

Det Kgl. Danske Videnskabernes Selskab.

Mathematisk-fysiske Meddelelser. **VIII**, 5.

SOME REMARKS
ON GENERALISATIONS OF ALMOST
PERIODIC FUNCTIONS

BY

A. BESICOVITCH AND H. BOHR

(CAMBRIDGE)

(COPENHAGEN)



KØBENHAVN

HOVEDKOMMISSIONÆR: ANDR. FRED. HØST & SØN, KGL. HOF-BOGHANDEL

BIANCO LUNOS BOGTRYKKERI

1927

Introduction.

The theory of almost periodic functions had its origin in the problem, what functions $f(x) = u(x) + iv(x)$ can be decomposed, in the interval $-\infty < x < \infty$, into pure oscillations, i. e. into oscillations of the form $e^{i\lambda x}$. The problem stated in these general terms has evidently no definite meaning until the notion of "decomposition" has been strictly defined and this of course can be done in many different ways.

The first and most primitive way of interpreting the term would perhaps be to regard as decomposable only those functions which can be represented as the sum of a finite number of oscillations:

$$(1) \quad s(x) = \sum_{\nu=1}^N a_{\nu} e^{i\lambda_{\nu} x}.$$

We shall denote by A the class of such functions $s(x)$. But at the first attempt to develop the theory of functions of this class we see that the definition is too narrow. Indeed the class A is not "closed" to limit processes, so that when working only with functions of the class A we should have to exclude from the start those operations which involve the idea of continuity in "functional space".

We must therefore close the class A, that is we must extend it to a larger class C(A) consisting of all functions $f(x)$ (including the functions of A itself) which are the

limits of sequences of functions $s(x)$ of A. But here again the content of the class $C(A)$ depends on the kind of the limit process employed.

The simplest limit process is that of ordinary convergence: $f(x)$ is a limit-function of the class A if there is a sequence

$$s_1(x), s_2(x), \dots, s_n(x), \dots$$

of functions of A such that for every x

$$(2) \quad f(x) = \lim_{n \rightarrow \infty} s_n(x).$$

But the class $C(A)$ to which this limit process leads is found to be too wide, in the sense that practically none of the characteristic properties of the functions $s(x)$ (those relating to oscillations) are conserved. In fact, as BESICOVITCH [1]¹ has shown, the above class $C(A)$ includes all bounded continuous functions.

The same is true even when we demand that the convergence shall be uniform in every finite interval.

We are therefore led to consider only limit processes which involve some sort of uniformity in the whole interval $-\infty < x < \infty$.

In the theory of a. p. (almost periodic) functions developed by BOHR in his papers in Acta Mathematica [1, 2, 3] the limit process employed was that of ordinary uniform convergence in the whole infinite interval $-\infty < x < \infty$. The class $C(A)$ corresponding to this limit process is the narrowest possible closure of the class A. But, as we shall see, the theory of larger closures $C(A)$ derived from more general processes of uniform convergence can be treated simply as generalisation of the theory of the above

¹ The list of papers quoted is given as an appendix.

closure $C(A)$ in the sense that we can extend very many results directly, without repeating the arguments of their proof.

In § 1 of this paper we give a short outline of a part of the theory of a. p. functions. § 2 contains some general remarks on the generalisation of the theory to larger classes. Finally in § 3 we give in full detail an application of the principles laid down in § 2 to a particular generalisation of a. p. functions; for this purpose we choose the class of summable functions discussed by STEPANOFF in his interesting paper in Math. Ann. [1], where such generalisations were studied for the first time.

§ 1.

For convenience and also in order to bring out as clearly as possible the similarity between the definitions and proofs of this paragraph and these of § 3 we introduce the following notation:

By
(3)
$$U\text{-}\lim f_n(x) = f(x)$$

we mean that $f_n(x)$ tends to $f(x)$ uniformly in the whole interval $-\infty < x < \infty$. We shall call the upper bound of the difference $|f(x) - g(x)|$ for $-\infty < x < \infty$ the U -distance between the functions $f(x)$ and $g(x)$ and shall denote it by $D_U[f(x), g(x)]$; thus (3) can be written in the equivalent form

$$D_U[f_n(x), f(x)] \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

By $D_U[f(x)]$ we mean $D_U[f(x), 0]$, i. e. the upper bound of $|f(x)|$ in the interval $-\infty < x < \infty$.

A being as before the class of all finite sums

$$s(x) = \sum_{\nu=1}^N a_{\nu} e^{i\lambda_{\nu}x}$$

where the λ_{ν} are real and different, and the a_{ν} arbitrary complex numbers, we denote by $C_U(A)$ the class of all functions $f(x)$ for each of which there is a sequence $s_n(x)$ of functions of A such that $f(x) = U\text{-lim } s_n(x)$.

We proceed to establish some properties of functions of the class $C_U(A)$, which follow directly from this definition.

1°. Every function $f(x)$ of $C_U(A)$ is bounded for $-\infty < x < \infty$.

For, given $f(x)$ we can choose an $s(x)$ so that

$$D_U[f(x), s(x)] < 1;$$

since $s(x)$ is plainly bounded, the result follows from the inequality

$$D_U[f(x)] \leq D_U[f(x), s(x)] + D_U[s(x)].$$

2°. Every function $f(x)$ of $C_U(A)$ is uniformly continuous in the whole interval $-\infty < x < \infty$.

For, given ε we can choose an $s(x)$ such that

$$D_U[f(x), s(x)] < \frac{\varepsilon}{3}.$$

$s(x)$ is evidently uniformly continuous; we can therefore choose δ so that

$$D_U[s(x+h), s(x)] < \frac{\varepsilon}{3} \quad \text{for } |h| < \delta.$$

From the inequality

$$D_U[f(x+h), f(x)] \leq D_U[s(x+h), s(x)] + 2D_U[f(x), s(x)]$$

it now follows that for $|h| < \delta$

$$D_U[f(x+h), f(x)] < \varepsilon.$$

3°. The sum and the product of two functions of $C_U(A)$ are again functions of $C_U(A)$.

This follows at once from the fact that the sum and the product of two functions $s(x)$ are functions $s(x)$.

In particular we observe that if $f(x)$ belongs to $C_U(A)$, then so does $f(x)e^{i\lambda x}$ for every real value of λ .

4°. Every function $f(x)$ of $C_U(A)$ has a mean value $M\{f(x)\}$, i. e.

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{\gamma}^{\gamma+T} f(x) dx$$

exists; this is even true uniformly in γ .

The property is obvious for any function $s(x)$ (the mean value in this case being the constant term in $s(x)$). Choosing $s(x)$ so that $D_U[f(x), s(x)]$ is "small" the result follows in the usual way from the inequality

$$\left| \frac{1}{T} \int_{\gamma}^{\gamma+T} f(x) dx - \frac{1}{T} \int_{\gamma}^{\gamma+T} s(x) dx \right| \leq D_U[f(x), s(x)].$$

From the same inequality we see that if

$$(4) \quad U\text{-lim } s_n(x) = f(x)$$

then

$$M\{s_n(x)\} \rightarrow M\{f(x)\}, \quad \text{as } n \rightarrow \infty,$$

and further, for any real value of λ ,

$$(5) \quad M\{s_n(x)e^{-i\lambda x}\} \rightarrow M\{f(x)e^{-i\lambda x}\}, \quad \text{as } n \rightarrow \infty.$$

For (4) implies

$$U\text{-lim } s_n(x)e^{-i\lambda x} = f(x)e^{-i\lambda x}.$$

5°. For any function $f(x)$ of $C_U(A)$ the mean value

$$M\{f(x)e^{-i\lambda x}\} = a(\lambda)$$

differs from zero for at most an enumerable set of values of λ .

Let $s_1(x), s_2(x), \dots$ be a sequence of functions of A such that

$$U\text{-lim } s_n(x) = f(x).$$

For each function $s_n(x)$ the mean value $M\{s_n(x)e^{-i\lambda x}\}$ differs from zero only for a finite set of values of λ , namely those occurring as the exponents in the polynomial expression for $s_n(x)$. It follows that, except in at most an enumerable set of λ 's, $M\{s_n(x)e^{-i\lambda x}\}$ is zero for every n , and so by (5) that $a(\lambda) = M\{f(x)e^{-i\lambda x}\}$ is zero except for at most an enumerable set of values of λ .

We can now denote these λ 's by

$$\lambda_1, \lambda_2, \dots$$

and the corresponding $a(\lambda)$ by

$$A_1, A_2, \dots$$

We express this symbolically by writing

$$(6) \quad f(x) \sim \sum A_\nu e^{i\lambda_\nu x}$$

and we call the series (finite or infinite) on the right the Fourier series of the function $f(x)$.

The Fourier series of a function $s(x)$ of A evidently coincides with its polynomial expression (1). From (5) we see that, as $n \rightarrow \infty$, the polynomial expression of $s_n(x)$ goes over by a "formal" limit passage into the Fourier series of $f(x)$; this already shows that Fourier series are likely to play an important part in the study of functions of $C_U(A)$.

We next consider a property of functions of $C_U(A)$ of a different kind. We call a real number $\tau = \tau(\epsilon)$ a

translation number of the function $f(x)$ belonging to ε if

$$D_U[f(x+\tau), f(x)] \leq \varepsilon.$$

Corresponding to any given function $s(x) = \sum_1^N a_\nu e^{i\lambda_\nu x}$

of the class A there exist for any given $\varepsilon > 0$ an infinite number of translation numbers, and the set of these numbers $\tau = \tau(\varepsilon)$ is even "relatively dense" in the sense that any interval of a certain length $l = l(\varepsilon)$ contains at least one such number $\tau(\varepsilon)$. This is an immediate consequence of the BOHL-WENNBERG theorem on diophantine approximations (see f. inst. Bohr [1], p. 120), which states that for an arbitrarily small δ the N diophantine inequalities¹

$$|\lambda_\nu x| < \delta \pmod{2\pi} \quad (\nu = 1, 2, \dots, N)$$

have relatively dense solutions with respect to x . In the ordinary way (i. e. by approaching the function $f(x)$ by a function $s(x)$ such that $D_U[f(x), s(x)]$ is small) we see that the functions $f(x)$ of $C_U(A)$ also possess the above property, namely that for every $\varepsilon > 0$ the ε -translation numbers exist and are relatively dense.

Functions which are continuous in $-\infty < x < \infty$ and possess this property are said to be almost periodic; we have just seen that every function of A, and even every function of $C_U(A)$, is almost periodic.

The main result of the theory of a. p. (almost periodic) functions is that the converse of the last statement is also true; every a. p. function is a function of $C_U(A)$, so that

The class $C_U(A)$ is identical with the class of a. p. functions.

¹ By $|a| < b \pmod{c}$, where a, b, c are real and b and c positive, we mean that there exists an integral n such that $|a - nc| < b$.

Naturally we shall not enter the proof of this theorem which would involve the development of almost the whole theory of a. p. functions. In proving the identity of the class $C_U(A)$ with the class of a. p. functions Bohr had first to show that the latter class possesses all the above properties 1°, . . . , 5° of the class $C_U(A)$; and though this was not the main difficulty of the investigation, it nevertheless involved considerations of a different character from the immediate deductions employed above in establishing these properties for the class $C_U(A)$. The main difficulty to be overcome was the proof of the "fundamental theorem" (Parseval's theorem) namely

$$M\{|f(x)|^2\} = \sum |A_\nu|^2,$$

from which follows as an immediate corollary the "uniqueness theorem":

An a. p. function is uniquely determined by its Fourier series, i. e. two different a. p. functions cannot have the same Fourier series.

When the identity of the class $C_U(A)$ and the class of a. p. functions has been established the question naturally arises: given an a. p. function $f(x)$, actually to find a sequence of functions $s_n(x)$ such that

$$U\text{-lim } s_n(x) = f(x).$$

A method of obtaining such a sequence of functions was given by Bohr in his second paper in Acta Math.; his sums $s_n(x)$ contained as exponents only exponents from the Fourier series of $f(x)$, a fact of importance in the extension of the theory to functions of a complex variable. An essentially simpler method of obtaining such approximation functions $s_n(x)$ was given by BOCHNER [1], who

succeeded in extending the Fejér summation method of classical Fourier theory to the class of a. p. functions. Like Bohr, he started from the representation of the "Fourier exponents" \mathcal{A}_ν with the help of a "base" $\alpha_1, \alpha_2, \dots$. By a base we mean a sequence of linearly independent¹ positive numbers $\alpha_1, \alpha_2, \dots$ (which generally is enumerable but in particular cases may be finite) such that every the α 's \mathcal{A}_ν may be expressed as a finite linear form in exponent with rational coefficients,

$$\mathcal{A}_\nu = r_{\nu,1} \alpha_1 + r_{\nu,2} \alpha_2 + \dots + r_{\nu,q_\nu} \alpha_{q_\nu}.$$

Fejér in his summation of Fourier series of pure periodic functions $f(x)$, with period 2π , used as approximation sums the expressions

$$s_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+t) H_n(t) dt = M\{f(x+t) H_n(t)\},$$

where the "kernel" $H_n(t)$ was given by

$$H_n(t) = \sum_{\nu=-n}^n \left(1 - \frac{|\nu|}{n}\right) e^{-i\nu t} = \frac{1}{n} \left(\frac{\sin n \frac{t}{2}}{\sin \frac{t}{2}} \right)^2.$$

Bochner replaced Fejér's simple kernel by a finite product of such kernels

$$H(t) = H_{n_1}(\beta_1 t) \dots H_{n_p}(\beta_p t) = \sum_{\substack{-n_1 \leq \nu_1 \leq n_1 \\ \dots \\ -n_p \leq \nu_p \leq n_p}} \left(1 - \frac{|\nu_1|}{n_1}\right) \dots \left(1 - \frac{|\nu_p|}{n_p}\right) e^{-i(\nu_1 \beta_1 + \dots + \nu_p \beta_p) t}$$

¹ $\alpha_1, \alpha_2, \dots$ are said to be linearly independent if no equation of the form

$$r_1 \alpha_1 + r_2 \alpha_2 + \dots + r_N \alpha_N = 0$$

holds, where $N \geq 1$ and the r 's are rational, not all naught.

where the β 's are linearly independent numbers. This composite kernel has the same characteristic properties as the Fejér kernel — it is always positive and its mean-value $M\{H(t)\}$ is equal to 1 (the constant term in the polynomial expansion of $H(t)$ being 1 on account of the linear independence of the β 's). Bochner considers an expression of the form

$$M\left\{f(x+t)H_{n_1}\left(\frac{\alpha_1}{N_1!}t\right)\dots H_{n_p}\left(\frac{\alpha_p}{N_p!}t\right)\right\}$$

which, since

$$f(x+t) \sim \sum A_\nu e^{iA_\nu x} \cdot e^{iA_\nu t},$$

is equal to the finite sum

$$s(x) = \sum_{\substack{-n_1 \leq \nu_1 \leq n_1 \\ \dots \\ -n_p \leq \nu_p \leq n_p}} \left(1 - \frac{|\nu_1|}{n_1}\right) \dots \left(1 - \frac{|\nu_p|}{n_p}\right) A_\nu e^{iA_\nu x}$$

where

$$(7) \quad A_\nu = \frac{\nu_1}{N_1!} \alpha_1 + \dots + \frac{\nu_p}{N_p!} \alpha_p$$

(and A_ν is to be interpreted as zero when the linear combination (7) of α 's is not an exponent in the Fourier series of $f(x)$). Bochner's result is:

The sum $s(x)$ tends uniformly to $f(x)$, as $p \rightarrow \infty$, $N_1 \rightarrow \infty$, $N_2 \rightarrow \infty$, ... and $\frac{n_1}{N_1!} \rightarrow \infty$, $\frac{n_2}{N_2!} \rightarrow \infty$, ... , in other words, provided the limit process is carried out in such a way that every exponent A_ν occurs sooner or later in $s(x)$.

As we shall have to prove in § 3 an analogous theorem for a more general class of functions, we next give a proof of Bochner's theorem which, though perhaps not so elegant as Bochner's own, is better adapted to generalisation. Like

Bochner's, our proof depends on the fundamental theorem. We start from the following property of a. p. functions, which (see Bohr, II p. 110) is a fairly direct consequence of the fundamental theorem:

"To any ε corresponds an integer M and a positive number η such that every solution t of the system of diophantine inequalities

$$|\mathcal{A}_1 t| < \eta \pmod{2\pi}, \dots, |\mathcal{A}_M t| < \eta \pmod{2\pi}$$

is a translation number $\tau(\varepsilon)$ of the given a. p. function".

Since each exponent \mathcal{A}_p can be represented as a finite linear form in the α 's with rational coefficients, we have the immediate corollary:

To any ε correspond integers P and Q and a positive number $\delta < \pi$ such that every solution t of the system of diophantine inequalities

$$(8) \quad \left| \frac{\alpha_1}{Q} t \right| < \delta \pmod{2\pi}, \dots, \left| \frac{\alpha_P}{Q} t \right| < \delta \pmod{2\pi}$$

is a translation number $\tau\left(\frac{\varepsilon}{2}\right)$ of the given a. p. function.

Let I_1 denote the set of all values of t satisfying (8) and I_2 the set of all other values of t . We can write the kernel

$$H(t) = H_{n_1}\left(\frac{\alpha_1}{N_1!} t\right) \dots H_{n_p}\left(\frac{\alpha_p}{N_p!} t\right)$$

as the sum of two kernels

$$H(t) = H'(t) + H''(t)$$

where

$$\begin{aligned} H'(t) &= H(t), & H''(t) &= 0 & \text{for } t \in I_1, \\ H'(t) &= 0, & H''(t) &= H(t) & \text{for } t \in I_2. \end{aligned}$$

We may further assume that $p > P$ and N_1, \dots, N_p all $> Q$, where P and Q are the integers (depending on ε) which occur in (8).

From

$$s(x) - f(x) = M \{ (f(x+t) - f(x)) \Pi(t) \}$$

we have

$$|s(x) - f(x)| \leq \overline{M} \{ |f(x+t) - f(x)| \Pi'(t) \} \\ + \overline{M} \{ |f(x+t) - f(x)| \Pi''(t) \}$$

where $\overline{M} \{ g(t) \}$ for a positive function $g(t)$ (which need not necessarily possess a mean value) denotes

$\limsup \frac{1}{2T} \int_{-T}^T g(t) dt$ as $T \rightarrow \infty$. Thus

$$(9) \quad \begin{cases} D_U [s(x), f(x)] \leq \overline{M} \{ D_U [f(x+t), f(x)] \Pi'(t) \} \\ \quad + \overline{M} \{ D_U [f(x+t), f(x)] \Pi''(t) \} \end{cases}$$

where the distance D_U on the right hand side is taken with respect to x, t remaining fixed.

Since for all values of t belonging to I_1

$$D_U [f(x+t), f(x)] \leq \frac{\varepsilon}{2}$$

while for all other values of t

$$\Pi'(t) = 0$$

we have at once

$$(10) \quad \begin{cases} \overline{M} \{ D_U [f(x+t), f(x)] \Pi'(t) \} \\ \leq \frac{\varepsilon}{2} \overline{M} \{ \Pi'(t) \} \leq \frac{\varepsilon}{2} M \{ \Pi(t) \} = \frac{\varepsilon}{2}. \end{cases}$$

Writing $K = D_U [f(x)]$, we have for the second term on the right hand side of (9)

$$(11) \quad \overline{M} \{ D_U [f(x+t), f(x)] \Pi''(t) \} \leq 2K \overline{M} \{ \Pi''(t) \}.$$

Let $I_{2,q}$ ($q = 1, 2, \dots, P$) denote the set of values t , at which the q -th of the inequalities (8) is not satisfied.

This set $I_{2,q}$ consists of the whole t -axis with the exception of small intervals of length $2l_q = 2\frac{\delta Q}{\alpha_q}$ with their centres at the zeros of $\sin\frac{\alpha_q t}{2Q}$. Since $N_q > Q$ and $N_q!$ is thus a multiple of Q , we observe that the zeros of $\sin\frac{\alpha_q t}{2N_q!}$ are included among these of $\sin\frac{\alpha_q t}{2Q}$. Putting

$$H_{nq}^*\left(\frac{\alpha_q}{N_q!}t\right) = H_{nq}\left(\frac{\alpha_q}{N_q!}t\right) \text{ for } t \in I_{2,q}, \text{ and } = 0 \text{ for } t \text{ outside } I_{2,q}$$

we evidently have for $-\infty < t < \infty$

$$(12) \quad H''(t) \leq \sum_{q=1}^P H_{nq}^* \cdot H_{n_1} H_{n_2} \dots H_{n_{q-1}} H_{n_{q+1}} \dots H_{n_p}$$

since for any t in I_2 at least one of the inequalities (8) is not satisfied and thus at least one of the P term of this sum is equal to $H(t)$. Therefore

$$\bar{M}\{H''(t)\} \leq \sum_{q=1}^P M\{H_{nq}^* \cdot H_{n_1} H_{n_2} \dots H_{n_{q-1}} H_{n_{q+1}} \dots H_{n_p}\}.$$

As the product $H_{n_1} \dots H_{n_{q-1}} H_{n_{q+1}} \dots H_{n_p}$ is a finite sum $s(x) = \sum \alpha_\nu e^{i\lambda_\nu x}$ with the constant term 1 whose other exponents do not satisfy any relation of the form $\lambda_\nu + r\alpha_q = 0$ with rational r , the mean value under the sign of summation is simply equal to the mean value $M\{H_{nq}^*\}$, both being evidently equal to the constant term in the ordinary Fourier series for the pure periodic function H_{nq}^* (with period $2\omega_q = 2\pi\frac{N_q!}{\alpha_q}$). Thus

$$(13) \quad \bar{M}\{H''(t)\} \leq \sum_{q=1}^P M\left\{H_{nq}^*\left(\frac{\alpha_q}{N_q!}t\right)\right\}.$$

Further

$$\begin{aligned}
 M \left\{ H_{n_q}^* \left(\frac{\alpha_q}{N_q!} t \right) \right\} &\leq \frac{1}{\omega_q} \int_{l_q}^{\omega_q} H_{n_q} \left(\frac{\alpha_q}{N_q!} t \right) dt \\
 &= \frac{1}{\omega_q} \int_{l_q}^{\omega_q} \frac{1}{n_q} \frac{\sin^2 \frac{n_q \alpha_q t}{2 N_q!}}{\sin^2 \frac{\alpha_q t}{2 N_q!}} dt \leq \frac{1}{n_q \omega_q} \int_{l_q}^{\omega_q} \frac{1}{\left(\frac{1}{2} \cdot \frac{\alpha_q t}{2 N_q!} \right)^2} dt \\
 &< \frac{16 (N_q!)^2}{n_q \omega_q \alpha_q^2} \int_{l_q}^{\omega_q} \frac{dt}{t^2} = \frac{16 (N_q!)^2}{n_q \omega_q \alpha_q^2 l_q} = \frac{16}{\pi \delta Q} \cdot \frac{N_q!}{n_q} < \frac{6}{\delta Q} \frac{N_q!}{n_q}.
 \end{aligned}$$

Thus by (13)

$$\bar{M} \{ H''(t) \} < \frac{6}{\delta Q} \sum_{q=1}^P \frac{N_q!}{n_q}.$$

On putting

$$n_q > \frac{24 KP}{\delta Q \varepsilon} N_q! \quad (q = 1, \dots, P)$$

we have therefore

$$2 K \bar{M} \{ H''(t) \} < \frac{\varepsilon}{2}$$

and thus finally by (9), (10), (11)

$$D_U[s(x), f(x)] < \varepsilon$$

for $p > P$, $N_q > Q$, $n_q > CN_q!$ ($q = 1, \dots, P$), where P , Q , C depend only on ε

Q. E. D.

§ 2.

A natural way of generalising the theory of a. p. functions is to use, in closing the class A, a limit process more general than the simple U -lim employed in § 1. Let G denote such a limit process. Then the first and main problem which arises is to determine the generalised almost periodic properties which characterise the class $C_G(A)$, the closure of A by the limit process G .

But once this problem has been solved in the original theory its solution for generalised classes may be reduced to considerations concerning only the limit process G and its effect on almost-periodicity. Stepanoff himself proceeded along these lines in the paper already quoted and thus escaped entering once more into the difficulties involved in establishing the fundamental theorem or the uniqueness theorem. It is however possible to go further in this way. We have only to take into account and use to the full the fact that the U -lim process employed in the original theory, while general enough to bring out the main properties of almost periodicity, is at the same time the "narrowest" of all limit processes of the kind described in the introduction. It is the narrowest in the sense that the closure of the class A by any limit process G coincides with the closure by G of the class $C_U(A)$ already closed by the U -process, i. e. in symbols

$$C_G(A) = C_G(C_U(A)).$$

Thus

$$C_G(A) = C_G(\text{a. p. functions})$$

which shows that the problem of characterising the class $C_G(A)$ by almost periodic properties is equivalent to the investigation of the effect of the limit process G on ordinary almost periodicity.

Once the character of the almost periodicity corresponding to a given process G has been determined, the next main question is to find an "algorithm" which, applied to a function $f(x)$ possessing this type of almost periodicity, will lead to a sequence $s_n(x)$ of functions of A which approach the function $f(x)$ in the sense of the given limit process G , i. e. for which

$$G\text{-lim } s_n(x) = f(x), \quad \text{as } n \rightarrow \infty.$$

The natural method of treating this question will generally be by establishing the existence of the Fourier series of $f(x)$ and applying to it a suitable method of summation.

We now carry out an investigation of the type just outlined, obtaining finally what is perhaps the most natural generalisation of the class of a. p. functions, namely the class of integrable (L) functions considered by Stepanoff. The very fact that this class of functions — defined by Stepanoff himself through generalised almost periodic properties — can also be characterised as the closure of the class of (ordinary) a. p. functions by a certain limit process (the S -process given below) has already been pointed out by Bochner [2].

§ 3.

The functions which we have to consider in this section are assumed to be defined almost everywhere in the whole interval $-\infty < x < \infty$ and to be integrable (L) over any finite interval. We begin by introducing a notation analogous to that of § 1. We say that the function $f(x)$ is the S -limit (Stepanoff limit) of the sequence $f_n(x)$ ($n = 1, 2, \dots$) and write

$$(14) \quad f(x) = S\text{-lim } f_n(x)$$

if

$$\text{Upper bound } \int_x^{x+1} |f(\xi) - f_n(\xi)| d\xi \rightarrow 0, \quad \text{as } n \rightarrow \infty^1.$$

By the S -distance $D_S[f(x), g(x)]$ between the functions $f(x)$ and $g(x)$ we mean

¹ Evidently the definition of the S -limit will not be altered if we replace $\int_x^{x+1} | | d\xi$ by $\int_x^{x+k} | | d\xi$, where k is an arbitrary positive constant.

$$\text{Upper bound } \int_x^{x+1} |f(\xi) - g(\xi)| d\xi;$$

$$-\infty < x < \infty$$

thus the equation (14) is equivalent to

$$D_S [f(x), f_n(x)] \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Further by $D_S [f(x)]$ we mean $D_S [f(x), 0]$, i. e. the upper bound of $\int_x^{x+1} |f(\xi)| d\xi$ in the interval $-\infty < x < \infty$.

This section is devoted to the study of the class $C_S(A)$, i. e. the closure of the class A by the S -limit process.

As in § 1 we begin by establishing a number of properties of the functions of $C_S(A)$ which follow directly from their definition as S -limits of functions $s(x)$ of A . The deduction of these properties is on the same lines as before, the only difference being that the U -processes of § 1 are here replaced by S -processes. Our notation is essentially that of Bochner [2].

1°. Every function $f(x)$ of $C_S(A)$ is "S-bounded" in $-\infty < x < \infty$, i. e. $D_S [f(x)]$ is finite.

The property follows from the inequality

$$D_S [f(x)] \leq D_S [f(x), s(x)] + D_S [s(x)]$$

in the same way as in § 1 ($s(x)$, being bounded, is a fortiori S -bounded).

2°. Every function $f(x)$ of $C_S(A)$ is "S-uniformly continuous" in $-\infty < x < \infty$, i. e.

$$D_S [f(x+h), f(x)] < \varepsilon \quad \text{for } |h| < \delta.$$

In the same way as before the property follows from the inequality

$$D_S [f(x+h), f(x)] \leq D_S [s(x+h), s(x)] + 2D_S [f(x), s(x)].$$

3,1^o. Evidently the sum of two functions of $C_S(A)$ is again a function of $C_S(A)$.

3,2^o. If $f(x)$ belongs to $C_S(A)$, then the product $f(x)e^{i\lambda x}$ belongs to $C_S(A)$ for every real λ .

For

$$D_S [f(x)e^{i\lambda x}, s(x)e^{i\lambda x}] = D_S [f(x), s(x)].$$

(The property does not hold for the product of any two functions of $C_S(A)$, since such a product may not be integrable).

4^o. Every function $f(x)$ of $C_S(A)$ has a mean value

$$M\{f(x)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{\gamma}^{\gamma+T} f(x) dx,$$

(and the limit even exists uniformly in γ).

This follows from the inequality

$$\begin{aligned} & \left| \frac{1}{T} \int_{\gamma}^{\gamma+T} f(x) dx - \frac{1}{T} \int_{\gamma}^{\gamma+T} s(x) dx \right| \leq \frac{1}{T} \int_{\gamma}^{\gamma+[T]+1} |f(x) - s(x)| dx \\ & \leq \frac{[T]+1}{T} D_S [f(x), s(x)] \leq 2 D_S [f(x), s(x)] \quad \text{for } T \geq 1. \end{aligned}$$

From the same inequality we see that if

$$S\text{-lim } s_n(x) = f(x)$$

then

$$M\{s_n(x)\} \rightarrow M\{f(x)\}$$

and, more generally,

$$M\{s_n(x)e^{i\lambda x}\} \rightarrow M\{f(x)e^{i\lambda x}\}$$

as $n \rightarrow \infty$.

5^o. Repeating word for word the argument of § 1 we can now establish the existence of a Fourier series, write

$$f(x) \infty \sum A_\nu e^{iA_\nu x},$$

in the same sense as before, and assert that when $S\text{-lim } s_n(x) = f(x)$ the Fourier series of $f(x)$ may be obtained from the expressions of $s_n(x)$ by a formal limit passage.

We now pass to the main part of the theory of the class $C_S(A)$, the investigation of the almost periodic properties characteristic of the class.

We shall in the future refer to the translation numbers defined in § 1 as U -translation numbers. We now introduce another kind of translation numbers, defined as follows: A number τ is said to be an S -translation number of the function $f(x)$ belonging to ε , if

$$D_S [f(x + \tau), f(x)] \leq \varepsilon.$$

We call $f(x)$ an S . a. p. function if for every positive ε there exists a relatively dense set (in the sense of § 1) of S -translation numbers $\tau(\varepsilon)$ of $f(x)$.

Theorem. *The class $C_S(A)$ is identical with the class of S . a. p. functions.*

In proving the identity of the class $C_S(A)$ with the class of S . a. p. functions we shall, so far as the latter class is concerned, use nothing but its definition; we do not even need to begin, as Stepanoff did in his development of the theory, by establishing the elementary properties, just deduced for $C_S(A)$ as immediate consequences of its definition. In accordance with § 2 we base our proof on the fact that the class $C_S(A)$ is identical with the class

$$C_S(\text{a. p. functions}).$$

For convenience we shall in future denote an a. p. function by $\sigma(x)$.

1. That every function $f(x)$ of $C_S(A)$ is an S. a. p. function, is obvious. We have only to choose $\sigma(x)$ in the inequality

$$D_S[f(x+\tau), f(x)] \leq D_S[\sigma(x+\tau), \sigma(x)] + 2D_S[f(x), \sigma(x)]$$

so that $D_S[f(x), \sigma(x)] < \frac{\varepsilon}{3}$, to ensure that every U -translation number of $\sigma(x)$ (which is a fortiori an S -translation number of $\sigma(x)$) belonging to $\frac{\varepsilon}{3}$ shall be an S -translation number of $f(x)$ belonging to ε .

2. We now proceed to the proof of the converse result, namely that every S. a. p. function is the S -limit of a sequence of functions $\sigma(x)$. For this purpose we consider the functions

$$\varphi_\delta(x) = \frac{1}{\delta} \int_x^{x+\delta} f(\xi) d\xi$$

already studied by Stepanoff. We shall prove first that $\varphi_\delta(x)$ (which is evidently continuous) is an a. p. function $\sigma(x)$, and secondly that

$$f(x) = S\text{-lim } \varphi_\delta(x), \quad \text{as } \delta \rightarrow 0.$$

We may suppose $\delta < 1$.

To prove the first statement we observe that every S -translation number of $f(x)$ which belongs to $\varepsilon\delta$ is also a U -translation number of $\varphi_\delta(x)$ belonging to ε . Given such an S -translation number $\tau = \tau(\varepsilon\delta)$ we have in fact for every x

$$\begin{aligned} |\varphi_\delta(x+\tau) - \varphi_\delta(x)| &= \frac{1}{\delta} \left| \int_x^{x+\delta} (f(\xi+\tau) - f(\xi)) d\xi \right| \\ &\leq \frac{1}{\delta} \int_x^{x+\delta} |f(\xi+\tau) - f(\xi)| d\xi \leq \varepsilon. \end{aligned}$$

To prove the second statement we first observe that any S -translation number τ of $f(x)$ belonging to ε is also an S -translation number of $\varphi_\delta(x)$ belonging to 2ε . For we have for every x_0

$$\begin{aligned} \int_{x_0}^{x_0+1} |\varphi_\delta(x+\tau) - \varphi_\delta(x)| dx &\leq \frac{1}{\delta} \int_{x_0}^{x_0+1} dx \int_x^{x+\delta} |f(\xi+\tau) - f(\xi)| d\xi \\ &\leq \frac{1}{\delta} \int_{x_0}^{x_0+1+\delta} |f(\xi+\tau) - f(\xi)| d\xi \int_{\xi-\delta}^{\xi} dx \leq \int_{x_0}^{x_0+2} |f(\xi+\tau) - f(\xi)| d\xi \leq 2\varepsilon. \end{aligned}$$

What we have to prove is that $D_S[f(x), \varphi_\delta(x)] < \varepsilon$ for $\delta < \delta_0(\varepsilon)$, i. e. that for every x_0

$$(15) \quad \int_{x_0}^{x_0+1} |f(x) - \varphi_\delta(x)| dx < \varepsilon \quad \text{for } \delta < \delta_0(\varepsilon).$$

Let $l = l\left(\frac{\varepsilon}{4}\right)$ be a number such that every interval of length l contains an S -translation number $\tau = \tau\left(\frac{\varepsilon}{4}\right)$ of $f(x)$ which (as we have just seen) is also an S -translation number, belonging to $\frac{\varepsilon}{2}$, of every $\varphi_\delta(x)$. Corresponding to the value of x_0 in (15) we select a translation number $\tau\left(\frac{\varepsilon}{4}\right)$ so that the point $x_0 + \tau$ lies in the interval $(0, l)$. Then

$$\begin{aligned} \int_{x_0}^{x_0+1} |f(x) - \varphi_\delta(x)| dx &\leq \int_{x_0}^{x_0+1} |f(x) - f(x+\tau)| dx \\ &+ \int_{x_0}^{x_0+1} |f(x+\tau) - \varphi_\delta(x+\tau)| dx + \int_{x_0}^{x_0+1} |\varphi_\delta(x+\tau) - \varphi_\delta(x)| dx \\ &\leq \frac{\varepsilon}{4} + \int_{x_0+\tau}^{x_0+\tau+1} |f(x) - \varphi_\delta(x)| dx + \frac{\varepsilon}{2} \leq \frac{3}{4}\varepsilon + \int_0^{l+1} |f(x) - \varphi_\delta(x)| dx. \end{aligned}$$

Thus the proof will be complete when we have established the following simple proposition, which we state as an

independent lemma since it may be of use in other problems involving the smoothing of an integrable function.

LEMMA. Let $f(x)$ be a function integrable (L) in a finite interval (a, b) , and let

$$\varphi_\delta(x) = \frac{1}{\delta} \int_x^{x+\delta} f(\xi) d\xi \quad (a < x < b - \delta).$$

Then, for any $\beta < b$,

$$\lim_{\delta \rightarrow 0} \int_a^\beta |f(x) - \varphi_\delta(x)| dx = 0.$$

We denote the number $b - \beta$ by d . Then $\varphi_\delta(x)$ is defined in the interval (a, β) for every $\delta \leq d$. By the theorem on the differentiation of a Lebesgue integral we have that $\varphi_\delta(x) \rightarrow f(x)$ as $\delta \rightarrow 0$ almost everywhere in (a, β) .

We first prove that to any $\varepsilon > 0$ corresponds an $\eta > 0$ such that

$$(16) \quad \int_E |\varphi_\delta(x)| dx < \varepsilon$$

for all values of $\delta (\leq d)$ and for every set $E \subset (a, \beta)$ such that $mE < \eta$. We write

$$\int_E |\varphi_\delta(x)| dx \leq \frac{1}{\delta} \int_E dx \int_x^{x+\delta} |f(\xi)| d\xi = \frac{1}{\delta} \iint_G |f(\xi)| dx d\xi$$

where G denotes the two-dimensional set of points in the x, ξ plane

$$x \in E, \quad x < \xi < x + \delta$$

whose measure

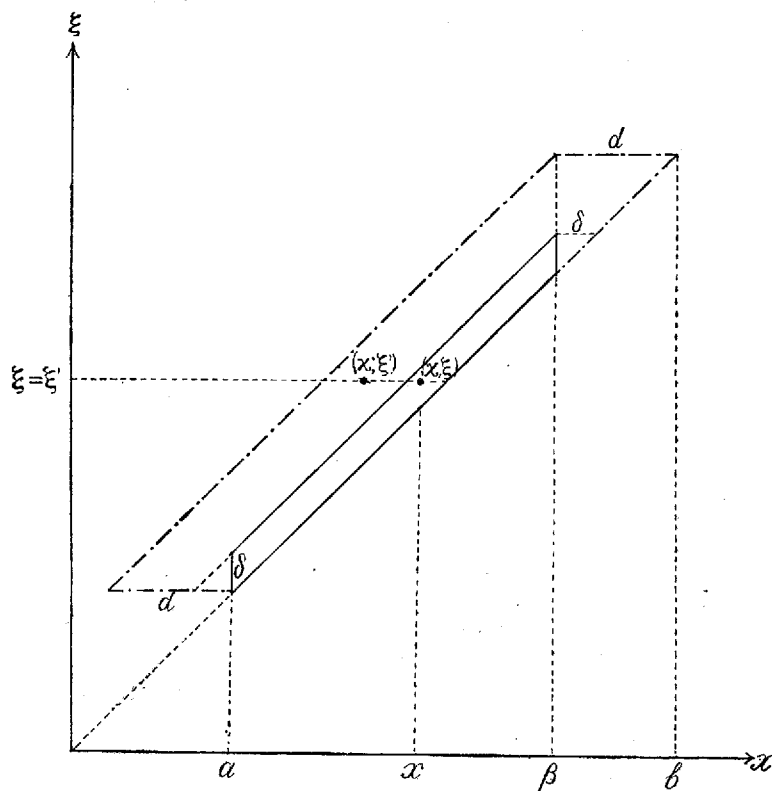
$$mG = \delta \cdot mE.$$

On making the simple transformation

$$\xi = \xi', \quad x' - \xi' = \frac{d}{\delta} (x - \xi)$$

the set G goes over to a set G' in the x', ξ' plane, of measure

$$(17) \quad mG' = \frac{d}{\delta} \cdot mG = d \cdot mE,$$



which evidently lies inside the constant (independent of δ) parallelogram P whose vertices are the points $(a-d, a)$, (a, a) , (b, b) , $(b-d, b)$. And

$$(18) \quad \frac{1}{\delta} \iint_G |f(\xi)| dx d\xi = \frac{1}{d} \iint_{G'} |f(\xi')| dx' d\xi'.$$

Since the function $f(\xi')$, regarded as a function of the two variables x', ξ' , is integrable in the parallelogram P ,

the integral on the right can be made less than ε by taking mE sufficiently small ($< \eta$), for by (17) this makes mG' "small".

(16) being established, the proof of the lemma may be completed in the usual way: We choose $\eta' \leq \eta$ so that $\int_E |f(x)| dx < \varepsilon$ for $mE < \eta'$, and then chose δ_0 so small that for $\delta < \delta_0$ the set $F = F_\delta$ of points of (a, β) at which $|f - \varphi_\delta| < \frac{\varepsilon}{\beta - a}$ is of measure $> \beta - a - \eta'$, and consequently the complementary set $C(F)$ of measure $< \eta'$. Then for $\delta < \delta_0$

$$\begin{aligned} \int_a^\beta |f(x) - \varphi_\delta(x)| dx &\leq \int_F |f(x) - \varphi_\delta(x)| dx + \int_{C(F)} |f(x)| dx \\ &\quad + \int_{C(F)} |\varphi_\delta(x)| dx < \frac{\varepsilon}{\beta - a} (\beta - a) + \varepsilon + \varepsilon = 3\varepsilon. \end{aligned}$$

When the identity of the class $C_S(A)$ with the class of S. a. p. functions has been established the next problem which arises is to find a simple algorithm which, applied to an arbitrary S. a. p. function $f(x)$, gives a sequence of finite sums $s_n(x)$ tending to $f(x)$ in the sense of the limit process characteristic of the class, i. e. such that

$$S\text{-}\lim s_n(x) = f(x), \quad \text{as } n \rightarrow \infty.$$

The above proof evidently provides a mean of constructing such a sequence. For the above functions $\varphi_\delta(x) = \frac{1}{\delta} \int_x^{x+\delta} f(\xi) d\xi$, being ordinary a. p. functions, can in accordance with § 1 be approached within an arbitrarily small U -distance (and a fortiori within an arbitrarily small S-

distance) by Fejér sums of their Fourier series. Therefore by first taking δ sufficiently small and then approximating sufficiently closely to $\varphi_\delta(x)$ by a Fejér sum, we shall obtain finite sums whose S -distance from $f(x)$ can be made arbitrarily small. This process can be regarded as an algorithm on the Fourier series of $f(x)$ itself, since the Fourier series of $\varphi_\delta(x)$ involved can be obtained at once from this series by a formal integration:

$$(19) \quad \varphi_\delta(x) \propto \sum A_\nu \frac{e^{iA_\nu \delta} - 1}{iA_\nu \delta} e^{iA_\nu x}$$

$\left(\frac{e^{iA_\nu \delta} - 1}{iA_\nu \delta} \text{ stands for } 1 \text{ when } A_\nu = 0 \right)$. This follows from the obvious equation

$$M \left\{ \frac{1}{\delta} \int_x^{x+\delta} f(\xi) d\xi \cdot e^{-i\lambda x} \right\} = M \{ f(x) e^{-i\lambda x} \} \frac{e^{i\lambda \delta} - 1}{i\lambda \delta}$$

already used by Bohr ([1], p. 62) for a. p. functions and by Stepanoff for S. a. p. functions.

The above algorithm for the construction of a sequence $s_n(x)$ is complicated by the presence of the parameter δ . This complication can be avoided; we shall show that *we may use as approximating sums the Fejér sums of $f(x)$ itself*, instead of the Fejér sums of the functions $\varphi_\delta(x)$.

For this purpose we first prove that the relation given in § 1 between the Fourier exponents and the translation numbers of an a. p. function holds also for an S. a. p. function; in other words that:

“If $\alpha_1, \alpha_2, \dots$ is a base of the Fourier exponents A_ν of an S. a. p. function $f(x)$; then to any $\varepsilon > 0$ correspond integers P, Q and a positive number $\delta < \pi$ such that every solution of the system of diophantine inequalities

$$\left| \frac{\alpha_1}{Q} t \right| < \delta \pmod{2\pi}, \dots, \left| \frac{\alpha_p}{Q} t \right| < \delta \pmod{2\pi}$$

is an S -translation number $\tau\left(\frac{\varepsilon}{2}\right)$ of $f(x)$ ".

We take a $\varphi_\delta(x)$ such that $D_S[f(x), \varphi_\delta(x)] < \frac{\varepsilon}{6}$. Then every S -translation number of $\varphi_\delta(x)$ (and a fortiori every U -translation number of $\varphi_\delta(x)$) belonging to $\frac{\varepsilon}{6}$ is also an S -translation number of $f(x)$ belonging to $\frac{\varepsilon}{2}$. We have now only to apply the theorem of § 1 (with $\frac{\varepsilon}{6}$ in place of $\frac{\varepsilon}{2}$) to the (ordinary) a. p. function $\varphi_\delta(x)$, which by (19) also has the sequence $\alpha_1, \alpha_2, \dots$ as a base for its Fourier exponents.

Using the notation of § 1 we write

$$\begin{aligned} s(x) &= M \left\{ f(x+t) \prod_{n_1} \left(\frac{\alpha_1}{N_1!} t \right) \dots \prod_{n_p} \left(\frac{\alpha_p}{N_p!} t \right) \right\} \\ &= \sum_{\substack{-n_1 \leq \nu_1 \leq n_1 \\ \dots \\ -n_p \leq \nu_p \leq n_p}} \left(1 - \frac{|\nu_1|}{n_1} \right) \dots \left(1 - \frac{|\nu_p|}{n_p} \right) A_\nu e^{i A_\nu x} \end{aligned}$$

where

$$A_\nu = \frac{\nu_1}{N_1!} \alpha_1 + \dots + \frac{\nu_p}{N_p!} \alpha_p$$

(and A_ν , as before, denotes zero when $\frac{\nu_1}{N_1!} \alpha_1 + \dots + \frac{\nu_p}{N_p!} \alpha_p$ is not one of the Fourier exponents of $f(x)$). We shall prove, just as in § 1, that

$$S\text{-lim } s(x) = f(x)$$

provided only $p, N_1, N_2, \dots, n_1, n_2, \dots$ tend to ∞ together in such a way that $\frac{n_1}{N_1!} \rightarrow \infty, \frac{n_2}{N_2!} \rightarrow \infty, \dots$

For any x at which $f(x)$ is defined (and so for almost all values of x) we have

$$s(x) - f(x) = M\{ (f(x+t) - f(x)) \Pi(t) \}$$

and therefore

$$|s(x) - f(x)| \leq M\{ |f(x+t) - f(x)| \Pi(t) \};$$

the mean value on the right exists, since $|f(x+t) - f(x)|$ is plainly an S. a. p. function of t . Thus for every x

$$(20) \quad \int_x^{x+1} |s(\xi) - f(\xi)| d\xi \leq \int_x^{x+1} M\{ |f(\xi+t) - f(\xi)| \Pi(t) \} d\xi.$$

We shall show that the last integral is less than or equal to

$$(21) \quad M\{ \Pi(t) \int_x^{x+1} |f(\xi+t) - f(\xi)| d\xi \};$$

(this mean value also exists since $\int_x^{x+1} |f(\xi+t) - f(\xi)| d\xi$ is even an ordinary a. p. function of t). We have

$$(22) \quad \left\{ \begin{aligned} & \frac{1}{2T} \int_{-T}^T \Pi(t) dt \int_x^{x+1} |f(\xi+t) - f(\xi)| d\xi \\ & = \int_x^{x+1} d\xi \frac{1}{2T} \int_{-T}^T |f(\xi+t) - f(\xi)| \Pi(t) dt. \end{aligned} \right.$$

As $T \rightarrow \infty$ the left side of (22) and therefore also the right side tends to (21). But the limit of the right hand side is, by a theorem of Fatou¹ greater than or equal to

$$\begin{aligned} & \int_x^{x+1} d\xi \lim \frac{1}{2T} \int_{-T}^T |f(\xi+t) - f(\xi)| \Pi(t) dt \\ & = \int_x^{x+1} d\xi M\{ |f(\xi+t) - f(\xi)| \Pi(t) \}. \end{aligned}$$

¹ If $f_n(x) \geq 0$ in (a, b) and $f_n(x) \rightarrow f(x)$ then

$$\lim \inf. \int_a^b f_n(x) dx \geq \int_a^b f(x) dx.$$

Thus by (20)

$$\int_x^{x+1} |s(\xi) - f(\xi)| d\xi \leq M \left\{ \Pi(t) \int_x^{x+1} |f(\xi+t) - f(\xi)| d\xi \right\}.$$

Dividing the t -axis, as in § 1, into two sets I_1, I_2 and writing $\Pi(t) = \Pi'(t) + \Pi''(t)$ as before we have

$$D_S[s(x), f(x)] \leq \bar{M} \left\{ D_S[f(x+t), f(x)] \Pi'(t) \right\} \\ + \bar{M} \left\{ D_S[f(x+t), f(x)] \Pi''(t) \right\}.$$

This inequality differs from the inequality (9) of § 1 only in having D_U replaced by D_S . Since further the distance

$$D_S[f(x+t), f(x)]$$

is $\leq \frac{\varepsilon}{2}$ in I_1 and $\leq 2K$ in I_2 (K is $D_S[f(x)]$), the rest of the proof is word for word the same as that of § 1.

In conclusion we may remark that the above "summation theorem" implies the "uniqueness theorem" (see Stepanoff) which states that *an S. a. p. function is uniquely determined by its Fourier series*. For if two functions are both S -limits of the same sequence (Fejér sums of the given Fourier series) their S -distance must be zero and consequently they are equivalent, i. e. equal almost everywhere.

MEMOIRS REFERRED TO.

A. BESICOVITCH:

- [1]. "On generalized almost periodic functions", Proc. London Math. Soc. (2), 25 (1926), p. 495—512.

S. BOCHNER:

- [1]. "Beiträge zur Theorie der fastperiodischen Funktionen", Math. Ann. 96 (1926), p. 119—147.
[2]. "Properties of Fourier series of almost periodic functions", Proc. London Math. Soc. (2), 26 (1927), p. 433—452.

H. BOHR:

- [1], [2], [3]. "Zur Theorie der fastperiodischen Funktionen" I, II, III, Acta Math. 45 (1924), p. 29—127, 46 (1925), p. 101—214, 47 (1926), p. 237—281.

W. STEPANOFF:

- [1]. "Über einige Verallgemeinerungen der fastperiodischen Funktionen", Math. Ann. 95 (1926), p. 473—498.

