The Homogeneous Chaos
Author(s): Norbert Wiener
Source: *American Journal of Mathematics*, Vol. 60, No. 4 (Oct., 1938), pp. 897-936
Published by: The Johns Hopkins University Press
Stable URL: http://www.jstor.org/stable/2371268
Accessed: 03/03/2009 11:31

THE HOMOGENEOUS CHAOS.

By Norbert Wiener.

1. Introduction. Physical need for theory. Statistical mechanics may be defined as the application of the concepts of Lebesgue integration to mechanics. Historically, this is perhaps putting the cart before the horse. Statistical mechanics developed through the entire latter half of the nineteenth century before the Lebesgue integral was discovered. Nevertheless, it developed without an adequate armory of concepts and mathematical technique, which is only now in the process of development at the hands of the modern school of students of integral theory.

In the more primitive forms of statistical mechanics, the integration or summation was taken over the manifold particles of a single homogeneous dynamical system, as in the case of the perfect gas. In its more mature form, due to Gibbs, the integration is performed over a parameter of distribution, numerical or not, serving to label the constituent systems of a dynamical ensemble, evolving under identical laws of force, but differing in their initial conditions. Nevertheless, the study of the mode in which this parameter of distribution enters into the individual systems of the ensemble does not seem to have received much explicit study. The parameter of distribution is essentially a parameter of integration only. As such, questions of dimensionality are indifferent to it, and it may be replaced by a numerical variable α with the range $(0, 1)$. Any transformation leaving invariant the probability properties of the ensemble as a whole is then represented by a measure-preserving transformation of the interval $(0, 1)$ into itself.

Among the simplest and most important ensembles of physics are those which have a spatially homogeneous character. Among these are the homogeneous gas, the homogeneous liquid, the homogeneous state of turbulence. In these, while the individual systems may not be invariant under a change of origin, or, in other words, under the translation of space by a vector, the ensemble as a whole is invariant, and the individual systems are merely permuted without change of probability. From what we have said, the translations of space thus generate an Abelian group of equi-measure transformations of the parameter of distribution.

One-dimensional groups of equi-measure transformations have become well known to the mathematicians during the past decade, as they lie at the root of Birkhoff's famous ergodic theorem.[1] This theorem asserts that if we have given a set S of finite measure, an integrable function $f(P)$ on S, and a one-parameter Abelian group T^λ of equi-measure transformations of S into itself, such that

$$(1) \qquad T^\lambda \cdot T^\mu = T^{\lambda+\mu} \qquad (-\infty < \lambda < \infty, \; -\infty < \mu < \infty),$$

then for all points P on S except those of a set of zero measure, and provided certain conditions of measurability are satisfied,

$$(2) \qquad \lim_{A\to\infty} \frac{1}{A} \int_0^A f(T^\lambda P)\, d\lambda$$

will exist. Under certain more stringent conditions, known as metric transitivity, we shall have

$$(3) \qquad \lim_{A\to\infty} \frac{1}{A} \int_0^A f(T^\lambda P)\, d\lambda = \int_S f(P)\, dV_P$$

almost everywhere. The ergodic theorem thus translates averages over an infinite range, taken with respect to λ, into averages over the set S of finite measure. Even without metric transitivity, the ergodic theorem translates the distribution theory of λ averages into the theory of S averages.

In the most familiar applications of the ergodic theorem, S is taken to be a spatial set, and the parameter λ is identified with the time. The theorem thus becomes a way of translating time averages into space averages, in a

[1] Cf. Eberhard Hopf, "Ergodentheorie," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 5. See particularly § 14, where further references to the literature are given.

manner which was postulated by Gibbs without rigorous justification, and which forms the entire basis of his methods. Strictly speaking, the space averages are generally in phase-space rather than in the ordinary geometrical space of three dimensions. There is no reason, however, why the parameter λ should be confined to one taking on values of the time for its arguments, nor even why it should be a one-dimensional variable. We are thus driven to formulate and prove a multidimensional analogue of the classical Birkhoff theorem.

In the ordinary Birkhoff theorem, the transformations T^λ are taken to be one-one point transformations. Now, the ergodic theorem belongs fundamentally to the abstract theory of the Lebesgue integral, and in this theory, individual points play no rôle. In the study of chaos, individual values of the parameter of integration are equally unnatural as an object of study, and it becomes desirable to recast the ergodic theorem into a true Lebesgue form. This we do in paragraph 2.

Of all the forms of chaos occurring in physics, there is only one class which has been studied with anything approaching completeness. This is the class of types of chaos connected with the theory of the Brownian motion. In this one-dimensional theory, there is a simple and powerful algorithm of phase averages, which the ergodic theorem readily converts into a theory of averages over the transformation group. This theory is easily generalized to spaces of a higher dimensionality, without any very fundamental alterations. We shall show that there is a certain sense in which these types of chaos are central in the theory, and allow us to approximate to all types.

Physical theories of chaos, such as that of turbulence, or of the statistical theory of a gas or a liquid, may or may not be theories of equilibrium. In the general case, the statistical state of a chaotic system, subject to the laws of dynamics, will be a function of the time. The laws of dynamics produce a continuous transformation group, in which the chaos remains a chaos, but changes its character. This is at least the case in those systems which can continue to exist indefinitely in time without some catastrophe which essentially changes their dynamic character. The study, for example, of the development of a state of turbulence, depends on an existence theory which avoids the possibility of such a catastrophe. We shall close this paper by certain very general considerations concerning the demands of an existence theory of this sort.

2. Definition. Types of Chaos. A *continuous homogeneous chaos* in n dimensions is a scalar or vector valued measurable function $\rho(x_1, \cdots, x_n; \alpha)$ of $x_1, \cdots, x_n; \alpha$, in which $x_1, \cdots, x_n$ assume all real values, while α ranges over $(0, 1)$; and in which the set of values of α for which

(4) $$\rho(x_1 + y_1, \cdots, x_n + y_n; \alpha) \text{ belongs to } S,$$

if it has a measure for any set of values of $y_1, \cdots, y_n$, has the same measure for any other set of values. In this paper, we shall confine our attention to scalar chaoses. A continuous homogeneous chaos is said to be *metrically transitive*, if whenever the sets of values of α for which $\rho(x_1, \cdots, x_n; \alpha)$ belongs to S and to S_1, respectively, have measures M and M_1, the set of values of α for which simultaneously

$$\rho(x_1, \cdots, x_n; \alpha) \text{ belongs to } S$$

and

$$\rho(x_1 + y_1, \cdots, x_n + y_n; \alpha) \text{ belongs to } S_1$$

has a measure which tends to M_1 as $y_1^2 + \cdots + y_n^2 \to \infty$.

If ρ is integrable, it determines the additive set-function

(5) $$\mathfrak{F}(\Sigma; \alpha) = \int \cdot \underset{\Sigma}{\cdot} \cdot \int \rho(x_1, \cdots, x_n; \alpha) dx_1 \cdots dx_n.$$

On the other hand, not every additive set-function may be so defined. This suggests a more general definition of a *homogeneous chaos*, in which the chaos is defined to be a function $\mathfrak{F}(\Sigma; \alpha)$, where α ranges over $(0, 1)$ and Σ belongs to some additively closed set Ξ of measurable sets of points in n-space. We suppose that if Σ and Σ_1 do not overlap,

(6) $$\mathfrak{F}(\Sigma + \Sigma_1; \alpha) = \mathfrak{F}(\Sigma; \alpha) + \mathfrak{F}(\Sigma_1; \alpha).$$

We now define the new point-set $\Sigma(y_1, \cdots, y_n)$ by the assertion

(7) $\Sigma(y_1, \cdots, y_n)$ contains $x_1 + y_1, \cdots, x_n + y_n$ when and only when Σ contains $x_1, \cdots, x_n$.

This leads to the definition of the additive set-function $\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)$ by

(8) $$\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha) = \mathfrak{F}(\Sigma(y_1, \cdots, y_n); \alpha).$$

If, then, for all classes S of real numbers,

(9) Measure of set of α's for which $\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)$ belongs to class S

is independent of $y_1, \cdots, y_n$, in the sense that if it exists for one set of these numbers, it exists for all sets, and has the same value, and if it is measurable in $y_1, \cdots, y_n$, we shall call $\mathfrak{F}$ a homogeneous chaos. The notion of metrical transitivity is generalized in the obvious way, replacing $\rho(x_1, \cdots, x_n; \alpha)$ by $\mathfrak{F}(\Sigma; \alpha)$, and $\rho(x_1 + y_1, \cdots, x_n + y_n; \alpha)$ by $\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)$.

The theorem which we wish to prove is the following:

THEOREM I. *Let $\mathfrak{F}(\Sigma; \alpha)$ be a homogeneous chaos. Let the functional*

$$\Phi\{\mathfrak{F}(\Sigma; \alpha)\} = g(\alpha) \tag{10}$$

be a measurable function of α, such that $\int_0^1 | g(\alpha) \log^+ | g(\alpha) | | d\alpha$ is finite. Then for almost all values of α,

$$\lim_{r\to\infty} \frac{1}{V(r)} \int \cdots \int_R \Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\} dy_1 \cdots dy_n \tag{11}$$

exists, where R is the interior of the sphere

$$y_1^2 + y_2^2 + \cdots + y_n^2 = r^2 \tag{12}$$

and $V(r)$ is its volume. If in addition, $\mathfrak{F}(\Sigma; \alpha)$ is metrically transitive,

$$\lim_{r\to\infty} \frac{1}{V(r)} \int \cdots \int_R \Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\} dy_1 \cdots dy_n = \int_0^1 \Phi\{\mathfrak{F}(\Sigma; \beta)\} d\beta \tag{13}$$

for almost all values of α.

3. Classical ergodic theorem. Lebesgue form. Theorem I is manifestly a theorem of the ergodic type. Let it be noticed, however, that we nowhere assume that the transformation of α given by $\beta = T\alpha$ when

$$\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \beta) = \mathfrak{F}(\Sigma; \alpha) \tag{14}$$

is one-one. This should not be surprising, as the ergodic theorem is fundamentally one concerning the Lebesgue integral, and in the theory of the Lebesgue integral, individual points play no rôle.

Nevertheless, in the usual formulation of the ergodic theorem, the expression $f(T^\lambda P)$ enters in an essential way. Can we give this a meaning without introducing the individual transform of an individual point?

The clue to this lies in the definition of the Lebesgue integral itself. If $f(P)$ is to be integrated over a region S, we divide S into the regions $S_{a,b}(f)$, defined by the condition that over such a region,

$$a < f(P) \leq b. \tag{15}$$

We now write

$$\int_S f(P) dV_P = \lim_{\epsilon\to 0} \sum_{-\infty}^{\infty} n\,\epsilon\, m(S_{(n-1)\epsilon, n\epsilon}(f)). \tag{16}$$

The condition that $f(P)$ be integrable thus implies the condition that it be measurable, or that all the sets $S_{a,b}(f)$ be measurable.

Now, if T is a measure-preserving transformation on S, the sets $TS_{a,b}(f)$

will all be measurable, and will have, respectively, the same measures as the sets $S_{a,b}(f)$. We shall define the function $f(TP) = g(P)$ by the conditions

$$TS_{a,b}(f) = S_{a,b}(g). \tag{17}$$

If T conserves relations of inclusion of sets, up to sets of zero measure, this function will clearly be defined up to a set of values of P of zero measure, and we shall have

$$\int_S f(TP)\,dV_P = \int_S f(P)\,dV_P. \tag{18}$$

We may thus formulate the original or discrete case of the Birkhoff ergodic theorem, as follows: *Let S be a set of points of finite measure. Let T be a transformation of all measurable sub-sets of S into measurable sub-sets of S, which conserves measure, and the relation between two sets, that one contains the other except at most for a set of zero measure. Then except for a set of points P of zero measure,*

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_0^N f(T^n P) \tag{19}$$

will exist.

The continuous analogue of this theorem needs to be formulated in a somewhat more restricted form, owing to the need of providing for the integrability of the functions concerned. It reads: *Let S be a set of points of finite measure. Let T^λ be a group of transformations fulfilling the conditions we have laid down for T in the discrete case just mentioned. Let $T^\lambda P$ be measurable in the product space of λ and of P. Then, except for a set of points P of zero measure,*

$$\lim_{A\to\infty} \frac{1}{A} \int_0^A f(T^\lambda P)\,d\lambda \tag{20}$$

will exist.

In the proofs of Birkhoff's ergodic theorem, as given by Khintchine and Hopf, no actual use is made of the fact that the transformation T is one-one, and the proofs extend to our theorem as stated here, without any change. The restriction of measurability, or something to take its place, is really necessary for the correct formulation of Khintchine's statement of the ergodic theorem, as Rademacher, von Neumann, and others have already pointed out.[2]

With the aid of the proper Lebesgue formulation of the ergodic theorem, the one-dimensional case of Theorem I follows at once. Actually more follows, as it is only necessary that g belong to L, instead of to the logarithmic class

[2] Cf. J. v. Neumann, *Annals of Mathematics*, 2, vol. 33, p. 589, note 11.

with which we replace it. To prove Theorem I in its full generality, we must establish a multidimensional ergodic theorem.

4. Dominated ergodic theorem. Multidimensional ergodic theorem. As a lemma to the multidimensional ergodic theorem, we first wish to establish the fact that if the function $f(P)$ in the ergodic theorem satisfies the condition

$$\int_S |f(P)| \log^+ |f(P)| \, dV_P < \infty, \tag{21}$$

then the expressions (19) and (20) not only exist, but the limits in question will be approached under the domination of a summable function of P. We shall prove this in the discrete case, for the sake of simplicity, but the result goes over without difficulty to the continuous case.

Let T be an equimeasure transformation of the set S of finite measure into itself, in the generalized sense of the last paragraph, and let W be a measurable sub-set of S, with the characteristic function $W(P)$. Let U be the set of all points P for which some $T^{-j}P$ belongs to W. Let $i(P)$, when P belongs to W, be defined as the smallest positive number n such that $T^{-n}P$ belongs to W, and let us call it the *index* of P. Every point of W, except for a set of measure 0, will have a finite index, since if we write W_∞ for the set of points without a finite index, no two sets $T^m W_\infty$ and $T^n W_\infty$ can overlap, while they all have the same measure and their sum has a finite measure. Thus except for a set of zero measure, we may divide W into the sets W_p, each consisting of the points of W of index p. It is easy to show that if $1 \leq p < \infty$, $1 \leq p' < \infty$, $0 \leq j < p$, $0 \leq j' < p'$, the sets $T^{-j}W_p$ and $T^{-j'}W_{p'}$ can not overlap over a set of positive measure unless $j = j'$, $p = p'$. Similarly, the sets $T^{-p}W_p$ and $T^{-p'}W_{p'}$ can not overlap over sets of positive measure, unless $p = p'$, and represent a dissection of W, except for a set of zero measure. Let us put W_{pq} for the logical product of W_p and $T^p W_q$; W_{pqr} for the logical product of W_{pq} and $T^{p+q}W_r$; and so on. Then if k is fixed, the sets $T^{-j}W_{p_1 p_2 \ldots p_k}$ cover U (except for sets of zero measure) once as j goes from 0 to p_1, once as it goes from p_1 to $p_1 + p_2$, and so on; making just k times between 0, inclusive, and $p_1 + p_2 + \cdots + p_k$, exclusive. Thus if S_K is the set of all the points in all the $W_{p_1 p_2 \ldots p_k}$ for which $p_1 + p_2 + \cdots + p_k = K$, the total measure of all the S_K's for $2^{N+1} \geq K \geq 2^N$ can not exceed $km(S)/2^N$.

Now let P lie in $T^{-j}W_{p_1 p_2 \ldots p_k}$, where $0 \leq j < p_1 + p_2 + \cdots + p_k$. Let us consider the sequence of numbers a_j, where if $p_1 + \cdots + p_l \leq j < p_1 + \cdots + p_{l+1}$, a_j is the greatest of the numbers

$$\frac{1}{j - p_1 - \cdots - p_l + 1}, \quad \frac{2}{j - p_1 - \cdots - p_{l-1} + 1},$$
$$\frac{3}{j - p_1 - \cdots - p_{l-2} + 1}, \cdots, \frac{l+1}{j+1}.$$

Then a_j will be the largest of the numbers

$$W(P), \frac{W(P)+W(TP)}{2}, \cdots, \frac{W(P)+W(TP)+\cdots+W(T^jP)}{j+1}.$$

The sum $\sum_0^k a_j$, for a fixed K, will have its maximum value when $p_1 = p_2 = \cdots = p_{k-1} = 1$; $p_k = K + 1 - k$, when it will be $k + \sum_{k+1}^{k} k/j \leq k(1 + \log K/k)$. In this case the sequence of the a_j's will be

$$\underbrace{1, 1, \cdots, 1,}_{k \text{ times}} \frac{k}{k+1}, \frac{k}{k+2}, \cdots, \frac{k}{K}.$$

This remark will be an easy consequence of the following fact: let us consider the sequence

$$(22) \qquad \cdots, \frac{\lambda}{n-1}, \frac{\lambda}{n}, 1, \frac{1}{2}, \cdots, \frac{1}{\left[\frac{n}{\lambda}\right]}, \frac{\lambda+1}{\left[\frac{n}{\lambda}\right]+n+1}, \cdots$$

and the modified sequence

$$(23) \qquad \cdots, \frac{\lambda}{n-1}, 1, \frac{1}{2}, \cdots, \frac{1}{\left[\frac{n-1}{\lambda}\right]}, \frac{\lambda+1}{\left[\frac{n-1}{\lambda}\right]+n}, \frac{\lambda+1}{\left[\frac{n-1}{\lambda}\right]+n+1}, \cdots$$

where of course

$$\frac{\lambda+1}{n+\left[\frac{n}{\lambda}\right]} \leq \frac{1}{\left[\frac{n}{\lambda}\right]}; \quad \frac{\lambda+1}{n+\left[\frac{n}{\lambda}\right]+1} > \frac{1}{\left[\frac{n}{\lambda}\right]+1};$$

$$\frac{\lambda+1}{n-1+\left[\frac{n-1}{\lambda}\right]} \leq \frac{1}{\left[\frac{n-1}{\lambda}\right]}; \quad \frac{\lambda+1}{n+\left[\frac{n-1}{\lambda}\right]} > \frac{1}{\left[\frac{n-1}{\lambda}\right]+1}.$$

Apart from the arrangement, the terms of (22) will be the same as the terms of (23), except that in (23), $\frac{\lambda+1}{\left[\frac{n-1}{\lambda}\right]+n}$ replaces λ/n. Now,

$$\frac{\lambda+1}{\left[\frac{n-1}{\lambda}\right]+n} > \frac{\lambda}{n},$$

so that the sum of the terms in (23) is greater than that in (22).

Since the transforms of a given set have the same measure as the set, and the sets $T^{-j}S_K$ cover U exactly k times, we have

$$\begin{aligned}
&\frac{1}{k}\int_S \Big\{ W(P) + \max\left(W(P), \frac{W(P)+W(TP)}{2}\right) \\
&\quad + \max\left(W(P), \frac{W(P)+W(TP)}{2}, \frac{W(P)+W(TP)+W(T^2P)}{3}\right) \\
&\quad + \cdots + \max\Big(W(P), \frac{W(P)+W(TP)}{2}, \cdots, \\
&\qquad\qquad \frac{W(P)+\cdots+W(T^{k-1}P)}{k}\Big)\Big\} dV_P \\
&\leq \sum_{K=0}^{\infty} m(S_K)\left(1+\log\frac{K}{k}\right) \\
&\leq m(W)\left(1+\log\frac{2^N}{k}\right) + \sum_{j=1}^{\infty}\frac{km(S)}{2^{N+j}}\left(1+\log\frac{2^{N+j}}{k}\right) \\
&\leq \left(1+\log\frac{2^N}{k}\right)\left\{ m(W) + \text{const.}\ \frac{km(S)}{2^N}\right\},
\end{aligned}$$

the constant being absolute. If we now put

$$N = \left[\log\frac{km(S)}{m(W)} \Big/ \log 2\right],$$

this last expression is dominated by

$$\text{const.}\ m(W)\left(1+\log\frac{m(S)}{m(W)}\right).$$

Since the constant is independent of k, we see that the Cesàro average of

$$\max\left(W(P), \frac{W(P)+W(TP)}{2}, \cdots, \frac{W(P)+\cdots+W(T^{k-1}P)}{k}\right)$$

is dominated by the same term. Thus

$$\begin{aligned}
&\int_S \max\left(W(P), \cdots, \frac{W(P)+\cdots+W(T^{k-1}P)}{k}\right) dV_P \\
&\qquad \leq \text{const.}\ m(W)\left(1+\log\frac{m(S)}{m(W)}\right),
\end{aligned}$$

and it follows by monotone convergence that there exists a function $W^*(P)$ such that

$$\int_S W^*(P)\,dV_P \leq \text{const.}\ m(W)\left(1+\log\frac{m(S)}{m(W)}\right), \tag{24}$$

and for all positive m,

$$\frac{W(P)+\cdots+W(T^mP)}{m+1} \leq W^*(P). \tag{25}$$

Now let $f(P)$ be a function such that

$$\int_S |f(P)| \log^+ |f(P)| \, dV_P < \infty. \tag{21}$$

Let $W^{(N)}$ be the set of points such that

$$2^N \leq |f(P)| \leq 2^{N+1}, \tag{26}$$

and let $f^{(N)}(P)$ be the function equal to $f(P)$ over this set of points, and 0 elsewhere. Then there exists a function $f^{(N)*}(P)$, such that

$$\frac{f^{(N)}(P) + \cdots + f^{(N)}(T^m P)}{m+1} \leq f^{(N)*}(P) \quad (m = 0, 1, 2, \cdots), \tag{27}$$

and

$$\int_S f^{(N)*}(P) \, dV_P \leq \text{const.}\, m(W^{(N)}) \left(1 + \log^+ \frac{m(S)}{m(W^{(N)})}\right) 2^N. \tag{28}$$

Hence if

$$\sum_{-\infty}^{\infty} m(W^{(N)}) \left(1 + \log^+ \frac{m(S)}{m(W^{(N)})}\right) 2^N < \infty \tag{29}$$

and

$$f^*(P) = \sum_{-\infty}^{\infty} f^{(N)*}(P), \tag{30}$$

then

$$\frac{f(P) + \cdots + f(T^m P)}{m+1} \leq f^*(P) \qquad (m = 0, 1, 2, \cdots), \tag{31}$$

and

$$\int_S f^*(P) \, dV_P \leq \text{const.} \sum m(W^{(N)}) \left(1 + \log^+ \frac{m(S)}{m(W^{(N)})}\right) 2^N. \tag{32}$$

However,

$$\begin{aligned} &\sum_{-\infty}^{\infty} m(W^{(N)}) \left(1 + \log^+ \frac{m(W^{(N)})}{m(S)}\right) 2^N \\ &\leq \sum_{-\infty}^{\infty} \int_{W^{(N)}} |f(P)| \, dV_P \left(1 + \log^+ \frac{2^{N+1} m(S)}{\int_S |f(P)| \, dV_P}\right) \\ &\leq \int_S |f(P) \left(1 + \log^+ \frac{2\,|f(P)|\, m(S)}{\int_S |f(Q) \, dV_Q}\right) dV_P \\ &\leq \int_S |f(P)| \, dV_P \left(1 + \log^+ \frac{2m(S)}{\int_S |f(P)| \, dV_P}\right) \\ &\quad + \int_S |f(P)| \log^+ |f(P)| \, dV_P. \end{aligned} \tag{33}$$

Thus $\int_S f^*(P)\,dV_P$ has an upper bound which is less than a function of $\int_S |f(P)|\,dV_P$ and $\int_S |f(P)| \log^+ |f(P)|\,dV_P$, tending to 0 as they both tend to 0. This establishes our theorem of the existence of a uniform dominant.

There is a sense in which (21) *is a best possible condition.* That is, if

$$\psi(x) = o(\log^+ x), \tag{34}$$

the condition

$$\int_S (\psi(|f(P)|)\,|f(P)|\,dV_P < \infty \tag{35}$$

is not sufficient for the existence of a uniform dominant. For let S be a set of measure 1, subdivided into mutually exclusive sets S_n, of measures respectively 2^{-n}. Let S_n be divided into mutually exclusive sets $S_{n,1}, \cdots, S_{n,\nu_n}$, all of equal measures. Let T transform $S_{n,k}$ into $S_{n,(k+1)}$ $(k < \nu^n)$, and S_{n,ν_n} into $S_{n,1}$. Let $f(P)$ be defined by

$$f(P) = a_n > 0 \text{ on } S_{n_1}, \quad (n = 1, 2, \cdots); \quad f(P) = 0 \text{ elsewhere.} \tag{36}$$

Then the smallest possible uniform dominant of

$$\frac{1}{N+1} \sum_0^N f(T^n P) \tag{37}$$

is

$$f^*(P) = \frac{a_n}{k} \text{ on } S_{n,k}, \tag{38}$$

and we have

$$\int_S f^*(P)\,dV_P = \sum \frac{a_n \Omega(\log \nu_n) 2^{-n}}{\nu_n}. \tag{39}$$

Thus if

$$\nu_n = 2^{2n} a_n, \qquad a_n = \Omega(n!), \tag{40}$$

the function $f^*(P)$ will belong to L if and only if

$$\infty > \sum \frac{a_n \Omega(\log a_n) 2^{-n}}{\nu_n} > \text{const.} \int_S f(P) \log^+ f(P)\,dV_P. \tag{41}$$

While we have proved the dominated ergodic theorem merely as a lemma for the multidimensional ergodic theorem, the theorem, and more particularly, the method by which we have proved it, have very considerable independent interest. We may use these methods to deduce von Neumann's mean ergodic theorem from the Birkhoff theorem; or vice versa, we may deduce the Birkhoff theorem, at least in the case of a function satisfying (21), from the von

Neumann theorem. These facts however are not relevant to the frame of the present paper, and will be published elsewhere.

We shall now proceed to the proof of the multidimensional ergodic theorem, which we shall establish in the two-dimensional case, although the method is independent of the number of dimensions. Let $T_1^\lambda T_2^\mu$ be a two-dimensional Abelian group of transformations of the set S (of measure 1) into itself, in the sense in which we have used this term in paragraph 3. Let $T_1^\lambda T_2^\mu P$ be measurable in λ, μ, and P. We now introduce a new variable x, ranging over $(0, 1)$, and form the product space Σ of P and x. We introduce the one-parameter group of transformations of this space, T^ρ, by putting

$$(42) \qquad T^\rho(P, x) = (T_1^{\rho \cos 2\pi x} T_2^{\rho \sin 2\pi x} P, x)$$

The expressions $T^\rho(P, x)$ will be measurable in ρ and (P, x), and the transformations T^ρ will all preserve measure on Σ. Thus by the ergodic theorem, for almost all points (P, x) of Σ, if $f(P) = f(P, x)$ belongs to L, the limit

$$(43) \quad \lim_{A\to\infty} \frac{1}{A} \int_0^A f(T^\rho(P, x))\, d\rho = \lim_{A\to\infty} \frac{1}{A} \int_0^A f(T_1^{\rho \cos 2\pi x} T_2^{\rho \sin 2\pi x} P)\, d\rho$$

will exist. If condition (26) is satisfied, it will follow by dominated convergence that the limit

$$(44) \qquad \lim_{A\to\infty} \frac{1}{A} \int_0^A d\rho \int_0^1 f(T_1^{\rho \cos 2\pi x} T_2^{\rho \sin 2\pi x} P)\, dx$$

will exist for almost all points (P, x), and hence for almost all points P.

For the moment, let us assume that $f(P)$ is non-negative. Then there is a Tauberian theorem, due to the author,[3] which establishes that the expression (44) is equivalent to

$$(45) \qquad \lim_{A\to\infty} \frac{1}{\pi A^2} \int_0^A \rho\, d\rho \int_0^{2\pi} f(T_1^{\rho \cos \theta} T_2^{\rho \sin \theta} P)\, d\theta.$$

The only point of importance which we must establish in order to justify this Tauberian theorem is that

$$(46) \qquad \int_0^1 \rho^{1+iu} d\rho = \frac{1}{2 + iu} \neq 0$$

for real values of u. Since every function $f(P)$ satisfying (21) is the difference of two non-negative functions satisfying this condition, (45) is established in the general case.

It will be observed that we have established our multidimensional ergodic

---

[3] N. Wiener, "Tauberian theorems," *Annals of Mathematics*, 2, vol. 33, p. 28.

theorem on the basis of assumption (21), and not on that of the weaker assumption that $f(P)$ belongs to L. What the actual state of affairs may be, we do not know. At any rate, all attempts to arrive at a direct analogue of the Khintchine proof for one dimension have broken down. The one-dimensional proof makes essential use of the fact that the difference of two intervals is always an interval, while the difference between two spheres is not always a sphere.

The precise statement of the multidimensional ergodic theorem is the following: *Let S be a set of points of measure 1, and let $T_1^{\lambda_1}T_2^{\lambda_2}\cdots T_n^{\lambda_n}$ be an Abelian group of equimeasure transformations of S into itself, in the sense of paragraph 3. Let $T_1^{\lambda_1}\cdots T_n^{\lambda_n}P$ be measurable in $\lambda_1,\cdots,\lambda_n$; P. Let R be the set of values of $\lambda_1,\cdots,\lambda_n$ for which*

$$\lambda_1^2+\lambda_2^2+\cdots+\lambda_n^2\leq r^2 \tag{47}$$

and let $V(r)$ be its volume. Let $f(P)$ satisfy the condition (21). Then for almost all values of P,

$$\lim_{r\to\infty}\frac{1}{V(r)}\int\underset{R}{\cdots}\int f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\,d\lambda_1\cdots d\lambda_n \tag{48}$$

exists.

That part of Theorem I which does not concern metric transitivity is an immediate corollary.

5. Metric transitivity. Space and Phase Averages in a Chaos. If the function $f(P)$ is positive, clearly

$$\begin{aligned}
&\underset{\lambda_1^2+\cdots+\lambda_n^2\leq r^2-\mu_1^2-\cdots-\mu_n^2}{\int\cdots\int} f(T_1^{\lambda}\cdots T_n^{\lambda}P)\,d\lambda_1\cdots d\lambda_n\\
&\leq\underset{\lambda_1^2+\cdots+\lambda_n^2\leq r^2}{\int\cdots\int} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}T_1^{\mu_1}\cdots T_n^{\mu_n}P)\,d\lambda_1\cdots d\lambda_n\\
&\leq\underset{\lambda_1^2+\cdots+\lambda_n^2\leq r^2+\mu_1^2+\cdots+\mu_n^2}{\int\cdots\int} f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\,d\lambda_1\cdots d\lambda_n.
\end{aligned} \tag{49}$$

Hence

$$\begin{aligned}
&\lim_{r\to\infty}\frac{1}{V(r)}\int\underset{R}{\cdots}\int f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}T_1^{\mu_1}\cdots T_n^{\mu_n}P)\,d\lambda_1\cdots d\lambda_n\\
&=\lim_{r\to\infty}\frac{1}{V(r)}\int\underset{R}{\cdots}\int f(T_1^{\lambda_1}\cdots T_n^{\lambda_n}P)\,d\lambda_1\cdots d\lambda_n,
\end{aligned} \tag{50}$$

and expression (48) has the same value for P and all its transforms under the group $T_1^{\lambda_1}\cdots T_n^{\lambda_n}$. The condition of positivity is clearly superfluous. Thus in case expression (48) does not almost everywhere assume a single value, there will be two classes S_1 and S_2 of elements of S, each of positive measure, and each invariant under all the transformations $T_1^{\lambda_1} \cdots T_n^{\lambda_n}$.

A condition which will manifestly exclude such a contingency is that if S_1 and S_2 are two sub-sets of S of positive measure, and ϵ is a positive quantity, there always exists a transformation $T = T_1^{\lambda_1} \cdots T_n^{\lambda_n}$, such that

$$\left| \frac{mS_1(TS_2)}{mS_2} - mS_1 \right| < \epsilon. \tag{51}$$

From this it will immediately follow that if a chaos is metrically transitive in the sense of paragraph 3, the group of transformations of the α space generated by translations of the chaos will have the property we have just stated, and under the assumptions of Theorem I,

$$\lim_{r\to\infty} \frac{1}{V(r)} \int \cdot_R \cdot \cdot \int \Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\} dy_1 \cdots dy_n \tag{52}$$

will exist and have the same value for almost all values of α.

If almost everywhere

$$\lim_{r\to\infty} \frac{1}{V(r)} \int \cdot_R \cdot \cdot \int \Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\} dy_1 \cdots dy_n = A, \tag{53}$$

and

$$\frac{1}{V(r)} \int \cdot_R \cdot \cdot \int \Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\} dy_1 \cdots dy_n < g(\alpha) \tag{54}$$

where $g(\alpha)$ belongs to L, then by dominated convergence,

$$A = \lim_{r\to\infty} \frac{1}{V(r)} \int \cdot_R \cdot \cdot \int dy_1 \cdots dy_n \int_0^1 d\alpha \Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\} = \int_0^1 \Phi\{\mathfrak{F}(\Sigma; \alpha)\} d\alpha. \tag{55}$$

That is, the average of $\Phi\{\mathfrak{F}(\Sigma; \alpha)\}$, taken over the finite phase space of α, is almost everywhere the same as the average of $\Phi\{\mathfrak{F}_{y_1, \ldots, y_n}(\Sigma; \alpha)\}$ taken over the infinite group space of points $y_1, \cdots, y_n$. This completes the establishment of Theorem I, and gives us a real basis for the study of the homogeneous chaos.[4]

6. Pure one-dimensional chaos. The simplest type of pure chaos is that which has already been treated by the author in connection with the Brownian motion. However, as we wish to generalize this theory to a multi-

[4] The material of this chapter, in the one-dimensional case, has been discussed by the author with Professor Eberhard Hopf several years ago, and he wishes to thank Professor Hopf for suggestions which have contributed to his present point of view.

plicity of dimensions, instead of referring to existing articles on the subject, we shall present it in a form which emphasizes its essential independence of dimensionality.

The type of chaos which we shall consider is that in which the expression $\mathfrak{F}(\Sigma;\alpha)$ has a distribution in α dependent only on the measure of the set Σ; and in which, if Σ_1 and Σ_2 do not overlap, the distributions of $\mathfrak{F}(\Sigma_1,\alpha)$ and $\mathfrak{F}(\Sigma_2,\alpha)$ are independent, in the sense that if $\phi(x,y)$ is a measurable function, and either side of the equation has a sense,

$$(56)\qquad \int_0^1\int_0^1 \phi(\mathfrak{F}(\Sigma_1;\alpha),\mathfrak{F}(\Sigma_2,\beta))\,d\alpha d\beta = \int_0^1 \phi(\mathfrak{F}(\Sigma_1,\alpha),\mathfrak{F}(\Sigma_2,\alpha))\,d\alpha.$$

We assume a similar independence when n non-overlapping sets $\Sigma_1,\cdots,\Sigma_n$ are concerned. It is by no means intuitively certain that such a type of chaos exists. In establishing its existence, we encounter a difficulty belonging to many branches of the theory of the Lebesgue integral. The fundamental theorem of Lebesgue assures us of the possibility of adding the measures of a denumerable assemblage of measurable sets, to get the measure of their sum, if they do not overlap. Accordingly, behind any effective realization of the theory of Lebesgue integration, there is always a certain denumerable family of sets in the background, such that all measurable sets may be approximated by denumerable combinations of these. This family is not unique, but without the possibility of finding it, there is no Lebesgue theory.

On the other hand, a theory of measure suitable for the description of a chaos must yield the measure of any assemblage of functions arising from a given measurable assemblage by a translational change of origin. This set of assemblages is essentially non-denumerable. Any attempt to introduce the notion of measure in a way which is invariant under translational changes of origin, without the introduction of some more restricted set of measurable sets, which does not possess this invariance, will fail to establish those essential postulates of the Lebesgue integral which deal with denumerable sets of points. There is no way of avoiding the introduction of constructional devices which seem to restrict the invariance of the theory, although once the theory is obtained it may be established in its full invariance.

Accordingly, we shall start our theory of randomness with a division of space, whether of one dimension or of more, into a denumerable assemblage of sub-sets. In one dimension, this division may be that into those intervals whose coördinates are terminating binary numbers, and in more dimensions, into those parallelepipeds with edges parallel to the axes and with terminating binary coördinates for the corner points. We then wish to find a self-consistent distribution-function for the mass in such a region, dependent only on the volume, and independent for non-overlapping regions.

This problem does not admit of a unique solution, although the solution becomes essentially unique if we adjoin suitable auxiliary conditions. Among these conditions, for example, is the hypothesis that the distribution is symmetric, as between positive and negative values, has a finite mean square, and that the measure of the set of α's for which $\mathfrak{F}(S;\alpha) > A$ is a continuous function of A.[5] Without going into such considerations, we shall assume directly that the measure of the set of instances in which the value of $\mathfrak{F}(S;\alpha)$ in a region S of measure M lies between a and $b > a$, is

$$\frac{1}{\sqrt{2\pi M}}\int_a^b \exp\left(-\frac{u^2}{2M}\right)du. \tag{57}$$

The formula

$$\frac{1}{\sqrt{2\pi M_1 M_2}}\int_{-\infty}^{\infty} \exp\left(-\frac{u^2}{2M_1}-\frac{(v-u)^2}{2M_2}\right)du = \frac{1}{\sqrt{M_1+M_2}}\exp\left(-\frac{v^2}{2(M_1+M_2)}\right) \tag{58}$$

shows the consistency of this assumption.

The distributions of mass among the sets of our denumerable assemblage may be mapped on the line segment $0 \leq \alpha \leq 1$, in such a way that the measure of the set of instances in which a certain contingency holds will go into a set of values of α of the same measure. This statement needs a certain amount of elucidation. To begin with, the only sets of instances whose measures we know are those determined by

$$\begin{aligned} a_1 &\leq \mathfrak{F}(S_1;\alpha) \leq b_1 \\ a_2 &\leq \mathfrak{F}(S_2;\alpha) \leq b_2 \\ &\cdots\cdots\cdots \\ a_n &\leq \mathfrak{F}(S_n;\alpha) \leq b_n; \end{aligned} \tag{59}$$

where $S_1, S_2, \cdots, S_n$ are to be found among our denumerable set of subdivisions of space. However, once we have established a correspondence between the measures of these specific sets of contingencies and their corresponding sets of values of α, we may use the measure of any measurable set of values of α to define the measure of its corresponding set of contingencies.

The correspondence between sets of contingencies and points on the line $(0,1)$ is made by determining a hierarchy of sets of contingencies

$$\begin{aligned} a_1^{(m,n)} &\leq \mathfrak{F}(S_1^{(m,n)};\alpha) \leq b_1^{(m,n)} \\ &\cdots\cdots\cdots\cdots \\ a_\nu^{(m,n)} &\leq \mathfrak{F}(S_\nu^{(m,n)};\alpha) \leq b_\nu^{(m,n)}. \end{aligned} \tag{60}$$

---

[5] Cf. the recent investigations of Cramér and P. Lévy.

Let us call such a contingency $C_{m,n}$. If m is fixed, let all the contingencies $C_{m,n}$ $(n = 1, 2, \cdots)$ be mutually exclusive, and let them be finite in number. Let us be able to write

$$C_{m,n} = \sum_{k=1}^{N} C_{m+1,n_k}. \tag{61}$$

If S' is one of our denumerable sets of regions of space, let every $C_{m,n}$ with a sufficiently large index m be included in a class determined by a set of conditions concerning the mass on S' alone, and restricting it to a set of values lying in an interval (c, d), corresponding to an integral of form (57) and of arbitrarily small value. (Here d may be ∞, or c may be $-\infty$.) Let us put $C_{1,1}$ for the entire class of all possible contingencies, and let us represent it by the entire interval $(0, 1)$. Let us assume that $C_{m,n}$ has been mapped into an interval of length corresponding to its probability, in accordance with (57), and let this interval be divided in order of the sequence of their n_k's into intervals corresponding respectively to the component C_{m+1,n_k}'s, and of the same measure. Except for a set of points of zero measure, every point of the segment $(0, 1)$ of α will then be determined uniquely by the sequence of the intervals containing it and corresponding to the contingencies $C_{m,n}$ for successive values of m. This sequence will then determine uniquely (except in a set of cases corresponding to a set of values of α of zero measure) the value of $\mathfrak{F}(S_n; \alpha)$ for every one of our original denumerable set of sets S_n.

So far, everything that we have said has been independent of dimensionality. We now proceed to something belonging specifically to the one-dimensional case. If the original sets S_n are the sets of intervals with binary end-points, of such a form that they may be written in the binary scale

$$(d_1 d_2 \cdots d_k \cdot d_{k+1} \cdots d_l,\ d_1 \cdots d_k \cdot d_{k+1} \cdots d_l + 2^{-l-1}) \tag{62}$$

where $d_1, \cdots, d_l$ are digits which are either 0 or 1, then any interval whatever of length not exceeding $2^{-\mu}$ and lying in (a, b) (where a and b are integers) may be written as the sum of not more than two of the $(b-a)2^{\mu+2}$ intervals of form (62) lying in (a, b) and of length $2^{-\mu-1}$, not more than two of the $(b-a)2^{\mu+2}$ intervals of length $2^{-\mu-2}$, and so on. The probability that the value of $|\mathfrak{F}(S_n; \alpha)|$ should exceed A, or in other words, the measure of the set of α's for which it exceeds A, is

$$\frac{2}{\sqrt{2\pi m(S_n)}} \int_A^\infty \exp\left(-\frac{u^2}{2m(S_n)}\right) du = o\{\exp(-Am(S_n)^{-\frac{1}{2}})\}. \tag{63}$$

Now let us consider the total probability that the value of $|\mathfrak{F}(S_n : \alpha)|$ should exceed $2^{-(\mu+1)(\frac{1}{2}-\epsilon)}$ for any one of the $(b-a)2^{\mu+1}$ intervals of length

$2^{-\mu-1}$, or $2^{-(\mu+2)(\frac{1}{2}-\epsilon)}$ for any one of the $(b-a)2^{\mu+2}$ intervals of length $2^{-\mu-2}$, or so on. This probability can not exceed

$$(64) \qquad \sum_{k=1}^{\infty} (b-a)2^{\mu+k}\, o(\exp(-2^{(\mu+k)\epsilon})) = o(2^{\mu}\exp(-2^{(\mu+1)\epsilon})).$$

On the other hand, the sum of $|\mathfrak{F}(S_n;\alpha)|$ for all the $2+2+\cdots$ intervals must in any other case be equal to or less than

$$(65) \qquad 2\sum_{k=1}^{\infty} 2^{-(\mu+k)(\frac{1}{2}-\epsilon)} = O(2^{-\mu(\frac{1}{2}-\epsilon)}).$$

Thus there is a certain sense in which over a finite interval, and except for a set of values of α of arbitrarily small positive measure, the total mass in a sub-interval of length $\leq 2^{-\mu}$ tends uniformly to 0 with $2^{-\mu}$. On this basis, we may extend the functional $\mathfrak{F}(S;\alpha)$ to all intervals S. It is already defined for all intervals with terminating binary end-points. If (c,d) is any interval whatever, let $c_1, c_2, \cdots$ be a sequence of terminating binary numbers approaching c, and let $d_1, d_2, \cdots$ be a similar sequence approaching d. Then except for a fixed set of values of α of arbitrarily small measure,

$$(66) \qquad \lim_{m,n\to\infty} |\mathfrak{F}((c_n, d_n)+\alpha) - \mathfrak{F}((c_m, d_m);\alpha)| = 0,$$

and we may put

$$(67) \qquad \mathfrak{F}((c,d);\alpha) = \lim_{m\to\infty} \mathfrak{F}((c_m, d_m);\alpha).$$

Formula (57), and the fact that $\mathfrak{F}(S_1;\alpha)$ and $\mathfrak{F}(S_2;\alpha)$ vary independently for non-overlapping intervals S_1 and S_2, will be left untouched by this extension.

Thus if Σ is an interval and T^{λ} a translation through an amount λ, we can define $\mathfrak{F}(T^{\lambda}\Sigma;\alpha)$, and it will be equally continuous in λ over any finite range of λ except for a set of values of α of arbitrarily small measure. From this it follows at once that it is measurable in λ and α together. Furthermore, we shall have

(68) Measure of set of α's for which $\mathfrak{F}(T^{\lambda}\Sigma;\alpha)$ belongs to C =
Measure of set of α's for which $\mathfrak{F}(\Sigma;\alpha)$ belongs to C.

Thus $\mathfrak{F}(\Sigma;\alpha)$ is a homogeneous chaos. We shall call it *the pure chaos.*

If $\Phi\{\mathfrak{F}(\Sigma;\alpha)\}$ is a functional dependent on the values of $\mathfrak{F}(\Sigma;\alpha)$ for a finite number of intervals Σ_n, then if λ is so great that none of the intervals Σ_n overlaps any translated interval $T^{\lambda}\Sigma_m$, $\Phi\{\mathfrak{F}(\Sigma;\alpha)\}$ will have a distribution entirely independent of $\Phi\{\mathfrak{F}(T^{\lambda}\Sigma;\alpha)\}$. As every measurable functional may be approached in the L sense by such a functional, we see at once that $\mathfrak{F}(\Sigma;\alpha)$ is metrically transitive.[6]

[6] Except that the method of treatment has been adapted to the needs of § 7, the

7. Pure multidimensional chaos. In order to avoid notational complexity, we shall not treat the general multidimensional case explicitly, but shall treat the two-dimensional case by a method which will go over directly to the most general multidimensional case. If our initial sets S_n are the rectangles with terminating binary coördinates for their corners and sides parallel to the axes, and we replace (a, b) by the square with opposite vertices (p, q) and $(p + r,\ q + r)$, an argument of exactly the same sort as that which we have used in the last paragraph will show that except in a set of cases of total probability not exceeding

$$(69) \qquad \text{const.} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} 2^{\mu+k} 2^{l}\, o(\exp(-2^{(\mu+k+l)\epsilon})) = o(2^{\mu} \exp(-2^{(\mu+1)\epsilon}))$$

the sum of $|\mathfrak{F}(S_n;\alpha)|$ for a denumerable set of binary rectangles with base $\leq 2^{-\mu}$, of the form (62), and adding up to make a vertical interval lying in the square (p, q), $(p + r, q + r)$, must be equal to or less than

$$(65) \qquad 2 \sum_{k=1}^{\infty} 2^{-(\mu+k)(\frac{1}{2}-\epsilon)} = O(2^{-\mu(\frac{1}{2}-\epsilon)}).$$

If we now add this expression up for all the base intervals of type (62) necessary to exhaust a horizontal interval of magnitude not exceeding $2^{-\mu}$, we shall again obtain an expression of the form (65). It hence follows that if we take the total mass on the coördinate rectangles within a given square, this will tend to zero uniformly with their area, except for a set of values of α of arbitrarily small measure. From this point the two-dimensional argument, and indeed the general multidimensional argument, follows exactly the same lines as the one-dimensional argument. It is only necessary to note that if

$$a_n \to a,\ b_n \to b,\ c_n \to c,\ d_n \to d$$

then the rectangles (a_m, b_m), (c_m, d_m) and (a_n, b_n), (c_n, d_n) differ at most by four rectangles of small area.[7]

From this point on, we shall write $\mathfrak{P}(S;\alpha)$ for a pure chaos, whether in one or in more dimensions.

8. Phase averages in a pure chaos. If $f(P)$ is a measurable step-function, the definition of

$$(70) \qquad \int f(P)\, d_P \mathfrak{P}(S;\alpha)$$

results of this section have previously been demonstrated by the author. (*Proceedings of the London Mathematical Society*, 2, vol. 22 (1924), pp. 454-467).

[7] Here we represent a rectangle by giving two opposite corners.

is obvious, for it reduces to the finite sum

$$\sum_{1}^{N} f_n \mathfrak{P}(S_n;\alpha) \tag{71}$$

where f_n are the N values assumed by $f(P)$, and S_n respectively are the sets over which these values are assumed. Let us notice that

$$\begin{aligned}\int_0^1 d\alpha \left| \int f(P)\, d_P \mathfrak{P}(S;\alpha) \right|^2 &= \sum_{m=1}^{N} \sum_{n=1}^{N} \int_0^1 d\alpha f_m \bar{f}_n \mathfrak{P}(S_m;\alpha)\mathfrak{P}(S_n;\alpha) \\ &= \sum_{1}^{N} | f_n |^2 \int_0^1 (\mathfrak{P}(S_n;\alpha))^2 d\alpha \\ &= \sum_{1}^{N} | f_n |^2 \frac{1}{\sqrt{2\pi m(S_n)}} \int_{-\infty}^{\infty} u^2 \exp\left(-\frac{u^2}{2m(S_n)}\right) du \\ &= \sum_{1}^{N} | f_n |^2 m(S_n) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-(u^2/2)} du \\ &= \sum_{1}^{N} | f_n |^2 m(S_n) = \int | f(P) |^2 dV_P,\end{aligned} \tag{72}$$

the integral being taken over the whole of space. In other words, the transformation from $f(P)$ as a function of P, to $\int f(P)\, d_P \mathfrak{P}(S;\alpha)$ as a function of α, retains distance in Hilbert space.[8] Such a transformation, by virtue of the Riesz-Fischer theorem, may always be extended by making limits in the mean correspond to limits in the mean. Thus both in the one-dimensional and in the many-dimensional case, we may *define*

$$\int f(P)\, d_P \mathfrak{P}(S;\alpha) = \underset{n\to\infty}{\mathrm{l.\,i.\,m.}} \int f_n(P)\, d_P \mathfrak{P}(S;\alpha) \tag{73}$$

where $f(P)$ is a function belonging to L^2, and the sequence $f_1(P), f_2(P), \cdots$ is a sequence of step-functions converging in the mean to $f(P)$ over the whole of space. The definition will be unambiguous, except for a set of values of α of zero measure.

If S is any measurable set, we have

$$\begin{aligned}\int_0^1 d\alpha \{\mathfrak{P}(S;\alpha)\}^n &= \frac{1}{\sqrt{2\pi m(S)}} \int_{-\infty}^{\infty} u^n \exp\left(-\frac{u^2}{2m(S)}\right) du \\ &= (m(S))^{n/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^n e^{-(u^2/2)} du \\ &\begin{cases} = 0 \text{ if } n \text{ is odd} \\ = (m(S))^{n/2}(n-1)(n-3)\cdots 1 \text{ if } n \text{ is even.}\end{cases}\end{aligned} \tag{74}$$

---

[8] Cf. Paley, Wiener, and Zygmund, *Mathematische Zeitschrift*, vol. 37 (1933), pp. 647-668.

This represents $(m(S))^{n/2}$, multiplied by the number of distinct ways of representing n objects as a set of pairs. Remembering that if $S_1, S_2, \cdots, S_{2n}$ are non-overlapping, their distributions are independent, we see that if the sets $\Sigma_1, \Sigma_2, \cdots, \Sigma_{2n}$ are either totally non-overlapping, or else such that when two overlap, they coincide, we have [9]

$$(75) \qquad \int_0^1 \mathfrak{P}(\Sigma_1;\alpha)\cdots\mathfrak{P}(\Sigma_n;\alpha)\,d\alpha = \Sigma\Pi\int_0^1 \mathfrak{P}(\Sigma_j;\alpha)\,\mathfrak{P}(\Sigma_k;\alpha)\,d\alpha,$$

where the product sign indicates that the $2n$ terms are divided into n sets of pairs, j and k, and that these factors are multiplied together, while the addition is over all the partitions of $1, \cdots, 2n$ into pairs. If $2n$ is replaced by $2n+1$, the integral in (75) of course vanishes.

Since $\mathfrak{P}(S;\alpha)$ is a linear functional of sets of points, and since both sides of (75) are linear with respect to each $\mathfrak{P}(\Sigma_k;\alpha)$ separately, (75) still holds when $\Sigma_1, \Sigma_2, \cdots, \Sigma_{2n}$ can be reduced to sums of sets which either coincide or do not overlap, and hence holds for all measurable sets.

Now let $f(P_1, \cdots, P_n)$ be a measurable step-function: that is, a function taking only a finite set of finite values, each over a set of values $P_1, \cdots, P_n$ which is a product-set of measurable sets in each variable P_k. Clearly we may define

$$(76) \qquad \int\cdots\int f(P_1,\cdots,P_n)\,d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha)$$

in a way quite analogous to that in which we have defined (70), and we shall have

$$(77) \qquad \int_0^1 d\alpha\int\cdots\int f(P_1,\cdots,P_n)\,d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha)$$
$$= \Sigma\int\cdots\int f(P_1,P_1,P_2,P_2,\cdots,P_n,P_n)\,dV_{P_1}\cdots dV_{P_n}$$

where the summation is carried out for all possible divisions of the $2n$ P's into pairs. Similarly in the odd case

$$(78) \qquad \int_0^1 d\alpha\int\cdots\int f(p_1,\cdots,P_{2n+1})\,d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha) = 0.$$

We may apply (77) to give a meaning to

$$(79) \qquad \int_0^1 d\alpha\,\left|\int\cdots\int f(P_1,\cdots,P_n)\,d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha)\right|^2.$$

If $f(\mathrm{P}_1, \cdots, P_n)$ is a measurable step-function, and

---

[9] Cf. Paley, Wiener, and Zygmund, *loc. cit.*, formula (2.05).

$$(80) \qquad |f(P_1,\cdots,P_n)| \leq |f_1(P_1)\cdots f_n(P_n)|\,; \qquad \int |f_k(P)|^2 dV_P \leq A \qquad (k=1,2,\cdots)$$

we shall have

$$(81) \qquad \int_0^1 d\alpha \left| \int\cdots\int f(P_1,\cdots P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha)\right|^2 \leq A^n(2n-1)(2n-3)\cdots 1.$$

If now $f(P_1,\cdots,P_n)$ is an integrable function satisfying (80), but not necessarily a step-function, let

$$(82) \quad f(\nu;P_1,\cdots,P_n) = \frac{1}{\nu}\operatorname{sgn} f(P_1,\cdots,P_n)\,[\nu f(P_1,\cdots,P_n)\operatorname{sgn} f(P_1,\cdots,P_n)].$$

Clearly almost everywhere

$$(83) \qquad \int\cdots\int f(\nu;P_1,\cdots,P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha) \leq \prod_1^n \int f_k(P)\, d_P\mathfrak{P}(S;\alpha)$$

and

$$(84) \qquad \overline{\lim_{\mu,\nu\to\infty}} \Big| \int\cdots\int f(\mu;P_1,\cdots,P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha) - \int\cdots\int f(\nu;P_1,\cdots,P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\, d_{P_2}\mathfrak{P}(S;\alpha)\Big| \leq \prod_1^n \int f_k(P)\, d_P\mathfrak{P}(S;\alpha) \left[\epsilon + \sum_1^n \left\{ \frac{\int_R f_k(P)\, d_P\mathfrak{P}(S;\alpha)}{\int f_k(P)\, d_P\mathfrak{P}(S;\alpha)} \right\}\right]$$

where R represents the exterior of a sphere of arbitrarily large volume. Let it be noted that both the numerator and the denominator of this fraction have Gaussian distributions, but that the mean square value of the numerator is arbitrarily small. Thus except for a set of values of α of arbitrarily small measure, the right side of expression (84) is arbitrarily small, so that we may write

$$(85) \qquad \lim_{\mu,\nu\to\infty} \left\{ \int\cdots\int f(\mu;P_1,\cdots,P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha) - \int\cdots\int f(\nu;P_1,\cdots,P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha) \right\} = 0.$$

Thus by dominated convergence,

$$(86) \qquad \lim_{\mu\to\infty} \int\cdots\int f(\mu;P_1,\cdots,P_n)\, d_{P_1}\mathfrak{P}(S;\alpha)\cdots d_{P_n}\mathfrak{P}(S;\alpha)$$

exists for almost all values of α, and we may write it by definition

$$\int \cdots \int f(P_1, \cdots, P_n) d_{P_1} \mathfrak{P}(S; \alpha) \cdots d_{P_n} \mathfrak{P}(S; \alpha). \tag{87}$$

This will clearly be unique, except for a set of values of α of zero measure. There will then be no difficulty in checking (77), (78), and (81).

9. Forms of chaos derivable from a pure chaos. Let us assume that $f(P)$ belongs to L^2, or that $f(P_1, \cdots, P_n)$ is a measurable function satisfying (80). Let us write $\widehat{PQ}$ for the vector in n-space connecting the points P and Q. Then the function

$$\int \cdots \int f(\widehat{PP_1}, \cdots, \widehat{PP_n}) d_{P_1} \mathfrak{P}(S; \alpha) \cdots d_{P_n} \mathfrak{P}(S; \alpha) = F(P; \alpha) \tag{88}$$

is a metrically transitive differentiable chaos. This results from the fact that $\mathfrak{P}(S; \alpha)$ is a metrically transitive chaos, and that a translation of P generates a similar translation of all the points P_k. The sum of a finite number of functions of the type (88) is also a metrically transitive differentiable chaos. To show that $F(P; \alpha)$ is measurable in P and α simultaneously, we merely repeat the argument of (83)–(86) with both P and α as variables.

We shall call a chaos such as (88) a *polynomial chaos homogeneously of the n-th degree*, and a sum of such chaoses a *polynomial chaos* of the degree of its highest term. In this connection, we shall treat a constant as a chaos homogeneously of degree zero.

By the multidimensional ergodic theorem, if Φ is a functional such that

$$\int_0^1 |\Phi(F(P; \alpha))| \log^+ |\Phi(F(P; \alpha))| \, d\alpha < \infty \tag{89}$$

we shall have

$$\lim_{r \to \infty} \frac{1}{V(r)} \int_R \Phi(F(P; \alpha)) dV_P = \int_0^1 \Phi(F(P; \alpha)) d\alpha \tag{90}$$

for almost all values of α. Since the distribution of $F(P; \alpha)$ is dominated by the product of a finite number of independent Gaussian distributions, we even have

$$\int_0^1 | F(P; \alpha) |^n d\alpha < \infty \tag{91}$$

for all positive integral values of n. In a wide class of cases this enables us to establish a relation of the type of (89).

In formula (90), we have an algorithm for the computation of the right-hand side. For example, if

$$(92)\qquad F(P;\alpha)=\int f(\widehat{PP_1})\,d_{P_1}\mathfrak{P}(S;\alpha),$$

and $P+Q$ is the vector sum of P and Q, we have for almost all α,

$$(93)\qquad \lim_{r\to\infty}\frac{1}{V(r)}\int_R F(P+Q;\alpha)\bar{F}(Q;\alpha)\,dV_Q=\int f(P+Q)\bar{f}(Q)\,dV_Q,$$

the integral being taken over the whole of space; if

$$(94)\qquad F(P;\alpha)=\int\cdots\int f(\widehat{PP_1},\widehat{PP_2})\,d_{P_1}\mathfrak{P}(S;\alpha)\,d_{P_2}\mathfrak{P}(S;\alpha),$$

we have almost always

$$(95)\qquad \lim_{r\to\infty}\frac{1}{V(r)}\int_R F(P+Q;\alpha)\bar{F}(Q;\alpha)\,dV_Q$$
$$=\left|\int f(Q,Q)\,dV_Q\right|^2+\int\int f(P+Q,P+M)\bar{f}(Q,M)\,dV_Q dV_M$$
$$+\int\int f(P+Q,P+M)\bar{f}(M,Q)\,dV_Q dV_M;$$

and if

$$(96)\qquad F(P;\alpha)=\int\int\int f(\widehat{PP_1},\widehat{PP_2},\widehat{PP_3})\,d_{P_1}\mathfrak{P}(S;\alpha)\,d_{P_2}\mathfrak{P}(S;\alpha)\,d_{P_3}\mathfrak{P}(S;\alpha),$$

we have almost everywhere

$$(97)\qquad \lim_{r\to\infty}\frac{1}{V(r)}\int_R F(P+Q;\alpha)\bar{F}(Q;\alpha)\,dV_Q$$
$$=\int\int\int \{f(Q,Q,P+M)\bar{f}(M,S,S)$$
$$+f(Q,Q,P+M)\bar{f}(S,M,S)+f(Q,Q,P+M)\bar{f}(S,S,M)$$
$$+f(Q,P+M,Q)\bar{f}(M,S,S)+f(Q,P+M,Q)\bar{f}(S,M,S)$$
$$+f(Q,P+M,Q)\bar{f}(S,S,M)+f(P+M,Q,Q)\bar{f}(M,S,S)$$
$$+f(P+M,Q,Q)\bar{f}(S,M,S)+f(P+M,Q,Q)\bar{f}(S,S,M)$$
$$+f(P+Q,P+M,P+S)\bar{f}(Q,M,S)$$
$$+f(P+Q,P+M,P+S)\bar{f}(Q,S,M)$$
$$+f(P+Q,P+M,P+S)\bar{f}(M,Q,S)$$
$$+f(P+Q,P+M,P+S)\bar{f}(M,S,Q)$$
$$+f(P+Q,P+M,P+S)\bar{f}(S,Q,M)$$
$$+f(P+Q,P+M,P+S)\bar{f}(S,M,Q)\}\,dV_Q dV_M dV_S.$$

We have similar results in the non-homogeneous case. Thus if

$$(98)\qquad F(P;\alpha)=A+\int f(\widehat{PP_1})\,d_{P_1}\mathfrak{P}(S;\alpha)$$
$$+\int\int g(\widehat{PP_1},\widehat{PP_2})\,d_{P_1}\mathfrak{P}(S;\alpha)\,d_{P_2}\mathfrak{P}(S;\alpha),$$

we have almost everywhere

$$
\begin{aligned}
(99) \qquad & \lim_{r\to\infty} \frac{1}{V(r)} \int_R F(P+Q;\alpha)\bar{F}(Q;\alpha)\,dV_Q \\
& = A\int \bar{g}(Q,Q)\,dV_Q + \bar{A}\int g(Q,Q)\,dV_Q \\
& + \int f(P+Q)\bar{f}(Q)\,dV_Q + \left|\int g(Q,Q)\,dV_Q\right|^2 \\
& + \int\int g(P+Q,P+M)\bar{g}(Q,M)\,dV_Q dV_M \\
& + \int\int g(P+Q,P+M)\bar{g}(M,Q)\,dV_Q dV_M.
\end{aligned}
$$

10. Chaos theory and spectra.[10] The function

$$
(100) \qquad \lim_{r\to\infty} \frac{1}{V(r)} \int_R F(P+Q)\bar{F}(Q)\,dV_Q = G(P)
$$

occupies a central position in the theory of harmonic analysis. If it exists and is continuous for every value of P, the function $F(P)$ is said to have an n-dimensional spectrum. To define this spectrum, we put

$$
(101) \qquad F_r(P) = \begin{cases} F(P) & \text{on } R; \\ 0 & \text{elsewhere.} \end{cases}
$$

It is then easy to show by an argument involving considerations like those of (49) that if

$$
(102) \qquad G_r(P) = \frac{1}{V(r)} \int_\infty F_r(P+Q)\bar{F}_r(Q)\,dV_Q,
$$

the integral being taken over the whole of space, then we have

$$
(103) \qquad G(P) = \lim_{r\to\infty} G_r(P).
$$

Since, if O is the point with zero coördinates, by the Schwarz inequality,

$$
(104) \qquad |G(P)| \leq G(O),
$$

the limit in (103) is approached boundedly.

If now we put

$$
(105) \qquad \phi_r(U) = (2\pi)^{-(n/2)} V(r)^{-\frac{1}{2}} \operatorname*{l.\,i.\,m.}_{s\to\infty} \int_S F_r(P) e^{iU\cdot P}\,dV_P
$$

where S is the interior of a sphere of radius s about the origin, the n-fold Parseval theorem will give us

---

[10] Cf. N. Wiener, "Generalized harmonic analysis," *Acta Mathematica*, vol. 55 (1930).

$$(106) \qquad | \phi_r(U)|^2 = (2\pi)^{-n} \int_\infty G_r(P) e^{iU \cdot P} dV_P.$$

If $M(U)$ is a function with an absolutely integrable Fourier transform, we shall have

$$(107) \quad \int_{-\infty}^{\infty} | \phi_r(U)|^2 M(U) dV_U = (2\pi)^{-n} \int_\infty G_r(P) dV_P \int_\infty M(U) e^{iU \cdot P} dV_U,$$

and hence

$$(108) \quad \lim_{r\to\infty} \int_{-\infty}^{\infty} | \phi_r(U)|^2 M(U) dV_U = (2\pi)^{-n} \int G(P) dV_P \int_\infty M(U) e^{iU \cdot P} dV_U$$

which will always exist. Let us put

$$(109) \qquad \mathfrak{M}\{M(U)\} = (2\pi)^{-n} \int_\infty G(P) dV_P \int_\infty M(U) e^{iU \cdot P} dV_U.$$

If S is any set of points of finite measure, and $S(P)$ is its characteristic function, let us put

$$(110) \qquad \overline{\mathfrak{M}}(S) = \underset{M(U) \geq S(U)}{\text{l. u. b.}} \mathfrak{M}(M(U)),$$

and

$$(111) \qquad \underline{\mathfrak{M}}(S) = \underset{M(U) \leq S(U)}{\text{g. l. b.}} \mathfrak{M}(M(U)).$$

If $\underline{\mathfrak{M}}(S)$ and $\overline{\mathfrak{M}}(S)$ have the same value, we shall write it $\mathfrak{M}(S)$, and shall call it the *spectral mass* of F on S. It will be a non-negative additive set-function of S, and may be regarded as determining the spectrum of F.

If $f(P_1, \cdots, P_n)$ satisfies (80) and $F(P;\alpha)$ is defined as in (88), we know that for any given P,

$$(112) \quad G(P;\alpha) = \lim_{r\to\infty} \frac{1}{V(r)} \int_R F(P+Q;\alpha) \bar{F}(Q;\alpha) = \int_0^1 F(P,\beta) \bar{F}(0,\beta) d\beta$$

for almost all values of α. This alone is not enough to assure that $F(\mathrm{P}, \alpha)$ has a spectrum for almost all values of α, as the sum of a non-denumerable set of sets of zero measure is not necessarily of zero measure. On the other hand, except for a set of values of α of zero measure, $G(P, \alpha)$ exists for all points P with rational coördinates.

We may even extend this result, and assert that if

$$(113) \qquad F_\theta(P;\alpha) = \frac{1}{V(\theta)} \int_{\text{length of } \widehat{PS} \leq \theta} F(S;\alpha) dV_S$$

and

$$(114) \qquad G_\theta(P;\alpha) = \lim_{r\to\infty} \frac{1}{V(r)} \int_R F_\theta(P+Q;\alpha) \bar{F}_\theta(Q;\alpha) dV_Q,$$

then except for a set of values of α of zero measure, $G_\theta(P;\alpha)$ exists for all points P with rational coördinates and all rational parameters θ, and it is easily proved that for almost all values of α, as θ tends to 0 through rational values,

$$\lim_{\theta\to 0}\lim_{r\to\infty}\frac{1}{V(r)}\int_R |F_\theta(Q;\alpha)-F(Q;\alpha)|^2\,dV_Q=0. \tag{115}$$

Now, by the Schwarz inequality,

$$\begin{aligned}(116)\quad &\left|\frac{1}{V(r)}\int_R F_\theta(P+Q;\alpha)\bar{F}_\theta(Q;\alpha)dV_Q\right.\\ &\left.-\frac{1}{V(r)}\int_R F_\theta(P_1+Q;\alpha)\bar{F}_\theta(Q;\alpha)dV_Q\right|\\ &\leq\left\{G_\theta(O;\alpha)\left(\frac{1}{V(r)}\int_R |F_\theta(P+Q;\alpha)-F_\theta(P_1+Q;\alpha)|^2\,dV_Q\right)\right\}^{\frac12}\\ &\leq\left\{G_\theta(O;\alpha)\left(\frac{1}{V(r)}\int_R dV_Q\left(\frac{1}{V(\theta)}\left[\int_{|\widehat{P+Q,S}|\leq\theta}-\int_{|\widehat{P+Q,S}|\leq\theta}\right]|F(S;\alpha|^2\,dV_S\right)\right.\right.\\ &\quad\times\left.\left.\left(\frac{1}{V(\theta)}\left[\int_{|\widehat{P+Q,S}|\leq\theta}-\int_{|\widehat{P_1+Q,S}|\leq\theta}\right]dV_S\right)\right\}^{\frac12}\right.\\ &\leq G_\theta(O;\alpha)O(|\widehat{PP_1}|^{\frac12}).\end{aligned}$$

It thus follows that if (114) exists for a given θ and all P's with rational coördinates, it exists for that θ and all real P's whatever. We may readily show that

$$G_\theta(O;\alpha)\leq G(O;\alpha). \tag{117}$$

By another use of the Schwarz inequality,

$$\begin{aligned}(118)\quad &\left|\frac{1}{V(r)}\int_R F_\theta(P+Q;\alpha)\bar{F}_\theta(Q;\alpha)dV_Q\right.\\ &\qquad\left.-\frac{1}{V(r)}\int_R F(P+Q;\alpha)\bar{F}(Q;\alpha)dV_Q\right|\\ &\leq\frac{1}{V(r)}\int_R |F_\theta(P+Q;\alpha)-F(P+Q;\alpha)|\;|F_\theta(Q;\alpha)|\,dV_Q\\ &\quad+\frac{1}{V(r)}\int_R |F(P+Q;\alpha)|\;|F_\theta(Q;\alpha)-F(Q;\alpha)|\,dV_Q\\ &\leq\{G(O;\alpha)\frac{1}{V(r)}\int_R |F_\theta(P+Q;\alpha)-F(P+Q;\alpha)|^2\,dV_Q\}^{\frac12}\\ &\quad+\{G(O;\alpha)\frac{1}{V(r)}\int_R |F_\theta(Q;\alpha)-F(Q;\alpha)|^2\,dV_Q\}^{\frac12}.\end{aligned}$$

Combining (115) and (118), we see that except for a set of values of α of zero measure, we have for all P,

$$G(P;\alpha)=\lim_{\theta\to 0}G_\theta(P;\alpha). \tag{119}$$

We thus have an adequate basis for spectrum theory. This will extend, not merely to functions $F(P, \alpha)$ defined as in (88), but to finite sums of such functions. It will even extend to the case of any differentiable chaos $F(P; \alpha)$, for which $F(P+Q; \alpha)\bar{F}(Q; \alpha)$ is an integrable function of α, and for which (115) holds. For a metrically transitive chaos, this latter will be true if

$$\lim_{\theta \to 0} \int_0^1 | F_\theta(P; \alpha) - F(P; \alpha)|^2 \, d\alpha = 0. \tag{120}$$

Under this assumption, we have proved that $F(P)$ has a spectrum, and the same spectrum, for all values of α.

This enables us to answer a question which has been put several times, as to whether there is any relation between the spectrum of a chaos and the distribution of its values. There is no unique relation of the sort. The function

$$\int g(P+Q)\bar{g}(Q)\,dV_Q, \tag{121}$$

where g belongs to L^2, may be so chosen as to represent any Fourier transform of a positive function of L, and if $f(P, Q, M)$ is a bounded step-function, the right-hand side of (97) will clearly be the Fourier transform of a positive function of L. In particular, let $f(P_1, P_2, P_3) = f(P_1)f(P_2)f(P_3)$,

$$F_1(P; \alpha) = \int f_1(\widehat{PP_1})\,d_{P_1}\mathfrak{P}(S; \alpha) \tag{122}$$

and choose $f_1(Q)$ in such a way that

$$\int f_1(P+Q)\bar{f_1}(Q)\,dV_Q = \text{right-hand side of (97)}. \tag{123}$$

Then

$$\int_0^1 (F_1(P; \alpha))^{2n} d\alpha = \left(\int_\infty | f_1(P)|^2 \, dV_P\right)^n (2n-1)(2n-3)\cdots 1 \tag{124}$$

and if $F(P, \alpha)$ is defined as in (96),

$$\int_0^1 (F(P; \alpha))^{2n} d\alpha = \left(\int | f(P)|^2 dV_P\right)^{3n} (6n-1)(6n-3)\cdots 1) \tag{125}$$

so that for all but at most one value of n,

$$\int_0^1 (F_1(P; \alpha)^{2n} d\alpha \neq \int_0^1 (F(P; \alpha))^{2n} d\alpha \tag{126}$$

and we obtain in F and F_1 two chaoses with identical spectra but different distribution functions. On the other hand, if

$$\int_\infty | f_1(P)|^2 \, dV_P = \int_\infty | f_2(P)|^2 \, dV_P, \tag{127}$$

the chaoses

$$\int f_1(\widehat{PP_1})\, d_{P_1}\mathfrak{P}(S;\alpha) \tag{128}$$

and

$$\int f_2(\widehat{PP_1})\, d_{P_1}\mathfrak{P}(S;\alpha) \tag{129}$$

will have the same distribution functions, but may have very different spectra.

11. The discrete chaos.[11] Let us now divide the whole of Euclidean n-space dichotomously into sets $S_{m,n}$, such that every two sets S_{m_1,n_1} and S_{m_2,n_2} have the same measure, and that each $S_{m,n}$ is made up of exactly two non-overlapping sets $S_{m+1,k}$. Let us divide all these sets into two categories, "occupied," and "empty." Let us require that the probability that a set be empty depend only on its measure, and that the probability that two non-overlapping sets be empty be the product of the probabilities that each be empty. Let us assume that both empty and occupied sets exist. Let every set contained in an empty set be empty, while if a set be occupied, let at least one-half always be occupied. We thus get an infinite class of schedules of emptiness and occupiedness, and methods analogous to those of paragraph 6 may be used to map the class of these schedules in an almost everywhere one-one way on the line $(0,1)$ of the variable α, in such a way that the set of schedules for which a given finite number of regions are empty or occupied will have a probability equal to the measure of the corresponding set of values of α.

By the independence assumption, the probability that a given set $S_{m,n}$ be empty must be of the form $e^{-Am(S_{m,n})}$. If $S_{m,n}$ is divided into the 2^ν intervals $S_{m+\nu,n_1}, \cdots, S_{m+\nu,n_{2^\nu}}$ at the ν-th stage of sub-division, the probability that just one is occupied and the rest are empty is

$$2^\nu(1 - \exp(-Am(S_{m,n})/2^\nu)\exp\left(-\frac{2^\nu - 1}{2^\nu}Am(S_{m,n})\right). \tag{130}$$

This contingency at the $\nu + 1$-st stage is a sub-case of this contingency at the ν-th stage. If we interpret probability to mean the same thing as the measure of the corresponding set of α's, then by monotone convergence, the probability that at every stage, all but one of the subdivisions of $S_{m,n}$ are empty, while the remaining one is occupied, will be the limit of (130), or

$$Am(S_{m,n})\exp(-Am(S_{m,n})). \tag{131}$$

[11] The ideas of this paragraph are related to discussions the author has had with Professor von Neumann, and the main theorem is equivalent to one enunciated by the latter.

Such a series of stages of subdivision will have as its occupied regions exactly those which contain a given point.

The probability that the occupied regions are exactly those which contain two points is the probability that each half of $S_{m,n}$ contain exactly one point, plus the probability that one-half is empty, and that in the occupied half, each quarter will contain exactly one point, plus and so on. This will be

$$(132)\quad \left\{ \frac{Am(S_{m,n})}{2} \exp\left(-\frac{Am(S_{m,n})}{2}\right) \right\}^2 + 2\exp\left(-\frac{Am(S_{m,n})}{2}\right) \left\{ \frac{Am(S_{m,n})}{2} \exp\left(-\frac{Am(S_{m,n})}{4}\right)^2 + \cdots = \exp(-Am(S_{m,n})m(S_{m,n}))^2(\tfrac{1}{4}+\tfrac{1}{8}+\cdots) = \frac{(Am(S_{m,n}))^2}{2} \exp(-Am(S_{m,n})).$$

If the probability that the occupied regions are exactly those containing $k-1$ points is

$$\frac{(Am(S_{m,n}))^{k-1}}{(k-1)!} \exp(-Am(S_{m,n}))$$

then a similar argument will show that the probability that the occupied regions are exactly those containing k points will be

$$(133)\quad \sum_{j=1}^{k-1} \frac{1}{j!} \frac{1}{(k-j)!} (Am(S_{m,n}))^k \exp(-Am(S_{m,n})) \left(1 + \frac{1}{2^{k-1}} + \frac{1}{4^{k-1}} + \cdots\right)$$
$$= \frac{1}{k!}(2^k - 2)\left(\frac{1}{1-\frac{1}{2^{k-1}}}\right)(Am(S_{m,n}))^k \exp(-Am(S_{m,n}))$$
$$= \frac{1}{k!}(Am(S_{m,n}))^k \exp(-Am(S_{m,n})).$$

Thus by mathematical induction, the probability that the occupied regions are exactly those containing k points will be

$$\frac{1}{k!}(Am(S_{m,n}))^k \exp(-Am(S_{m,n}))$$

and the sum of this for all values of k will be

$$(134)\qquad \sum_0^\infty \frac{1}{k!}(Am(S_{m,n}))^k \exp(-Am(S_{m,n})) = 1.$$

In other words, except for a set of contingencies of probability zero, the occupied regions will be exactly those containing a given finite number of points.

We may proceed at once from the fact that the probability that a set S_1 contains exactly k points is

$$\frac{1}{k!}(Am(S_1))^k e^{-Am(S_1)}$$

while the probability that the non-overlapping set S_2 contains exactly k points is

$$\frac{1}{k!}(Am(S_2))^k e^{-Am(S_2)}$$

to the fact that the probability that the set $S_1 + S_2$ contains exactly k points is

$$(135) \qquad \sum_0^k \frac{1}{j!}\frac{1}{(k-j)!}(Am(S_1))^j(Am(S_2))^{k-j}e^{-Am(S_1+S_2)} = \frac{1}{k!}(Am(S_1+S_2))^k e^{-Am(S_1+S_2)}.$$

From this, by monotone convergence, it follows at once that the probability that any set S which is the sum of a denumerable set of our fundamental regions $S_{m,n}$ should contain exactly k points is

$$\frac{1}{k!}(Am(S))^k e^{-Am(S)}.$$

It is then easy to prove this for all measurable sets S.

We are now in a position to prove that the additive functional $\mathfrak{D}(S;\alpha)$, consisting in the number of points in the region S on the basis of the schedule corresponding to α, is a homogeneous metrically transitive chaos. The rôle which continuity filled in paragraphs 6 and 7, of allowing us to show that $\mathfrak{F}_{y_1,\ldots,y_n}(S;\alpha)$ was measurable in $y_1,\cdots,y_n$ and α, is now filled by the fact that the probability that any of the points in a region lie within a very small distance of the boundary, is for any Jordan region the probability that a small region be occupied, and is small. The metric transitivity of the chaos results as before from the independence of the distribution in non-overlapping regions.

The discrete or Poisson chaos which we have thus defined is the chaos of an infinite random shot pattern, or the chaos of the gas molecules in a perfect gas in statistical equilibrium according to the old Maxwell statistical mechanics. It also has important applications to the study of polycrystalline aggregates, and to similar physical problems.

Two important formulae are

$$(136) \qquad \int_0^1 \mathfrak{D}(S;\alpha)\,d\alpha = e^{-Am(S)}\sum_1^\infty \frac{k}{k!}(Am(S))^k = Am(S),$$

and

$$(137)\quad \int_0^1 (\mathfrak{D}(S;\alpha))^2 d\alpha = e^{-Am(S)} \sum_1^\infty \frac{k^2}{k!} (Am(S))^k = (Am(S))^2 + Am(S).$$

Let it be noted that if we define

$$(138)\qquad \int f(P)\, d_P \mathfrak{D}(S;\alpha)$$

for a measurable step-function $f(P)$ as in (70), by

$$(139)\qquad \sum_1^N f_n \mathfrak{D}(S_n;\alpha),$$

(72) is replaced by

$$(140)\quad \int_0^1 d\alpha \,\Big| \int f(P)\, d_P \mathfrak{D}(S;\alpha) - A\int_\infty f(Q)\, dV_Q \Big|^2$$
$$= \sum_{m=1}^N \sum_{n=1}^N \int_0^1 d\alpha\, f_m \bar{f}_n \mathfrak{D}(S_m;\alpha)\mathfrak{D}(S_n;\alpha)$$
$$- 2\Re\{\bar{A}\sum_{m=1}^\infty \int_0^1 d\alpha\, f_m \mathfrak{D}(S_m;\alpha)\int_\infty f(Q)\, dV_Q + \Big| A\int_\infty f(Q)\, dV_Q\Big|^2\}$$
$$= \sum_{m=1}^N |f_m|^2 Am(S_m)$$
$$= A\int_\infty |f(P)|^2\, dV_P.$$

Thus the transformation from $f(P)$ as a function of P, to

$$(141)\qquad \int f(P)\, d_P \mathfrak{D}(S;\alpha) - A\int_\infty f(Q)\, dV_Q$$

as a function of α, retains distance in Hilbert space, apart from a constant factor, and if $f(P)$ belongs to L and L^2 simultaneously, and $\{f_n(P)\}$ is a sequence of step-functions converging in the mean both in the L sense and in the L^2 sense to $f(P)$, we may *define*

$$(142)\quad \int f(P)\, d_P \mathfrak{D}(S;\alpha) = A\int_\infty f(Q)\, dV_Q$$
$$+ \underset{n\to\infty}{\text{l. i. m.}} \Big(\int f_n(P)\, d_P \mathfrak{D}(S;\alpha) - A\int_\infty f_n(Q)\, dV_Q \Big).$$

As in the case of (73), this definition is substantially unique. We may prove the analogue of (93) in exactly the same way as (93) itself, and shall obtain

$$(143)\quad \lim_{r\to\infty} \frac{1}{V(r)} \int_R \Big\{ \int_\infty (f(P+\widehat{Q)M};\alpha)\, d_M \mathfrak{D}(S;\alpha) - A\int_\infty f(M)\, dV_M \Big\}$$
$$\times \Big\{ \int_\infty \bar{f}(\widehat{QM};\alpha)\, d_M \mathfrak{D}(S;\alpha) - A\int_\infty \bar{f}(M)\, dV_M \Big\}\, dV_Q$$
$$= A^2 \int_\infty f(P+Q)\bar{f}(Q)\, dV_Q.$$

As we may see by appealing to the theory of spectra, one interpretation of this in the one-dimensional case is the following: *If a linear resonator be set into motion by a haphazard series of impulses forming a Poisson chaos, the effect, apart from that of a constant uniform stream of impulses, will have the same power spectrum as the energy spectrum of the response of the resonator to a single impulse.*

12. The weak approximation theorem for the polynomial chaos. We wish to show that the chaoses of paragraph 9 are in some sense everywhere dense in the class of all metrically transitive homogeneous chaoses. We shall show that if $\mathfrak{F}(S;\alpha)$ is any homogeneous chaos in n dimensions, there is a sequence $\mathfrak{F}_k(S;\alpha)$ of polynomial chaoses as defined in paragraph 9, such that if $S_1, \cdots, S_\nu$ is any finite assemblage of bounded measurable sets in n-space selected from among a denumerable set, and

$$\int_0^1 |\mathfrak{F}(S_\lambda;\alpha)|^\mu \, d\alpha < \infty \qquad (\lambda = 1, 2, \cdots, \nu) \tag{144}$$

is finite, then

$$\int_0^1 \mathfrak{F}(S_1;\alpha) \cdots \mathfrak{F}(S_\nu;\alpha)\, d\alpha = \lim_{n\to\infty} \int_0^1 \mathfrak{F}_n(S_1;\alpha) \cdots \mathfrak{F}_n(S_\nu;\alpha)\, d\alpha. \tag{145}$$

We first make use of the fact that if the probability that a quantity u be greater in absolute value than A, be less than

$$\frac{2}{\sqrt{2\pi B}} \int_A^\infty e^{-(u^2/2B)} du, \tag{146}$$

then if $\psi(u)$ is any even measurable function bounded over $(-\infty, \infty)$, we may find a polynomial $\psi_\epsilon(u)$, such that the mean value of

$$|\psi(u) - \psi_\epsilon(u)|^n, \tag{147}$$

which will be

$$\frac{1}{\sqrt{2\pi B}} \int_{-\infty}^\infty |\psi(u) - \psi_\epsilon(u)|^n \, e^{-(u^2/2B)} du, \tag{148}$$

is less than ϵ. Since it is well known that if $\phi(u)$ is a continuous function vanishing outside a finite interval, and

$$\sum_1^\infty A_n H_n(u) e^{-(u^2/2)} \tag{149}$$

is the series for $\phi(u)$ in Hermite functions, then we have uniformly

$$\phi(u) = \lim_{t \to 1-0} \sum_1^\infty A_n t^n H_n(u) e^{-(u^2/2)}, \tag{150}$$

to establish the existence of $\psi_\epsilon(u)$, we need only prove it in the case in which

$$\psi(u) = u^k e^{-Cu^2} \tag{151}$$

for an arbitrarily small value of C: as for example for $C = 1/4n_1B$. We shall then have

$$\left| \psi(u) - u^k \sum_0^N \frac{(cu^2)^k}{k!} \right| \leq |u|^k \sum_0^\infty \frac{(cu^2)^k}{k!} = |u|^k e^{u^2/4n_1B}, \tag{152}$$

so that by dominated convergence, and if we take N large enough, we may make

$$\frac{1}{\sqrt{2\pi B}} \int_{-\infty}^{\infty} | \psi(u) - \psi_\epsilon(u) |^n e^{-(u^2/2B)} du < \epsilon \qquad (n \leq n_1). \tag{153}$$

Now let

$$\psi_K(u) = \begin{cases} 0 & (|u| < K); \\ 1 & (|u| \geq K); \end{cases} \tag{154}$$

and let us put

$$\mathfrak{G}(P;\alpha) = \frac{1}{V(r)} \mathfrak{P}(S;\alpha) \quad (S = \text{interior of } |\widehat{QP}| \leq r). \tag{155}$$

The chaos

$$\mathfrak{L}(P;\alpha) = \psi_K(\mathfrak{G}(P;\alpha)) \tag{156}$$

may then be approximated by polynomial chaoses in such a way as to approximate simultaneously to all polynomials in $\mathfrak{L}(P;\alpha)$ by corresponding polynomials in the approximating chaoses. Since the distribution of the values of $\mathfrak{G}(P;\alpha)$ will be Gaussian, with a root mean square value proportional to a power of r, and $\mathfrak{G}(P;\alpha)$ will be independent in spheres of radius η about two points P_1 and P_2 more remote from each other than $2r + 2\eta$, it follows that if we take K to be large enough, we may make the probability that $\mathfrak{L}(P;\alpha)$ differs from 0 between two spheres of radii respectively $r + \eta$ and H about a given point where it differs from 0, as small as we wish.

We now form the new chaos

$$\int_{|\widehat{PQ}| < x} \mathfrak{L}(Q;\alpha) dV_Q, \tag{157}$$

which we may also approximate, with all its polynomial functionals, by a sequence of polynomial chaoses. The use of polynomial approximations tending boundedly to a step function over a finite range will show us that this is also true of the chaos determined by

(158) $$\psi_\gamma\Big(\int_{|\widehat{PQ}|<x} \mathcal{L}(Q;\alpha)\,dV_Q = \mathfrak{M}(P;\alpha).$$

By a proper choice of the parameters, this can be made to have arbitrarily nearly all its mass uniformly distributed over regions arbitrarily near to arbitrarily small spheres, all arbitrarily remote from one another, except in an arbitrarily small fraction of the cases. We then form

(159) $$\frac{1}{(2\pi k)^{n/2}}\int_\infty (\mathfrak{M}(Q;\alpha)+\delta)\exp\left(-\frac{|\widehat{PQ}|^2}{2k}\right)dV_Q = \mathfrak{N}(P;\alpha)$$

where δ is taken to be very small. This chaos again, as far as all its polynomial functionals are concerned, will be approximable by polynomial chaoses. Since it is bounded away from 0 and ∞, and since over such a range the function $1/x$ may be approximated uniformly by polynomials, it follows that in our sense,

(160) $$1/\mathfrak{N}(P;\alpha)$$

is approximable by polynomial chaoses.

If $\varpi(P)$ is any measurable function for which arbitrarily high moments are always finite, it is easy to show that

(161) $$\frac{1}{(2\pi k)^{n/2}}\int_\infty \varpi(Q)(\mathfrak{M}(Q;\alpha)+\delta)\exp\left(-\frac{|\widehat{PQ}|^2}{2k}\right)dV_Q = \mathfrak{W}(P;\alpha)$$

is approximable by polynomial chaoses. Multiplying expressions (160) and (161), it follows that

(162) $$\mathfrak{W}(P;\alpha)/\mathfrak{N}(P;\alpha) = \mathfrak{U}(P;\alpha)$$

is approximable by polynomial chaoses.

If A is a large enough constant, depending on the choice of the constant ϵ, we have

(163) $$\frac{1}{(2\pi k)^{n/2}}\int_{|P|>A}\exp\left(-\frac{|P|^2}{2k}\right)dV_P$$
$$= \int_A^\infty x^n e^{-(x^2/2k)}\,dx\Big/\int_0^\infty x^n e^{-(x^2/2k)}\,dx$$
$$< \frac{1}{(2\pi k)^{n/2}}\exp\left(-\frac{(A-\epsilon)^2}{2k}\right).$$

Thus by the proper choice of the parameters of $\mathfrak{M}(P;\alpha)$, if we take k small enough and then δ small enough, the chaos (162) will consist as nearly as we wish, from the distribution standpoint, of an infinite assemblage of convex cells of great minimum dimension, in each of which the function $\varpi(P)$ is repeated, with the origin moved to some point remote from the boundary.

Now let $\mathfrak{F}(S;\alpha)$ be a metrically transitive homogeneous chaos. Let us form

$$\mathfrak{F}(r;S;\alpha)=\frac{1}{V(r)}\int_R \mathfrak{F}_{x_1,\ldots,x_n}(S;\alpha)\,dx_1\cdots dx_n. \tag{164}$$

Clearly by the fundamental theorem of the calculus, over any finite region in $(x_1,\cdots,x_n)$, we shall have for almost all points and almost all values of α,

$$\mathfrak{F}(S;\alpha)=\lim_{r\to 0}\mathfrak{F}(r;S;\alpha); \tag{165}$$

and if (144) holds, it is easy to show that

$$\int_0^1 |\,\mathfrak{F}(r;S;\alpha)|^n d\alpha < \text{const.} \tag{166}$$

From this it follows that

$$\lim_{r\to 0}\int_0^1 |\,\mathfrak{F}(r;S;\alpha)-\mathfrak{F}(S;\alpha)|^n d\alpha=0 \tag{167}$$

and by the ergodic theorem, except for a set of values of α of zero measure, as r tends to 0 through a denumerable set of values,

$$\lim_{r\to 0}\frac{1}{V(r)}\int_R |\,\mathfrak{F}_{x_1,\ldots,x_n}(r;S;\alpha) - \mathfrak{F}_{x_1,\ldots,x_n}(S;\alpha)|^n dx_1\cdots dx_n=0. \tag{168}$$

With this result as an aid, enabling us to show that the distribution of $\mathfrak{F}(S;\alpha)$ is only slightly affected by averaging within a small sphere with a given radius, or even within any small region near enough to a small sphere with a given radius, we may proceed as in (161) and (162) and form the chaos

$$\mathfrak{F}_k(S;\alpha)=\frac{1}{\mathfrak{N}(\mathfrak{P};\alpha)(2\pi k)^{n/2}}\int_\infty \mathfrak{F}_{x_1,\ldots,x_n}(S;\beta) \times(\mathfrak{M}(x_1,\cdots,x_n;\alpha)+\delta)\exp\frac{(-\sum_1^n x_j^2)}{2k}\,dx_1\cdots dx_n. \tag{169}$$

For almost all β, in each of the large cells of this chaos, (169) will have as nearly as we wish the same distribution as some $\mathfrak{F}_{x_1, \ldots, x_n}(S;\alpha)$, where $(x_1, \cdots, x_n)$ lies in the interior of the cell, remote from the boundary. These cells may so be determined that except for those filling an arbitrarily small proportion of space, all are convex regions with a minimum dimension greater than some given quantity.

To establish (143), it only remains to show that the average of a quantity depending on a chaos over a large cell tends to the same limit as its average over a large sphere. To show this, we only need to duplicate the argument of paragraph 4, where we prove the multidimensional ergodic theorem, for large pyramids with the origin as a corner, instead of for large spheres about the origin. We may take the shapes and orientations of these pyramids to form a denumerable assemblage, from which we may pick a finite assemblage which will allow us to approach as closely as we want to any cell for which the ratio of the maximum to the minimum distance from the origin within it does not exceed a given amount. It is possible to show that by discarding cells whose measure is an arbitrarily small fraction of the measure of all space, the remaining cells will have this property.

13. The physical problem. The transformation of a chaos. The statistical theory of a homogeneous medium, such as a gas or liquid, or a field of turbulence, deals with the problem, given the statistical configuration and velocity distribution of the medium at a given initial time, and the dynamical laws to which it is subject, to determine the configuration at any future time, with respect to its statistical parameters. This of course is not a problem in the first instance of the history of the individual system, but of the entire ensemble, although in proper cases it is possible to show that almost all systems of the ensemble do actually share the same history, as far as certain specified statistical parameters are concerned.

The dynamical transformations of a homogeneous system have the very important properties, that they are independent of any choice of origin in time or in space. Leaving the time variable out of it, for the moment, the simplest space transformations of a homogeneous chaos $\mathfrak{F}(S;\alpha)$ which have this property are the polynomial transformations which turn it into

$$
\begin{aligned}
&K_0 + \int K_1(x_1 - y_1, x_2 - y_2, \cdots, x_n - y_n)\,\mathfrak{F}_{y_1, \ldots, y_n}(S;\alpha)\,dy_1 \cdots dy_n \\
&+ \cdots \qquad\qquad (170) \\
&+ \int \cdots \int K_\nu(x_1 - y_1^{(1)}, \cdots, x_n - y_n^{(1)}, \cdots, x_1 - y_1^{(\nu)}, \cdots, \\
&\quad x_n - y_n^{(\nu)})\,\mathfrak{F}_{y_1^{(1)}, \ldots, y_n^{(1)}}(S;\alpha) \cdots \\
&\qquad\qquad \mathfrak{F}_{y_1^{(\nu)}, \ldots, y_n^{(\nu)}}(S;\alpha)\,dy_1^{(1)} \cdots dy_n^{(\nu)}.
\end{aligned}
$$

These are a sub-class of the general class of polynomial transformations

$$
\begin{aligned}
(171)\qquad & K_0 + \int K_1(x_1, \cdots, x_n; y_1, \cdots, y_n) \mathfrak{F}_{y_1, \ldots, y_n}(S; \alpha)\, dy_1 \cdots dy_n \\
& + \cdots \\
& + \int \cdots \int K_\nu(x_1, \cdots, x_n; y_1^{(1)}, \cdots, y_n^{(1)}; \cdots; y_1^{(\nu)}, \cdots, y_n^{(\nu)}) \\
& \quad \times \mathfrak{F}_{y_1^{(1)}, \ldots, y_n^{(1)}}(S; \alpha) \cdots \mathfrak{F}_{y_1^{(\nu)}, \ldots, y_n^{(\nu)}}(S; \alpha)\, dy_1^{(1)} \cdots dy_n^{(\nu)}.
\end{aligned}
$$

If a transformation of type (171) is invariant with respect to position in space, it must belong to class (170). On the other hand, in space of a finite number of dimensions and in any of the ordinary spaces of an infinite number of dimensions, polynomials are a closed set of functions, and hence every transformation may be approximated by a transformation of type (171).

A polynomial transformation such as (170) of a polynomial chaos yields a polynomial chaos. If then we can approximate to the state of a dynamical system at time 0 by a polynomial chaos, and approximate to the transformation which yields its status at time t by a polynomial transformation, we shall obtain for its state at time t, the approximation of another polynomial chaos. The theory of approximation developed in the last section will enable us to show this.

On the other hand, the transformation of a dynamical system induced by its own development is infinitely subdivisible in the time, and except in the case of linear transformations, this is not a property of polynomial transformations. Furthermore, when these transformations are non-linear, they are quite commonly not infinitely continuable in time. For example, let us consider the differential equation

$$
(172)\qquad \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0.
$$

This corresponds to the history of a space-distribution of velocity transferred by particles moving with that velocity. Its solutions are determined by the equation

$$
(173)\qquad u(x, t) = u(x - tu(x, t), 0),
$$

or if ψ is the inverse function of $u(x, 0)$,

$$
(174)\qquad x - tu(x, t) = \psi(u(x, t)).
$$

Manifestly, if two particles with different velocities are allowed to move long enough to allow their space-time paths to cross, $u(x, t)$ will cease to exist as a single-valued function. This will always be the case for *some* value of t if $u(x, 0)$ is not constant, and for almost all values of t and α if it is a polynomial chaos.

By Lagrange's formula, (174) may be inverted into

$$u(x, t) = \sum_{0}^{\infty} \frac{(-t)^n}{n!} \cdot \frac{\partial^n}{\partial x^n} \{(u(x, 0))^n\}. \tag{175}$$

In a somewhat generalized sense, the partial sums of this formally represent polynomial transformations of the initial conditions. However, it is only for a very special sort of bounded initial function, and for a finite value of the time, that they converge. It is only in this restricted sense that the polynomial transformation represents a true approximation to that given by the differential equation.

It will be seen that the useful application of the theory of chaos to the study of particular dynamical chaoses involves a very careful study of the existence theories of the particular problems. In many cases, such as that of turbulence, the demands of chaos theory go considerably beyond the best knowledge of the present day. The difficulty is often both mathematical and physical. The mathematical theory may lead inevitably to a catastrophe beyond which there is no continuation, either because it is not the adequate presentation of the physical facts; or because after the catastrophe the physical system continues to develop in a manner not adequately provided for in a mathematical formulation which is adequate up to the occurrence of the catastrophe; or lastly, because the catastrophe does really occur physically, and the system really has no subsequent history. The hydrodynamical investigations required in the case of turbulence are directly in the spirit of the work of Oseen and Leray, but must be carried much further.

The study of the history of a mechanical chaos will then proceed as follows: we first determine the transformation of the initial conditions generated by the dynamics of the ensemble. We then determine under what assumptions the initial conditions admit of this transformation for either a finite or an infinite interval of time. Then we approximate to the transformation for a given range of values of the time by a polynomial transformation. Then, having regard to a definition of distance between two functions determined by the transformation, we approximate to the initial chaos by a polynomial chaos. Next we apply the polynomial transformation to the polynomial chaos, and obtain an approximating polynomial chaos at time t. Finally, we apply our algorithm of the pure chaos to determine the averages of the statistical parameters of this chaos, and express these as functions of the time.

The results of such an investigation belong to a little-studied branch of statistical mechanics: the statistical mechanics of systems not in equilibrium. To study the classical, equilibrium theory of statistical mechanics by the methods of chaos theory is not easy. As yet we lack a method of representing all forms of homogeneous chaos, which will tell us by inspection when two differ merely by an equimeasure transformation of the parameter of distribution. In certain cases, in which the equilibrium is stable, the study of the history of a system with an arbitrary initial chaos will yield us for large values of t an approximation to equilibrium, but this will often fail to be so, particularly in the case of differentiable chaoses, or the only equilibrium may be that in which the chaos reduces to a constant.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.