THE ERGODIC THEOREM

By Norbert Wiener

1. Ergodic theory has its roots in statistical mechanics. Both in the older Maxwell theory and in the later theory of Gibbs, it is necessary to make some sort of logical transition between the average behavior of all dynamical systems of a given family or ensemble, and the historical average behavior of a single system. This transition was not carried out with any rigor until the theory of Lebesgue measure had been developed. The fundamental theorems are those due to Koopman, von Neumann, and Carleman, on the one hand, and to Birkhoff, on the other. A careful account of them and their proofs is to be found in Eberhard Hopf's *Ergodentheorie* (Berlin, 1937) in the series *Ergebnisse der Mathematik und ihrer Grenzgebiete.* The bibliography of that monograph is so complete that it relieves me from all need of furnishing one on my own account. It is important to point out, however, that Birkhoff's first paper (Proc. Nat. Acad. Sci. (1931)) contains Theorem IV of this paper.

The fundamental theorems of ergodic theory, which emerged in the epoch 1931–32, are recognized as theorems pertaining in the first instance to the abstract theory of the Lebesgue integral. Much of this paper will be devoted to new proofs of these known results and to their proper orientation mutually and with respect to the rest of analysis. They have several variant forms, but perhaps the simplest form of the von Neumann theorem reads:

Theorem I. *Let S be a measurable set of points of finite measure. Let T be a transformation of S into itself, which transforms every measurable subset of S into a set of equal measure, and whose inverse has the same property. Let $f(P)$ be a function defined over S and of Lebesgue class L^2. Then there exists a function $f_1(P)$, also belonging to L^2 and such that*

$$\lim_{N\to\infty} \int_S \left| f_1(P) - \frac{1}{N+1} \sum_{n=0}^{N} f(T^n P) \right|^2 dV_P = 0. \tag{1.01}$$

Birkhoff's theorem reads:

Theorem II. *Let S and T be as in Theorem* I. *Let $f(P)$ be a function defined over S and of Lebesgue class L. Then, except for a set of points P of zero measure,*

$$f_1(P) = \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} f(T^n P) \tag{1.02}$$

will exist and belong to L.

These theorems have continuous analogues. These are, respectively:

Received December 7, 1938; an invited address before the American Mathematical Society at its meeting held April 7–8 in conjunction with the centennial celebration of Duke University.