$

for all sufficiently large values of s , say $s \geq s_0$. By adding the inequalities obtained from (38) when s is given a sufficiently large number of consecutive values $s = s_0, s_0 + 1, \dots$, we arrive at a contradiction to the modified inequality (34). This completes the proof of the theorem.

Like the preceding theorems, Theorem VI may be applied to analytic functions of special order and type, and gives theorems of a kind treated, for instance, by Pólya [10], but here again the constant e in the theorem does not seem to be best possible. However, Theorem VI is applicable to non-analytic, quasi-analytic functions. Choosing for instance $m_n = n! (\log n)^n$, it follows that

$$\liminf \frac{\log N_n}{\log n} \leq e^{el}.$$

REFERENCES

1. S. Banach, *Théorie des opérations linéaires*, Warszawa 1932.
2. Th. Bang, *Om quasi-analytiske Funktioner*, Thesis, København 1946.
3. Th. Bang, *Metric spaces of infinitely differentiable functions*, Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, 1, 375.
4. T. Carleman, *Fonctions quasi analytiques*, Paris 1926.
5. L. Carleson, *On Bernstein's approximation problem*, Proc. Amer. Math. Soc. 2 (1951), 953-961.
6. W. Gontcharoff, *Recherches sur les dérivées successives des fonctions analytiques*, Ann. Sci. École Norm. Sup. 47 (1930), 1-78.
7. I. I. Hirschman, *On the distributions of zeros of functions belonging to certain quasi-analytic classes*, Amer. J. Math. 72 (1950), 396-406.

8. S. Mandelbrojt, *Séries de Fourier et classes quasi-analytiques de fonctions*, Paris 1935.
9. S. Mandelbrojt, *Analytic functions and classes of infinitely differentiable functions*, Rice Inst. Pamphlet 29, No. 1 (1942).
10. G. Pólya, *On the zeros of the derivatives of a function and its analytic character*, Bull. Amer. Math. Soc. 49 (1943), 178–191.
11. J. M. Whittaker, *Interpolatory function theory*, Cambridge 1935.

UNIVERSITY OF COPENHAGEN, DENMARK