On a recent generalisation of G. D. Birkhoff's ergodic theorem.

By FREDERICK RIESZ in Budapest.

I.

Quite recently, in a joint paper, N. DUNFORD and D. S. MILLER have given a most surprising generalisation of Birkhoff's famous theorem[1]). The purpose of the present paper is to work out a simple proof of the generalised theorem, running on the same lines as that one given by the author, a few years ago, for the original theorem[2]). The latter was based upon the following elementary lemma on numerical sequences.

Lemma A. *Given a real, finite sequence $a_1, a_2, \ldots, a_n$ and an integer $m \leqq n$, let us consider, if there are any, the sums $a_k + a_{k+1} + \ldots + a_l$ of positive value, formed of at most m successive terms of the given sequence. Then the sum of all the a_k figuring as beginning terms in one at least of the said sums, is itself positive.*

For the sake of completeness, let us recall the main lines of our argument. First, to prove the lemma, let a_{k_1} be the first of the "beginning" terms and $a_{k_1} + \ldots + a_{l_1}$ the shortest of the positive sums starting from a_{k_1}. Then all terms of this sum are figuring amongst our beginning terms. If not so, if, say, a_h, $k_1 < h \leqq l_1$, would not figure amongst them, then $a_h + \ldots + a_{l_1} \leqq 0$ and so $a_{k_1} + \ldots + a_{h-1} > 0$ and this sum were shorter than that running from a_{k_1} to a_{l_1}. The same procedure, applied on the remaining terms $a_{l_1+1}, \ldots$, leads to a sum $a_{k_2} + \ldots + a_{l_2} > 0$ and so on, until all the beginning terms have come in, which concludes the proof.

Next, consider a measurable set Ω, of finite or infinite measure, measure and integral being defined as by LEBESGUE or, more generally, with respect to a distribution of positive masses.

---

1) N. DUNFORD and D. S. MILLER, On the ergodic theorem, *Transactions American Math. Society*, **60** (1946), pp. 538—549.

2) F. RIESZ, Sur quelques problèmes de la théorie ergodique, *Matematikai és Fizikai Lapok*, 49 (1942), pp. 34—62 (in Hungarian, with an abstract in French); Sur la théorie ergodique, *Commentarii Math. Helvetici*, 17 (1944), pp. 221—239.

Let T be a mapping of Ω into itself, not necessarily one-to-one and let us suppose T to be measure-preserving in the sense that, for any measurable set e and its map Te, the set $T^{-1}(Te)$ of points P whose images belong to Te, is of the same measure as Te. Then starting witha summable function $f_1(P)$ and putting $f_k(P)=f_1(T^{k-1}P)$, Birkhoff's theorem states that *the arithmetical mean of* $f_1, \ldots, f_n$ *converges, almost everywhere* (briefly a. e.) *to a summable function* $\varphi(P)$, *invariant with respect to the mapping* T.

Let us remind that Birkhoff's original statement deals only with the one-to-one case, but there is no real difficulty in passing to the more general one. As to Ω, we restrict ourselves, in the first two sections, to the case of finite measure; in fact, our argument for the case of infinite measure seems, at first sight, not to apply immediately to the generalisation to follow, dealing with no measure-preserving mappings.

Birkhoff's original proof was based, in substance, upon the following fact, applied in a slightly different form and which he has established by means of an ingenious dismemberment of the intervening point sets, but which is also an easy consequence of Lemma A.

Lemma B[3]). *Let be* E *an invariant set of points* P *for which the least upper bound of the arithmetical means* $\frac{1}{n}(f_1(P)+\ldots+f_n(P))$ *is positive. Then*

$$\int_E f_1(P) \geqq 0. \tag{1}$$

Proof. Let $E^{(m)}$ be the set of points P of E for which one at least of the sums

$$\sum_1^l f_k(P) \qquad (l \leqq m)$$

is positive. The sets $E^{(m)}$, increasing with m, are finally filling up the set E. So, instead of (1), it suffices to verify the similar inequality

$$\int_{E^{(m)}} f_1(P) \geqq 0. \tag{2}$$

To this effect, apply lemma A to the finite sequence $f_1(P), \ldots, f_{n+m}(P)$ (with $n+m$ in place of n), forming for each P the sum of the "beginning" terms. This sum being $\geqq 0$, the same is true for its integral and so

$$\sum_1^{n+m} \int_{E_k} f_k(P) \geqq 0, \tag{3}$$

[3]) As a matter of fact, this lemma, due to Yosida and Kakutani and called by its authors the "maximal ergodic theorem", differs from that of Birkhoff only that it deals with the least upper bound in place of the upper limit.

where E_k means the set of points P for which $f_k(P)$ is a beginning term.

Again, obviously, for $k \leqq n$ we have $TE_k = E_{k-1}$, $E_k = T^{-1}E_{k-1}$ and as T preserves the measure and so the integral, the first n integrals in (3) are equal and as $E_1 = E^{(m)}$, their common value is the same which figures in (2). On the other hand, the last m integrals in (3) are dominated by the integral of $|f_1(P)|$ over Ω. Therefore, from (3),

$$n\int_{E^{(m)}} f_1(P) + m\int_{\Omega} |f_1(P)| \geqq 0$$

whence, m being fixed, (2) follows as $n \to \infty$. So Lemma B is proved.

Frome here on, the proof of Birkhoff's theorem runs on the usual line. For any pair of values α, β where $\alpha > \beta$, consider the set $E_{\alpha\beta}$ of all points for which

$$\overline{\lim} \frac{1}{n} \sum_1^n f_k(P) > \alpha$$

and, at the same time,

$$\underline{\lim} \frac{1}{n} \sum_1^n f_k(P) < \beta.$$

The set $E_{\alpha\beta}$ is manifestly invariant with respect to T, so that it may play the part of E; moreover, applying Lemma B to the functions $f_1(P) - \alpha$ and $\beta - f_1(P)$ in place of $f_1(P)$, respectively, the corresponding set E is identical with $E_{\alpha\beta}$. So, by Lemma B,

$$\int_{E_{\alpha\beta}} (f_1(P) - \alpha) \geqq 0, \qquad \int_{E_{\alpha\beta}} (\beta - f_1(P)) \geqq 0$$

and adding, we have

$$\int_{E_{\alpha\beta}} (\beta - \alpha) \geqq 0.$$

As $\alpha > \beta$, this means that $E_{\alpha\beta}$ has to be of measure 0. Finally let α, β run over all rational pairs; then $\sum E_{\alpha\beta}$, sum of a denumerable sequence of nullsets, is itself a nullset and so the limit $\varphi(P)$ of the above mean exists a. e. Since, moreover,

$$\int_{\Omega} \left| \frac{1}{n} \sum_1^n f_k(P) \right| \leqq \frac{1}{n} \int_{\Omega} \sum_1^n |f_k(P)| = \int_{\Omega} |f_1(P)|,$$

the limit $\varphi(P)$ is summable on account of Fatou's theorem and so finite a. e.

The invariance of $\varphi(P)$ is, as well known, an immediate consequence of the equation

$$T\left[\frac{1}{n}(f_1 + \ldots + f_n)\right] = \frac{n+1}{n}\left[\frac{1}{n+1}(f_1 + \ldots + f_{n+1})\right] - \frac{f_1}{n}.$$

Thus the proof is concluded. To say a few words about the "integral" form of the theorem in which the iterated mappings T^k are replaced by a group T_t depending upon a continuous parameter t, it is well known how by means of a simple artifice, due to E. HOPF and KHINTCHINE, the corresponding problem may be reduced to the discrete case. It is, however, not without interest to observe that the above argument may be adapted so as to lead directly to the "integral" formulation. One has only to use, instead of Lemma A, the following lemma, equally easy to prove. *Given, for $a \leqq t \leqq b$, a summable function $g(t)$ and a length d, $0 < d \leqq b-a$, let be e the set of points t_0, if there are any, for which there exists an $h < d$ such that the integral of $g(t)$ from t_0 to t_0+h is positive. Then the integral of $g(t)$ over the set e is* $\geqq 0$. This lemma is but a corollary of the following when applied to the integral $G(t)$ of $g(t)$.

Let $G(t)$ be continuous for $a \leqq t \leqq b$ and for a given d, $0 < d \leqq b-a$, consider the set e of all interior points t for which there exists a t', $t < t' < t+d$, so that $G(t) < G(t')$. Then e is an open set, composed of intervals (a_k, b_k) for each of which $G(a_k) \leqq G(b_k)$.

This form of our lemma is but a slight modification of another, used by the author in 1932 to prove the differentiability a. e. of monotone functions as well as a much important inequality of HARDY and LITTLEWOOD[4]). Its proof runs on the same lines.

II.

Now let us turn over to our proper subject, dealing with the generalisation of Birkhoff's theorem by DUNFORD and MILLER. Instead of supposing the mapping to be measure-preserving, they put *the more general hypothesis that, for a fixed constant K and for any measurable set e,*

$$\frac{1}{n}\sum_{1}^{n}|T^{-k}e| \leqq K|e| \tag{4}$$

where $|e|$ denotes the measure of e. *From there, at least when Ω is of finite measure, they draw the same conclusion as* BIRKHOFF. In fact, they prove that the said conclusion i. e. the actual, pointwise convergence a. e. of the arithmetical mean of $f_1, \ldots, f_n$ is a consequence of a sort of mean convergence and that the latter is equivalent to the hypothesis (4).

---

[4]) F. RIESZ, Sur l'existence de la dérivée des fonctions d'une variable réelle et des fonctions d'intervalle, *Verhandlungen des internationalen Math. Kongresses Zürich* 1932, vol. 1, pp. 258—269; Sur un théorème de maximum de MM. Hardy et Littlewood, *Journal London Math. Society*, 7 (1932), pp. 1—13. See also G. H. HARDY, J. E. LITTLEWOOD and G. PÓLYA, *Inequalities* (Cambridge, 1934), p. 293.

In the present paper, I wish to show how the same result may be obtained by means of a slight modification of our above argument.

To tell the whole truth, DUNFORD and MILLER are going farther; in fact, they establish the same conclusion for a group of mappings depending on several parameters, like it has been done by DUNFORD for von Neumann's mean ergodic theorem and by N. WIENER for Birkhoff's theorem[5]). The present argument seems not to apply to this more general case. So we have to restrict ourselves to the case of a single T and its iterates or to that of a group T_t depending on a single parameter t. We content ourselves to expose our argument for the former case, letting the latter to the reader.

In addition to Lemma A, we need another lemma, dealing with infinite numerical sequences.

Lemma C. *Let $a_1, a_2, \ldots$ be an infinite sequence of positive terms and assume that for a fixed constant K,*

$$\frac{1}{n-h}(a_{h+1}+\ldots+a_n)\leqq Ka_h \tag{5}$$

for any n and any $h<n$. Then $a_n=o(n)$.

(Observe that the evaluation $a_n=O(n)$ is obvious but what we need is the fact that $a_n=o(n)$.)

Proof. Adding $K(a_{h+1}+\ldots+a_n)$ to both sides of (5) and dividing by K, we get

$$a_h+\ldots+a_n\geqq\left(1+\frac{1}{(n-h)K}\right)(a_{h+1}+\ldots+a_n).$$

Putting $h=2,3,\ldots,n-1$ and multiplying, we have

$$a_2+\ldots+a_n\geqq a_n\prod_{h=2}^{n-1}\left(1+\frac{1}{(n-h)K}\right).$$

Confronting this with the case $h=1$ of (5) we get

$$a_n\leqq\frac{(n-1)Ka_1}{\prod\limits_{h=2}^{n-1}\left(1+\frac{1}{(n-h)K}\right)}\leqq\frac{(n-1)K^2a_1}{1+\frac{1}{2}+\ldots+\frac{1}{n-2}}=O\left(\frac{n}{\log n}\right)=o(n)$$

and so Lemma C is proved.

Let us apply it to

$$a_n=\int\limits_{\Omega}|f_n(P)|.$$

For these a_n, the validity of (5) follows at once from our hypothesis (4)

---

[5]) N. DUNFORD, A mean ergodic theorem, *Duke Math. Journal*, 5 (1939), pp. 635–646; N. WIENER, The ergodic theorem, *ibidem*, pp. 1—18.

and so when this hypothesis is fulfilled then

$$\int_{\Omega} |f_n(P)| = o(n).$$

Now we are in a position to ask what Lemma B goes over into under the broader hypothesis. From the above proof of the lemma let us preserve the meaning of the sets $E, E^{(m)}$ and E_k, and nothing in our argument has to be changed until we come to the inequality (3). From here on, writing (3) in the form

$$\sum_1^n \int_{E_k} f_k(P) + \sum_{k=1}^m \int_{E_{n+k}} f_{n+k}(P) = \Sigma' + \Sigma'' \geqq 0,$$

consider first the behaviour of Σ' when m increases. Then the set $E_1 = E^{(m)}$ is also increasing and it is going over into E for $m \to \infty$. So, for m sufficiently large, the integral of $|f_1(P)|$ over $E - E_1$ is less than an arbitrary fixed $\varepsilon > 0$. By hypothesis (4), as for $k = 1, 2, \ldots, n$, $E - E_k = T^{1-k}(E - E_1)$, it follows immediately that

$$\sum_2^n \int_{E-E_k} |f_k(P)| \leqq (n-1) K \int_{E-E_1} |f_1(P)| < n K \varepsilon.$$

Therefore Σ' differs from

$$\sum_1^n \int_E f_k(P)$$

by less than $\varepsilon(1 + nK)$. Having now fixed ε and so m too, let n go to infinity; then the terms of Σ'' the absolute values of which are dominated by the corresponding integrals of $|f_k|$ over Ω, are, by Lemma C, of order $o(n+k) = o(n+m) = o(n)$ and so $\Sigma'' = m o(n) = o(n)$. Summing up,

$$\frac{1}{n} \sum_1^n \int_E f_k(P) \geqq -\varepsilon \frac{1+nK}{n} + o(1)$$

where $\varepsilon > 0$ is arbitrary and so finally

$$\underline{\lim} \frac{1}{n} \sum_1^n \int_E f_k(P) \geqq 0. \tag{6}$$

This is the result that has to take over the part played by Lemma B. In fact, let as above, $E_{\alpha\beta}$ be the set for which the upper limit of the arithmetical mean of $f_1(P), \ldots, f_n(P)$ is $> \alpha$ and its lower limit $< \beta$; then by replacing $f_1(P)$ by $f_1(P) - \alpha$ and by $\beta - f_1(P)$ resp. like above, and the set $E_{\alpha\beta}$, evidently invariant with respect to T, taking over the part of Ω, the same set will play, in both cases, also the part of E and so, by (6),

$$\underline{\lim}\frac{1}{n}\sum_{1}^{n}\int_{E_{\alpha\beta}}(f_k(P)-\alpha)\geqq 0,\quad \underline{\lim}\frac{1}{n}\sum_{1}^{n}\int_{E_{\alpha\beta}}(\beta-f_k(P))\geqq 0$$

and hence, adding,

$$(\beta-\alpha)|E_{\alpha\beta}|\geqq 0.$$

As $\alpha>\beta$, this implies that $|E_{\alpha\beta}|=0$.

The proof concludes like the above one of Birkhoff's theorem.

III.

Finally, I should like to say a few words about the case when Ω is *of infinite measure*. As we know, Birkhoff's theorem holds true in this case; to prove it, we only have, before integrating $f_1(P)-\alpha$ over $E_{\alpha\beta}$, to ascertain that the constant is a summable function i. e. that $E_{\alpha\beta}$ is of finite measure. Let us assume that $\alpha>0$; if not, then $\beta<0$ and we had to reason on $\beta-f_1(P)=-f_1(P)-(-\beta)$ instead of $f_1(P)-\alpha$.

Let E' be a subset of $E_{\alpha\beta}$, of finite measure, but otherwise arbitrary. Let $e_1(P)$ be its characteristic function. Apply Lemma B to the function $g_1(P)=f_1(P)-\alpha e_1(P)$ in place of $f_1(P)$; then, for the set E corresponding to $g_1(P)$, we have

$$\int_E g_1(P)\geqq 0 \tag{7}$$

and so, as manifestly $E'\subset E$,

$$\int_E f_1(P)\geqq \alpha\int_E e_1(P)=\alpha|E'|. \tag{8}$$

Therefore

$$|E'|\leqq\frac{1}{\alpha}\int_\Omega|f_1(P)|;$$

thus, for all subsets E' of finite measure of $E_{\alpha\beta}$, a common bound is established; therefore $E_{\alpha\beta}$ itself is of finite measure.

Now, after having recalled the argument for Birkhoff's problem, we try to extend it to the more general one, dealt with in Section II. There the integral in (7) has to be replaced by

$$\underline{\lim}\frac{1}{n}\sum_{1}^{n}g_k(P)$$

and so, by (8),

$$\underline{\lim}\frac{1}{n}\int_E\sum_{1}^{n}f_k(P)\geqq\alpha\ \underline{\lim}\frac{1}{n}\sum_{1}^{n}\int_E e_k(P)= \tag{9}$$

$$=\alpha\ \underline{\lim}\frac{1}{n}\left\{|E'|+\sum_{1}^{n-1}|T^{-k}E'|\right\}.$$

By our hypothesis (4), the mean value on the left, and therefore its lower limit too, are dominated by

$$\frac{1+(n-1)K}{n}\int_{\Omega}|f_1(P)|\leqq C\int_{\Omega}|f_1(P)|$$

where C is independent of n. On the other hand, however, if we want to get an upper estimate of $|E'|$, we cannot do, as it seems, without an additional hypothesis. The simplest one that suggests itself, is *the counterpart of hypothesis* (4), *namely that for a certain constant* $K'>0$ *and for any measurable set* e

$$\frac{1}{n}\sum_{1}^{n}|T^{-k}e|\geqq K'|e|.$$

In fact, assuming the latter hypothesis, the right hand side of (9) is greater than $C_1|E'|$ where $C_1>0$ is independent of n. So finally

$$|E'|\leqq\frac{C}{C_1}\int_{\Omega}|f_1(P)|$$

and from there the finiteness of $|E'_{\alpha\beta}|$ follows like above in Birkhoff's case.

(Received June 20, 1947.)