

UNIVERSAL
LIBRARY

OU_156748

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 519.1

Accession No. 51900

Sy 98A

Author Symposia in applied mathematics.

Title

-ics. Proceedings.

Applied probability - 1957.

This book should be returned on or before the date last marked below.

PROCEEDINGS OF
SYMPOSIA IN APPLIED MATHEMATICS

VOLUME VII

APPLIED PROBABILITY

McGRAW-HILL BOOK COMPANY, INC.

NEW YORK TORONTO LONDON

1957

FOR THE AMERICAN MATHEMATICAL SOCIETY

80 WATERMAN STREET, PROVIDENCE, RHODE ISLAND

PROCEEDINGS OF THE
SEVENTH SYMPOSIUM IN APPLIED MATHEMATICS
OF THE AMERICAN MATHEMATICAL SOCIETY

Held at the Polytechnic Institute of Brooklyn
April 14-15, 1955

COSPONSORED BY
THE OFFICE OF ORDNANCE RESEARCH

L. A. MacColl
EDITOR

EDITORIAL COMMITTEE

R. V. Churchill
A. E. Heins
F. J. Murray

Prepared by the American Mathematical Society under Contract
No. DA-19-020-ORD-3538 with the Ordnance Corps, U.S. Army.

Copyright © 1957 by the McGraw-Hill Book Company, Inc. Printed in the United States of America. All rights reserved except those granted to the United States Government. Otherwise, this book, or parts thereof, may not be reproduced in any form without permission of the publishers.

Library of Congress Catalog Card Number 50-1183

CONTENTS

EDITOR'S PREFACE	v
Brownian Motion Depending on n Parameters: The Particular Case $n = 5$	1
By PAUL LÉVY	
A New Look at the First Boundary-value Problem	21
By J. L. DOOB	
On Boundaries Defined by Stochastic Matrices (Abstract)	35
By WILLIAM FELLER	
On the Application of Functional Calculus to the Statistical Theory of Turbulence	41
By EBERHARD HOPF	
Stochastic Processes of Astronomical Interest	51
By GUIDO MÜNCH	
The Singularity in the Spectrum of Homogeneous Turbulence	67
By G. K. BATCHELOR	
Probability in Classical Physics	73
By MARK KAC	
Infinite Models in Physics.	87
By S. M. ULAM	
Quantum Theory and the Foundations of Probability.	97
By B. O. KOOPMAN	
INDEX	103

EDITOR'S PREFACE

The Seventh Symposium in Applied Mathematics, sponsored by the American Mathematical Society and the Office of Ordnance Research, and devoted to *Mathematical Probability and Its Applications*, was held at the Polytechnic Institute of Brooklyn on April 14 and 15, 1955. This volume contains the papers (one in abstract form) which were presented at the Symposium.

Prolonged consideration by the members of the Program Committee, under the chairmanship of Dr. H. W. Bode, resulted in the decision that the Symposium should be concerned with three principal themes, *viz.*, *The Theory of Diffusion*, *The Theory of Turbulence*, and *Probability in Classical and Modern Physics*. However, it was the intention of the Committee that these terms should be interpreted broadly and that the speakers should avail themselves of considerable freedom in determining the actual contents of their papers. In particular, it was understood that the term "theory of diffusion" was to be interpreted so as to cover a wide variety of relations between probability and differential equations.

The three themes were dealt with in the order in which they have been mentioned, and the papers appear here in the order in which they were given.

Many individuals have participated, directly and indirectly, in the work of preparing this volume. The editor wishes to express here his sincere thanks to all of these collaborators. The advice and encouragement given by Professor R. V. Churchill, Chairman of the Editorial Committee for the Proceedings of Symposia in Applied Mathematics, has been particularly helpful. All who participated in the Symposium are indebted to the McGraw-Hill Book Company, Inc., which, beginning with the Proceedings of the Symposium on Elasticity, has undertaken the task of bringing the Proceedings of these Symposia on Applied Mathematics to the scientific public in book form.

L. A. MACCOLL

Editor

BROWNIAN MOTION DEPENDING ON n PARAMETERS: THE PARTICULAR CASE $n = 5$

BY

PAUL LÉVY

INTRODUCTION

Let E_n be the Euclidean n -dimensional space, A and B points of E_n , and $X(A)$ the Brownian function, *i.e.*, the Gaussian random function (r.f.) defined up to an additive constant by the formula

$$(1) \quad X(B) - X(A) = \xi[r(A, B)]^{\frac{1}{2}},$$

in which $r(A, B)$ is the distance between A and B , and ξ is a reduced Gaussian variable.

Let us now suppose that a set $\mathcal{E} \subset E_n$ is given and that the values of $X(A)$ are known at every $A \in \mathcal{E}$. At the other points $X(A)$ has a conditional distribution, which we shall write in the canonical form $m + \sigma\xi$, where $m = E\{X(A)|\mathcal{E}\}$ and $\sigma = \sigma\{X(A)|\mathcal{E}\}$ are the conditional expectation and the conditional standard deviation. If we obtain new information, *i.e.*, if $\mathcal{E}$ increases, σ is nondecreasing.

Now let Ω_t be the surface of a sphere with center O and radius t , and let $\bar{M}(t)$ be the mean value of $X(A)$ on Ω_t . Although $\bar{M}(t)$ is not a well-defined r.f., $M(t) = \bar{M}(t) - X(O)$ is a well-defined Gaussian r.f. Thus, to redefine $M(t)$, it is sufficient to know its covariance $\Gamma_n(t_1, t_2)$.

Let $\mathcal{E}_t$ denote the entire region outside of Ω_t , that is, $r(O, A) \geq t$, and let

$$(2) \quad \sigma\{X(O)|\mathcal{E}_t\} = c_n t^{\frac{1}{2}}, \quad \sigma\{X(O)|\Omega_t\} = c'_n t^{\frac{1}{2}}.$$

Since σ is a nonincreasing function of $\mathcal{E}$, $c_n t^{\frac{1}{2}}$ is the minimum of $\sigma\{X(O)|\mathcal{E}\}$ if $\mathcal{E} \subset \Omega_t$. Since E_n is a section of E_{n+1} , one obviously has

$$(3) \quad c_n \leq c'_n, \quad c_n \geq c_{n+1}, \quad c'_n \geq c'_{n+1}.$$

Then, in the limiting case $n = \infty$ (Hilbert space), c_n and c'_n have limits c_ω and c'_ω . These limits are

$$(4) \quad c_\omega = 0, \quad c'^2_\omega = 1 - 2^{-\frac{1}{2}}.$$

It is easy to find the value of c'_ω but not easy to obtain the values of c_n and to prove that $\lim c_n = 0$. Since $m\{X(O)|\mathcal{E}_t\}$ is obviously a linear function of the values of $\bar{M}(u)$ in $(0, t)$, we were led to consider the continuation to the left of $\bar{M}(u)$. Thus the first problem is to define the covariance $\Gamma_n(t_1, t_2)$.

If n is even, the covariance is an elliptic integral, and the study of $M(t)$ appears to be very difficult. If n is odd ($n = 2p + 1$), then Γ_n is a rational function of $t = \min(t_1, t_2)$ and $t' = \max(t_1, t_2)$. Thus one can obtain an

explicit expression for $M(t)$ and solve all problems concerning the continuation of $\tilde{M}(t)$ to the left and to the right.

But the subject is wide, and this is only the first of several papers.† The second will be presented at the next Berkeley Symposium. In this first paper, the particular case $n = 5$ will be completely described. The case $n = 3$ is too simple to give satisfactory insight into the general theory. The value $n = 5$ presents for the first time a very curious circumstance: $M(t)$ satisfies two different stochastic differential equations; the first determines the continuation to the right, and the second the continuation to the left. These two equations are related to two analytic expressions for $M(t)$.

The author was very surprised to notice these facts, and thinks it worthwhile to begin with a preliminary chapter in which a quite analogous circumstance appears in a very simple case. After this chapter, it will be easier to follow the solution of our problem in the case $n = 5$. Brief indications on the general case $n = 2p + 1$ will be given in Sec. 2.3 and in the footnotes.

1. A SIMPLE CLASS OF GAUSSIAN RANDOM FUNCTIONS

1.1. The Greek letters ξ and η , with or without subscripts, will always denote reduced Gaussian variables. The well-known Brownian (or Bachelier-Wiener) function will be denoted by

$$(5) \quad X(t) = \int_0^t \xi_u(du)^{\frac{1}{2}}, \quad (t > 0).$$

This formula has the same meaning as

$$(6) \quad \delta X(t) = \xi_t(dt)^{\frac{1}{2}},$$

where ξ_t depends on t and dt and the Cauchy condition $X(0) = 0$. Then, if $X_u = \lambda \xi_u$, with λ a constant, we may write

$$\int_0^t f(u) X_u(du)^{\frac{1}{2}} = \lambda \int_0^t f(u) dX(u) = \lambda f(t) X(t) - \lambda \int_0^t X(u) df(u).$$

Now let

$$(7) \quad \Psi(t) = \int_0^t (tX_u + uY_u)(du)^{\frac{1}{2}}, \quad (t > 0),$$

where X_u and Y_u are joint Gaussian variables, with $E(X_u) = E(Y_u) = 0$, and

$$(8) \quad E(X_u^2) = a, \quad E(X_u Y_u) = b, \quad E(Y_u^2) = c,$$

where a, b, c are independent of u , and necessarily

$$(9) \quad a \geq 0, \quad c \geq 0, \quad b^2 \leq ac;$$

moreover, unless $\Psi(t)$ is identically zero,

$$(9a) \quad a + c > 0.$$

† The most important results, without proofs, have already been given in C. R. Acad. Sci. Paris vol. 239 (1954) pp. 1181, 1584; vol. 240 (1955) pp. 1043, 1308. This is the first part of the detailed statement.

If $b^2 = ac$, we may write $\Psi(t)$ in the simpler form

$$(10) \quad \Psi(t) = \int_0^t (\lambda t + \mu u) \xi_u(du)^{\frac{1}{2}}, \quad (\lambda^2 = a, \lambda\mu = b, \mu^2 = c).$$

Obviously, if either ξ_u is replaced by $-\xi_u$ or λ and μ are replaced by $-\lambda$ and $-\mu$, then we have the same random function.† Then we shall assume

$$(11) \quad s = \lambda + \mu \geq 0,$$

without loss of generality. In the case (10) we have

$$\delta\Psi(t) = \lambda dt \int_0^t \xi_u(du)^{\frac{1}{2}} + st\xi_t(dt)^{\frac{1}{2}}, \quad (dt > 0),$$

and, if we use (5), then

$$(12) \quad \delta\Psi(t) = \lambda X(t) dt + st \delta X(t)$$

even with $dt < 0$. This formula and the Cauchy condition $\Psi(0) = 0$ are equivalent to (10).

If $s = 0$, then $\Psi(t)$ is differentiable, with derivative $\lambda X(t)$. In this case, $\Psi(t)$, is a Markovian function of order 2, *i.e.*, the two joint functions $\Psi(t)$ and $\Psi'(t)$ form a Markovian system. In the general case, $\Psi(t)$ and $X(t)$ form a Markovian function. Also $\Psi(t)$ is a Markovian function in the trivial cases $\lambda = 0$ and $\mu = 0$ and also if $k = \lambda/(\lambda + \mu) = -1$ or -2 . [This will result from formula (28).]

1.2. A Gaussian function is defined if its covariance is known. If

$$t = \min(t_1, t_2), \quad t' = \max(t_1, t_2),$$

then the covariance of $\Psi(t)$ can be written

$$\begin{aligned} \gamma(t_1, t_2) &= \int_0^t E[(tX_u + uY_u)(t'X_u + uY_u)] du \\ &= t' \int_0^t (at + bu) du + \int_0^t (btu + cu^2) du. \end{aligned}$$

Finally, setting

$$(13) \quad \alpha = a + \frac{b}{2}, \quad \beta = \frac{b}{2} + \frac{c}{3},$$

we obtain

$$(14) \quad \gamma(t_1, t_2) = \alpha t^2 t' + \beta t^3, \quad [t = \min(t_1, t_2), t' = \max(t_1, t_2)].$$

In the case (10), the values of α and β are

$$(15) \quad \alpha = \lambda^2 + \frac{\lambda\mu}{2}, \quad \beta = \frac{\lambda\mu}{2} + \frac{\mu^2}{3}.$$

† We consider a random function (r.f.) as well defined if we know the set of all possible functions and the probability distribution in that set. If one of these functions is considered apart from the others, we shall say "an individual function is known (or given)." Thus, if ξ_u is replaced by $-\xi_u$, we change the individual function but the r.f. is the same.

One might have thought the $\Psi(t)$ form a family of r.f. depending on three parameters, a, b, c . We now see that we have only two parameters, α and β .

1.3. THEOREM 1. 1°. *The function $\gamma(t_1, t_2)$ in (14) is a covariance if, and only if,*

$$(16) \quad 3\alpha + \beta \geq 0, \quad 3\beta + \alpha \geq 0.$$

2°. *In this case $\gamma(t_1, t_2)$ is the covariance of a function $\Psi(t)$, which can, in general, be represented in the form (7) depending on one parameter and also represented in two distinct ways in the form (10). The only exception is the case*

$$(17) \quad (3\alpha + \beta)(3\beta + \alpha) = 0,$$

where $\Psi(t)$ has only one representation, which may be written in the form (10).

Now, to exclude the case $\Psi(t) \equiv 0$, we may add the condition

$$(18) \quad \alpha + \beta = \frac{\gamma(t, t)}{t^3} = E \left[\frac{\Psi^2(t)}{t^3} \right] > 0.$$

Then at least one of the numbers $3\alpha + \beta$ and $3\beta + \alpha$ [the sum of which is $4(\alpha + \beta)$] is positive.

Proof. 1°. If $\gamma(t_1, t_2)$ is the covariance of a r.f. $\Phi(t)$, $e^{-3u}\Phi(e^{2u})$ has the covariance

$$(19) \quad \alpha e^{-|u|} + \beta e^{-3|u|}, \quad (u = u_1 - u_2);$$

and, conversely, if this function is a covariance, $\gamma(t_1, t_2)$ is a covariance. Now, by a theorem of A. Khintchine, an even function of $u_1 - u_2$ that has a continuous Fourier transform is a covariance (in fact a stationary covariance) if, and only if, the Fourier transform

$$(20) \quad \frac{\alpha}{1+u^2} + \frac{3\beta}{9+u^2} = \frac{3(3\alpha + \beta) + (\alpha + 3\beta)u^2}{(1+u^2)(9+u^2)}$$

is never negative. Thus the conditions (16) are necessary and sufficient. Q.E.D.

2°. We have to compute a, b, c from (9) and (13). Using (13) to eliminate b and c , the last equation (9) is

$$(21) \quad b^2 - ac = a^2 - (5\alpha + 3\beta)a + 4\alpha^2 \leq 0.$$

We deduce from (16) that

$$(22) \quad D = (5\alpha + 3\beta)^2 - 16\alpha^2 = 3(3\alpha + \beta)(\alpha + 3\beta) \geq 0.$$

Therefore, the equation

$$(23) \quad x^2 - (5\alpha + 3\beta)x + 4\alpha^2 = 0$$

has real roots, a_1 and $a_2 \geq a_1$, with $b^2 \leq ac$ if, and only if, $a \in [a_1, a_2]$. From

$$(24) \quad a_1 a_2 = 4\alpha^2 \geq 0, \quad a_1 + a_2 = \frac{1}{2}(\alpha + 3\beta) + \frac{3}{2}(3\alpha + \beta) \geq 0$$

it follows that $a_2 \geq a_1 \geq 0$. If a is chosen in $[a_1, a_2]$, and if b and c are computed from (13), two possibilities arise: either $a = a_1 = 0$, and then $\alpha = 0$ and $c = 3\beta \geq 0$, or else $a > 0$, and then $c \geq 0$ is deduced from $ac \geq b^2$. In both cases, the conditions (9) and (16) are fulfilled.

It is obvious that we get the form (10) if, and only if, $b^2 = ac$, that is, $a = a_1$ or a_2 . We have one solution if $D = 0$ and two if $D > 0$. Q.E.D.

1.4. We shall now consider only the form (10) of $\Psi(t)$. We shall write the letters λ , μ , Ψ , or X with the subscript i ($i = 1$ or 2) if we wish to specify that $a = \lambda^2$ has the value a_i . From

$$(25) \quad s^2 = (\lambda + \mu)^2 = \alpha + 3\beta,$$

and from (11), we deduce that s has only one value (in the general case, we have $a + 2b + c = \alpha + 3\beta$). We introduce k , K , and s by

$$(26) \quad \lambda = ks, \quad \mu = (1 - k)s, \quad \beta = K\alpha.$$

We find from (15)

$$K = \frac{3\lambda\mu + 2\mu^2}{6\lambda^2 + 3\lambda\mu} = \frac{3k(1 - k) + 2(1 - k)^2}{6k^2 + 3k(1 - k)} = \frac{2 - k - k^2}{3(k + k^2)},$$

that is,

$$(27) \quad k^2 + k = \frac{2}{3K + 1}.$$

We have real values k_1 and k_2 of k if

$$1 + \frac{8}{3K + 1} = \frac{3(K + 3)}{3K + 1} \geq 0.$$

This condition is another form of (22), and k_1 and k_2 are linked by the involutive relation

$$(28) \quad k_1 + k_2 = -1.$$

The double points $-\frac{1}{2}$ and ∞ of this involution* correspond respectively to the values -3 and $-\frac{1}{3}$ of K (if $k = \infty$, $s = \lambda + \mu = 0$; if $k = -\frac{1}{2}$, $3\lambda + \mu = 0$).

Since $a = \lambda^2 = k^2s^2$ and $a_2 \geq a_1$, and taking into account (28), we have

$$k_2^2 - k_1^2 = k_1 - k_2 \geq 0,$$

and

$$(29) \quad k_1 \geq -\frac{1}{2} \geq k_2.$$

1.5. Consider now the equation (12). We deduce

$$(30) \quad E\{[\delta\Psi(t)]^2\} = s^2t^2|dt| + o(dt), \quad (dt \rightarrow 0)$$

and have a new proof of the following fact: if the r.f. $\Psi(t)$ is defined, s^2 has only one possible value. The same conclusion holds if an individual $\Psi(t)$ is given in a small interval (t_1, t_2) . This can be deduced easily from well-known almost

sure (a.s.) properties of $X(t)$.† Since λ has two possible values (unless $K = -3$ or $-\frac{1}{3}$), there is no hope of deducing this parameter from an individual $\Psi(t)$.

Formula (12) leads us to another important conclusion when we write it successively with the subscripts 1 and 2 and suppose $\Psi_1(t) = \Psi_2(t)$. Then

$$(31) \quad st \delta[X_2(t) - X_1(t)] = [\lambda_1 X_1(t) - \lambda_2 X_2(t)] dt + o(dt).$$

Thus the difference $X_2(t) - X_1(t)$ is differentiable (even if $s = 0$; in this case it is zero).

1.6. If an individual $X(u)$ is given in $(0, T)$, all the ξ_u are known in $(0, T)$, and the corresponding $\Psi(t)$ is given by (10) or (12). The converse problem is very important. The solution is given by the following theorem.

THEOREM 2. Suppose an individual $\Psi(t)$ is given in $[0, T]$. 1°. If

$$k = k_1 > k_2,$$

and $t \in (0, T)$, then only one $X(t)$ corresponds to the given $\Psi(t)$. 2°. The same conclusion holds if $k_1 = k_2$. 3°. If $k = k_2 < k_1$, $X(t)$ is not known but depends on a Gaussian variable c , with positive standard deviation $\sigma(c)$.

Proof. In all three cases, (12) may be written in the form

$$(32) \quad s \delta[t^k X(t)] = t^{k-1} \delta\Psi(t).$$

Thus, if $s \neq 0$, and if $X^*(t)$ is a particular solution of this equation, the general solution is

$$(33) \quad X(t) = X^*(t) + ct^{-k},$$

and c is obviously either a known number or a Gaussian random variable with positive standard deviation $\sigma(c)$. As the given values of $\Psi(t)$ and the unknown values of $X(t)$ are joint Gaussian variables, there exists no other possibility.

1°. If $k = k_1 > k_2$, then $k > -\frac{1}{2}$, and

$$(34) \quad t^k X(t) \xrightarrow{\text{a.s.}} 0, \quad (t > 0).$$

This is a Cauchy condition, from which we deduce

$$(35) \quad st^k X(t) = \int_0^t u^{k-1} d\Psi(u) = t^{k-1} \Psi(t) - (k-1) \int_0^t u^{k-2} \Psi(u) du.$$

Since here $s \neq 0$, the first statement is proved.

2°. $k_1 = k_2$ implies $s = 0$ or $k = -\frac{1}{2}$. If $s = 0$, we have already deduced from (12) that $\lambda X(t)$ is the derivative of $\Psi(t)$. If $k = -\frac{1}{2}$, the above Cauchy condition holds in a generalized sense. One has

$$(36) \quad U^{-1} \int_0^U e^u X(e^{-2u}) du \xrightarrow{\text{a.s.}} 0, \quad (U \rightarrow +\infty).$$

This condition yields c .

† For instance, if $n \rightarrow \infty$, $2^n[X(t + 2^{-n}) - X(t)]^2$ tends a.s. to 1 in the Cesaro sense. Then $2^n[\Psi(t + 2^{-n}) - \Psi(t)]^2$ tends a.s. to $s^2 t^2$ in the same sense.

3°. The value of $\sigma(c)$ will be found in Sec. 1.7 [see formula (44)]. It is positive, and this will complete the proof of the present theorem.

Another proof may be deduced from a theorem which may be briefly stated as follows. *If a function $f(t)$, given in $(0, T)$, is a possible function $X(t)$, then $f(t) + ct^\alpha$ ($c \neq 0$) is also a possible function $X(t)$ if and only if $\alpha > \frac{1}{2}$. Thus, if $k < -\frac{1}{2}$, that is, $k = k_2 < k_1$, we have no way to choose one of the functions $X(t)$ in (33) and say that it is the correct one.*

COROLLARY 1. 1°. If $k = k_1$ (even if $k_1 = k_2$), we have exactly the same information, whether an individual $\Psi(t)$ is given in $[0, T]$ or the corresponding $X(t)$ is given in this interval.

2°. If $k = k_2 < k_1$, we have more information when $X(t)$ is given in $[0, T]$ than when $\Psi(t)$ is given in this interval.

1.7. The continuation of $\Psi(t)$ to the right. 1°. As a preliminary problem, suppose an individual $X(u)$ given in $(0, t)$. The problem is to write $\Psi(t')$, ($t' > t$), in the canonical form $m + \sigma\xi$, m being the conditional expectation and σ the standard deviation. Since in

$$\Psi(t') = \int_0^t (\lambda t' + \mu u) \xi_u(du)^{\frac{1}{2}} + \int_t^{t'} (\lambda t' + \mu u) \xi_u(du)^{\frac{1}{2}}$$

the ξ_u in the first integral are known [$\xi_u(du)^{\frac{1}{2}} = \delta X(u)$] and in the second integral they are independent of the given individual $X(u)$, we have

$$(37) \quad m = \int_0^t (\lambda t' + \mu u) dX(u),$$

$$(38) \quad \sigma\xi = \int_t^{t'} (\lambda t' + \mu u) \xi_u(du)^{\frac{1}{2}}.$$

It follows from (38) that

$$(39) \quad \sigma^2 = \int_t^{t'} (\lambda t' + \mu u)^2 du = (t' - t) \left[\lambda^2 t'^2 + \lambda \mu t' (t' + t) + \frac{\mu^2}{3} (t'^3 + tt' + t^2) \right]$$

As expected, if $t' - t = dt$, then $m - \Psi(t) + \sigma\xi$ has the form (12).

2°. Suppose now an individual $\Psi(u)$ given in $(0, t)$. If $\Psi(u)$ is given in the form $\Psi_1(u)$, we have exactly the same information as if $X_1(u)$ were given, and the preceding conclusion holds: $m + \sigma\xi$ is again the canonical form of $\Psi(t')$, ($t' > t$), and m may be written

$$(40) \quad m = \Psi(t) + \lambda(t' - t)X(t).$$

Then, taking into account (35), we have

$$(41) \quad \Psi(t') - \Psi(t) = k \frac{t' - t}{t} \int_0^t \left(\frac{u}{t} \right)^{k-1} d\Psi(u) + \sigma\xi,$$

where σ is given by (39).

3°. If now $k_1 > k_2$ and we use the form $\Psi_2(t)$ of $\Psi(t)$, then $X(t) = X_2(t)$. Since this function is not known, the number m defined by (37) is not the conditional expectation of $\Psi(t')$. However, m and $\sigma\xi$ are still independent, (40) holds, and we may suppose $T = t$ in (33). Thus we have

$$(42) \quad \sigma^2 = \lambda_2^2 \sigma^2(c) t^{-2k_2} (t' - t)^2 + \sigma_2^2,$$

where $\sigma^2 = \sigma_1^2$ and σ_2^2 are computed from (39), written with the subscripts 1 and 2. From $\lambda = ks$ and $\mu = (1 - k)s$, we deduce

$$(43) \quad \sigma^2 = (t' - t)s^2 \left[\frac{1 + k + k^2}{3} t'^2 + \frac{1 + k - 2k^2}{3} tt' + (1 - 2k + k^2)t^2 \right].$$

Also, from $k_1 + k_2 = -1$, $k_1^2 - k_2^2 = -(k_1 - k_2)$, we deduce

$$\sigma_1^2 - \sigma_2^2 = (k_1 - k_2)s^2 t(t' - t)^2.$$

Using (42), we have finally

$$(44) \quad \sigma^2(c) = (k_1 - k_2)k_2^{-2}t^{k_2-k_1} > 0, \quad (k_2 - k_1 = 1 + 2k_2 < 0).$$

Thus the last statement of Theorem 2 is proved. Moreover, if $u < t$, then the unknown part of $X_2(u)$ is $\sigma(c)\xi u^{-k_1}$, and its standard deviation, $\sigma(c)u^{-k_1}$, is a decreasing function of t , as foreseen, since, as t increases, we have more information.

1.8. The continuation of $\Psi(t)$ to the left. 1°. The formulas (32) and (33) hold if an individual $\Psi(t)$ is given in (T, ∞) . But here, as $t \rightarrow \infty$, then $t^k X(t) \xrightarrow{\text{a.s.}} 0$ if, and only if, $k = k_2 < -\frac{1}{2}$ (instead of $k = k_1 > -\frac{1}{2}$). Thus we have a theorem which is quite like Theorem 2, but the subscripts 1 and 2 are interchanged. If $k = k_2$ (even in the limiting cases $k = -\frac{1}{2}$ or ∞), $X(t) = X_2(t)$ is known in (T, ∞) , and if $k < -\frac{1}{2}$, it is given by

$$(45) \quad st^k X_2(t) = \int_{\infty}^t u^{k-1} d\Psi(u).$$

On the contrary, if $k = k_1 > -\frac{1}{2}$, $X(t) = X_1(t)$ is a r.f., and may be written in the form (33), with c again a Gaussian variable.

2°. Since $X_2(t)$ depends only on $d\Psi(u)$, the preceding results hold if the difference $\Psi(t) - \Psi(T)$, instead of $\Psi(t)$, is known in (T, ∞) . Then $X_2(t)$ is known in (T, ∞) , and not only are the ξ_u with subscripts $u > T$ known in $X_2(t)$, but also the others are linked by

$$J_0 = \int_0^T \xi_u(du)^{\frac{1}{2}} = X_2(T).$$

The continuation problem is not so simple as in Sec. 1.7. We have to write $\Psi(t)$, ($t \leq T$) in the canonical form $m + \sigma\xi$. From the general theory of joint Gaussian variables, we know that $m = ctJ_0$ and that c is given by the

condition that σ^2 should be minimum. One has

$$\begin{aligned}\sigma^2 &= E\{[\Psi(t) - cJ_0]^2\} = \int_0^t [(\lambda - c)t + \mu u]^2 du + c^2 t^2 \int_t^T du \\ &= (\lambda - c)^2 t^3 + (\lambda - c)\mu t^3 + \frac{\mu^2}{3} t^3 + c^2 t^2 (T - t) \\ &= c^2 t^2 T - (2\lambda + \mu)ct^3 + \left(\lambda^2 + \lambda\mu + \frac{\mu^2}{3}\right)t^3.\end{aligned}$$

Thus the minimum is obtained with

$$(46) \quad c = \left(\lambda + \frac{\mu}{2}\right) \frac{t}{T}, \quad m = \left(\lambda + \frac{\mu}{2}\right) \frac{t^2}{T} X_2(T),$$

and its value is

$$(47) \quad \begin{aligned}\sigma^2 &= \left(\lambda^2 + \lambda\mu + \frac{\mu^2}{3}\right)t^3 - \left(\lambda + \frac{\mu}{2}\right)^2 \frac{t^4}{T} \\ &= \frac{\mu^2}{12} t^3 + \left(\lambda + \frac{\mu}{2}\right)^2 t^3 \left(1 - \frac{t}{T}\right).\end{aligned}$$

If $t = 0$, then $m = \sigma = 0$. If $t = T$, then the canonical form for $\Psi(T)$ is

$$(48) \quad \Psi(T) = \left(\lambda + \frac{\mu}{2}\right) T X_2(T) + \sigma \xi, \quad \left(\sigma^2 = \frac{\mu^2}{12} T^3\right),$$

and, if $\lambda = 0, \mu = 1$,

$$(49) \quad J_1 = \int_0^T u \xi_u (du)^{\frac{1}{2}} = \frac{T}{2} X_2(T) + \frac{T}{2} \left(\frac{T}{3}\right)^{\frac{1}{2}} \xi,$$

from which one easily returns to (48).

All these formulas hold, if $X(T)$ is known, for λ_i, μ_i, X_i , either with $i = 1$ or with $i = 2$. However, we have supposed only that an individual $\Psi(t) - \Psi(T)$ is given in (T, ∞) and that $X_2(t)$ is given by (45) but not $X_1(t)$ unless $\lambda_1 = \lambda_2$. Thus these formulas give the canonical forms of $\Psi(t)$ and $\Psi(T)$ if $i = 2$ but not if $i = 1$, unless $k_1 = k_2$.

3°. If $k = k_1 > -\frac{1}{2}$, we deduce from (33) and (48)

$$\Psi(T) = \frac{k_1 + 1}{2} s T X^*(T) + \frac{k_1 + 1}{2} c s T^{1-k_1} + \sigma_1 \xi,$$

and, since we have necessarily the same $\sigma = \sigma_2$ as in (47), thus

$$(k_1 + 1)^2 \frac{s^2}{4} \sigma^2(c) = \sigma_2^2 - \sigma_1^2 = (k_1 - k_2) \frac{s^2 T^2}{4},$$

and finally

$$(50) \quad \sigma^2(c) = \frac{(k_1 - k_2)}{(k_1 + 1)^2} T^2 = \frac{k_1 - k_2}{k_2^2} T^2,$$

$$(51) \quad \sigma^2[X_1(t)] = \frac{k_1 - k_2}{k_2^2} t^{-2k_1} T^2, \quad (t \geq T).$$

As foreseen, this value increases with T .

4°. The same method may be used for other problems. If, for instance, only $X(t) - X(T)$ is given in (T, ∞) , we have no information on the ξ_u with subscripts $u < t$, and the problem becomes quite trivial. On the contrary, if $\mu \neq 0$, and if $\Psi(t)$ is given in (T, ∞) , then $X_2(t)$, ($t \geq T$), J_0 and J_1 are known, and the conditional expectation of $\Psi(t)$, ($t < T$) has the form $c_0 J_0 + c_1 J_1$. The calculation is more complicated than in 2° but not difficult. If $\mu = 0$, then $\Psi(t) = \lambda t X(t)$, and this problem is trivial.

1.9. Application of the Fourier-Wiener series. It is known that the complex Brownian function

$$Z(t) = \frac{X(t) + iY(t)}{2^{\frac{1}{2}}} = \int_0^t \xi_u(du)^{\frac{1}{2}}$$

may be defined in $(0, 2\pi)$ by the Paley-Wiener formula

$$(52) \quad Z(t) = (2\pi)^{-\frac{1}{2}} \left[\zeta t + \sum_{n \neq 0} n^{-1} \zeta_n (e^{int} - 1) \right],$$

where $\zeta = 2^{-\frac{1}{2}}(\xi + i\eta)$ and the $\zeta_n = 2^{-\frac{1}{2}}(\xi_n + i\eta_n)$ are independent. Integrating from 0 to t , we get

$$(53) \quad \int_0^t (t-u) \xi_u(du)^{\frac{1}{2}} = (2\pi)^{-\frac{1}{2}} \left[\zeta \frac{t^2}{2} + \sum_{n \neq 0} \frac{\zeta_n}{in^2} (e^{int} - 1 - int) \right].$$

Now, from

$$(54) \quad \int_0^t (\lambda t + \mu u) \xi_u(du)^{\frac{1}{2}} = stZ(t) - \mu \int_0^t (t-u) \xi_u(du)^{\frac{1}{2}},$$

and from (52) and (53), we obtain the integral on the left side of (54) [the real part of which is $\Psi(t)$] in the form of a series.

1.10. Generalization. Instead of $\Psi(t)$, let us consider the r.f.

$$(55) \quad \int_0^t (t^k X_{0,u} + t^{k-1} u X_{1,u} + \cdots + u^k X_{k,u}) (du)^{\frac{1}{2}}$$

and the particular case

$$(56) \quad \int_0^t (\lambda_0 t^k + \lambda_1 t^{k-1} u + \cdots + \lambda_k u^k) \xi_u(du)^{\frac{1}{2}}.$$

Although the integral (55), where the $X_{k,u}$ are defined by formulas like (8), introduces $(h+1)(h+2)/2$ parameters, only $h+1$ of them are essential, and every r.f. of the form (55) has, in general, several distinct representations of the form (56). If the condition (11) is replaced by $\Sigma \lambda_r \geq 0$, one has at first sight 2^{h-1} distinct representations. But they may introduce complex numbers, and the problem is to know how many are real; perhaps never more than two.

These random functions are related to the function $M(t)$ in the space E_{2h+3} in the same way as the functions (7) and (10) are related to $M(t)$ in the case $n = 5$. We shall see in Chap. 2 that $M(t)$ has two distinct expressions, con-

nected with two distinct stochastic differential equations (see Sec. 2.8). The same result probably holds for the more general r.f. defined by formula (56).

2. THE RANDOM FUNCTION $M(t)$

2.1. Definitions and general remarks. Let us consider, as in the Introduction, the Euclidean space E_n and the Brownian function $X(A)$, defined up to an additive constant by the formula

$$(57) \quad X(B) - X(A) = \xi[r(A, B)]^\frac{1}{2},$$

where $r(A, B)$ is the distance between A and B .

We consider now a family of closed surfaces Ω_t , depending on a parameter t , and denote by $\bar{M}(t)$, $M(t)$, $U_t(B)$ and $\rho_n(t, t')$ the mean values of $X(A)$, $X(A) - X(O)$, $r(A, B)$, and $U_r(A)$ as A describes Ω_t . Since obviously

$$2E\{[X(A) - X(O)][X(B) - X(O)]\} = r(O, A) + r(O, B) - r(A, B),$$

the covariance of $M(t)$ is

$$(58) \quad \Gamma_n(t, t') = E[M(t)M(t')] = \frac{1}{2}[U_t(O) + U_{t'}(O) - \rho_n(t, t')].$$

General theorems on $\Gamma(t, t')$ are easy to deduce, especially if n is an odd number, from the relation of $U_t(B)$ with the Newtonian potential. In this paper we shall consider only the particular case in which Ω_t is a sphere with center O and radius t , and especially the case $n = 5$. Thus we need no general theory.

When Ω_t is the considered sphere, we have

$$(59) \quad \Gamma_n(t, t') = \frac{1}{2}[t + t' - \rho_n(t, t')],$$

and, setting

$$(60) \quad I_h = \int_0^{\pi/2} \sin^h \theta \, d\theta, \quad r^2 = t^2 + t'^2 - 2tt' \cos \theta,$$

we have

$$(61) \quad \rho_n = \frac{1}{2I_{n-2}} \int_0^\pi r \sin^{n-2} \theta \, d\theta,$$

so that, if $n = 5$,

$$(62) \quad \rho_5 = \frac{3}{4} \int_0^\pi r \sin^3 \theta \, d\theta.$$

We see at once, taking $\cos \theta$ or r as parameter, that if n is an even number, ρ_n is an elliptic integral. It is only when n is an odd number that ρ_n and Γ_n are elementary functions. Then we can easily obtain explicit expressions for $M(t)$. As already stated, we shall consider here only the case $n = 5$, and write $\Gamma(t, t')$ instead of $\Gamma_5(t, t')$.

2.2. The covariance $\Gamma(t_1, t_2)$. In this paper t_1 and t_2 will denote two non-negative numbers, and we shall set

$$t = \min(t_1, t_2), \quad t' = \max(t_1, t_2).$$

Then $t' \geq t$, and $\Gamma(t_1, t_2) = \Gamma(t, t')$.

From (62), choosing r as a new parameter and taking account of

$$r dr = t t' \sin \theta d\theta,$$

we deduce

$$\begin{aligned} \rho_5 &= \frac{3}{16t^3t'^3} \int_{t'-t}^{t'+t} [-(t'^2 - t^2)^2 + 2(t^2 + t'^2)r^2 - r^4] r^2 dr \\ &= \frac{3}{16t^3t'^3} \left[-(t'^2 - t^2)^2 \frac{r^3}{3} + 2(t^2 + t'^2) \frac{r^5}{5} - \frac{r^7}{7} \right]_{t'-t}^{t'+t}, \end{aligned}$$

and finally

$$(63) \quad \rho_5 = t' + \frac{2t^2}{5t'} - \frac{t^4}{35t'^3} \dagger$$

Then we deduce from (59)

$$(64) \quad \Gamma(t_1, t_2) = \frac{t}{2} - \frac{t^2}{5t'} + \frac{t^4}{70t'^3}.$$

If we put $t = e^{2u}$, then $M(t)t^{-\frac{1}{2}}$ is a Gaussian stationary function, with covariance

$$\frac{\Gamma(t_1, t_2)}{(t_1 t_2)^{\frac{1}{2}}} = \frac{1}{2} e^{-|u|} - \frac{1}{5} e^{-3|u|} + \frac{1}{70} e^{-7|u|}, \quad (u = u_1 - u_2).$$

This function and its first four derivatives are continuous, but the fifth derivative has a jump at the point $u = 0$.‡

2.3. The continuity theorem for $M(t)$. Consider the function

$$(65) \quad cI^p X(t) = c \int_0^t \frac{(t-u)^p}{p!} dX(u) = c \int_0^t \frac{(t-u)^p}{p!} \xi_u(du)^\dagger,$$

the p th derivative of which is $cX(t)$. Its covariance is

$$\frac{c^2}{(p!)^2} \int_0^t (t-u)^p (t'-u)^p du,$$

and when the sign of $t_2 - t_1$ changes, the difference between this function written for $t_2 > t_1$ and the analytical continuation of the function written for

† If $n = 2p + 1$, one has more generally

$$\Gamma(t_1, t_2) = \frac{t}{2} - c_1 \frac{t^2}{t'} + c_2 \frac{t^4}{t'^3} - \dots + (-1)^p c_p \frac{t^{2p}}{t'^{2p-1}},$$

and it is easy to deduce $c_1, c_2, \dots, c_p$ from a continuity theorem on $U_t(A)$.

‡ This is a particular case of the continuity theorem: if $n = 2p + 1$, the difference

$$\Gamma_n(t_1, t_2) - (-1)^p \frac{|t_2 - t_1|^n}{4n! t_1^p t_2^p}, \quad (t_1, t_2 > 0),$$

and all its derivatives of order $\leq n$ are continuous.

$t_2 < t_1$ is

$$\begin{aligned} -\frac{c^2}{(p!)^2} \int_t^{t'} (t-u)^p (t'-u)^p du &= (-1)^{p-1} \frac{c^2}{(p!)^2} (t'-t)^n \int_0^1 v^p (1-v)^p dv \\ &= (-1)^{p-1} \frac{c^2}{n!} (t'-t)^n, \quad (n = 2p + 1). \end{aligned}$$

Thus, for a suitable value of c , this covariance has precisely the analytical character of $\Gamma_n(t_1, t_2)$ up to derivatives of order n . We are led to expect that when $n = 2p + 1$, then $M(t)$ has the same analytical character as the r.f. (65), i.e., it has a.s. continuous derivatives up to order p , and the p th derivative, as in the case of $cX(t)$, is not differentiable. It is easy to deduce the value of c from a foregoing remark (see footnote† on p. 12).

We do not need a complete proof here. The statements lead to a heuristic method for obtaining the Gaussian r.f. $M(t)$. This done, it will be easy to find the covariance of the function obtained and verify that it is the correct expression for $M(t)$. In this paper we shall carry out the case $n = 5$, ($p = 2$).

Let us notice that the continuity theorem is connected with a theorem of M. Loève, which states that if the covariance of a r.f. $M(t)$ is continuous up to the order $2p$, then $M(t)$ has a.s. at least $p - 1$ continuous derivatives. This theorem is not sufficient for our purpose. Only in the limiting case $n = \infty$ (E_n is then the Hilbert space) does it enable us to conclude that the derivatives of all orders of $M(t)$ exist a.s. and are continuous.

2.4. The functions $\Phi(t)$ and $M(t)$. Let

$$\Phi(t) = t^3 M(t). \dagger$$

Then the covariance of $\Phi(t)$ is

$$(66) \quad t_1^3 t_2^3 \Gamma(t_1, t_2) = \frac{1}{2} t^4 t'^3 - \frac{1}{8} t^5 t'^2 + \frac{1}{70} t^7,$$

and the covariance of $\Phi''(t)$ is

$$(67) \quad \frac{\partial^4 [t_1^3 t_2^3 \Gamma(t_1, t_2)]}{\partial t_1^2 \partial t_2^2} = 36 t^2 t' - 8 t^4.$$

We recognize the covariance of a function $\Psi(t)$, ‡ and we get at once the formula

$$(68) \quad \Phi_i''(t) = 2 \int_0^t [u + \lambda_i(t-u)] \xi_u(3du)^{\dagger}, \quad (i = 1, 2; \lambda_1 = 2, \lambda_2 = -3),$$

which gives two distinct expressions for $\Phi''(t)$. Since $\Phi(0) = \Phi'(0) = 0$, we deduce

† In the general theory, $\Phi(t) = t^{n-2} M(t)$.

‡ The reader who has not read Sec. 1 need only verify that the r.f.'s (68) have the covariance (67).

$$\begin{aligned}\Phi_i(t) &= \int_0^t \left[u(t-u)^2 + \frac{\lambda_i}{3} (t-u)^3 \right] \xi_u(3du)^{\frac{1}{2}} \\ &= \int_0^t \left[t(t-u)^2 + \left(\frac{\lambda_i}{3} - 1 \right) (t-u)^3 \right] \xi_u(3du)^{\frac{1}{2}},\end{aligned}$$

that is,

$$\begin{aligned}(69) \quad \Phi_1(t) &= \int_0^t [t(t-u)^2 - \tfrac{1}{3}(t-u)^3] \xi_u(3du)^{\frac{1}{2}} \\ &= \tfrac{1}{3} \int_0^t (2t^3 - 3tu^2 + u^3) \xi_u(3du)^{\frac{1}{2}},\end{aligned}$$

$$\begin{aligned}(70) \quad \Phi_2(t) &= \int_0^t [t(t-u)^2 - 2(t-u)^3] \xi_u(3du)^{\frac{1}{2}} \\ &= \int_0^t (-t^3 + 4t^2u - 5tu^2 + 2u^3) \xi_u(3du)^{\frac{1}{2}},\end{aligned}$$

and finally

$$\begin{aligned}(71) \quad M_1(t) &= \int_0^t \left[\left(1 - \frac{u}{t}\right)^2 - \tfrac{1}{3} \left(1 - \frac{u}{t}\right)^3 \right] \xi_u(3du)^{\frac{1}{2}} \\ &= \int_0^t \left(\frac{2}{3} - \frac{u}{t} + \frac{u^3}{3t^3} \right) \xi_u(3du)^{\frac{1}{2}},\end{aligned}$$

$$\begin{aligned}(72) \quad M_2(t) &= \int_0^t \left[\left(1 - \frac{u}{t}\right)^2 - 2 \left(1 - \frac{u}{t}\right)^3 \right] \xi_u(3du)^{\frac{1}{2}} \\ &= \int_0^t \left(-1 + 4 \frac{u}{t} - 5 \frac{u^2}{t^2} + 2 \frac{u^3}{t^3} \right) \xi_u(3du)^{\frac{1}{2}}\end{aligned}$$

2.5. Important remarks. 1°. If

$$(73) \quad F(t, u) = \sum_0^p f_h(t) \varphi_h(u),$$

then the covariance of the r.f.

$$(74) \quad M(t) = \int_0^t F(t, u) \xi_u(du)^{\frac{1}{2}}$$

is

$$\begin{aligned}(75) \quad \int_0^t F(t, u) \sum_0^p f_h(t') \varphi_h(u) du &= \sum_0^p f_h(t') g_h(t), \\ [g_h(t) &= \int_0^t F(t, u) \varphi_h(u) du, \quad t' \geq t].\end{aligned}$$

† If $n = 2p + 1$, formula (71) becomes

$$M_1(t) = k_p \int_0^t Q_p\left(\frac{u}{t}\right) \xi_u(du)^{\frac{1}{2}},$$

with

$$Q_p(u) = \int_u^1 (1-v^2)^{p-1} dv, \quad k_p^2 Q_p(0) = p.$$

The author thinks that a second expression $M_2(t)$ of $M(t)$ exists for every $p > 1$. But he could not find an expression as simple as that for $M_1(t)$. When $n = 7$, ($p = 3$),

$$M_2(t) = \int_0^t \left[-1 + 3 \frac{u}{t} - 8 \frac{u^2}{t^3} + 9 \frac{u^4}{t^4} - 3 \frac{u^5}{t^5} \right] \xi_u\left(\frac{5}{3} du\right)^{\frac{1}{2}}.$$

Thus, if the function $F(t, u)$ in (74), where t is a parameter, has for every fixed u the form

$$(76) \quad \sum_0^p \alpha_h f_h(t),$$

then the covariance $\Gamma(t', t)$, considered as a function of t' , ($t' \geq t$), has the same form.

In other words, since, if the functions $f_h(t)$ have derivatives up to the order $p + 1$, (76) is the general solution of a linear differential equation

$$(77) \quad \varepsilon_t[y(t)] = 0,$$

we have the following theorem: if

$$(78) \quad \varepsilon_t[F(t, u)] = 0, \quad (t \geq u),$$

then $F(t, u)$ has the form (73), and

$$(79) \quad \varepsilon_t[\Gamma(t, u)] = 0, \quad (t \geq u).$$

2°. In the case of the function $M_1(t)$ defined by (71), the form (76) and the equation (77) become

$$(80) \quad \alpha_0 + \frac{\alpha_1}{t} + \frac{\alpha_2}{t^3},$$

$$(81) \quad \varepsilon_t[y(t)] \equiv t^2 \frac{d^3 y}{dt^3} + 7t \frac{d^2 y}{dt^2} + 8 \frac{dy}{dt} = 0.$$

The integrand $F_1(t, u)$ in $M_1(t)$ and the covariance $\Gamma(t, u)$ both have the form (80) if $t \geq u$, and are solutions of (81). But $M_2(t)$ has the same covariance; nevertheless the integrand $F_2(t, u)$ in $M_2(t)$ is not a solution of (81). Thus

THEOREM 3. (78) implies (79), yet (79) does not imply (78).

The fact that $F_1(t, u)$, but not $F_2(t, u)$, is a solution of (81) is very important. This is related to the results of Sec. 1. If an individual $M(u)$ is given in $(0, t)$, then $\Phi(u)$ and $\Phi''(u)$ are known, and the ξ_u in $\Phi_1'(t)$, which are the same as in $M_1(t)$, are known. But the ξ_u in $\Phi_2''(t)$ and in $M_2(t)$ are unknown.

We shall give in Sec. 2.6 a new proof of this property of $M_1(t)$, without reference to Sec. 1. This second proof is useful for the generalization of the present theory to the case $n = 2p + 1 > 5$.

2.6. The continuation of $M(t)$ to the right. 1°. **THEOREM 4.** If an individual $M(u)$ is given in $(0, t)$, and if $t' \geq t$, the canonical form $m + \sigma \xi$ of $M(t')$ is given by

$$(82) \quad m = m_1(t, t') \equiv \int_0^t F_1(t', u) \xi_u (3du)^{\frac{1}{3}},$$

$$(83) \quad \sigma \xi = \sigma_1 \xi \equiv \int_t^{t'} F_1(t', u) \xi_u (3du)^{\frac{1}{3}},$$

$$(84) \quad F_1(t', u) = \left(1 - \frac{u}{t'}\right)^2 - \frac{1}{3} \left(1 - \frac{u}{t'}\right)^3 = \frac{2}{3} - \frac{u}{t'} + \frac{u^3}{3t'^3};$$

and obviously $F_1(t', u)$ cannot be replaced by $F_2(t', u)$ that is, the integrand in $M_2(t)$.

Proof. If ε_t has the meaning (81), we obviously have

$$(85) \quad \varepsilon_t[F_1(t', u)] = 0, \quad (u \leq t \leq t'),$$

$$(86) \quad \varepsilon_t[m_1(t, t')] = 0, \quad (t \leq t'),$$

$$(87) \quad \varepsilon_t[\Gamma(u, t')] = 0, \quad (u \leq t \leq t'),$$

and we deduce from (83)

$$(88) \quad \sigma_1^2 = 3 \int_t^{t'} F_1^2(t', u) du = O[(t' - t)^5].$$

In the above, $m_1(t, t')$ is the solution of (86) that fulfills the Cauchy condition

$$m_1(t, t') = M(t) + (t' - t)M'(t) + \frac{(t' - t)^2}{2} M''(t) + o[(t' - t)^2], \quad (t' \downarrow t).$$

Now, from (87), we deduce

$$(89) \quad E\{M(u)\varepsilon_t[m(t, t')]\} = 0,$$

a formula which holds for every u in $(0, t)$ and $(t' \geq t)$. Since $M(t)$ is a Gaussian r.f., it follows that the conditional expectation $m = m(t, t')$ is necessarily a linear functional of the given $M(u)$, and (89) implies

$$(90) \quad \varepsilon_t[m(t, t')] = 0.$$

Moreover, $M(t') = m + \sigma\xi$ having a.s. two derivatives, we have

$$\sigma = o[(t' - t)^2]$$

and

$$m(t, t') = M(t) + (t' - t)M'(t) + \frac{(t' - t)^2}{2} M''(t) + o[(t' - t)^2].$$

Thus m and m_1 are solutions of the same Euler equation (81) with the same Cauchy conditions, so that $m = m_1$. Q.E.D.

2°. From (83), (84), and (88), setting $u = t' - v$, we deduce

$$\sigma^2 = 3 \int_0^{t'-t} \left(\frac{v^2}{t'^2} - \frac{v^3}{3t'^3} \right)^2 dv = \frac{3(t' - t)^5}{5t'^4} - \frac{(t' - t)^6}{3t'^5} + \frac{(t' - t)^7}{21t'^7},$$

that is,

$$(91) \quad \sigma^2 = (t' - t)^5 \frac{33t'^2 + 25tt' + 5t^2}{105t'^6},$$

and if $t' - t = dt \downarrow 0$, then

$$(92) \quad \sigma^2 = \frac{3t}{5} \left(\frac{dt}{t} \right)^5 + O(dt^6).$$

3°. Now, we deduce at once from (81) and (90) that

$$m(t, t + dt) - M(t) = M'(t) dt + M''(t) \frac{dt^2}{2} - [8M'(t) + 7tM''(t)] \frac{dt^3}{6t^2} + O(dt^4),$$

and, from $M(t') = m + \sigma\xi$ and from (92), that

$$(93) \quad \delta M(t) = M'(t) dt + \frac{M''(t)}{2} dt^2 - [8M'(t) + 7tM''(t)] \frac{dt^3}{6t^2} + \left(\frac{3}{5}\right)^{\frac{1}{2}} \frac{\xi}{t^2} (dt)^{\frac{1}{2}}, \quad (dt > 0).$$

This formula is obviously a stochastic differential equation of the third order; the term in dt^3 is essential.

We see by this equation that $\{M(t), M'(t), M''(t)\}$ and also $\{M'(t), M''(t)\}$ are Markovian systems. The r.f. $M'(t)$ is a solution of

$$(94) \quad \delta M'(t) = M''(t) dt - [8M'(t) + 7tM''(t)] \frac{dt^2}{2t^2} + 2 \frac{\xi}{t^2} (dt)^{\frac{1}{2}}.$$

This equation may be obtained easily by the same method as used for (93). It is also easy to write the equations having $\Phi(t)$ or $\Phi'(t)$ as solutions.

4°. The integration of (93) is easy. If $M(t)$, $M'(t)$, and $M''(t)$ are given (Cauchy conditions), we have to deduce m from

$$\begin{aligned} m(t, t') - M(t) &= \int_0^t \left[\frac{t' - t}{tt'} u - \frac{t'^3 - t^3}{3t^3 t'^3} u^3 \right] \xi_u (3du)^{\frac{1}{2}}, \\ M'(t) &= \int_0^t \left[\frac{u}{t^2} - \frac{u^3}{t^4} \right] \xi_u (3du)^{\frac{1}{2}}, \\ tM''(t) &= \int_0^t \left[-\frac{2u}{t^3} + \frac{4u^3}{t^5} \right] \xi_u (3du)^{\frac{1}{2}} \end{aligned}$$

[formulas deduced from (71) and (82)]. We have

$$m(t, t') = M(t) + \frac{t(5t'^2 - tt' - t^2)}{3t'^3} (t' - t)M'(t) + \frac{t^2(2t' + t)}{3t'^3} \frac{(t' - t)^2}{2} M''(t),$$

and the canonical form $M(t') = m(t, t') + \sigma\xi$, where σ is given by (91).

It is also easy to obtain the canonical forms of $M'(t)$ and $M''(t)$, or to deduce from (81) the analogous equation corresponding to $\Phi(t)$. Obviously, the canonical forms of $M'(t)$ and $M''(t)$, with the Cauchy condition $M(t) = 0$, are sufficient to redefine the r.f. $M(t)$.

5°. More generally, when $F(t, u)$ in (74) is an arbitrary function, we shall say that formula (74) is the canonical representation of $M(t)$ if the canonical form $m + \sigma\xi$ of $M(t')$ ($t' > t$), when $M(u)$ is given in $(0, t)$, is given by

$$m = \int_0^t F(t', u) \xi_u (du)^{\frac{1}{2}},$$

as in Theorem 4. Every continuous Gaussian r.f. has at most one canonical form. Using this definition, we may state Theorem 4 as follows: *The*

canonical representation of the function $M(t)$ considered in this chapter is $M_1(t)$ and not $M_2(t)$.

The canonical representation of $M(t)$ may be also defined by the following property: if the ξ_u are given in $(0, t)$, $M(u)$ is known in this interval. The converse holds if and only if $M(t)$ is defined by its canonical representation.

2.7. The continuation of $M(t)$ to the left. If an individual Brownian function $X(A)$ is given everywhere in $\mathcal{E}_t$ [see formula (2) in the Introduction], then $\bar{M}(u)$ is known in (t, ∞) ; $M(u) - M(t)$, $M'(u)$, $M''(u)$ are known; but $M(t)$ is unknown, and the problems of putting $X(O)$ or $M(t)$ in canonical form are the same.

From the Markovian character of the vector r.f. $\{M'(t), M''(t)\}$, we deduce that this problem depends only on the known values of $M'(t)$ and $M''(t)$, and not on the other values of $M'(u)$ in (t, ∞) . For the solution, we may use either $M_1(t)$ or $M_2(t)$. If we use $M_1(t)$, the conditions are

$$\begin{aligned} t^2 M'(t) &= \int_0^t \left[u - \frac{u^3}{t^2} \right] \xi_u (3du)^{\frac{1}{2}}, \\ 2t M'(t) + t^2 M''(t) &= 2 \int_0^t \frac{u^3}{t^3} \xi_u (3du)^{\frac{1}{2}}, \end{aligned}$$

or

$$\begin{aligned} (95) \quad J_3 &\equiv \int_0^t \frac{u^3}{t^3} \xi_u (3du)^{\frac{1}{2}} = \frac{t^2}{2} M''(t) + t M'(t), \\ J_1 &\equiv \int_0^t \frac{u}{t} \xi_u (3du)^{\frac{1}{2}} = \frac{t^2}{2} M''(t) + 2t M'(t). \end{aligned}$$

Then the canonical form of

$$J_0 = \int_0^t \xi_u (3du)^{\frac{1}{2}}$$

is $\lambda J_1 + \mu J_3 + \sigma \xi$, where λ and μ are defined by the condition that

$$\sigma^2 = 3t \int_0^1 (1 - \lambda u - \mu u^3)^2 du = 3t \left(\frac{\lambda^2}{3} + \frac{2\lambda\mu}{5} + \frac{\mu^2}{7} - \lambda - \frac{\mu}{2} + 1 \right)$$

is minimal. Then

$$\begin{aligned} \lambda &= \frac{45}{16}, & \mu &= -\frac{35}{16}, & \sigma^2 &= \frac{27}{8}t, \\ J_0 &= \frac{5}{16}(9J_1 - 7J_3) + \sigma \xi, \end{aligned}$$

and we deduce from (71) and (95) that

$$(96) \quad M(t) = \frac{7I_1 - 9I_3}{8} + c\xi t^{\frac{1}{2}} = \frac{1}{8}[5M'(t) - t^2 M''(t)] + c\xi t^{\frac{1}{2}},$$

with

$$(97) \quad c^2 = \frac{4\sigma^2}{9t} = \frac{3}{16}$$

This value of c is important; $ct^\dagger$ is the minimum of $\sigma\{X(O)\}$ if nothing is known about the values of $X(A)$ inside of the sphere Ω_t .†

In a similar way, we could obtain the canonical form of $M(t_0)$, ($0 < t_0 < t$), when $M'(t)$ and $M''(t)$, or when $M(t)$, $M'(t)$, and $M''(t)$ are given. The calculation is more complicated but without difficulty.

2.8. Final remarks; the second stochastic differential equation for $M(t)$.
1°. We saw in Secs. 1.6 and 1.8 that $\Psi_1(t)$ and $\Psi_2(t)$ are connected by the Cauchy conditions at the points $t = 0$ and $t = \infty$, respectively. The same conditions exist for $M_1(t)$ and $M_2(t)$. Indeed, if an individual $M(t)$ is given, then $\Phi(t) = t^3 M(t)$ and $\Phi''(t) = \Psi(t)$ are known. Thus the ξ_u in $M_1(t)$ [the same as in $\Psi_1(t)$] are known (for every $u < t$) whenever $M(t')$ is known in $(0, t)$, and the ξ_u in $M_2(t')$ (for every $u > t$) are known whenever $M(t')$ is known in (t, ∞) .

Now, what happens if only $\bar{M}(t')$ is given in (t, ∞) ? Then $\Phi''(t) = \Psi_0(t) + ct$, where $\Psi_0(t)$ is a known function and c is a Gaussian variable. By (12), we see that the addition of the term ct to $\Psi(t)$ has the same effect as the addition of c to $\lambda X(t)$, and nothing is changed about the ξ_u ($u > t$). Then the ξ_u with subscripts $u > t$ in $M_2(t')$ are known.

2°. We shall now apply the same method as in Sec. 2.6. Let

$$(98) \quad \mathfrak{F}_t[y] = t^3 \frac{d^3 y}{dt^3} - 4t^2 \frac{d^2 y}{dt^2} + 8 \left(t \frac{dy}{dt} - y \right).$$

Then $\Gamma(t, t_0)$, when $t_0 < t$, is a solution of

$$(99) \quad \mathfrak{F}_{t_0}[\Gamma(t, t_0)] = 0. \ddagger$$

Let $m(t_0, t)$ be the conditional expectation of $M(t_0)$, ($t_0 \leq t$), when $\bar{M}(t')$ [or $M(t') - M(t)$] is given in (t, ∞) . We deduce from equation (99)

$$E\{[M(t') - M(t)]\mathfrak{F}_{t_0}[m(t_0, t)]\} = 0.$$

Since $m(t_0, t)$ is a linear functional of the given $M(t') - M(t)$, this equation implies

$$(100) \quad \mathfrak{F}_{t_0}[m(t_0, t)] = 0.$$

† In the general case $n = 2p + 1$, we have

$$2c_n^2 = \frac{1 \cdot 3 \cdot 5 \cdots (2p-1)}{2 \cdot 4 \cdot 6 \cdots (2p)} = \frac{2}{\pi} \int_0^{\pi/2} \sin^{n-1} \theta \, d\theta = \frac{2}{\pi} I_{n-1},$$

and $c_\infty = 0$, as was stated in the Introduction. Probably the formula $I_{n-1} = \pi c_n^2$ holds also if n is even.

‡ This equation may be deduced from (81). If $f(t)$ is a solution of (81), then $tf(1/t)$ is a solution of (98). The same remark holds for the study of $M(t)$ in the $(2p+1)$ -dimensional space.

Coming back to $M(t)$, if $t_0 = t + dt$, ($dt < 0$), we have

$$(101) \quad \delta M(t) = M'(t) dt + M''(t) \frac{dt^2}{2} \\ + [t^2 M''(t) - 2t M'(t) + 2M(t)] \frac{2dt^3}{3t^3} + \sigma \xi_t,$$

where σ^2 obviously has the same value $3(dt)^5/(5t^4)$ as in (93). Only the third-order term changes according to whether $dt > 0$ or $dt < 0$. In both cases, the Cauchy conditions are $M(t)$, $M'(t)$, $M''(t)$, and the differential equation gives the definition of the Markovian stochastic process of order 3.

Obviously, if t is replaced by u in (101), the ξ_u are the same as in $M_2(t)$ [formula (72)]. The ξ_u provide new information when $M(t)$, first given in (t, ∞) , is now given at the point $t + dt$, ($dt < 0$).

ÉCOLE POLYTECHNIQUE,
PARIS

A NEW LOOK AT THE FIRST BOUNDARY-VALUE PROBLEM

BY

J. L. DOOB

1. Introduction. The first boundary-value problem, sometimes called the Dirichlet problem, as discussed in this paper, is the problem of finding a function u which belongs to a specified linear class and has a specified boundary function f . A more precise statement of the problem will be given below. The problem has been so generalized in recent years that the relation between u and f is now determined by the method of obtaining u , given f , rather than by a specified limit relationship. It will be shown below, however, how the limit relationships holding in the simplest classical case, when the function u is harmonic and the domain considered is a plane disk, can be extended to very general cases. Since the full mathematical details will be published elsewhere, the main stress in this paper will be on motivation. To avoid the compulsion of choosing either imprecise statements or lengthy hypotheses, the hypotheses stated here will be far stronger than necessary. In fact they will be so strong that some applications, *e.g.*, the application to the first boundary-value problem on a discrete lattice, will be excluded.

The functions considered will be defined on a compact metric space $R \cup R'$. Here R is an open dense set of the space, with no points in common with R' , so that R' is the boundary of R . (In general, boundaries will be denoted by primes.) The first boundary-value problem will be considered on R , with boundary function defined on R' . It will not be necessary to consider $R \cup R'$ as a subset of a larger space, with R' the boundary of R relative to this larger space.

We now define the class of functions to be considered. It is supposed that there is a distinguished class of open sets, called *regular sets*, with the following properties. *The closure of a regular set is a subset of R . Every open subset D of R is the union of a monotone sequence of regular sets whose closures are subsets of D .* The functions considered are defined by their properties on regular sets. If D is a regular set, and if $z \in D$, it is supposed that there is a measure $\mu(z, D, \cdot)$ of Borel subsets of D' , with $\mu(z, D, D') = 1$, having the properties M1, M2, M3 listed below.

M1. *If f is a nonnegative Borel measurable function, defined on D' , the function u defined on D by the integral*

$$(1) \quad u(z) = \int_{D'} f(\xi) \mu(z, D, d\xi)$$

is finite-valued and continuous on an open set D_1 if it is finite-valued on a dense subset of D_1 . If f is continuous, u has the boundary function f in the sense that u is continuous on $D \cup D'$, if defined as f on D' .

The function defined by (1), for any integrand f , will be described as the regular average of f on D' , relative to D . A function u defined on an open subset of R will be called *regular* if it is continuous, and if, whenever D is a regular set whose closure is in the domain of definition of u , u on D is equal to the regular average of itself on D' , relative to D . Our second condition is now the following.

M2. *The average in (1) defines a function which is regular on any open set on which it is finite.*

A further hypothesis is needed to ensure that every function regular in a neighborhood of each point of its domain of definition is regular. This hypothesis will be formulated in terms of a covering concept which we now define. For each z, D , the intersection of all the closed subsets of D' , of $\mu(z, D, \cdot)$ measure 1 will be called the support of this measure function. The support is closed and has measure 1. Let R_1 be an open subset of R , and suppose that, to each point z of R_1 there is assigned a neighborhood $N(z)$ of itself, whose closure is a subset of R_1 . Then we shall say that *the point z' of R_1 covers the point z'' of R_1 , relative to R_1 and the assigned neighborhood system* if there is a finite succession $z_1 = z', \dots, z_{n+1} = z''$ of points of R_1 , $n \geq 1$, and of regular sets $D_1, \dots, D_n$, such that $D_j \subset N(z_j)$, and that z_{j+1} is a point of the support of the measure function $\mu(z_j, D_j, \cdot)$, for $j \leq n$. Our final condition can now be formulated.

M3. *Let D be a regular set, let z be a point of D , and let A be the set of points covered by z , relative to D and any specified neighborhood system. Then some point of D' is a limit point of A .*

In the application to harmonic functions, M3 is satisfied, and in fact in that case $A = D$. In the application to solutions of the heat equation, M3 is also satisfied, but A and D are never equal.

It will be necessary to define subregular and superregular functions also. A subregular function is a function u , defined on an open subset B of R , and satisfying the following conditions:

(a) $-\infty \leq u < \infty$; u is finite on a set dense in B .

(b) u is upper semicontinuous.

(c) *If D is a regular set whose closure is a subset of B , u on D is less than, or equal to, the regular average of u on D' relative to D .* (Alternatively, we could require that, if u_1 is a function defined and continuous on $D \cup D'$, regular on D , then $u \leq u_1$ on D if this inequality is valid on D' .)

A superregular function is defined as the negative of a subregular function.

In its simplest form, the first boundary-value problem can now be stated as follows. Let D be an arbitrary open subset of R , and let f be a continuous function defined on D' . A function u is to be found, defined and regular on D , with boundary function f . In particular, if D is regular, the problem is solvable, and in fact the solution is simply the regular average of f on D' relative to D . In any case there can be at most one solution. In fact the difference between two solutions is regular on D and has the boundary function

0. It is then easily seen that a form of the usual maximum principle, valid under our stated hypotheses, implies that this difference vanishes identically. If the problem is always solvable, whatever the continuous boundary function f , the value $u(z)$ of the solution at the point z defines a bounded linear functional of f (uniform topology), and therefore can be put in the form (1). Moreover the measure $\mu(z, D, \cdot)$ involved is a measure of Borel subsets of D' , and it can be shown that the family of measures has the properties M1 and M2. Thus, if desired, the set D can be adjoined to the class of regular sets.

In the discussion to be given below, we shall usually take D as R itself. This specialization involves no loss of generality and has certain conceptual advantages. In fact, in many applications, it is R that is given, and R' can then be specified in a variety of ways, depending on the desired degree of boundary ramification.

Before proceeding further, we note as an example the special case in which R is an open set in Euclidean N -space, and regularity of a function means harmonicity. In this case, the regular sets can be taken as any open sets with sufficiently smooth boundaries. More precisely, the regular sets are the open sets all of whose boundary points are regular in the usual potential-theoretic sense. This is the maximal class of regular sets. It would suffice for our conditions to take a much more restricted class. If R is not bounded, a topology can be introduced which will give R a compact closure. The measure $\mu(z, D, \cdot)$ is known as the harmonic measure on D' relative to D , at z . For D a sphere, this measure has a density (relative to ordinary surface measure on D') which is the Poisson kernel used in the classical integral formula expressing a harmonic function on a sphere in terms of its values on the boundary. We shall refer to the example of harmonic functions simply as the *harmonic case*.

2. The Perron-Wiener-Brelot method. We now sketch the Perron-Wiener-Brelot (PWB) method of solving the first boundary-value problem. This method began as a method for solving, in the harmonic case, the original problem stated in Sec. 1 but has now become the essential defining element of a generalized problem. We shall consider the problem for functions defined on R itself, having already justified this apparent specialization. We apply the method, as is now customary, assuming that the assigned boundary function f on R' is an arbitrary function, not even necessarily finite-valued. We assume, of course, that R' is not empty. The PWB method now proceeds as follows.

The lower [upper] class of functions is defined as the class of functions on R , including the function $-\infty$ [$+\infty$] and every subregular [superregular] function on R , bounded from above [below], whose upper [lower] limit at each point of R' is at most [least] the value of f at the point. The upper [lower] limit of functions in the lower [upper] class is called the lower [upper] solution. With the help of the weak form of the maximum principle that can be shown to be available here, it is shown that the upper solution is greater than, or

equal to, the lower solution. If these two solutions are finite and equal, it can be shown that they are regular, and the function u on R obtained in this way is called the first boundary-value problem solution on R , as determined by f . In this case the function f is called *resolutive*. If every continuous function on R' is resolutive, R is called resolutive. If R is resolutive, it does not follow that the first boundary-value-problem solution determined by a continuous function has that function as a boundary function in the usual sense. In fact this lack of a simple relationship between boundary function and solution has led to a division of boundary points into regular and irregular points, in the harmonic case, according as the solution has or has not the assigned boundary limit at the boundary point in question, for every continuous boundary function. If f is not even continuous but is resolutive, the relationship between f and the corresponding solution, aside from that explicit in the PWB method, is not at all clear. In the classical cases of interest, all open sets have been proved resolutive.

In applying the PWB method, we now see that the following problems are encountered. *How can one characterize the resolutive boundary functions and corresponding solutions? In what sense, if any, does a PWB solution have its determining boundary function as a limit at the boundary?* One can of course argue that these problems have no real interest, because they are created by the method rather than by the original problem, but this objection is only partially valid. In the first place, if the original problem has a solution, it is a PWB solution also. In the second place, the generalization furnished by the PWB method is a natural one, and the other ways used to generalize the original problem have in the past always led to PWB solutions. The PWB method has, however, one real defect. The class of resolutive boundary functions is a linear class, closed under monotone convergence (when dominated by a member of the class). But it is apparently not true that the maximum of two resolutive functions is resolutive, in particular that $|f|$ is resolutive whenever f is, and this defect prevents an immediate representation of the general solution in the form (1) with $D = R$, using a Daniell integral. It is possible to generalize the method to overcome this defect, however. (The details of this generalization will be presented elsewhere, but are sketched in Sec. 6.)

In the following sections, the first italicized question will be answered in part, and the second will be answered completely.

If it is known that R is resolutive, we can repeat an argument used already. To the continuous boundary function f corresponds the solution u , and if $z \in R$, $u(z)$ is a bounded linear functional of f . Hence $u(z)$ can be put in the form

$$(2) \quad u(z) = \int_{R'} f(\xi) \mu(z, R, d\xi).$$

Here $\{\mu(z, R, \cdot), z \in R\}$ is a family of measures satisfying conditions M1 and M2 (with $D = R$). It can then be shown that an arbitrary function f_1 on

R' is resolutive if and only if there is a Borel measurable function f on R' such that $f = f_1$, almost everywhere with respect to every measure in the above family, and that f is integrable with respect to every measure in this family. In that case, the PWB solution determined by f is given by (2). In the classical studies, which have concentrated on the harmonic case, the main problem has been to prove that R is always resolutive, and the proof has involved potential theoretic ideas, notably the concept of capacity. Such ideas are not available in our general theory, and we must in fact forego any discussion of the resolvability of R .

3. Review of the harmonic case. Suppose that the regular functions are the harmonic functions in N -space, and let D_r be the sphere of radius r , center the origin. Then the following facts are well known.

If u is subharmonic on D_1 , and if u is bounded from above by a nonnegative harmonic function, then u has a finite limit along almost every radius to D'_1 . Here the measure of radii refers to the surface measure of the set of their end points on D_1 , where the surface measure is normalized to have value 1 for the whole surface. The stated condition on u can also be written in the form

$$\sup_{r < 1} \int_{D_r} |u(\xi)| \mu(O, D_r, d\xi) < \infty.$$

In this form we have used the notation of Sec. 2, as applied in this case to harmonic measure. We recall that the harmonic measure used here, that on the surface of a sphere relative to the interior, at the center, is proportional to ordinary surface measure.

Furthermore, if u is a function harmonic on D_1 , it can be expressed by the Poisson integral,

$$u(z) = \int_{D_1} f(\xi) \mu(O, D_1, d\xi),$$

where f is a Lebesgue measurable and integrable function on D'_1 , if, and only if, u , considered only on D'_r , defines a family of functions, depending on the parameter r , which is uniformly integrable relative to the corresponding measure family $\{\mu(O, D_r, \cdot), r < 1\}$. Finally, if u is given by the Poisson integral, u has the function f as a boundary limit along almost all radii. In view of later generalizations to be made, we remark that, because of the special properties of harmonic functions, not shared by regular functions, in general, the above uniformity condition imposed on u implies the validity of the apparently stronger condition that, if $\{B_n, n \geq 1\}$ is any monotone sequence of regular sets whose closures lie in D_1 and whose union is D_1 , then u on B'_n defines a sequence of functions uniformly integrable with respect to the sequence of measures $\{\mu(z_0, B_n, \cdot), n \geq 1\}$, for every point $z_0 \in D_1$. Here we consider only values of n so large that $z_0 \in B_n$. A similar apparently stronger form can be given to the condition imposed above on the subharmonic function u .

It is natural to ask whether these theorems in the harmonic case can be extended to our general case. The first and most obvious task is to find a

reasonable substitute for radial approach to the boundary, as used in the harmonic case for spherical domains. We first give a suggestive method for finding such a substitute in the harmonic case, for arbitrary domains, and then extend the method to the general case.

4. Trajectories to the boundary. For simplicity, we consider first harmonic functions defined on an open plane set D . We wish to find a method of approaching the boundary which is determined by the intrinsic characteristics of the problem, *i.e.*, by the intrinsic characteristics of the class of regular functions considered. From this point of view, the success of the system of radial approach, evidenced by the results stated in Sec. 3, when D is a disk, must be considered accidental. The system of radial approach is not even invariant under conformal mapping of D onto itself, although harmonic functions go into harmonic functions under such mapping. If D is not a disk but at least has a smooth boundary, one can hope to prove some of the desired theorems by replacing radial approach by any system of nontangential approach to the boundary, and this has been done. But again this partial success can be considered a lucky accident, and the idea of nontangential approach is not much help if D is an arbitrary open set! The problem is brought into clearest focus if D is an abstract Riemann surface and if again regularity means harmonicity. In this case D has no relative boundary, since it is not a part of a larger surface as it is in the plane case. But any system of approach to the boundary in the plane case that depends only on the intrinsic properties of harmonic functions on D must remain meaningful if D is a Riemann surface when such properties are all that one can use. An appropriate system of boundary approach must imply a corresponding boundary concept in the case of a Riemann surface, and therefore in all cases. We shall verify this somewhat cryptic aphorism below.

The system of radial approach to a circular boundary, in the harmonic case, has one suggestive feature. Measure is defined on classes of radii in such a way that the distribution of the meeting points of radii and perimeter D' is the uniform distribution, which is, of course, the harmonic measure $\mu(z_0, D, \cdot)$, if z_0 is the center of D , the initial point of the radial trajectories. Now if we wish to find suitable trajectories from an arbitrary initial point, in an arbitrary open set, it is natural to try to keep this property in the small. This suggests the possibility that it may be possible to define a system of trajectories with initial point z_0 in such a way that, if D_1 is a regular set which contains z_0 , the distribution of first meeting points of trajectories with D'_1 , as determined by the measure function defined on sets of trajectories, should be the harmonic measure $\mu(z_0, D_1, \cdot)$. One natural way to try to obtain such a trajectory system would be as follows. If z_0 is the initial point, and if D_1 is a disk, center z_0 , radius ϵ , choose z_1 at random on D'_1 (uniform distribution). Let D_2 be a disk, center z_1 , radius ϵ , and choose z_2 at random on D'_2 , and so on. In this way a random sequence of points is obtained. If z_{n-1} is joined to z_n by a linear segment, a continuous random trajectory is obtained. The trajectory is, of course, a random walk, with steps of length ϵ and a uniform distribution of directions

at each step. It is well known that, for small ϵ , the random walk is very nearly plane Brownian motion. This suggests that suitable trajectories from z_0 may be the Brownian-motion trajectories. [Instead of taking D_j as a small disk, with center z_{j-1} , one could as well take D_j as a small open set containing z_{j-1} , as long as z_j is chosen on D'_j , with distribution the harmonic measure $\mu(z_{j-1}, D_j, \cdot)$.] The point of this development is that, for the trajectories chosen, the instantaneous motion is governed by harmonic measure, which is about all that is available. All we have done so far, of course, is to suggest why probability trajectories enter the first boundary-value problem in a natural way. We have not yet stated any theorems which justify their use.

In the general case, we can proceed as follows. Let R be as described in Sec. 1, and let $\{R_n, n \geq 1\}$ be a monotone sequence of regular sets, with union R , such that the closure of R_n is a subset of R_{n+1} . If z_0 is any point of R , let N be the first value of n for which $z_0 \in R_n$. Then choose $z_1 = \dots = z_{N-1} = z_0$. Choose z_N at random on R'_N , with the distribution $\mu(z_0, R_N, \cdot)$; then choose z_{N+1} on R'_{N+1} , with the distribution $\mu(z_N, R_{N+1}, \cdot)$, and so on. In this way a random walk on R is defined, which approaches R' . We do not have continuous paths, merely sequences. This seemingly unfortunate fact is inevitable, because we have not imposed enough hypotheses on the space to hope to be able to get continuous paths.

We have not formalized our trajectory systems in measure-theoretic language as yet. The relevant notation will be introduced in the next section. The formalization is very simple, since we are dealing with a Markov process, with transition probabilities given by the specified measure functions $\mu(z, D, \cdot)$.

5. Statement of results. We now come to the justification of our trajectory systems. Formally, as we have defined the trajectory systems, to each point z_0 of R there corresponds a measure space $\Omega(z_0)$, which may depend on z_0 , and a sequence of measurable functions on this space, $\{z_n, n \geq 0\}$, where z_n is a function with values on the set R'_n , as described in Sec. 4. Measurability of z_n means that the inverse image on $\Omega(z_0)$ of a Borel set on R_n under the map defined by z_n is measurable. If $z_n(\omega)$ is the value of z_n at the point ω of $\Omega(z_0)$, $\{z_n(\omega), n \geq 0\}$ is a sequence of points on R , which we shall call the ω trajectory. This sequence approaches R' , in the sense that only finitely many points of the sequence are in any compact subset of R . We can now talk of the measure of a set of trajectories, meaning the measure of the corresponding ω set.

Let u be a function on R . Then $\{u(z_n), n \geq 1\}$ is a sequence of functions on $\Omega(z_0)$. It turns out that, because of our definition of the trajectories, this sequence of functions has very simple properties if u is subregular or regular. In the language of probability theory, the function sequence is then a semimartingale or martingale, respectively. In view of the known properties of semimartingales and martingales the theorems stated in the harmonic case can be generalized as follows.

If u is subregular on R , bounded from above by a nonnegative regular function, then u has a finite limit along almost every trajectory from any point of

R. This condition on u is equivalent to the condition that, if $\{S_n, n \geq 1\}$ is some monotone sequence of regular sets, with union R ,

$$\sup_n \int_{S_n} |u(\zeta)| \mu(z, S_n, d\zeta) < \infty,$$

for all z in R . (This condition is satisfied for every S_n sequence if it is satisfied for a single one.) We denote by $\mathbf{H}$ the class of subregular functions satisfying this condition.

Let $\mathbf{D}$ be the class of functions u , regular on R , for which u on S'_n , as n varies, defines a sequence of functions uniformly integrable with respect to the measure sequence $\{\mu(z, S_n, \cdot), n \geq 1\}$, for all z in R . This uniform integrability condition is satisfied for every S_n sequence if it is satisfied for a single one. If $u \in \mathbf{D}$, then $u \in \mathbf{H}$ also, and u can be expressed in terms of its boundary limit on our trajectories. In fact, considering trajectories from z_0 , define $x(z_0)$, a function on the space $\Omega(z_0)$, by

$$x(z_0) = \lim_{n \rightarrow \infty} u[x(z_n)].$$

Then $u(z_0)$ is the integral of $x(z_0)$ on $\Omega(z_0)$, that is, in probability notation,

$$u(z_0) = \mathbf{E}\{x(z_0)\}.$$

In particular, it can be shown that, if f is a resolutive boundary function on R' , determining the PWB solution u , then $u \in \mathbf{D}$, and u has f as its boundary limit along our trajectories, in the following sense. (Note that the resolvitivity of R itself is not a hypothesis here.) Fix an initial point z_0 of trajectories, and let $R(\omega)$ be the set of limit points on R' of the ω trajectory. Then it can be shown that, if f is resolutive, f is constant on each set $R(\omega)$, for almost all ω . Without straining the notation too much, we then write $f[R(\omega)]$ for this constant value. It is also shown that, for almost all ω ,

$$\lim_{n \rightarrow \infty} u[z_n(\omega)] = f[R(\omega)],$$

so that u has the assigned boundary value as a limit along our trajectories. Finally, the above evaluation of u makes $u(z_0)$ an average of f on R' in the sense that

$$u(z_0) = \mathbf{E}\{f[R(\omega)]\}.$$

That is, somewhat roughly, the trajectories define a distribution on R' , by their convergence to boundary points, and $u(z_0)$ is the average of f relative to this distribution. This is the generalization of the Poisson integral in the harmonic case. The rough description of this expression for $u(z_0)$ will be made concrete in the next section. In the most important classical cases, $R(\omega)$ is a single point, for almost every ω . In particular, this is true if R is resolutive, and, in that case, the distribution of $R(\omega)$ is the measure $\mu(z_0, R, \cdot)$ defined above.

The trajectories we have used may be replaceable in particular cases by continuous paths. This is true, for example, as suggested in Sec. 4, in the har-

monic case, but the fact that the trajectories are not continuous in general, and cannot be made so, is actually no hindrance to the theory.

As we have already remarked, in many applications R and the class of regular functions are held fast, and R' is varied. This is what is done, for example, in the harmonic case, when the boundary is ramified. Changing R' may entirely change the class of PWB solutions. For example R' can always be made into a single point, since R is locally compact, and if this is done, the class of PWB solutions, and the resolutive boundary functions, will be constant functions on their respective domains. But no matter how R' is changed, the class of PWB solutions will always be a subclass of D , so that boundary limits will always exist along our trajectories, as stated above. And although the limit sets $R(\omega)$ may also change radically in nature, the results relative to these sets remain true. This remark suggests that the PWB method is non-intrinsic, in the sense that it is not intrinsic to the class of regular functions on R , and that it may be possible to define a fully intrinsic first boundary-value problem without explicit use of an assigned boundary! This will be done in Sec. 7. This remark on the PWB method is not a criticism. The method is, of course, tied to the nature of R' , because it was designed with the properties of regular functions in the neighborhoods of individual boundary points in mind. It is thus no criticism to remark that the properties of PWB solutions suggest the possibility of a generalized first boundary-value problem not involving any preassigned boundary. We shall see in the next section, however, that a simple generalization of the PWB method leads to a slightly larger class of solutions with desirable properties not possessed by the original class.

6. The generalized PWB method. Suppose that f is any function on R' , and that the PWB method is applied to define lower and upper classes of functions, relative to f , as described in Sec. 2. Then it is clear that, if u_1 is a function in the lower class, $u_1 \in H$, and if u_2 is a function in the upper class, $-u_2 \in H$. Hence u_1 and u_2 have boundary limits along trajectories, in the sense described in Sec. 5. In fact it can be shown that, choosing any initial point z_0 of trajectories, and using the notation of Sec. 5, ,

$$\begin{aligned}\lim_{n \rightarrow \infty} u_1[z_n(\omega)] &\leq \inf_{\zeta \in R(\omega)} f(\zeta) \\ \lim_{n \rightarrow \infty} u_2[z_n(\omega)] &\geq \sup_{\zeta \in R(\omega)} f(\zeta)\end{aligned}$$

for almost all ω . It is therefore natural to enlarge the lower and upper classes of functions, by admitting to the lower [upper] class every subregular [super-regular] function on R which is bounded from above [below] and satisfies the first [second] of the above equations, for almost all ω , and every initial point z_0 . The definitions of lower and upper solutions, resolutive of f , and so on, can then be made as usual, in terms of the enlarged classes. We shall call this method the stochastic Perron-Wiener-Brelot (SPWB) method. A SPWB first boundary-value-problem solution, corresponding to the SPWB-resolutive boundary function f , is in the class D , has f as boundary-limit function along

our trajectories, and is expressible in terms of f , in precisely the same way as a PWB solution. The class of SPWB-resolutive boundary functions has the properties of the PWB resolutive class, listed in Sec. 2, and in addition can be shown to have the property that, if f and g are SPWB-resolutive, their maximum is also. This last property can be used to show that there is, for each z in R , a measure $\mu(z, R, \cdot)$, with the following properties. Each measure is defined on a Borel field of Borel subsets of R' , independent of z , with value 1 on R' . If f_1 is a SPWB-resolutive boundary function, there is a Borel measurable function f on R' , integrable with respect to each measure in the family, and equal almost everywhere on R' to f_1 , with respect to every measure in the family. Moreover the SPWB solution determined by f_1 is given by

$$u(z) = \int_{R'} f(\xi) \mu(z, R, d\xi).$$

Conversely, if f_1 is a function on R' , and if there is a function f with the stated properties, then f_1 is SPWB-resolutive.

Thus the class of PWB solutions has been made part of a larger class, given by the usual elegant averages, and with the same boundary-limit properties as in the simplest case when R is PWB-resolutive. In particular, if R is PWB-resolutive, the classes of PWB- and SPWB-resolutive boundary functions coincide.

7. An intrinsic first boundary-value problem. Suppose that a trajectory system has been defined on R , as explained in Sec. 4. Let u be a Borel measurable function defined on R , and suppose that, using our customary notation for trajectory systems, $\lim_{n \rightarrow \infty} u(z_n) = x(z_0)$ exists with probability 1, as a finite or infinite limit, for each initial trajectory point z_0 . The family of random variables obtained in this way obviously has the property that each random variable $x(z_0)$ is defined in terms of z_n for $n \geq N$, for every N . Moreover, it can be shown that, if A is any linear Borel set, the probability that the value of $x(z)$ lies in A defines a regular function of z , on R . Any family $\{x(z), z \in R\}$ of random variables with these properties will be called a *stochastic boundary function*. It is to be understood here that there is a specified trajectory system and that $x(z)$ is a random variable defined on the measure space on which the trajectories with initial point z are defined. It can be shown that every stochastic boundary function is the stochastic boundary function of some function u on R , by which we mean that each random variable $x(z)$ of the stochastic boundary function is the limit of u (with probability 1) along trajectories from z . Note that R' is not involved in any way in this definition of a stochastic boundary function.

According to the theorems stated in Sec. 5, if $u \in \mathbf{H}$, u has a stochastic boundary function. Moreover, if $\mathbf{D}'$ is the class of stochastic boundary functions with the property that each random variable involved has a finite expectation, it can be shown that the stochastic boundary function of any function in the class $\mathbf{H}$ is in the class $\mathbf{D}'$. (The fact that the class $\mathbf{D}'$ is by definition

dependent on the nested sequence of regular sets used to define the trajectories will not affect the analysis to be made.) Conversely, to any stochastic boundary function $\{x(z), z \in R\}$ in the class $\mathbf{D}'$ corresponds a function u on R , by means of the definition

$$(3) \quad u(z) = \mathbf{E}\{x(z)\}.$$

It can be shown that this function is regular, is in the class $\mathbf{D}$, and has the given stochastic boundary function as its stochastic boundary function. In particular, every function in the class $\mathbf{D}$ has a stochastic boundary function, according to Sec. 5, and in turn is determined by this stochastic boundary function by way of (3).

In view of these facts, it is natural to define a generalized first boundary-value problem, in terms of R alone, in which what is given is a stochastic boundary function and what is to be found is a regular function with the specified stochastic boundary function. As usual in such studies, some condition must be imposed on u to ensure unique solutions, or else a method must be prescribed which leads to unique solutions. For example, it follows from our discussion that there is one and only one solution if the stochastic boundary function is in the class $\mathbf{D}'$ and if the solution is to be in the class $\mathbf{D}$. In fact the unique solution is given by (3). In order to avoid the *ad hoc* nature of a method in which we simply define the solution of the problem to be that given by (3), we describe a method of obtaining a solution which is more in the spirit of the PWB method.

Let $\{x(z), z \in R\}$ be any stochastic boundary function. We define lower and upper classes corresponding to this function as follows. A function v is in the lower [upper] class if it is identically $-\infty$ [$+\infty$] or if it is subregular [superregular], bounded from above [below], and if, using our customary notation for a trajectory system, with initial point z_0 ,

$$\lim_{n \rightarrow \infty} v(z_n) \leq x(z_0), \quad \left[\lim_{n \rightarrow \infty} v(z_n) \geq x(z_0) \right],$$

with probability 1, for every z_0 in R . The stated limits exist with probability 1 because $v[-v]$ is in the class $\mathbf{H}$. The remaining PWB-type definitions are formulated as usual. It can be shown that the given stochastic boundary function is resolute in the new sense if, and only if, it is in the class $\mathbf{D}'$, when the corresponding solution is given by (3).

We have seen that our previous formulations of the first boundary-value problem, *i.e.*, the PWB and SPWB methods, always lead to solutions in the class $\mathbf{D}$. Thus, if u is such a solution, as determined by the ordinary boundary function f , u has a stochastic boundary function $\{x(z), z \in R\}$, and, in the notation of our previous discussion,

$$x[z(\omega)] = f[R(\omega)],$$

with probability 1, for each z . Thus every PWB- or SPWB-resolute boundary function f corresponds to a stochastic boundary function, in the class $\mathbf{D}'$,

in such a way that the two determine the same first boundary-value-problem solution, according to the relevant determining methods. In a real sense, then, our present method is the strongest possible generalization of the previous ones. The class of solutions is maximal, and the method manufactures its own boundary, by the device of manufacturing its own "boundary functions" without reference to a boundary.

8. Applications. We have already remarked on the harmonic case, but a few more remarks may be illuminating. Suppose then that R is an open set in Euclidean N -space, and that the regular functions are the harmonic functions. Then the results of Sec. 5 are applicable if, for example, we define R' as the ordinary boundary of R and introduce a point at infinity to make $R \cup R'$ compact if R is unbounded. It is known that every open set is PWB-resolutive in the harmonic case. (See Brelot [1] for a full discussion of the PWB method in this case.) If R' has zero capacity, the first boundary-value problem becomes trivial, because the only functions in the class H are then constants, so that the only solutions in any of the senses we have defined are the constant functions. As already remarked, our discrete trajectories can be replaced by continuous ones in the harmonic case. (See Doob [4] for further details on this case.) In many applications, it is desirable to ramify the ordinary boundary of R , that is, to topologize R somewhat differently, and thereby be led to a boundary R' with the property that ordinary (relative) boundary points may be counted more than once, depending on the manner of approach. Brelot [2] has discussed the PWB method in detail in the ramified case. The approach we have given requires no change.

More generally, if R is a Riemann surface, or even more generally if R is what Brelot and Choquet [3] call a Greenian space, and again if the regular functions are the harmonic functions, the above remarks on the harmonic case remain valid. Note that, in this application, R' has no a priori boundary, in the sense of a relative boundary with respect to a larger space, but there are various ways in which a boundary may be defined and the PWB method applied, as shown by Brelot and Choquet. Their analysis depends to a considerable extent on the introduction of the capacity concept.

Generalizing the harmonic case in another direction, one can consider the solutions of the heat equation in N spatial dimensions. That is, R is now an open set in Euclidean $(N + 1)$ -space, and the regular functions are the solutions of the heat equation. In this case also, the sequential trajectories we have defined can be replaced by continuous curves. See Doob [5] for further details. It has not yet been proved that in this case all open sets are PWB-resolutive, although the assertion is probably true. At any rate, it is not difficult to find a class of regular sets suitable for our needs. When $N = 1$, for example, the regular sets can be taken as the bounded open sets whose boundaries are finite unions of polygons with no sides perpendicular to the time axis.

It is clear that, at least with suitable smoothness hypotheses on the coeffi-

cients, the results of this paper are applicable to solutions of general linear parabolic and elliptic second-order partial differential equations in Euclidean space of any number of dimensions. Tautz [6] has investigated the latter from a very general point of view, by first defining regular functions somewhat as in Sec. 1, applying the PWB method to solve the first boundary-value problem, and then imposing further hypotheses which ensure the resolvability of his domains and make the connection with differential equation classes. Our results are applicable to the classes of solutions he found.

BIBLIOGRAPHY

1. M. Brelot, *Familles de Perron et problème de Dirichlet*, Acta Univ. Szeged. vol. 9 (1939) pp. 133-153.
2. ———, *Le problème de Dirichlet "ramifié,"* Ann. Univ. Grenoble vol. 22 (1946) pp. 167-200.
3. ——— and G. Choquet, *Espaces et lignes de Green*, Ann. Inst. Fourier Grenoble vol. 3 (1951) pp. 199-263.
4. J. L. Doob, *Semimartingales and subharmonic functions*, Trans. Amer. Math. Soc. vol. 77 (1954), pp. 86-121.
5. ———, *A probability approach to the heat equation*, Trans. Amer. Math. Soc. vol. 80 (1955), pp. 216-280.
6. Georg Tautz, *Zur Theorie der ersten Randwertaufgabe*, Math. Nachr. vol. 2 (1949), pp. 279-303.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

ON BOUNDARIES DEFINED BY STOCHASTIC MATRICES¹

BY

WILLIAM FELLER

An operator A defined on a space E may in a natural way define a boundary for E . If A is unsymmetric, the adjoint operator may define a different boundary. The case of a denumerable space E is particularly instructive because of the absence of a topology defined independently of A . In this abstract we consider only the case of a denumerable E . The two boundaries need not be denumerable, and their topological structure may be complicated. These boundaries are useful for the formulation of lateral conditions and the construction of semigroups. These show surprisingly close analogies with harmonic functions, diffusion, etc.

1. The cone $\mathfrak{P}$. Let E be the set of points $1, 2, 3, \dots$. Boldface small letters will denote column vectors, that is, $\mathbf{x}$ stands for a vector with components $x(1), x(2), \dots$. We put $\|\mathbf{x}\| = \sup |x(i)|$. In particular $\mathbf{1}$ stands for the vector all of whose components equal 1.

Consider now a matrix Π with elements $\Pi(i, j)$ such that

$$(1) \quad \Pi(i, j) \geq 0, \quad \sum_j \Pi(i, j) \leq 1.$$

An equation of the form

$$(2) \quad \mathbf{y} = \Pi \mathbf{x}, \quad \text{that is,} \quad y(i) = \sum_j \Pi(i, j)x(j),$$

states that the value of $y(i)$ is an average of the values $x(j)$. Such an equation therefore has many properties in common with the relation

$$(3) \quad v = \Delta u,$$

where Δ is the Laplacian. In particular, the solutions of

$$(4) \quad \mathbf{x} = \Pi \mathbf{x}$$

are close analogues of harmonic functions.

We denote by $\mathfrak{P}$ the totality of all nonnegative solutions $\mathbf{x}$ of (4) such that $\|\mathbf{x}\| \leq 1$.

Each bounded solution of (4) is a linear combination of solutions in $\mathfrak{P}$. The set $\mathfrak{P}$ is a *positive cone* in the sense that if $\mathbf{x}, \mathbf{y} \in \mathfrak{P}$, then also $p_1 \mathbf{x}_1 + p_2 \mathbf{x}_2 \in \mathfrak{P}$ whenever $p_i \geq 0$ and $p_1 + p_2 \leq 1$.

It is well known that such a cone is completely determined by its *extremal elements* in the sense of Krein and Mil'man. We shall now describe these extremal elements and their probabilistic interpretation. It will be seen that

¹ Abstract.

the extremal elements determine a *boundary* for E which can be of a very complicated structure.

2. Extremal elements and sojourn probabilities. For an arbitrary set $\Gamma \subset E$ write

$$(5) \quad \Pi(i, \Gamma) = \sum_{j \in \Gamma} \Pi(i, j).$$

It is easily seen that the quantity $\Pi^n(i, E)$ decreases monotonically as $n \rightarrow \infty$, and we put

$$(6) \quad s_E(i) = \lim_{n \rightarrow \infty} \Pi^n(i, E).^2$$

Clearly $0 \leq s_E \leq 1$, and $s_E = 1$ if, and only if, $\Pi(i, E) = 1$, that is, if Π is stochastic. If $s_E = 0$, then $\mathfrak{P}$ is empty, and (4) has no bounded solutions. Under any circumstances s_E is the *maximal element* in $\mathfrak{P}$, that is,

$$(7) \quad \Pi s_E = s_E, \quad \text{and} \quad x \leq s_E \quad (\text{for each } x \in \mathfrak{P}).$$

Probabilistically the matrix Π defines a random walk (or Markov chain), and $s_E(i)$ is the probability that the processes starting at i will not terminate after finitely many steps.

We now define the *restriction of Π to the set Γ* by

$$(8) \quad \Pi_\Gamma(i, j) = \begin{cases} \Pi(i, j), & (\text{if } j \in \Gamma), \\ 0, & (\text{if } j \notin \Gamma). \end{cases}$$

By analogy with (6) we put

$$(9) \quad \sigma_\Gamma(i) = \begin{cases} \lim_{n \rightarrow \infty} \Pi_\Gamma^n(i, \Gamma), & (\text{if } i \in \Gamma), \\ 0, & (\text{if } i \notin \Gamma) \end{cases}$$

and

$$(10) \quad s_\Gamma(i) = \lim_{n \rightarrow \infty} \sum_j \Pi^n(i, j) \sigma_\Gamma(j), \quad (i \in E).$$

Of course, for most sets Γ we have $\sigma_\Gamma = s_\Gamma = 0$. If s_Γ (and therefore σ_Γ) does not vanish identically, then Γ is called a *sojourn set* and s_Γ the *sojourn probability corresponding to Γ* .

For our random walk, $\sigma_\Gamma(i)$ is the probability that, starting from $i \in \Gamma$, the system will never leave Γ . And $s_\Gamma(i)$ is the probability that, starting from an arbitrary i , the system will, after a finite number of steps, pass into Γ and stay there forever. If the chain is recurrent and irreducible, then E is the only sojourn set. If all states are transient, then the complement of any finite set is a sojourn set.

The sojourn probabilities coincide with the extremals of $\mathfrak{P}$ in the sense of Krein and Mil'man. Thus every $x \in \mathfrak{P}$ is essentially a positive linear combination of sojourn probabilities.

² Here and below, Π^n is the n th (matrix) power of Π .

3. Boundary points. If all states i are recurrent, then $\mathfrak{B}$ contains only scalar multiples of the solution **1**, and E has no boundary.³ In order not to burden our statements by obvious exceptions, we shall now assume that *all* states are transient. Analytically, this condition is equivalent to

$$(11) \quad \sum_n \Pi^n(i, i) < \infty, \quad (i = 1, 2, \dots).$$

Suppose then that Γ is a sojourn set. There exists a monotonically decreasing sequence of sets $\Gamma_n \subset \Gamma$ contracting to the empty set such that for each i

$$(12) \quad s_{\Gamma_n}(i) = s_{\Gamma}(i).$$

Probabilistically this means: the event "the system ultimately goes into Γ and stays forever in Γ " is equivalent to each of the events "the system ultimately stays in Γ_n ." In other words, if this event takes place, the system passes successively into $\Gamma_n, \Gamma_{n+1}, \dots$. Now $\Gamma_n \downarrow \emptyset$, and therefore the net $\{\Gamma_n\}$ may be described as an *escape route to infinity*.

To describe the simplest situation, consider the case where there exist exactly N linearly independent bounded solutions of (4). Then there exist exactly N *nonoverlapping sojourn sets* $\Gamma^{(1)}, \dots, \Gamma^{(N)}$, and

$$(13) \quad s_E = s_{\Gamma^{(1)}} + \dots + s_{\Gamma^{(N)}}.$$

Starting from an arbitrary point i , the process either stops after finitely many steps [with probability $1 - s_E(i)$], or it passes into one of the sets $\Gamma^{(k)}$ and ultimately remains restricted to this set. In the latter case the system moves along the escape route corresponding to $\Gamma^{(k)}$. In other words, there are N mutually exclusive escape routes to infinity, and the system is bound to move along one of them.

In this case we speak of a *boundary consisting of N isolated points*. If $\Gamma_1 \supset \Gamma_2 \supset \dots$ is an escape route, each Γ_k is a neighborhood of the corresponding boundary point.

The simplicity of the case just considered stems from the fact that with an *isolated* boundary point there is associated a solution $\mathbf{x} \in \mathfrak{B}$, that is, a *positive* probability of the system converging to it. Unfortunately, $\mathfrak{B}$ may contain a continuum of independent solutions, and the boundary may have a very complicated topological structure. To give a precise definition of the boundary in terms of $\mathfrak{B}$ it is necessary to consider $\mathfrak{B}$ as a lattice and define the boundary as the *maximal ideal space*. In this abstract we shall describe only the analogy with harmonic functions.

Let $\mathfrak{B}'$ be the cone of bounded positive functions harmonic, say, in the circular disk $r < 1$. Each $u \in \mathfrak{B}'$ can be represented by means of the Poisson integral in terms of its boundary values. A harmonic function u is an *extremal*

³ The set E corresponds to a closed manifold if the chain is irreducible and to a union of closed manifolds if the chain is reducible.

of $\mathfrak{P}'$ if its boundary values are 1 on a set Λ of the unit circle and 0 almost everywhere on the complement of Λ . Again such an extremal admits of a probabilistic interpretation: in the Wiener process (or ordinary diffusion) $u(P)$ represents the probability that a particle starting at P will ultimately be absorbed in Λ (rather than in the complement).

Thus the elements $u \in \mathfrak{P}'$ are in one-to-one correspondence with the sets of positive measure on the circle. To each *point* Q of the circle there corresponds an *unbounded* harmonic function, *viz.*, the radial derivative of the Green function $G(P, Q)$. In the trite case of a circular domain there is, of course, no point in defining the boundary in terms of $\mathfrak{P}'$. However, for more complicated domains, *e.g.*, when the boundary contains prime ends, it has been well known since the pioneer work of Martin that it is convenient to disregard the "ordinary" topology of the boundary and replace it by one more intimately connected with properties of the Δ -operator. Finally, in our case there is no *a priori* defined boundary, and we are compelled to introduce one in terms of Π .

4. Equivalent operators; unbounded solutions. Up to now we have considered only bounded solutions of (4). There exist also unbounded ones, but their study and their probabilistic meaning can be reduced to the bounded case. Let $\mathbf{z}$ be an arbitrary positive⁴ solution of $\Pi\mathbf{z} = \mathbf{z}$, and put

$$(14) \quad \Pi'(i, j) = \Pi(i, j) \frac{z(j)}{z(i)}.$$

Then Π' is a strictly stochastic matrix, that is, Π' satisfies the conditions (1) with strict equality at the second place. To each solution $\mathbf{x}$ of (4) the transformation

$$(15) \quad \mathbf{x} \rightarrow \mathbf{x}', \quad x'(i) = \frac{x(i)}{z(i)}$$

makes correspond a solution $\mathbf{x}'$ of $\mathbf{x}' = \Pi'\mathbf{x}'$. Clearly $\mathbf{x}'$ is bounded if, and only if, $\mathbf{x}$ is dominated by a scalar multiple of $\mathbf{z}$. If Π is strictly stochastic, then the relation (15) with $\mathbf{x} = \mathbf{1}$ shows that (14) is reversible and the relation between Π and Π' symmetric.

We see thus that (14) establishes an equivalence relation between strictly stochastic matrices and that there is a one-to-one correspondence between the solutions (4) for equivalent matrices. If in (14) the solution $\mathbf{z}$ is unbounded, the boundary of Π as defined in Sec. 3 can be enlarged by adding to it the boundary of Π' . In short, a bigger and better boundary may be defined by considering the *total* boundary of E relative to *all* solutions of $\Pi\mathbf{x} = \mathbf{x}$ or, what amounts to the same, relative to the set of all equivalent matrices Π' .

In this formulation the procedure is bound to appear artificial. Its use in the formulation of the appropriate lateral conditions for the Kolmogoroff differential equations shows that notions are natural and applicable. More-

⁴ Only obvious conventions are required if $z(i) = 0$ for certain i .

over, (14) is a close analogue of certain transformations frequently used for Sturm-Liouville operators and in diffusion theory. Probabilistically, Π' is obtained from Π by a conditioning process. The reader is asked to interpret (14) in the simple case where z is a sojourn probability.

5. The adjoint boundary. Up to now Π was considered as an operator by premultiplication on column vectors. The adjoint operation consists in postmultiplying row vectors by Π . To (4) there corresponds the equation

$$(16) \quad \mathbf{a} = \mathbf{a}\Pi, \quad \text{that is,} \quad a(j) = \sum_i a(i)\Pi(i,j).$$

For the case of a probability vector, that is, when $a(i) \geq 0$, $\sum a(i) = 1$, Kolmogoroff has introduced the matrix

$$(17) \quad \Pi^*(i,j) = \Pi(j,i) \frac{a(j)}{a(i)}$$

as the time reversion of Π . Clearly Π^* is a strictly stochastic matrix. Recent investigations by Derman, Harris, and Robbins show that it is useful and meaningful to consider (17) also when $\sum a(i) = \infty$.

We can now define the boundary of Π^* , and in this way endow E with two boundaries. Note that if $\mathbf{a}$ in (17) is replaced by another solution of (16), then Π^* is replaced by an equivalent stochastic matrix so that the new boundary of E is independent of the choice of $\mathbf{a}$.

In Sec. 3 we have seen that the random walk defined by Π moves *loward* the boundary as there defined. Probabilistically, therefore, this boundary is an *exit boundary*. For the matrix Π^* the direction of motion is reversed, and therefore the exit boundary for Π^* is an *entrance boundary*. In principle, the two boundaries occur in ordinary diffusion theory, say, in an interval $a \leq x \leq b$. In the classification introduced by the author, b may play the role of an exit boundary, an entrance boundary, of both (then it was called "regular"), or of neither ("natural"). In the spirit of the present theory a natural boundary is no boundary at all.

However, there is an essential and interesting difference between our case and all classical cases: usually the underlying space has a natural topology, and a point of the "ordinary" boundary may or may not belong to either of the two boundaries defined by an operator. Our space E is not endowed by any natural topology, and *there is absolutely no connection between the two boundaries*. Three points of the exit boundary together may form two points of the entrance boundary, etc. The more remarkable is the fact that in the case of differential equations our boundaries lead to boundary conditions in every respect analogous to classical boundary conditions.

6. The Kolmogoroff differential equations. Up to now we have considered only iterations of the operator Π . Instead of the integral exponents we now introduce a continuous time parameter, *i.e.*, we deal with matrices $P(t; i, j)$

which satisfy the conditions (1) and have the semigroup property

$$(18) \quad P(t+s) = P(t)P(s), \quad (t, s > 0).$$

As Π defined a random walk, so P defines a stochastic process in E .

Kolmogoroff has pointed out that an interesting class of matrices P satisfies the two infinite systems of differential equations,

$$(19) \quad P'(t) = -pP(t) + p\Pi P(t),$$

and

$$(20) \quad P'(t) = -P(t)p + P(t)p\Pi,$$

together with the initial condition $P(0) = \text{identity matrix}$. Here Π stands for the matrix (1), and p is a diagonal matrix with elements $p_1, p_2, \dots$. For our stochastic process the transition probabilities $\Pi(i, j)$ represent a road system which determines where the system can move to from i . The p_i regulate the speed of the movement, inasmuch as $1/p_i$ is the expected duration of an uninterrupted stay of the system at i .

The author showed in 1940 that there always exists a matrix P satisfying both (19) and (20) but there can exist infinitely many admissible common solutions to these differential equations. This phenomenon is neatly explained by the notion of the two boundaries. In fact, in connection with (19) and (20) we encounter all the phenomena known for parabolic partial differential equations. Thus we encounter the classical boundary conditions of the Sturm-Liouville theory: absorbing, reflecting, elastic barriers, etc. The semigroups and processes connected with (19) and (20) are the same as those of diffusion theory, except that the boundaries can be more complicated, so that we find a much greater variety. Of course, there are no normal or partial derivatives in a denumerable space, and the boundary conditions must be written in a new analytical form. However, this form carries over to partial differential equations, and we are able to view many known phenomena from a more comprehensive point of view.

PRINCETON UNIVERSITY,
PRINCETON, N.J.

ON THE APPLICATION OF FUNCTIONAL CALCULUS TO THE STATISTICAL THEORY OF TURBULENCE

BY

EBERHARD HOPF

We consider an incompressible fluid of density 1 and nonvanishing viscosity μ which occupies the entire x -space, $x = (x_1, x_2, x_3)$. An instantaneous velocity field of the fluid is denoted by $u = u(x)$, $u = (u_1, u_2, u_3)$,

$$(1) \quad \operatorname{div} u = 0.$$

In order to render the flow problem—determination of $u(x, t)$, $t > 0$, by means of the Navier-Stokes equations of flow if $u(x, 0) = u(x)$ is given—a reasonable one, certain conditions will have to be placed upon $u(x)$ at $x = \infty$ and at all times t . We imagine that sufficient conditions of this sort are roughly that averages over the x -space of quantities such as $u_\alpha(x)$ and $u_\alpha(x)u_\beta(x+a)$, a being an arbitrary vector, exist. Exact specification of the “admissible” fields u and precise formulation of this flow problem are matters probably beyond our present mathematical capabilities. The space of these admissible fields u will be called phase space of the fluid or u -space. We presume that the initial-value problem (given a point u in u -space, find $T^t u = u(x, t)$, $t > 0$, from the Navier-Stokes equations such that $u(x, 0) = u$) has a unique solution in u -space. Consider, at time $t = 0$, a cloud or ensemble of phases u , in analytical terms, a probability distribution $P(A)$ in u -space, A being a variable and essentially arbitrary part of u -space, and $P(A)$ being the probability that the phase is in A . The hydrodynamical phase motion T^t carries this ensemble into another one at time t with a well-determined probability distribution $P^t(B) = P(T^{-t}B)$, where $T^{-t}B$ is the part of u -space whose phases are carried into a part B after time t . Consider, furthermore, a number-valued functional $F(u)$ in u -space. Examples of such functionals are $u_\alpha(x^0)$, $u_\alpha(x^0)u_\beta(x^1)$, where x^0 , x^1 are fixed points in fluid space. The phase average or u -average of a functional $F(u)$, with respect to a definite moving phase distribution P^t , is defined by the Lebesgue-Stieltjes integral extended throughout u -space

$$(2) \quad \bar{F} = \int F(u) P^t(du).$$

It is the expectation of F with respect to the probability distribution P^t . It depends on the initial distribution $P = P^0$ and, in general, on t .

Theoretically, the motion of a phase ensemble is well determined by the individual phase motion $T^t u$ or, in other words, by the solutions of the Navier-Stokes equations. This literal description of the motion of an ensemble requires the knowledge of those solutions, and, hence, it is an important task of the statistical theory of turbulence to replace this description by an analytically more tractable one. A possible way to achieve this is to use the Fourier

transform or characteristic functional of P^t rather than the intractable P^t itself [2]. We consider, besides the solenoidal vector fields $u(x)$, entirely arbitrary sufficiently smooth vector fields $y = y(x)$ in x -space, which we suppose to vanish sufficiently rapidly at $x = \infty$. Their totality is called y -space. We define the scalar product of a field u and a field y by

$$(3) \quad (u, y) = \int u(x) \cdot y(x) dx = \int u_\alpha(x) y_\alpha(x) dx,$$

where $dx = dx_1 dx_2 dx_3$ and where the integration extends over the entire x -space. The exponential

$$\exp [i(u, y)]$$

is a functional of u in u -space which also depends on a parameter field $y = y(x)$. Consider, for fixed y , its average with respect to a moving phase ensemble P^t ,

$$(4) \quad \Phi(y, t) = \overline{\exp [i(u, y)]}.$$

This is a functional of y in y -space and, in general, a function of t . For fixed t , it represents the Fourier-Stieltjes transform or characteristic functional of the probability distribution P^t in u -space.

Functional derivatives of first order of a functional $\Phi(y)$,

$$(5) \quad \frac{\partial \Phi}{\partial y_\alpha(x) dx},$$

we define by the differential

$$\delta \Phi = \int_x \frac{\partial \Phi}{\partial y_\alpha(x) dx} \delta y_\alpha(x) dx.$$

(5) depends on a continuously variable point x and an index $\alpha = 1, 2, 3$. If (5) exists everywhere, *i.e.*, for every vector field $y(x)$, it is, of course, also a functional of y in y -space. For the case of a characteristic functional $\Phi(y)$ of a probability distribution in u -space, the functional derivatives of various orders, taken at the origin $y = 0$, that is, for the identically vanishing field $y = 0$, furnish the moments of various orders of this distribution (if they exist) or, in hydrodynamical language, the velocity-correlation functions. For instance, we obtain from (4)

$$\left. \frac{\partial \Phi}{\partial y_\alpha(x) dx} \right|_{y=0} = \overline{i u_\alpha(x)}$$

and

$$(6) \quad \left. \frac{\partial^2 \Phi}{\partial y_\alpha(x) dx \partial y_\beta(x') dx'} \right|_{y=0} = -\overline{u_\alpha(x) u_\beta(x')}.$$

The solenoidal character (1) of the basic fields u expresses itself in the property that

$$\frac{\partial}{\partial x_\alpha} \frac{\partial \Phi}{\partial y_\alpha(x) dx} \equiv 0$$

or in the equivalent property that $\Phi(y)$ depends only on the solenoidal part $\tilde{y}(x)$ in the representation

$$(7) \quad y(x) = \tilde{y}(x) + \text{grad } h(x), \quad \text{div } \tilde{y} = 0, \quad \tilde{y}(\infty) = 0$$

of the arbitrary field $y(x)$.

The author has shown [2] that, in consequence of the Navier-Stokes equations of individual flow, the Fourier transform (4) of any moving phase ensemble satisfies for all fields y the functional differential equation

$$(8) \quad \frac{\partial \Phi}{\partial t} = L_2(\Phi) + \mu L_1(\Phi),$$

where L_1 and L_2 are linear functional differential operators of first and second order, respectively,

$$(8a) \quad L_1(\Phi) = \int y_\alpha(x) \frac{\partial^2}{\partial x_\beta \partial x_\beta} \left(\frac{\partial \Phi}{\partial y_\alpha(x)} dx \right) dx,$$

and

$$(8b) \quad L_2(\Phi) = i \int \tilde{y}_\alpha(x) \frac{\partial}{\partial x_\beta} \left(\frac{\partial^2 \Phi}{\partial y_\beta(x) dx \partial y_\alpha(x) dx} \right) dx.$$

L_1 arises from the viscosity terms of the Navier-Stokes equations, L_2 from the inertia terms. The integrals are extended over the entire x -space.

The arbitrary field $y(x)$ plays the role of the argument or "independent variables" in Φ and in (8), and (8) represents a direct differential description of the evolution of a phase ensemble moving in accordance with the hydrodynamical laws. A more convenient form of the Φ -equation (8) is obtained by a transformation from the "independent variables" $y(x)$ to other "independent variables" $z(k)$. $z(k)$ is a vector field with complex components $z_\alpha(k)$ in a k -space, $k = (k_1, k_2, k_3)$, the wave-number space, which is related to the field $y(x)$ in x -space by the linear Fourier transformation

$$(9) \quad y(x) = (2\pi)^{-3} \int e^{ik \cdot x} z(k) dk$$

and its inverse

$$(9a) \quad z(k) = \int e^{-ix \cdot k} y(x) dx.$$

An exact specification of the admissible fields $y(x)$ and their transformed fields $z(k)$ cannot be attempted here. Suffice it to say that the fields $z(k)$ are largely arbitrary except that they vanish at $k = \infty$ sufficiently rapidly and that ($z^* = \text{conjugate of } z$)

$$\begin{aligned} z^*(k) &= (z_1^*(k), z_2^*(k), z_3^*(k)) \\ &= (z_1(-k), z_2(-k), z_3(-k)) = z(-k), \end{aligned}$$

which expresses the real-valuedness of $y(x)$. The convenience of the transformation (9) rests upon the classical fact that x -differentiations go over into

simple algebraic operations. So the decomposition (7) of a field into a solenoidal field $\tilde{y}(x)$ and a gradient field becomes

$$z(k) = \tilde{z}(k) + \varphi(k)k, \quad k \cdot \tilde{z} = 0,$$

φ being a scalar. In other words, $\tilde{z}$ is simply

$$(10) \quad \tilde{z} = z - \frac{k \cdot z}{z^2} k, \quad \text{or} \quad \tilde{z}_\alpha = \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) z_\beta,$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol. The functional (4) becomes a functional $\Phi(z, t)$ of a field $z(k)$. The solenoidal character of the velocity fields $u(x)$ expresses itself in the condition

$$(11) \quad k_\alpha \frac{\partial \Phi}{\partial z_\alpha(k) dk} \equiv 0$$

or in the equivalent condition that $\Phi(z)$ depends only on the solenoidal part $\tilde{z}(k)$ of $z(k)$.

If, by means of (9) and (9a), the functional derivatives of Φ with respect to the field y are expressed in terms of those with respect to the field z (the transformation is carried out in a purely formal manner), then the operators L_1 and L_2 in the Φ -equation (8) become

$$(12a) \quad L_1(\Phi) = \int_k k^2 z_\alpha(k) \frac{\partial \Phi}{\partial z_\alpha(k) dk} dk$$

and

$$(12b) \quad L_2(\Phi) = \int_{k'} \frac{\partial}{\partial z_\beta(k') dk'} \left[\int_{k''} \tilde{z}_\alpha(k' + k'') k''_\beta \frac{\partial \Phi}{\partial z_\alpha(k'') dk''} dk'' \right] dk',$$

the integrations being extended over all k -space. The Φ -equation is to be satisfied for every $z = z(k)$ and for every $t \geq 0$. Actually, the form of L_2 arrived at is not (12b) but the one with the functional differentiation in front of the square bracket carried out inside. The reason why the two first-order differentiations have been separated in (12b) is that, for most distributions considered in turbulence theory, the second derivatives

$$\frac{\partial^2 \Phi}{\partial z_\beta(k') dk' \partial z_\alpha(k'') dk''}$$

are Dirac functions of the two points k', k'' , zero for $k' + k'' \neq 0$ and in general infinite otherwise. The form (12b) completely avoids this trouble (see [2]). The direct derivation of (12a) and (12b) by means of (9) and (9a) is the one used by us originally (unpublished). In the author's paper [2] a different derivation of this k -space form of the Φ -equation is given. For reasons which we cannot go into here, this second derivation though also purely formal appears more convincing to us.

From now on we confine ourselves to phase ensembles with

$$\text{Average velocity} = 0,$$

in other words, to ensembles for which

$$\left. \frac{\partial \Phi}{\partial z_\alpha(k)} \frac{dk}{dk} \right|_{z=0} \equiv 0.$$

The double-correlation functions of the velocities are given by the formula (see [2])

$$(13) \quad \int_k \frac{\partial}{\partial z_\alpha(k)} \frac{dk}{dk} \left[\int_{k'} e^{-i(k \cdot x + k' \cdot x')} \frac{\partial \Phi}{\partial z_\beta(k')} \frac{dk'}{dk'} \right] dk \Big|_{z=0} = -\overline{u_\alpha(x) u_\beta(x')}.$$

From this formula another formula is derived for the correlations of velocity differences. The separation of the two first-order differentiations serves the same purpose as in (12b). (13) is, of course, the k -space analogue of (6). The formal Taylor series of Φ about $y = 0$ or $z = 0$ has the form

$$(14) \quad \Phi = 1 + \Phi_2 + \Phi_3 + \dots,$$

where Φ_n is a polynomial functional in y or z homogeneous of degree n . Φ_2 is a quadratic (Hermitian) form. From the power series for the exponential in (4) it follows that

$$\Phi_n = \frac{i^n}{n!} \overline{(u, y)^n}$$

and, in particular, that Φ_2 is negatively definite. In general, the polynomial Φ_n will have the form of an n -fold k -space integral.

In the important case where the phase ensemble has a probability distribution invariant under arbitrary translations in x -space — called homogeneous distribution in the theory of turbulence — $\Phi(z)$ is invariant under replacement of $z(k)$ by $e^{iak}z(k)$, a being an arbitrary real vector. In this case — we confine ourselves to it from here on — the integral representing $\Phi_n = \Phi_n(z)$ is extended only over the diagonal subspace $k^1 + \dots + k^n = 0$ of $(k^1, \dots, k^n)$ -space. For example, Φ_2 has the form

$$(15) \quad \Phi_2(z) = -\frac{1}{2} \int \Gamma_{\alpha\beta}(k) z_\alpha(k) z_\beta(-k) dk,$$

where, in virtue of $z^*(k) = z(-k)$, it may be assumed that

$$\Gamma_{\alpha\beta}^*(k) = \Gamma_{\beta\alpha}(k) = \Gamma_{\alpha\beta}(-k).$$

From $\Phi_2 \leq 0$ it follows that the Hermitian form $\Gamma_{\alpha\beta}(k) z_\alpha z_\beta^*$ is ≥ 0 for every k . If (14) and (15) are inserted into (13), it follows that

$$\overline{u_\alpha(x) u_\beta(x')} = \int \Gamma_{\alpha\beta}(k) e^{ik \cdot (x' - x)} dk.$$

In other words, $\Gamma_{\alpha\beta}$ is the spectral energy tensor of the theory of homogeneous turbulence. If it is desired to use the incompressibility, this can be done right

in the beginning, $\Phi = \Phi(\bar{z})$ and, hence,

$$\Phi_2(z) = -\frac{1}{2} \int f_{\alpha\beta}(k) \bar{z}_\alpha(k) \bar{z}_\beta(-k) dk.$$

By virtue of (10) and (15) this implies the parametric representation

$$\Gamma_{\alpha\beta}(k) = \left(\delta_{\nu\alpha} - \frac{k_\nu k_\alpha}{k^2} \right) \left(\delta_{\mu\beta} - \frac{k_\mu k_\beta}{k^2} \right) f_{\nu\mu}(k)$$

of $\Gamma_{\alpha\beta}$.

If, in addition to being homogeneous, the phase distribution is isotropic, or, in other words, if $\Phi(z)$ is unchanged by any rotation and reflection of the field $z(k)$ about $k = 0$, we must have

$$f_{\alpha\beta}(k) = \delta_{\alpha\beta} f(|k|), \quad (f \geq 0)$$

or

$$(16) \quad \Phi_2 = -\frac{1}{2} \int f(|k|) \bar{z}(k) \cdot \bar{z}(-k) dk.$$

This implies that

$$\Gamma_{\alpha\beta}(k) = \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) f(|k|)$$

and

$$\overline{u_\alpha(x) u_\beta(x')} = \int \left(\delta_{\alpha\beta} - \frac{k_\alpha k_\beta}{k^2} \right) f(|k|) e^{ik \cdot (x' - x)} dk.$$

$f(|k|)$ is the density of the distribution of the average energy

$$\frac{1}{2} \overline{u(x) \cdot u(x)}$$

over the wave-number space.

The only nontrivial solutions of the Φ -equation known to the author (see [2]) are time-independent solutions in the case of zero viscosity μ ,

$$(17) \quad \Phi(z) = W(\int \bar{z}(k) \cdot \bar{z}(-k) dk),$$

where $W(s)$ is continuously differentiable twice. The most interesting case is that where

$$(18) \quad W(s) = e^{-(\kappa^2/2)s}.$$

This is the Fourier transform of a Gaussian distribution of solenoidal vector fields $v(k)$, $k \cdot v(k) = 0$, in k -space with mean values zero, isotropic and with no correlation between distinct parts of k -space — we hope the mathematical reader forgives us the use of this loose language. Actually, the individual values of $v(k)$ are here in general meaningless, and the actual random variable is the additive set function represented by the indefinite integral $\int v dk$. The Taylor expansion of $W(s)$ about $s = 0$ shows that, if $W(0) = 1$,

$$\Phi_2 = W'(0) \int \bar{z}(k) \cdot \bar{z}(-k) dk$$

and, in virtue of (16), that the density of average energy in this case is con-

stant throughout k -space (equidistribution of energy over the different wave numbers). Gaussian distributions with nonconstant $f(|k|)$ cannot be time-independent solutions of the Φ -equation with $\mu = 0$ [3].

Experimental results appear to indicate that the Gaussian solution is approximately realized through part of the spectrum of highly turbulent flow. But even if the Gaussian solution is actually realized, it says itself hardly anything about physically significant quantities such as the average dissipation per unit mass

$$\epsilon = \mu \int k^2 f(|k|) dk.$$

Full knowledge about the relations between ϵ , μ , and the double correlations, for instance, must come from a certain presently unknown solution

$$\Phi(z; \mu, \epsilon) \quad \text{or} \quad \Phi(z, t; \mu, \epsilon)$$

of the Φ -equation. In a famous paper Kolmogoroff was able to deduce part of this relation and, in particular, its limit form as $\mu \rightarrow 0$ from a highly plausible ergodic hypothesis (the reader will find a clear exposition of this universal equilibrium theory in Batchelor's outstanding book [1]). This limit form tells us that

$$f(|k|) = \text{const.} \cdot |k|^{-5},$$

which is at variance with the equidistribution of the Gaussian solution. One can at present only speculate about the apparent discrepancy. Possibly the two ranges of the spectrum in which these two laws are approximately valid tend to separate as $\mu \rightarrow 0$.

It seems that the idea of a normal distribution cannot be altogether discarded. As far as approximate procedures are concerned it seems natural to try instead of a power series a Bruns-Charlier series

$$(19) \quad \Phi = e^{\Phi_2} (1 + \Phi_3 + \dots)$$

in order to satisfy the Φ -equation. On equating polynomials of equal degree on both sides we find equations

$$(20) \quad \begin{aligned} \frac{\partial \Phi_2}{\partial t} &= L_2(\Phi_3) + \mu L_1(\Phi_2), \\ \frac{\partial \Phi_3}{\partial t} &= \frac{1}{2} L_2(\Phi_2^2) + L_2(\Phi_4) + \mu L_1(\Phi_3), \end{aligned}$$

etc. The equation for $\partial \Phi_n / \partial t$ always contains Φ_{n+1} on its right. A first approximation would be to neglect Φ_4 in the second equation. The two equations then constitute a complete system of integrodifferential equations for Φ_2 and Φ_3 . The same two equations are obtained if a power series (14) is inserted into the Φ -equation and if then in the resulting equation $\partial \Phi_3 / \partial t = \dots$ the polynomial Φ_4 is replaced by $\Phi_2^2/2$, in other words, if in this equation the exact relation between fourth and second moments valid for a normal distribution is used. In this latter form the idea has been tried repeatedly since 1941.

However, this is not the place to discuss the results in detail, and we may be forgiven for referring only to the most recent paper [4] on the subject. Let it suffice that (19) has a chance of success only if as $\mu \rightarrow 0$ the remainder $\Phi_3 + \dots$ becomes negligible. However, this can hold only in a restricted sense, and the question which quantities can be obtained in this way asymptotically correctly is a highly intricate one. The full theory of the Kolmogoroff range of the spectrum of turbulence cannot be expected to come out of the Bruns-Charlier expansion.

A different approach toward an exact theory of the quasi-equilibrium range of the spectrum of turbulence was started by us three years ago. The idea was to separate the variables in the Φ -equation. We tried this first in one particular way assuming that

$$(21) \quad \Phi(z, t) = \Psi(w), \quad w(k) = \sqrt{f(|k|, t)} z(k),$$

where $f(|k|, t)$ is the spectral energy density associated with the phase ensemble. In other words, we assumed that $\Phi(z, t)$ depends on the time t only through the mediation of the spectral distribution of energy. This assumption leads to a split of the Φ -equation into a functional differential equation for the time-independent universal functional Ψ and an integrodifferential equation $\partial f / \partial t = \dots$, which resembles somewhat the famous approximate equation designed for f by Heisenberg on certain physical assumptions. Unfortunately, the particular setup (21) failed. An extension of the Φ -theory including space-time correlations of the velocities which we developed thereafter gave us a simple reason why (21) is bound to fail. We feel, however, that the failure of the particular setup (21) is no reason against the general idea of separating the double correlations out of the Φ -equation. This can be tried in different ways, and the author knows no offhand reason why this idea should be unsound. Unfortunately, such attempts are laborious as long as we are not in possession of a simplifying formal technique and are bound to take time.

A rough comparison with the formal mathematical structure of classical statistical mechanics of conservative systems of many mutually interacting particles should not be out of place if we try to assess the possibilities of the Φ -approach to statistical hydrodynamics. There is, of course, the familiar analogy between the largeness of the number of particles and the largeness of the over-all Reynolds number with the associated largeness of the range of eddy sizes present in the turbulent fluid. How does classical statistical mechanics reach its goal, *viz.*, the evaluation of the statistical averages of physical quantities? It starts, first of all, from a known invariant volume measure in the phase space of the system. This measure is a simple time-independent solution of Liouville's differential equation for the propagation of a phase distribution. The dynamical system is postulated (ergodic hypothesis) to have at any energy level essentially no other invariant measure than the invariant measure induced on the energy surface by the volume measure.

The next step consists in the evaluation of averages based on this measure

and, in particular, in the determination of the simplified asymptotic form which these complicated expressions have if the number of particles is sufficiently large.

In statistical hydrodynamics such a procedure cannot be followed unless an analogous relevant measure or solution of the "Liouville equation" is found in the hydrodynamical case. Our Φ -equation is a particular form of this Liouville equation. If we consider for a moment a finite-dimensional differential system

$$\frac{du}{dt} = U(u), \quad u = (u_1, \dots, u_n), \quad U = (U_1, \dots, U_n),$$

then the Liouville equation for a moving probability distribution $P^t(A)$ in u -space is, if we write it in a modernized form,

$$(22) \quad \frac{d}{dt} \int_u \varphi(u) P^t(du) = \int_u \varphi'_\alpha(u) U_\alpha(u) P^t(du).$$

This is to hold for every continuously differentiable function $\varphi(u)$. If P^t has a smooth density with respect to the ordinary volume in u -space, then this is equivalent to the ordinary differential form of the Liouville equation for the density function. For the validity of (22) for all φ it suffices that it be valid for certain particular functions φ , for instance, for the functions

$$(23) \quad \varphi(u) = e^{i(y,u)}, \quad y = (y_1, \dots, y_n), \quad (y,u) = y_\alpha u_\alpha$$

with an arbitrary parameter point y . If the right-hand sides U_α of the differential system are polynomials of the u_β , then the validity of (22) with (23) for every y can be shown to imply that

$$\Phi(y,t) = \int_u e^{i(y,u)} P^t(du)$$

satisfies a certain linear partial differential equation. Our Φ -equation is just this equation for the hydrodynamical case—the Navier-Stokes equations are of polynomial character. The use of the Φ -form of Liouville's equation avoids, of course, the direct use of the measure P or P^t . Instead of forming averages by integration with respect to this measure one gets them by functional differentiation out of Φ . In this way, one gets first the velocity correlations. But it is a fact that all the physically important averages can be obtained from the correlations by x -differentiations and -integrations. To the asymptotic simplification of the averages in classical statistical mechanics of systems of very many particles there should correspond a more simple asymptotic form of the hydrodynamic functional Φ for a phase distribution with very large over-all Reynolds number. At present the determination of the exact asymptotic form of Φ meets with considerable difficulties. Of course, to determine the exact relevant Φ if the Reynolds number is not large, in other words, regardless of the value of μ (and ϵ), is a much more complicated problem. One should

contrast this with the case of a conservative system where the relevant measure happens to be known regardless of the number of particles. We shall make no attempt here to formulate the problem more completely or to speculate about the nature of the present difficulties.

The question whether the assumption we made in the beginning, *viz.*, the assumption of the uniqueness of the phase motion $T^t u$, u given, is justified we hope to take up at another occasion. It may be a crucial one in the case considered here, *viz.*, the case of a fluid filling the entire three-dimensional space, even if the viscosity is not zero.

BIBLIOGRAPHY

1. G. K. Batchelor, *The theory of homogeneous turbulence*, Cambridge University Press, London, 1953.
2. E. Hopf, *Statistical hydromechanics and functional calculus*, J. Rational Mech. and Anal. vol. 1 (1952) pp. 87–123.
3. ——— and E. W. Titt, *On certain special solutions of the Φ -equation of statistical hydrodynamics*, J. Rational Mech. and Anal. vol. 2 (1953) pp. 587–591.
4. I. Proudman and W. H. Reid, *On the decay of a normally distributed and homogeneous turbulent velocity field*, Philos. Trans. Roy. Soc. London Ser. A vol. 247 (1954) pp. 163–189.

GRADUATE INSTITUTE FOR MATHEMATICS AND MECHANICS, INDIANA UNIVERSITY,
BLOOMINGTON, IND.

STOCHASTIC PROCESSES OF ASTRONOMICAL INTEREST

BY

GUIDO MÜNCH

1.1. Introduction. In recent times astronomers have paid a great deal of attention to problems related to the distribution of diffuse matter in interstellar space.¹ The observations have revealed that both the density and the motion of the interstellar matter are affected by irregular fluctuations, superimposed on their systematic dependence on position determined by the large-scale features of the gravitational potential of the Galactic system. Pending the development of a physical theory for the turbulent state of motion present in the interstellar medium, which would account for such fluctuations, the observed phenomena have been studied in terms of simplified models involving a number of parameters of immediate physical meaning. Because of its simplicity, the model first advanced by V. A. Ambarzumian [1–5] has played a prominent role in the study of the subject. This model visualizes the interstellar matter as if it were concentrated in discrete units (which are called clouds), distributed in space at random, independently of each other, and with linear dimensions small compared with the average distance separating any two neighboring clouds. This distance results in orders of magnitude smaller than the dimensions of the whole system, in such a way that we postulate for the number of clouds $N(s)$ intersected by a line of sight to a point at the distance s a distribution according to a Poisson law of variance νs . Depending on the nature of the particular problem under study, we may attach to the clouds additional attributes, as optical or kinematic properties. The aim of the study is then to derive from suitable observations information regarding such attributes. While many are the problems which have been studied on the basis of Ambarzumian's model, it would appear evident that it cannot describe situations where the linear dimensions of the clouds enter explicitly.²

The astronomical context in which the stochastic processes to be considered later arise is the origin of the fluctuations in surface brightness observed along the Milky Way. This problem may be studied in Ambarzumian's model very simply, since only the absorbing properties of the clouds and their distribution along the line of sight enter, independently of their motions. The Milky Way phenomenon arises from the integration along the line of sight of light radiated by stars and the obscuring effects of the intervening dark clouds. The stars are distributed in space nearly uniformly at random, in the sense of Poisson's law, the average distance between stars being 10^{19} cm. The clouds,

¹ See [7], [19], and [20] for further references to the extensive astronomical literature on the subject.

² The astronomical basis of Ambarzumian's model has been critically analyzed in [6] and [15].

on the other hand, have a mean separation around $4 \cdot 10^{20}$ cm. Since the line of sight to a low galactic latitude point may extend to $3 \cdot 10^{21}$ cm., it is clear that, because the number of clouds likely to occur is small, considerable fluctuation will be observed in the amounts of radiation received in various directions. For our purposes we may consider then the stars to be spread out into a continuous medium emitting isotropically with constant emissivity $4\pi\eta$ (energy radiated per unit volume and unit time). Let the number $q < 1$ measure the fractional energy transmitted by a cloud along a pencil of radiation. If we suppose q to be the same for every cloud, the contribution $\Delta I \Delta\omega$ to the observed intensity I (energy received per unit time, unit area, and unit solid angle) arising from an element of volume subtending a solid angle $\Delta\omega$ at the observer and extending between distances s_{i-1} and s_i then satisfies the inequalities

$$(1) \quad \eta q^{N(s_i)}(s_i - s_{i-1}) < \Delta I < \eta q^{N(s_{i-1})}(s_i - s_{i-1}).$$

The summation over all intervals $\Delta s_i = s_i - s_{i-1}$ covering $(0, s)$, in the limit as $\Delta s_i \rightarrow 0$, leads us to consider the integral

$$(2) \quad I(s) = \int_0^s \eta q^{N(s)} ds,$$

to which a precise meaning as a stochastic integral in a definite sense could be attached.

In terms of the dimensionless variables

$$(3) \quad U = \frac{I\nu}{\eta} \quad \text{and} \quad t = \nu s,$$

we write (2) in the form

$$(4) \quad U_q(t) = \int_0^t q^{N(t')} dt',$$

$N(t)$ being a random variable taking integer values n with probability

$$(5) \quad p_n(t) = e^{-t} \frac{t^n}{n!}, \quad (n = 0, 1, 2, \dots).$$

It will be found necessary to consider also the random function

$$(6) \quad U_q(t, n) = \int_0^t q^{N(t', n)} dt',$$

where $n = N(t)$, an integer supposed known, represents the number of clouds occurring in $(0, t)$, but their position in $(0, t')$ is governed by the binomial law

$$(7) \quad p_m(t', n) = \Pr \{N(t', n) = m\} = \binom{n}{m} \left(\frac{t'}{t}\right)^m \left(1 - \frac{t'}{t}\right)^{n-m}$$

From the point of view of the applications it is desirable to generalize the definition (4) and consider the case when q is not a constant but a parameter

taking random values in (0,1) according to an a priori given probability law $\varphi(q)$. We write then

$$(8) \quad U_{\varphi}(t) = \int_0^t \left[\prod_{i=1}^{N(t')} q_i \right] dt',$$

where the q_i 's are independent random variables, all with the same probability distribution function $\varphi(q)$, and their number $N = N(t')$ is a Poisson variable of variance t' .

In this review the advances so far made in the study of the properties of the random functions (4), (6), and (8) will be summarized. No further references will be made to the related astronomical questions, but the functions will be considered in the belief that they may have some interest from the purely mathematical point of view.

Although (4) is only a particular case of (8), because it has not been possible to carry the study of $U_{\varphi}(t)$ as far as that of $U_q(t)$ we shall consider both of them separately.

2.1. The process $U_q(t)$. Let $F(u, t)$ denote the probability distribution function (pr.d.f.) of the random variable defined by (4). Then

$$(9) \quad \begin{aligned} F(u + 0, t) &= \Pr \{ U_q(t) < u \}, \\ F(u - 0, t) &= \Pr \{ U_q(t) \leq u \}. \end{aligned}$$

Since $q < 1$, it is clear that always $U_q(t) < t + 0$ and

$$(10) \quad F(t + 0, t) = 1.$$

On the other hand, we may have $U_q(t) = t$ in the event $N(t) = 0$, which has probability $p_0 = e^{-t}$, and

$$(11) \quad F(t - 0, t) = 1 - e^{-t}.$$

We see thus that $F(u, t)$ is discontinuous at $u = t$. At every point u in the interval $\Omega = (0, t)$, open to the right, $F(u, t)$ is continuous.

Let $t_1 \in \Omega$ and rewrite (5) in the form³

$$(12) \quad U(t) = U(t_1) + q^{N(t_1)} \int_{t_1}^t q^{N(t') - N(t_1)} dt'.$$

Since $N(t)$ is an integer-valued random variable with independent increments, with probability 1 we have

$$(13) \quad N(t') - N(t_1) = N(t' - t_1),$$

and we may introduce $t' - t_1$ as a new variable of integration in (12), obtaining

$$(14) \quad U(t) = U(t_1) + q^{N(t_1)} U(t - t_1).$$

³ For simplicity in the notation we shall omit the subscripts q and φ in $U_q(t)$ and $U_{\varphi}(t)$ whenever from the context it appears clear to which function reference is being made.

Let now the probability distribution for the joint occurrence of the events $U(t) < u$ and $U(t_1) < u_1$ be denoted by

$$(15) \quad F_2(u, t; u_1, t_1) = \Pr \{U(t) < u, U(t_1) < u_1\}.$$

We may obtain a relation between $F(u, t)$ and $F_2(u, t; u_1, t_1)$ through the auxiliary pr.d.f.

$$(16) \quad G_n(u, t) = \Pr \{U_q(t) < u | N(t) = n\}$$

by noticing that from the definition of conditional probability we have

$$(17) \quad F_2(u, t; u_1, t_1) = \sum_{n=0}^{\infty} p_n(t_1) \int_0^{u_1} \Pr \{U(t) < u | U(t_1) = u'_1, N(t_1) = n\} d_{u'_1} G_n(u'_1, t_1).$$

But (14) gives

$$(18) \quad \Pr \{U(t) < u | U(t_1) = u'_1, N(t_1) = n\} = \Pr \left\{ U(t - t_1) < \frac{u - u'_1}{q^n} \right\},$$

and thus (17) may be written⁴ [15]

$$(19) \quad F_2(u, t; u_1, t_1) = \sum_{n=0}^{\infty} p_n(t_1) \int_0^{u_1} F\left(\frac{u - u'_1}{q^n}, t - t_1\right) d_{u'_1} G_n(u'_1, t_1).$$

For $u_1 = t_1 + 0$ and arbitrary $t_1 < t$, (19) gives

$$(20) \quad F(u, t) = \sum_{n=0}^{\infty} p_n(t_1) \int_0^{t_1} F\left(\frac{u - u'_1}{q^n}, t - t_1\right) d_{u'_1} G_n(u'_1, t_1),$$

a relation between values of $F(u, t)$ for two different values of t . In the limit as $t_1 \rightarrow t$, (17) degenerates to the trivial relation

$$(21) \quad F(u, t) = \sum_{n=0}^{\infty} p_n(t) G_n(u, t).$$

3.1. The differential equation for $F(u, t)$. To evaluate $F(u, t)$ we may follow two different procedures. Either we find first the sequence $\{G_n(u, t)\}$ and then carry out the summation indicated in (21), or alternatively we derive from (20) a differential equation by considering its limit as $t_1 \rightarrow 0$. We postpone to a later paragraph the evaluation of $\{G_n(u, t)\}$ and now follow the latter method, for which we need to know only G_0 and G_1 , functions with values that may be

⁴ It is plausible to expect that a general expression for

$$F_n(u, t; u_1, t_1; u_n, t_n) = \Pr \{U(t) < u, U(t_1) < u_1, \dots, U(t_n) < u_n\}$$

may be found in terms of $F(u, t)$ and $G_n(u, t)$.

written down immediately from the definitions (6) and (7):

$$(22) \quad G_0(u, t) = \begin{cases} 1, & (\text{for } u \geq t), \\ 0, & (\text{for } u < t) \end{cases}$$

and

$$(23) \quad G_1(u, t) = \begin{cases} 1, & (\text{for } u > t), \\ \frac{u - qt}{(1 - q)t}, & (\text{for } qt \leq u \leq t), \\ 0, & (\text{for } u < qt). \end{cases}$$

Taking into account (22) and (23), together with

$$(24) \quad \begin{aligned} p_0(t_1) &= 1 - t_1 + o(t_1), & p_1(t_1) &= t_1 + o(t_1) \\ p_n(t_1) &= o(t_1), & & (\text{for } n \geq 2), \end{aligned}$$

from (20) we derive

$$(25) \quad F(u, t) = (1 - t_1)F(u - t_1, t - t_1) + \int_{qt}^t F\left(\frac{u - u'_1}{q}, t - t_1\right) \frac{du'_1}{1 - q}.$$

If the values of F appearing in the right-hand side of (23) are expressed as series in powers of u'_1 and t_1 and again quantities $o(t_1)$ are neglected in the result, we obtain

$$(26) \quad \left(1 + \frac{\partial}{\partial u} + \frac{\partial}{\partial t}\right) F(u, t) = F\left(\frac{u}{q}, t\right)$$

at those points where the derivatives have a meaning. Since, according to (10) and (11), $F(u, t)$ is discontinuous at $u = t$, it follows from (26) that $\partial F / \partial u$ is discontinuous at $u = qt$. However, if the u -derivative appearing in (26) is understood taken only in the left-handed neighborhood of $u = qt$, it also holds true at this point. Quite generally it is seen that the discontinuity of $F(u, t)$ at $u = t$ implies that the k th derivative of $F(u, t)$ is discontinuous at $u = q^k t$, but the function itself and its derivatives of lower order are continuous.

We shall show in the next section how equation (26), together with the boundary conditions

$$(27a) \quad F(u, t) = 1, \quad (\text{for } u > t),$$

$$(27b) \quad F(u, t) = 0, \quad (\text{for } u = 0),$$

and the requirement that $F(u, t)$ is to be continuous at every $u \in \Omega$, completely determines $F(u, t)$.

3.2. Solution of the differential equation for $F(u, t)$. Let Ω be covered by the sequence $\{\Omega_k\}$ of subintervals defined by the inequalities

$$(28) \quad q^k t \leq u < q^{k-1} t, \quad (k = 1, 2, \dots).$$

When $u \in \Omega_1$, $u/q > t$, and according to (27a), $F(u/q, t) = 1$. From (26), then,

$$(29) \quad \left(1 + \frac{\partial}{\partial u} + \frac{\partial}{\partial t}\right) F(u, t) = 1, \quad (u \in \Omega_1),$$

the solution of which may be expressed in terms of an arbitrary function $\phi(x)$ in the form

$$(30) \quad F(u, t) = 1 - e^{-t}\phi(t - u), \quad (u \in \Omega_1).$$

When $u \in \Omega_2$, $u/q \in \Omega_1$, and we may substitute in (26) for $F(u/q, t)$ the formal result (30). The resulting nonhomogeneous equation

$$(31) \quad \left(1 + \frac{\partial}{\partial u} + \frac{\partial}{\partial t}\right) F(u, t) = 1 - e^{-t}\phi\left(t - \frac{u}{q}\right), \quad (u \in \Omega_2)$$

may be formally integrated in terms of the arbitrary function ϕ , when we require the continuity of $F(u, t)$ at $u = qt$, in the form

$$(32) \quad F(u, t) = 1 - e^{-t}\phi(t - u) + e^{-t} \frac{q}{1 - q} \int_0^{t-u/q} \phi(x) dx, \quad (u \in \Omega_2).$$

Quite generally, the equation for $F(u, t)$ for $u \in \Omega_k$ may be formally integrated in terms of its values for $u \in \Omega_{k-1}$, and by induction it is easy to prove that

$$(33) \quad F(u, t) = 1 - e^{-t}\phi(t - u) - e^{-t} \sum_{j=1}^{k-1} \frac{Q_j}{(j-1)!} \int_0^{t-u/q^j} \left(t - \frac{u}{q^j} - x\right)^{j-1} \phi(x) dx, \quad (u \in \Omega_k),$$

where Q_j is a constant defined by

$$(34) \quad Q_j = - \frac{q^j}{1 - q^j} Q_{j-1} = (-1)^j \prod_{l=1}^j \frac{q^l}{1 - q^l}, \quad (j = 1, 2, \dots),$$

with $Q_0 = 1$.

The point $u = 0$ is defined by the limit of the sequence $\{\Omega_k\}$ as $k \rightarrow \infty$. By setting $u = 0$ in (33) and taking into account the boundary condition (27b), we obtain

$$(35) \quad \phi(t) + \sum_{j=1}^{\infty} \frac{Q_j}{(j-1)!} \int_0^t (t-x)^{j-1} \phi(x) dx = e^t,$$

which may be considered as an integral equation for the determination of ϕ . It may be readily solved by taking a Laplace transform, for we obtain

$$(36) \quad \mathfrak{L}_z[\phi(t)] = \frac{1}{z-1} \left[\sum_{j=0}^{\infty} Q_j z^{-j} \right]^{-1}.$$

Applying the identity [12]

$$(37) \quad \sum_{j=0}^{\infty} Q_j x^j = \prod_{j=1}^{\infty} (1 - q^j x),$$

we derive

$$(38) \quad \mathcal{L}_z[\phi(t)] = \frac{1}{z-1} \prod_{j=1}^{\infty} \left(1 - \frac{q^j}{z}\right)^{-1}.$$

The inverse Laplace transform of this function is immediate, as (38) has simple poles at $z = q^j$ ($j = 0, 1, 2, \dots$), and $z = 0$ is a limit point of poles. By integrating around contours containing a finite number of poles and taking the limit as $z = 0$ is approached, one obtains

$$(39) \quad \phi(t) = K \sum_{j=0}^{\infty} Q_j e^{q^j t},$$

where

$$(40) \quad K = \prod_{k=1}^{\infty} (1 - q^k)^{-1}.$$

Substituting (39) in (33) we finally obtain⁵

$$(41) \quad F(u, t) = 1 - K \sum_{j=0}^{\infty} Q_j e^{-(1-q^j)t} \sum_{l=0}^{k-1} Q_l q^{-lj} e^{-u/q^{l-j}}, \quad (u \in \Omega_k).$$

It may be noticed that the singularity of $\mathcal{L}_z[\phi(t)]$ at $z = 0$ implies that the derivatives of all orders with respect to u of $F(u, t)$ vanish identically at $u = 0$. Also, if we let $u \rightarrow t$ in (41), we obtain a result identical to that expressed in (11).

The form of (38) is quite suitable for the numerical evaluation of $F(u, t)$, and even for values of q as high as 0.8 only a few terms of the series suffice to determine $F(u, t)$, say, with four-figure accuracy.⁶

3.3. The limiting distribution $F(u, \infty)$. As $t \rightarrow \infty$, from (41) one would expect that

$$(42) \quad F(u, \infty) \equiv F(u) = 1 - K \sum_{l=0}^{\infty} Q_l e^{-u/q^l},$$

a result which may be verified by solving directly the differential equation⁷

$$(43) \quad F(u) + F'(u) = F\left(\frac{u}{q}\right),$$

which results from letting $t \rightarrow \infty$ in (20) and then taking a limit as $t_1 \rightarrow 0$.

⁵ This solution of equation (26) was first found in [10], where the form of $F(u, t)$ for $u \in \Omega_1$ was found directly from the definition of $U_q(t)$ and then iterated in successive Ω_k . The solution given above, published here for the first time, shows that the differential equation (26) completely determines $F(u, t)$.

⁶ In [10] the values of $F(u, t)$, for $t = 2$ and $t = 4$ and $q = 0.75$ and $q = 0.80$, are given in graphical form.

⁷ Equation (43) was first derived in [3]. Its solution was found in [8], where a table of numerical values of $F(u)$, for a number of values in the range $0.50 \leq q \leq 0.85$, is also given. Independently the same solution was found in [18].

4.1. The differential equation for $G_n(u, t)$. Let $G_n(t)$, the p.d.f. of the random function defined by (6), be written in the form

$$(44) \quad G_n(u, t) = \Pr \left\{ U_q(t_1, n) + q^{N(t_1, n)} \int_{t_1}^t q^{N(t', n) - N(t_1, n)} dt' < u \right\},$$

where $t_1 \in \Omega$. Since with probability 1

$$(45) \quad N(t', n) - N(t_1, n) = N(t' - t_1, n),$$

we may introduce $t' - t_1$ as a new variable of integration in the integral extending over (t_1, t) and rewrite

$$(46) \quad G_n(u, t) = \Pr \left\{ U_q(t_1, n) + q^{N(t_1, n)} \int_0^{t-t_1} q^{N(t', n)} dt' < u \right\}.$$

Now for sufficiently small t_1 , $N(t_1, n)$ may take only the values zero or 1 with probabilities $1 - nt_1/t$ or nt_1/t , respectively. Neglecting quantities $o(t_1)$, we thus have

$$(47) \quad G_n(u, t) = \left(1 - \frac{nt_1}{t}\right) \Pr \left\{ \int_0^{t-t_1} q^{N(t', n)} dt' < u - t_1 \right\} \\ + \frac{nt_1}{t} \Pr \left\{ \int_0^{t-t_1} q^{N(t', n-1)} dt' < \frac{u - \theta t_1}{q} \right\},$$

where $q \leq \theta \leq 1$. Notice that the exponent of q under the second integral sign in (47) has been written as $N(t', n-1)$ to take into account the certain occurrence of only $n-1$ clouds in (t_1, t) when it is known that $N(t_1, n) = 1$.

From (47) it follows that

$$(48) \quad G_n(u, t) = \left(1 - \frac{nt_1}{t}\right) G_n(u - t_1, t - t_1) + \frac{nt_1}{t} G_{n-1}\left(\frac{u - \theta t_1}{q}, t - t_1\right).$$

Expressing the values of G_n appearing in the right-hand side of (48) as series in powers of t_1 and neglecting terms $o(t_1)$, we find

$$(49) \quad G_n(u, t) + \frac{t}{n} \left(\frac{\partial}{\partial u} + \frac{\partial}{\partial t} \right) G_n(u, t) = G_{n-1} \left(\frac{u}{q}, t \right)$$

at those points where the derivatives have a meaning.

From the definition (6) we have

$$(50) \quad G_n(u, t) = \begin{cases} 1, & \text{(for } u > t), \\ 0, & \text{(for } u \leq q^n t), \end{cases}$$

and thus G_0 is a step function with a unit jump at $u = t$. All other functions of the sequence $\{G_n\}$ are continuous for $u \in \Omega$.

4.2. Solution of the differential recurrence relation satisfied by $G_n(u, t)$. We shall construct a set of functions $\{G_n\}$ which satisfy (49) and the normaliz-

ing conditions (50). It is convenient to define the auxiliary function

$$(51) \quad \Psi_n(u, t) = t^n [1 - G_n(u, t)],$$

which satisfies the simpler equation

$$(52) \quad \left(\frac{\partial}{\partial u} + \frac{\partial}{\partial t} \right) \Psi_n(u, t) = n \Psi_{n-1} \left(\frac{u}{q}, t \right)$$

and the normalizing conditions

$$(53a) \quad \Psi_n(u, t) = 0, \quad (\text{for } u > t),$$

$$(53b) \quad \Psi_n(u, t) = t^n, \quad (\text{for } u \leq q^n t).$$

Since Ψ_0 is discontinuous at $u = t$, from (52) it follows that if $\Psi_1(u, t)$ is to be continuous at every $u \in \Omega$, then necessarily $\partial \Psi_1 / \partial u$ is discontinuous at $u = qt$. However, a meaning may be attached to (52) at this point if $\partial \Psi_1 / \partial u$ is understood to be taken to the left.

To find the solution of (52) we follow a method similar to the one employed in Sec. 2.2. Since $u/q > t$ when $u \in \Omega_1$, according to (53a) we have

$$(54) \quad \left(\frac{\partial}{\partial u} + \frac{\partial}{\partial t} \right) \Psi_n(u, t) = 0, \quad (u \in \Omega_1).$$

In terms of an arbitrary function $\psi_n(x)$, we thus have

$$(55) \quad \Psi_n(u, t) = \psi_n(t - u).$$

When $u \in \Omega_2$, $u/q \in \Omega_1$, and we may use the values of $\Psi_{n-1}(u/q, t)$, as given by (55), in the right-hand side of (52). The resulting equation

$$(56) \quad \left(\frac{\partial}{\partial u} + \frac{\partial}{\partial t} \right) \Psi_n(u, t) = n \psi_{n-1} \left(t - \frac{u}{q} \right), \quad (u \in \Omega_2)$$

may be formally integrated in the form

$$(57) \quad \Psi_n(u, t) = \psi_n(t - u) + n Q_1 \int_0^{t-u/q} \psi_{n-1}(x) dx, \quad (u \in \Omega_2),$$

where the continuity of $\Psi_n(u, t)$ at $u = qt$ is taken into account. Quite generally, the equation for $\Psi_n(u, t)$ in $u \in \Omega_k$ may be formally integrated in terms of its values in $u \in \Omega_{k-1}$ and by induction it follows that

$$(58) \quad \Psi_n(u, t) = \psi_n(t - u) + \sum_{j=1}^{k-1} j \binom{n}{j} Q_j \int_0^{t-u/q^j} \left(t - \frac{u}{q^j} - x \right)^{j-1} \psi_{n-j}(x) dx, \\ (u \in \Omega_k, k = 2, 3, \dots, n).$$

According to the normalization condition (53b), for the particular value $u = q^n t$, this equation gives

$$(59) \quad t^n = \psi_n[t(1 - q^n)] \\ + \sum_{j=1}^{n-1} j \binom{n}{j} Q_j \int_0^{t(1-q^{n-j})} [t(1 - q^{n-j}) - x]^{j-1} \psi_{n-j}(x) dx, \\ (n = 2, 3, \dots)$$

a system of integral equations of a simple recursive type which is readily solved. From (55) and (53b) it follows that

$$(60) \quad \psi_1(x) = \frac{x}{1 - q},$$

and, hence, setting $n = 2$ in (59), we derive

$$(61) \quad \psi_2(x) = \frac{x^2}{(1 - q)(1 - q^2)}.$$

Quite generally, by induction it is proved that for any $n \geq 1$

$$(62) \quad \psi_n(x) = x^n \prod_{l=1}^n (1 - q^l)^{-1}.$$

Substituting this result in (58), it is found that

$$(63) \quad \Psi_n(u, t) = \sum_{j=0}^{k-1} \left(t - \frac{u}{q^j}\right)^n Q_j \prod_{l=1}^{n-j} (1 - q^l)^{-1}.$$

Since, according to (53), (55), and (60), this equation also holds true for $n = 0$ and 1, we may finally write from (51) and (63)

$$(64) \quad G_n(u, t) = 1 - \sum_{j=0}^{k-1} \left(1 - \frac{u}{q^j t}\right)^n Q_j \prod_{l=1}^{n-j} (1 - q^l)^{-1}, \quad (u \in \Omega_k, 1 \leq k \leq n).$$

We shall verify that (64), when substituted in (21), leads to an expression for $F(u, t)$ identical to (41). By applying the identity (37) we can write

$$(65) \quad \prod_{l=1}^{n-j} (1 - q^l)^{-1} = K \prod_{l=1}^{\infty} (1 - q^{l+n-j}) = K \sum_{r=0}^{\infty} Q_r q^{r(n-j)},$$

and substituting (65) in (64) and the result in (21), we find

$$(66) \quad F(u, t) = 1 - K e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{j=0}^{k-1} Q_j \left(1 - \frac{u}{q^j t}\right)^n \sum_{r=0}^{\infty} Q_r q^{r(n-j)} \\ = 1 - K e^{-t} \sum_{r=0}^{\infty} Q_r \sum_{j=0}^{k-1} Q_j q^{-rj} \sum_{n=0}^{\infty} \frac{1}{n!} \left[q^r \left(t - \frac{u}{q^j}\right) \right]^n,$$

an expression equivalent to (41).

It is noticed that the moment sequences of $G_n(u, t)$ have the values

$$(67) \quad \mu_i(t, n) = \int_0^\infty u^i d_u G_n(u, t) = \frac{i! n!}{(i+n)!} t^i \prod_{j=1}^n \frac{1 - q^{i+j}}{1 - q^j}, \quad (i = 0, 1, 2, \dots).$$

5.1. The moments of $U_q(t)$. We shall exhibit now the explicit relationship between the bivariate moments of $F_2(u, t; u_1, t_1)$

$$(68) \quad \mu_{rs}(t, t_1) = \int_0^\infty u_1^r \int_0^\infty u^s d_{uu_1} [F_2(u, t; u_1, t_1)], \quad (r, s = 0, 1, 2, \dots),$$

and the moments of $F(u, t)$

$$(69) \quad \mu_i(t) = \int_0^\infty u^i d_u F(u, t).$$

From (19) we obtain

$$(70) \quad d_{uu_1} F_2(u, t; u_1, t_1) = \sum_{n=0}^\infty p_n(t_1) d_u \left[F\left(\frac{u - u_1}{q^n}, t - t_1\right) \right] d_{u_1} G_n(u_1, t_1),$$

and thus (68) may be rewritten

$$(71) \quad \mu_{rs}(t, t_1) = \sum_{n=0}^\infty p_n(t_1) \int_0^\infty u_1^r d_{u_1} [G_n(u_1, t_1)] \int_0^\infty u^s d_u F\left(\frac{u - u_1}{q^n}, t - t_1\right).$$

If $q^{-n}(u - u_1) = x$ is introduced as a new variable of integration and $u^s = (u_1 + q^n x)^s$ is expressed as a polynomial, we obtain

$$(72) \quad \mu_{rs}(t, t_1) = \sum_{k=0}^s \binom{s}{k} \mu_{s-k}(t - t_1) \sum_{n=0}^\infty p_n(t_1) q^{n(s-k)} \int_0^\infty u_1^{r+k} d_{u_1} G_n(u_1, t_1).$$

The summation over the index n may be carried out explicitly by noticing that $G_n(u, t)$ is a homogeneous function of degree, zero in u and t :

$$(73) \quad \sum_{n=0}^\infty p_n(t_1) q^{n(s-k)} \int_0^\infty u_1^{r+k} d_{u_1} G_n(u_1, t_1) = e^{-(1-q^{s-k})t_1} q^{-(r+k)(s-k)} \mu_{r+k}(t_1 q^{s-k}),$$

and, accordingly, (72) becomes

$$(74) \quad \mu_{rs}(t, t_1) = \sum_{k=0}^s \binom{s}{k} \frac{e^{-(1-q^{s-k})t_1}}{q^{(r+k)(s-k)}} \mu_{r+k}(t_1 q^{s-k}) \mu_{s-k}(t - t_1).$$

The values of the moments $\mu_i(t)$ may be found from the explicit form of $F(u, t)$:

$$(75) \quad \mu_i(t) = i! \sum_{l=0}^i Q_l q^{-li} e^{-(1-q^l)t} \prod_{j=1}^{i-l} (1 - q^j)^{-1}, \quad (i = 1, 2, \dots).$$

6.1. Equation for the probability law of $U_\varphi(t)$. Denote by $W_\varphi(u, t) \equiv W(u, t)$ the pr.d.f. of the random function defined by (8),

$$(76) \quad W(u, t) = \Pr \left\{ \int_0^t \left(\prod_{i=1}^{N(t')} q_i \right) dt' < u \right\},$$

and let, as before, $t_1 \in (0, t)$. We may then write

$$(77) \quad W(u, t) = \Pr \left\{ U_\varphi(t_1) + \left(\prod_{i=1}^{N(t_1)} q_i \right) \int_{t_1}^t \left(\prod_{i=N(t_1)}^{N(t')} q_i \right) dt' < u \right\}.$$

Since the q_i 's are independent random variables and their number in (t_1, t) is given by a Poisson law, we have

$$(78) \quad \prod_{i=N(t_1)}^{N(t')} q_i = \prod_{i=1}^{N(t' - t_1)} q_i,$$

and we may introduce $t' - t_1$ as the variable of integration in the integral appearing in (77), to obtain

$$(79) \quad W(u, t) = \Pr \left\{ U_\varphi(t_1) + \left(\prod_{i=1}^{N(t_1)} q_i \right) U_\varphi(t - t_1) < u \right\}.$$

Now, if t_1 is sufficiently small, we may have $N(t_1) = 0$ or $N(t_1) = 1$. In the former event, with probability $1 - t_1 + o(t_1)$,

$$(80) \quad U_\varphi(t_1) = t_1, \quad \text{and} \quad \prod_{i=1}^{N(t_1)} q_i = 1,$$

while in the latter, with probability $t_1 + o(t_1)$, we have

$$(81) \quad \Pr \left\{ U_\varphi(t_1) + \left(\prod_{i=1}^{N(t_1)} q_i \right) U_\varphi(t - t_1) < u \mid N(t_1) = 1 \right\} \\ = \int_0^1 \Pr \{ \theta t_1 + q U_\varphi(t - t_1) < u \} d\varphi(q),$$

where $0 \leq \theta \leq 1$. By the definition of conditional probability, from (79) it follows that

$$(82) \quad W(u, t) = (1 - t_1) W(u - t_1, t - t_1) + t_1 \int_0^1 W\left(\frac{u - \theta t_1}{q}, t - t_1\right) d\varphi(q).$$

Expressing the values of W appearing in the right-hand side of (82) as power series in t_1 , and neglecting quantities $o(t_1)$, it follows that⁸

$$(83) \quad \left(1 + \frac{\partial}{\partial u} + \frac{\partial}{\partial t} \right) W(u, t) = \int_0^1 W\left(\frac{u}{q}, t\right) d\varphi(q),$$

⁸ Equation (83) was first derived in [7]. The particular case $t = \infty$, considered below in Sec. 6.3, had earlier been derived in [3].

at those points where $\partial W/\partial u$ is defined, and provided the behavior of $\varphi(q)$ near $q = 0$ is restricted sufficiently to make the integration term by term of the power series a valid procedure. At $q = 1$ the nature of the physical problem would require $\varphi(q)$ to be continuous.

From the definition (8) it follows that for all t

$$(84) \quad W(t+0, t) = 1, \quad W(t-0, t) = 1 - e^{-t}, \quad \text{and} \quad W(0, t) = 0.$$

Hence, in the limit as $t \rightarrow 0$, $W(u, t)$ is a step function with a unit jump at $u = 0$.

A proof for the existence of a solution of (83) has not yet been found for the case of a finite t and arbitrarily given $\varphi(q)$.⁹ However, as we shall presently show, the moment sequence of $W(u, t)$ and the limiting distribution as $t \rightarrow \infty$ are explicitly known.

6.2. The moment sequence of $W(u, t)$. Let

$$(85) \quad M_m(t) = \int_0^\infty u^m d_u F(u, t), \quad (m = 0, 1, 2, \dots)$$

denote the m th moment of $W(u, t)$. We may derive a differential equation for $M_m(t)$ from (83) when $W(u, t)$ has a second derivative with respect to u . For, multiplying by u^m the derivative with respect to u of (83), and integrating the result over all u , we obtain

$$(86) \quad \frac{dM_m}{dt} + (1 - \lambda_m)M_m = mM_{m-1}(t), \quad (m = 1, 2, \dots),$$

where

$$(87) \quad \lambda_m = \int_0^1 q^m d\varphi(q)$$

are the moments of $\varphi(q)$. From the known behavior of $W(u, t)$ at $t = 0$ it follows that

$$(88) \quad M_m(t)|_{t=0} = 0, \quad (m = 1, 2, \dots),$$

and this, together with (86), implies that

$$(89) \quad \left. \frac{d^j M_m}{dt^j} \right|_{t=0} = 0, \quad (j = 1, 2, \dots, m-1).$$

The solution of the system of linear equations (86) satisfying the conditions (88) and (89) may be explicitly found. By successive applications of (86) we derive

$$(90) \quad \left\{ \prod_{j=1}^m \left(\frac{d}{dt} + 1 - \lambda_j \right) \right\} M_m(t) = m!, \quad (m = 1, 2, \dots),$$

⁹ An operational procedure for solving (83) has been offered in [17], which essentially consists in a "formal identification" of the moment sequence of $W(u, t)$ with its Mellin transform. Such identification to be meaningful at all requires a suitable "interpretation" of the moment sequence. While this technique is not a rigorous procedure, for the singular case considered in Sec. 3.2, it reproduces the value of $F(u, t)$ given by (41).

the general solution of which is

$$(91) \quad M_m(t) = \sum_{k=1}^m A_k e^{-(1-\lambda_k)t} + m! \prod_{j=1}^m (1 - \lambda_j)^{-1}.$$

The m constants of integration are determined by the conditions (88) and (89), which give

$$(92) \quad \sum_{k=1}^m A_k = -m! \prod_{j=1}^m (1 - \lambda_j)^{-1}, \quad \sum_{k=1}^m (1 - \lambda_k)^j A_k = 0, \\ (j = 1, \dots, m-1).$$

The matrix of this linear system of equations is the Vandermonde matrix, the inverse of which may be written down explicitly [7]. The solution of (92) is thus found to be

$$(93) \quad A_k = (-1)^m m! (1 - \lambda_k)^{-1} \prod_{j \neq k}^{(1,m)} (\lambda_j - \lambda_k)^{-1} \\ = m! \prod_{j \neq k}^{(0,m)} (\lambda_k - \lambda_j)^{-1},$$

and, accordingly, quite generally from (91) we have

$$(94) \quad M_m(t) = m! \sum_{k=0}^m e^{-(1-\lambda_k)t} \prod_{j \neq k}^{(0,m)} (\lambda_k - \lambda_j)^{-1}, \\ (m = 1, 2, \dots).$$

An alternative procedure to evaluate $\{M_m(t)\}$ is to consider directly from the definition (8) the average

$$(95) \quad M_m(t) = \left\{ \left[\int_0^t \left(\prod_{i=1}^{N(t')} q_i \right) dt' \right]^m \right\}_{\text{av}},$$

which is carried out over all possible values of the q_i 's and their possible positions in $(0, t)$. From (95) we thus obtain

$$(96) \quad M_m(t) = \int_0^t \left\{ \prod_{i=1}^{N(t')} q_i^m \right\}_{\text{av}} M_{m-1}(t - t') dt' \\ = \int_0^t e^{-(1-\lambda_m)t'} M_{m-1}(t - t') dt',$$

which clearly is an integral form of (86).

In the particular case when $\varphi(q)$ is a step function with a unit jump at q , $\lambda_i = q^i$, and (94) reduces to a form equivalent to (75).

6.3. The limiting distribution $W(u, \infty)$. In the limit as $t \rightarrow \infty$, (83) coalesces into

$$(97) \quad W(u) + W'(u) = \int_0^1 W\left(\frac{u}{q}\right) d\varphi(q).$$

The distribution function $W(u)$ satisfying this equation for arbitrarily given $\varphi(q)$ has been constructed by following an iteration procedure. Formally we rewrite (97) in the form

$$(98) \quad W(u) = e^{-u} \int_0^1 d\varphi(q) \int_0^u e^{u'} W\left(\frac{u'}{q}\right) du'.$$

Starting with a "guess" distribution function $W_1(u)$, we construct a sequence of functions $\{W_n(u)\}$ by successive iterations in (98). It has been found [16] that such a sequence converges to the limit

$$(99) \quad W(u) = 1 - e^{-u} \sum_{k=0}^{\infty} L_k(u) \sum_{j=0}^k (-1)^j \binom{k}{j} \prod_{i=1}^{j+1} (1 - \lambda_i)^{-1},$$

where $L_k(u)$ stands for the Laguerre polynomial and λ_i is the i th moment of $\varphi(q)$ as defined by (87). The reason why the limiting function (99) exists is that each iteration performs an integral operation over $d\varphi(q)$ on the initial "guess" function $W_1(u)$ for an argument divided by q . Since $q < 1$, it is clear that in the limit as $n \rightarrow \infty$, only the value of $W_1(u)$ at $u = \infty$ matters, and this is set a priori equal to unity.

Since from (94)

$$(100) \quad M_m = \lim_{t \rightarrow \infty} M_m(t) = m! \prod_{i=1}^m (1 - \lambda_i)^{-1},$$

we may rewrite $W(u)$, from (99), in terms of its own moments, in the form

$$(101) \quad W(u) = 1 - e^{-u} \sum_{k=0}^{\infty} L_k(u) \sum_{j=0}^k \frac{(-1)^j}{(j+1)!} \binom{k}{j} M_{j+1}$$

Formally, it is easy to verify that (101) is a solution of the Stieltjes moment problem associated with the sequence $\{M_m\}$ defined in (100). It also may be shown that (101) leads to an expression identical to (42) for the case $\lambda_i = q^i$.

BIBLIOGRAPHY

1. V. A. Ambarzumian and S. G. Gordeladse, Bull. Abastumani Observ. No. 2 (1938) p. 37.
2. ———, Bull. Abastumani Observ. No. 4 (1940) p. 17.
3. ———, C. R. (Doklady) Acad. Sci. URSS. vol. 14 (1944) p. 223.
4. ———, Proc. Acad. Sci. Armenian SSR. vol. 1 (1944) pp. 9–13.

5. ———, Trans. Intern. Astron. Union vol. 7 (1950) pp. 452–455.
6. B. J. Bok, Harvard College Observ. Bull. No. 920 (1951).
7. S. Chandrasekhar and G. Münch, Astrophys. J. vol. 112, (1950) pp. 380–392.
8. ——— and ———, Astrophys. J. vol. 112 (1950) pp. 393–398.
9. ——— and ———, Astrophys. J. vol. 113 (1951) pp. 150–165.
10. ——— and ———, Astrophys. J. vol. 114 (1951) pp. 110–122.
11. ——— and ———, Astrophys. J. vol. 115 (1952) pp. 103–123.
12. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Clarendon Press, Oxford, 1945, p. 275.
13. B. E. Markarian, Contr. Burakan Observ. Acad. Sci. Armenian SSR. No. 1 (1946).
14. G. Münch and S. Chandrasekhar, Astrophys. J. vol. 115 (1952) pp. 94–102.
15. ———, Astrophys. J. vol. 116 (1952) pp. 575–586.
16. ———, Astrophys. J. vol. 121 (1955) pp. 291–299.
17. A. Ramakrishnan and P. M. Mathews, Astrophys. J. vol. 119 (1954) pp. 81–90.
18. G. E. Rusakov, Utchenia Zapiske Ser. Math. vol. 13 (1949) p. 53.
19. L. Spitzer, Astrophys. J. vol. 108 (1948) pp. 276–309.
20. B. Stromgren, Astrophys. J. vol. 108 (1948) pp. 242–275.

MOUNT WILSON AND PALOMAR OBSERVATORIES, CARNEGIE INSTITUTION OF WASHINGTON,
CALIFORNIA INSTITUTE OF TECHNOLOGY.
PASADENA, CALIF.

THE SINGULARITY IN THE SPECTRUM OF HOMOGENEOUS TURBULENCE¹

BY

G. K. BATCHELOR

My contribution to this Symposium concerns one aspect of the theory of homogeneous turbulence, *i.e.*, turbulence whose statistical properties are independent of position in an infinite fluid, and I regret to say that probability theory enters only as a source of terminology and of basic concepts such as the energy spectrum. However, it may be that mathematicians engaged in probability theory will be interested to hear about some of the special properties of the energy spectrum in the case of homogeneous turbulence and to see the statistical consequences of the fact that the stochastic variable under discussion (the vector velocity $\mathbf{u}$) satisfies the Navier-Stokes equation of motion. I have noticed, in the literature of probability theory, that mathematicians often wish to consider some particular stochastic process, as an illustrative model; such particular models should, of course, be chosen to be consistent with the requirements of the physical system concerned, and in this paper I shall describe one of the requirements in the case of turbulence. This requirement, which has been overlooked in the past, is that the energy spectrum have a singularity of a certain kind at the origin in wave-number space.

Let me attempt first to give a concise account of the background to my remarks about the singularity in the spectrum. I am talking about the statistical properties of a random variable $\mathbf{u}$, which is a function of position $\mathbf{x}$ and time t , stationary in $\mathbf{x}$ but not in t . We define a tensor covariance with separation $\mathbf{r}$

$$R_{ij}(\mathbf{r}) = \overline{u_i(\mathbf{x})u_j(\mathbf{x}')}, \quad \mathbf{x}' = \mathbf{x} + \mathbf{r}$$

and a tensor spectrum function

$$\Phi_{ij}(\boldsymbol{\kappa}) = \frac{1}{8\pi^3} \int R_{ij}(\mathbf{r}) e^{-i\boldsymbol{\kappa} \cdot \mathbf{r}} d\mathbf{r}$$

in the usual way. $\Phi_{ij}(\boldsymbol{\kappa})$ is the density, in three-dimensional wave-number space, of contributions to (twice) the energy tensor $\frac{1}{2}\overline{u_i u_j}$ when the velocity field $\mathbf{u}(\mathbf{x})$ is resolved into Fourier components. A simpler function, which specifies the distribution of total energy with respect to the magnitude of the wave-number vector, is

$$E(\boldsymbol{\kappa}) = \frac{1}{2} \int \Phi_{ii}(\boldsymbol{\kappa}) \kappa^2 d\Omega(\boldsymbol{\kappa}),$$

¹ This is essentially an abstract of a joint paper by Dr. Ian Proudman and the writer, entitled *The large-scale structure of homogeneous turbulence*, which has since been published (Philos. Trans. Roy. Soc. London Ser. A vol. 248 [1956] p. 369).

where $d\Omega(\kappa)$ denotes an element of solid angle in the direction of κ and the integration is over all such directions. The purpose of a good deal of the research on turbulence is to determine the properties of the spectrum tensor $\Phi_{ij}(\kappa)$, in view of its physical importance.

The velocity field in any one realization of the turbulent flow is governed by the equations

$$\frac{\partial u_i}{\partial x_i} = 0 \quad (\text{conservation of mass}),$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = - \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} \quad (\text{equation of momentum}),$$

where p and ν are the kinematic pressure and viscosity. In the case of homogeneous turbulence, the statistical properties of $\mathbf{u}$ as a function of $\mathbf{x}$ are regarded as being known at some initial instant t_0 , and the problem is to determine the subsequent development of the turbulence. Thus, we have a statistical initial-value problem, governed by the above differential equations for $\mathbf{u}$. It is not possible to write down differential equations in $\Phi_{ij}(\kappa)$ alone, because all the statistical parameters specifying $\mathbf{u}$ as a function of $\mathbf{x}$ at any one value of t are needed to determine the time derivatives of $\Phi_{ij}(\kappa)$ at that instant.

This initial-value problem has not been solved, but nevertheless a good deal of information about the spectrum tensor has been obtained by a variety of methods. These methods nearly all involve some assumption that gets around the mathematical difficulties of a complete solution. For instance, it has been assumed in several investigations—and this brings me close to the point of my paper—that the spectrum tensor $\Phi_{ij}(\kappa)$ is analytic in the neighborhood of $\kappa = 0$ or, more precisely, that a number of derivatives of $\Phi_{ij}(\kappa)$ with respect to κ exist at $\kappa = 0$. This assumption implies that a corresponding number of integral moments of $R_{ij}(\mathbf{r})$ converge, and was thought to be justified by the rough argument that any practical method of generating homogeneous turbulence—such as placing a regular array of bars in a uniform stream of air—involves a finite length scale, so that the statistical connection between the velocities at points at a distance apart large compared with this length scale ought to be zero for all practical purposes. I dare say that no mathematician would be happy with the assumption of an analytic form of the spectrum, but the hard-pressed theoretical physicist, faced with an almost intractable problem, is perhaps more ready to assent to plausible arguments in favor of such an assumption. At all events, it has usually been assumed that integral moments of $R_{ij}(\mathbf{r})$, and of several other mean values of velocity products, all converge, and a number of consequences have been deduced. I shall not take time to describe these consequences and their relation to experimental results, since this is a matter for specialists in the theory of turbulence. My concern here is with the assumption itself.

In some recent work, Dr. Ian Proudman and I have shown that in fact $R_{ij}(\mathbf{r})$ cannot decrease so rapidly as $r \rightarrow \infty$ that all integral moments converge,

and we have been able to determine explicitly the asymptotic form of $R_{ij}(\mathbf{r})$ that seems likely to occur with practical methods of generating homogeneous turbulence. We were led to undertake an investigation of the matter by the chance discovery, by Proudman and Reid,² that in one particular hypothetical case, *viz.*, that in which fourth-order cumulants of $\mathbf{u}(\mathbf{x})$ are instantaneously zero, some of the integral moments of the mean triple velocity product cannot continue to converge, whatever may have been the initial state of affairs. This showed that in one very particular, but kinematically possible, case the general assumption of convergent integral moments was wrong, and we felt that a reappraisal of the validity of the assumption was needed.

Now that we are wise after the event, it is easy to see that trouble arises from the presence of the pressure term in the Navier-Stokes equation. We are accustomed to think of this equation of motion as a purely differential equation, but when the pressure p is replaced by an expression in terms of the velocity $\mathbf{u}$, *viz.*,

$$p(\mathbf{x}) = \frac{1}{4\pi} \int \frac{\partial^2 u'_i u'_j}{\partial x'_i \partial x'_j} \frac{d\mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|},$$

an integral character is revealed. The gradient of p is a force which acts locally but which has a value determined by the whole instantaneous velocity field. The effect of this integral character of the pressure term in the equation of motion is to introduce statistical connections between distant points in the fluid.

The action of pressure forces can be seen more clearly in a simpler but analogous problem. Suppose that we have an infinite fluid of which only a finite central region is in motion at some initial instant. At subsequent instants the motion spreads to the surrounding fluid by the action of inertial, viscous, and pressure forces. The first two are proportional to the local velocity or one of its spatial gradients, and so they can never make the velocity fall off more slowly than exponentially with distance from the central region. However, the pressure force at any point is proportional to an integral of the velocity over the whole field, and exponential smallness of the velocity locally will not ensure exponential smallness of the rate of change of velocity due to pressure forces. Indeed, at the initial instant the pressure can be written as an expansion in powers of $1/x$, for large values of x :

$$p(\mathbf{x}) = \frac{1}{x} \int A(\mathbf{x}') d\mathbf{x}' - \frac{\partial 1/x}{\partial x_i} \int x'_i A(\mathbf{x}') d\mathbf{x}' + \frac{1}{2!} \frac{\partial^2 1/x}{\partial x_i \partial x_j} \int x'_i x'_j A(\mathbf{x}') d\mathbf{x}' + \dots,$$

where $A(\mathbf{x}')$ stands for

$$\frac{1}{4\pi} \frac{\partial^2 u_i(\mathbf{x}') u_j(\mathbf{x}')}{\partial x'_i \partial x'_j},$$

² Philo. Trans. Roy. Soc. London Ser. A vol. 247 (1954) p. 163.

and the order of magnitude of $p(\mathbf{x})$ as $x \rightarrow \infty$ will depend on the first integral moment of $A(\mathbf{x}')$ that is nonzero. The corresponding contribution to $\partial u_i / \partial t$ is then one order smaller in x . Hence a flow in which the velocity is exponentially small at large values of x is in general possible only instantaneously, the subsequent asymptotic form of the velocity being an integral power law in x^{-1} .

The relevance of this result to our problem of homogeneous turbulence lies in the analogy between the above finite region in which the motion initially occurs and the volume of fluid over which the velocity is significantly correlated with itself in homogeneous turbulence. The finite region of fluid in motion in the above simple problem is replaced by a single "eddy" in homogeneous turbulence, and the asymptotic form of the magnitude of the velocity is replaced by the asymptotic form of the velocity covariance. The dynamical equation for $R_{ij}(\mathbf{r})$ is of the form

$$\frac{\partial R_{ij}(\mathbf{r})}{\partial t} = \left(\frac{\partial \overline{pu_j'}}{\partial r_i} - \frac{\partial \overline{p'u_i}}{\partial r_j} \right) + \begin{array}{c} \text{other terms arising from} \\ \text{inertial and viscous forces,} \end{array}$$

and the explicit expression for $\overline{pu_j'}$ in terms of the velocity is found from the relation

$$\nabla^2 \overline{pu_j'} = - \frac{\partial^2 \overline{u_i u_k u_j'}}{\partial x_i \partial x_k}$$

to be

$$\overline{pu_j'} = \frac{1}{4\pi} \int \frac{\partial^2 \overline{u_i' u_k' u_j'}}{\partial x_i'' \partial x_k''} \frac{d\mathbf{x}''}{|\mathbf{x}'' - \mathbf{x}|}.$$

Hence, even if at some initial instant all cumulants of the velocity distribution decrease exponentially, $\overline{pu_j'}$ will in general have an asymptotic form that is an integral power of r^{-1} , so that $(\partial/\partial t)R_{ij}(\mathbf{r})$ will likewise have a power-law asymptotic form (one order smaller) in r^{-1} , the exact power depending on the first integral moment of $\partial^2 \overline{u_i' u_k' u_j'} / \partial x_i'' \partial x_k''$ that is nonzero at the initial instant.

Thus, it is clear that the old assumption that cumulants of the velocity distribution are exponentially small—or, equivalently, that the Fourier transforms, such as the spectrum tensor, are analytic near $\mathbf{\kappa} = 0$ —is not correct, in general. It is not correct because pressure forces, which represent action at a distance (in an incompressible fluid), build up long-range statistical connections which are larger than an exponential form. It is necessary now to consider what kinds of asymptotic forms of velocity cumulants are dynamically persistent (and which therefore have more than a passing significance) and to consider in particular what asymptotic forms are likely to occur when homogeneous turbulence is generated in the usual way by setting a regular array of rods across a uniform stream. The detailed action of the grid of rods on the air passing through it is a complicated process, but Dr. Proudman and I felt that the relevant aspects could be summarized by a simple hypothesis. This hypothesis, which I must ask the reader to take on trust here, is that for homoge-

neous turbulence generated in practice there is a virtual origin in time at which all integral moments of velocity cumulants converge. The important consequence of this hypothesis is that the effect of pressure forces is subsequently to develop asymptotic forms that are integral power laws. Moreover, the set of asymptotic forms that develops from this initial condition is the only dynamically persistent set of asymptotic forms that we have been able to find.

With this hypothesis, it is possible to show, from a consideration of all the time derivatives of $R_{ij}(\mathbf{r})$ at the initial instant, that $R_{ij}(\mathbf{r})$ subsequently becomes of order r^{-5} for large values of r ; the technique is simply to examine expansions, like that already given for the pressure p , of all terms in the dynamical expressions for $\partial R_{ij}/\partial t$, $\partial^2 R_{ij}/\partial t^2$, etc. It can also be shown that the vorticity covariance $\overline{\omega_i(\mathbf{x})\omega_j(\mathbf{x}')}$ becomes of order r^{-3} , although $\overline{u_i(\mathbf{x})\omega_j(\mathbf{x}')}$ becomes of order r^{-6} . Thus the asymptotic form of $R_{ij}(\mathbf{r})$ has a semi-irrotational property, which is due to the fact that pressure forces act only indirectly on vorticity and which is analogous to the completely irrotational character of the motion generated by pressure forces outside a disturbance confined initially to a finite region. It happens that this semi-irrotational property of $R_{ij}(\mathbf{r})$ allows the asymptotic form of $R_{ij}(\mathbf{r})$ when r is large to be determined explicitly. The details of the calculation are very long, and in any case are not very relevant here, so I shall simply quote the result:

$$R_{ij}(\mathbf{r}) = \frac{1}{4} C_{pqmn} \left(\delta_{ip} \nabla^2 - \frac{\partial^2}{\partial r_i \partial r_p} \right) \left(\delta_{jq} \nabla^2 - \frac{\partial^2}{\partial r_j \partial r_q} \right) \frac{\partial^2 r}{\partial r_m \partial r_n} + O(r^{-6})$$

when r is large, where C_{pqmn} is related to the fourth integral moment of the vorticity covariance in a known way. Other terms in the expression for $R_{ij}(\mathbf{r})$ are integral powers of r^{-1} .

Corresponding to this form of $R_{ij}(\mathbf{r})$ at large values of r , there is an expression for the spectrum tensor $\Phi_{ij}(\boldsymbol{\kappa})$ at small values of $\boldsymbol{\kappa}$. Either by taking the Fourier transform of the above asymptotic form, or by a direct method, it is found that

$$\Phi_{ij}(\boldsymbol{\kappa}) = \frac{1}{4\pi^2} C_{pqmn} \left(\delta_{ip} - \frac{\kappa_i \kappa_p}{\kappa^2} \right) \left(\delta_{jq} - \frac{\kappa_j \kappa_q}{\kappa^2} \right) \kappa_m \kappa_n + O(\kappa^3 \log \kappa)$$

when κ is small. Thus the spectrum tensor is of order κ^2 for small values of κ . The nonanalytic character of the leading term at $\boldsymbol{\kappa} = 0$ is associated with the direction of $\boldsymbol{\kappa}$; this was expected, in view of the essentially multidimensional character of the action of pressure forces in building up long-range statistical connections. The spectrum function $E(\kappa)$ obtained by integrating $\Phi_{ii}(\boldsymbol{\kappa})$ over the surface of a sphere in $\boldsymbol{\kappa}$ -space has a leading term of the form $\text{const} \cdot \kappa^4$, although this function also is not analytic near $\boldsymbol{\kappa} = 0$:

$$E(\kappa) = C\kappa^4 + O(\kappa^5 \log \kappa)$$

for small values of $\boldsymbol{\kappa}$.

The last integral moment of $R_{ij}(\mathbf{r})$ that converges is

$$\int r_k r_l R_{ij}(\mathbf{r}) d\mathbf{r},$$

which is only conditionally convergent. This integral, which in the special case of isotropic turbulence has usually been referred to as Loitsiansky's "invariant," has hitherto been believed to be constant during all stages of decay of the turbulence. However, an examination of the proof shows that it depends on the triple velocity product $\overline{u_i u_k u_j'}$ being of smaller order than r^{-4} when r is large. By an examination of the asymptotic form of $\overline{u_i u_k u_j'}$, similar to that carried out for $R_{ij}(\mathbf{r})$, we find

$$\overline{u_i u_k u_j'} = T_{iklm} \frac{\partial^3 1/r}{\partial r_j \partial r_l \partial r_m} + O(r^{-5})$$

when r is large, where T_{iklm} is proportional to the first integral moment of $\overline{u_i u_k u_l'}$. The consequence of this leading term's being of order r^{-4} is that $\int r_k r_l R_{ij}(\mathbf{r}) d\mathbf{r}$ is not constant during the decay, contrary to past belief. In the case of isotropic turbulence the above leading term in the expression for $R_{ij}(\mathbf{r})$ vanishes for reasons of symmetry, and $R_{ij}(\mathbf{r})$ becomes at least as small as order r^{-6} . The integral $\int r_k r_l R_{ij}(\mathbf{r}) d\mathbf{r}$ now converges absolutely, but $\overline{u_i u_k u_j'}$ is still of order r^{-4} , and $\int r_k r_l R_{ij}(\mathbf{r}) d\mathbf{r}$ is not invariant.

Finally, the results for the final period of decay, when viscous dissipation has suppressed all but the part of the spectrum in the neighborhood of zero wave number, had to be examined afresh. We found that the only definite prediction that can be made about the final period of decay is that the total energy ultimately diminishes as t^{-3} , which is also the result found experimentally.

For readers here, I imagine that the principal point of interest lies in the demonstration that a mathematical model of turbulence in which the spectrum is analytic near zero wave number is inadequate. The nonanalytic character of the spectrum is due to the fact, which has been overlooked by people working on turbulence hitherto, that the Navier-Stokes equation is in part an integral equation and contains a term which represents action at a distance and which can lead to the generation of long-range statistical connections in the fluid.

TRINITY COLLEGE,
CAMBRIDGE, ENGLAND.

PROBABILITY IN CLASSICAL PHYSICS

BY

MARK KAC

Introduction. The introduction of probability considerations has proved, over the past hundred years, to be one of the most fruitful and profound additions to the methodological structure of physics. The enormous success of statistical theories and the rich yield of new and important results have tended, and still tend, to obscure the fact that only a modest beginning has been made and that most fundamental problems still await solution.

The problems of statistical physics fall roughly into two main categories: (a) equilibrium problems and (b) problems of approach to equilibrium.

The problems under (a) offer no conceptual difficulties. Starting with the assumption of a canonical (or another) ensemble, the problems are immediately reduced to a purely analytic or combinatorial form.

The challenge of these problems lies in their enormous technical difficulty and in an almost complete lack of general methods of attack.

Problems under (b) are centered around the crucial questions of why and how probability considerations can be introduced into physics.

In addition to technical difficulties (which are formidable!) of solving equations like Boltzmann's equation, there are even more fundamental problems of understanding how these "statistical equations" are related to "deterministic," *e.g.*, dynamical, equations.

In what follows we shall give a brief review of both kinds of problems in the hope of familiarizing mathematicians with this vast and fruitful field.

PART I. EQUILIBRIUM PROBLEMS

1. General formulation. Suppose that a physical system is composed of n constituents and that each constituent is characterized by a variable η_k . The set $(\eta_1, \dots, \eta_n)$, which we shall call the "configuration," determines the system completely, and we assume that the total energy E of the system is a known function of the η 's

$$(1) \quad E \equiv E(\eta_1, \dots, \eta_n).$$

We now make the assumption that the probability (or probability density) of finding the system in the configuration $(\eta_1, \eta_2, \dots, \eta_n)$ is proportional to

$$(2) \quad e^{-\beta E},$$

where

$$(3) \quad \beta = \frac{1}{kT},$$

T being the absolute temperature and k the "gas constant."

It will be outside the scope of this review to justify this fundamental assumption. (As a matter of fact, no thoroughly satisfactory mathematical justification seems to exist.) Suffice it to say that the crucial problem is to calculate the so-called "partition function"

$$(4) \quad Q_n = \omega_n \sum_{\eta} e^{-\beta E},$$

where $\sum$ denotes an integral or a summation over all allowable configurations and ω_n a suitable normalizing constant depending on the problem.

Other quantities are of great physical interest, but for the purpose of this exposition we shall be mainly concerned with the partition function.

The main problem is always to find the so-called "free energy per constituent," which is defined as

$$(5) \quad F = \lim_{n \rightarrow \infty} \frac{1}{n} \log Q_n.$$

The existence of the limit can often be demonstrated, and although it is not always easy, the primary problem is to discuss the properties of F as a function of β and some other parameters. The other parameters are introduced through certain constraints placed upon the η 's, as will be seen later.

2. The Ising models. To appreciate the difficulties involved in the general problem we shall consider the so-called Ising model.

In its simplest form (the one-dimensional model) the model is a linear chain of $n = N$ "sites"; each site k has a "spin" η_k capable of assuming the values $+1$ and -1 . The energy is assumed to be given by the formula

$$(6) \quad E = J \sum_1^{n-1} \eta_k \eta_{k+1},$$

where J is a physical constant expressing the strength of coupling of neighboring spins.

The partition function (4) can now be written in the form

$$(7) \quad Q_n = 2^{-n} \sum_{\eta} \exp \left(-\frac{J}{kT} \sum_1^{n-1} \eta_k \eta_{k+1} \right), \quad (\omega_n = 2^{-n}),$$

and, by noting that the random variables $\eta_k \eta_{k+1}$ are independent, we get

$$Q_n = \left(\cosh \frac{J}{kT} \right)^{n-1},$$

and consequently

$$(8) \quad F = \log \cosh H, \quad H = \frac{J}{kT}.$$

Let us now consider the two-dimensional Ising model. This is an N by N square lattice ($N^2 = n$), and again with each lattice point P (site) we associate

a "spin" η_P . We assume further that the energy is given by the expression

$$(9) \quad E = J \sum_P^* \eta_P \eta_Q,$$

where the asterisk on Σ indicates that Q is a "nearest neighbor" of P , that is, a lattice point directly above (below) or directly to the right (left). In order that each point have four nearest neighbors it is customary (although not necessary) to place the lattice on a torus, thus making the lattice cyclic. [Had we adopted this device for the one-dimensional model, the energy would have been $J \sum_1^n \eta_k \eta_{k+1}$, $\eta_{n+1} = \eta_1$, and a slightly more complicated calculation would have given $Q_n = (\cosh H)^n + (\sinh H)^n$. The limiting result would have remained unaltered.]

The calculation of F in this case presents already formidable difficulties; and it was a great achievement when L. Onsager solved the problem in 1944, using an extremely ingenious algebraic method. The algebraic method has since been simplified by B. Kaufman, and a different method (which is largely combinatorial) has been developed by M. Kac and J. C. Ward. (See [1], where references to other works are given.)

The complexity of the methods contrasts sharply with the simplicity of the answer, which is

$$(10) \quad F = \frac{1}{2} \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \log[\cosh^2 4H - \sinh 4H(\cos \omega_1 + \cos \omega_2)] d\omega_1 d\omega_2,$$

where again

$$H = \frac{J}{kT}.$$

The most interesting feature of (10) is that F as a function of H (or T) has a singularity for $T \neq 0$. This reflects itself in the presence of a singularity of important thermodynamic functions. For example, the specific heat (per particle) which is given by the formula

$$(11) \quad C = kH^2 \frac{d^2 F}{dH^2}$$

has a logarithmic singularity.

The presence of the singularity has another profound effect. Consider the correlation coefficient ρ of two spins at the points A and B which by definition (noting that $\eta^2 = 1$) is

$$(12) \quad \rho(A, B) = \frac{\sum_{\eta} \eta_A \eta_B \exp(-H \sum^* \eta_P \eta_Q)}{\sum_{\eta} \exp(-H \sum^* \eta_P \eta_Q)}.$$

It was shown by Onsager that for $T > T_c$ (T_c = critical temperature, *i.e.*, temperature at which F has its singularity) $\rho(A, B)$ approaches 0 as the distance between A and B approaches ∞ , but for $T < T_c$ this is no longer the case! Thus for sufficiently low temperatures we have a tendency of spins to align. This extremely interesting mathematical phenomenon corresponds to the well-known physical fact that solids, below a certain critical temperature, exhibit magnetic properties.

For the three-dimensional Ising model no method has been found for calculating F and $\rho(A, B)$, and the problem appears to be of great difficulty and depth.

3. The spherical models. The mathematical phenomena exhibited by the Ising model are of sufficient interest in themselves to warrant further study. For this purpose T. H. Berlin and M. Kac [2] introduced a class of models which, although faulty in many physical respects, are amenable to a thorough mathematical discussion.

In these models we shall assume

$$(13) \quad E = J \sum_P^* \eta_P \eta_Q$$

but allow the η 's to vary continuously, subject only to the constraint

$$(14) \quad \sum_P \eta_P^2 = n.$$

The partition function is now defined by the formula

$$(15) \quad Q_n = \frac{1}{S_n(\sqrt{n})} \int_{S_n(\sqrt{n})} \exp\left(-\frac{E}{kT}\right) d\sigma,$$

where the integration is over the surface of the sphere (14) and $S_n(\sqrt{n})$ denotes also its numerical value.

For this model explicit results can be obtained for lattices of all dimensionalities. The calculations are done by the method of steepest descent (a familiar and powerful tool in statistical physics), and the results are as follows:

If d is the dimensionality of the lattice, set

$$(16) \quad f_d(z) = \frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} \log \left[z - \sum_1^d \cos \omega_j \right] d\omega_1 \cdots d\omega_d$$

and

$$(17) \quad g_d(z) = \frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} \frac{d\omega_1 \cdots d\omega_d}{z - \sum_1^d \cos \omega_j}.$$

The equation determining the saddle point z_* is

$$(18) \quad g_d(z_*) = 4H,$$

and the result is

$$(19) \quad F = -\frac{1}{2} - \frac{1}{2} \log 4H + 2Hz_* - \frac{1}{2}f_d(z_*).$$

For $d = 1$ and $d = 2$, (18) has always, *i.e.*, for all H and hence all T , a unique real solution, and F turns out to be an analytic function of H or T . For $d = 3$ there is a critical H_c (corresponding to a critical temperature T_c) such that for $H > H_c$ ($T < T_c$) equation (18) has no solution.

Although for $H > H_c$ there is no saddle point, the method of steepest descent can be properly modified to obtain the result

$$(20) \quad F = -\frac{1}{2} - \log 4H + 6H - \frac{1}{2}f_3(3).$$

Thus for $d = 3$, F as a function of H (or T) is not analytic but consists of two analytic "pieces."

Correspondingly, the correlation coefficient $\rho(A, B)$ goes to 0 as the distance between A and B goes to infinity for $d = 1$ and $d = 2$ but approaches

$$(21) \quad 1 - \frac{T}{T_c}$$

for $d = 3$ and $T < T_c$. For $T > T_c$ even for $d = 3$ the correlation coefficient approaches 0.

A comparison of (16) (for $d = 2$) with (10) exhibits a strong formal similarity whose origin is at the moment quite dark. It would be of considerable interest to elucidate this similarity and to understand the connection between the problems discussed here and the classical random-walk problems, where one again encounters the function $g_d(z)$.

4. The problem of the imperfect gas. The basic problem is to explain the familiar p - v diagram and, in particular, to predict on the basis of statistical mechanics the phenomenon of condensation.

In this case we have

$$(22) \quad E = \sum_{1 \leq i < j \leq n} W(|\mathbf{r}_i^* - \mathbf{r}_j|)$$

and

$$(23) \quad \omega_n = (n!)^{-1}.$$

$W(r)$ is the intermolecular potential, $\mathbf{r}_1, \dots, \mathbf{r}_n$ position vectors of centers of gas molecules (or atoms), and $|\mathbf{r}_j - \mathbf{r}_i|$ the distance between end points of position vectors $\mathbf{r}_i$ and $\mathbf{r}_j$.

About $W(r)$ one makes the following assumptions:

$$\begin{array}{ll} (a) & W(r) > 0, \quad 0 \leq r \leq \delta, \\ (b) & W(r) \rightarrow \infty, \quad r \rightarrow 0, \\ (c) & W(r) \leq 0, \quad r > \delta, \\ (d) & W(r) = 0, \quad r > d. \end{array}$$

Assumption (d) can be replaced by the requirement that $W(r)$ approach 0 with sufficient rapidity.

The partition function is now

$$(24) \quad Q_n = \frac{1}{n!} \int_V \cdots \int_V \exp\left(-\frac{E}{kT}\right) d\mathbf{r}_1 \cdots d\mathbf{r}_n,$$

where V is the volume of the container (each $\int$ is thus a triple integral).

Again we want to find

$$(25) \quad F = \lim_{n \rightarrow \infty} \frac{\log Q_n}{n}$$

but with the important restriction

$$(26) \quad \lim_{n \rightarrow \infty} \frac{V}{n} = v.$$

The fact that the limit (25) exists has been demonstrated by L. van Hove [3], who also proved that

$$F \equiv F(v, T)$$

has many properties which one expects on the basis of classical thermodynamics. The basic property of F , viz., that for $T < T_c$, ($T_c \equiv$ critical temperature) F as a function of v has a singularity (condensation!) and that

$$p = - \frac{\partial F}{\partial v}$$

is constant over an interval, has never been fully demonstrated.

Considerable progress was made by Ursell, Mayer, Uhlenbeck and Kahn, Yang and Lee, and others, but we are still very far from a complete solution.

Recently Uhlenbeck and his students tried to reduce the problem to a combinatorial one. They found interesting connections with the theory of Cayley trees and their generalizations, but again only partial results were obtained.

The simplest case is the case of the gas of "hard spheres."

In this case

$$\begin{aligned} W(r) &= \infty, & 0 \leq r \leq \delta, \\ W(r) &= 0, & r > \delta, \end{aligned}$$

and Q_n has a very simple interpretation.

In fact,

$$(27) \quad Q_n = \frac{V^n}{n!} P_n,$$

where P_n is the probability that if n points are chosen at random in V , no two of them will be closer than δ .

The one-dimensional version of this problem is very easy to solve.
In fact, let

$$\psi(x) = \begin{cases} 1, & 0 < x < \delta, \\ 0, & x > \delta, \end{cases}$$

and note that

$$(28) \quad P_n = \frac{1}{L^n} \int_0^L \cdots \int_0^L \prod_{1 \leq i < j \leq n} (1 - \psi(|x_j - x_i|)) dx_1 \cdots dx_n.$$

Since the integrand in (28) is symmetric, we have

$$(29) \quad P_n = \frac{n!}{L^n} \int_{0 \leq x_1 < x_2 < \cdots < x_n \leq L} \prod_{1 \leq i < j \leq n} (1 - \psi(|x_j - x_i|)) dx_1 \cdots dx_n,$$

and since for $0 \leq x_1 < x_2 < \cdots < x_n \leq L$ we have

$$(30) \quad \prod_{1 \leq i < j \leq n} (1 - \psi(|x_j - x_i|)) = \prod_{i=1}^{n-1} (1 - \psi(x_{i+1} - x_i)),$$

it follows that

$$P_n = \frac{n!}{L^n} R_n(L),$$

where

$$R_n(L) = \int_0^L dx_n \int_0^{x_n} dx_{n-1} (1 - \psi(x_n - x_{n-1})) \int_0^{x_{n-1}} dx_{n-2} (1 - \psi(x_{n-1} - x_{n-2})) \cdots \int_0^{x_2} dx_1 (1 - \psi(x_2 - x_1)).$$

Taking Laplace transforms, we get

$$\int_0^\infty e^{-sL} R_n(L) dL = \frac{1}{s} \left(\int_0^\infty e^{-sx} (1 - \psi_n(x)) dx \right)^{n-1} = s^{-n} e^{-(n-1)\delta s},$$

and consequently

$$R_n(L) = \frac{(L - (n-1)\delta)^{n-1}}{(n-1)!}.$$

Finally,

$$(31) \quad P_n = \frac{n}{L} \left(1 - \frac{(n-1)\delta}{L} \right)^{n-1}.$$

The simplicity of the result is due to formula (30), which in turn depends on the fact that points on a straight line are ordered in a natural way.

The natural ordering of points on a straight line reflects itself in the possibility of computing (at least in principle) more complicated integrals of the type

$$(32) \quad Q_n = \frac{1}{n!} \int_0^L \cdots \int_0^L \exp \left(-\frac{1}{kT} \sum_{1 \leq i < j \leq n} W(|t_j - t_i|) \right) dt_1 \cdots dt_n.$$

Let, for instance,

$$W(t) = \alpha e^{-\beta|t|},$$

and note that

$$(33) \quad Q_n = \frac{\exp(n\alpha/kT)}{n!} \int_0^L \cdots \int_0^L \exp\left(-\frac{\alpha}{2kT} \sum_{i,j=1}^n \exp(-\beta|t_j - t_i|)\right) dt_1 \cdots dt_n.$$

Let now $x(\tau)$ be the stationary, Markoffian, Gaussian process whose correlation function is

$$e^{-\beta|t|}.$$

It follows easily that

$$\exp\left(-\frac{\alpha}{2kT} \sum_{i,j=1}^n \exp(-\beta|t_j - t_i|)\right) = E \left\{ \exp\left(i \left(\frac{\alpha}{2kT}\right)^{\frac{1}{2}} \sum_{j=1}^n x(t_j)\right) \right\},$$

and hence

$$(34) \quad Q_n = \frac{\exp(n\alpha/kT)}{n!} E \left\{ \left(\int_0^L \exp\left(i \left(\frac{\alpha}{2kT}\right)^{\frac{1}{2}} x(\tau)\right) d\tau \right)^n \right\}.$$

Because the process $x(\tau)$ is Markoffian, the calculation of

$$E \left\{ \left(\int_0^L \exp\left(i \left(\frac{\alpha}{2kT}\right)^{\frac{1}{2}} x(\tau)\right) d\tau \right)^n \right\}$$

can be reduced to solving a differential equation of parabolic type.

We are concerned here with the extension of the theory of distributions of functionals of the form

$$\int_0^L V(x(\tau)) d\tau,$$

Such a theory was given by the author in case $x(\tau)$ is a Wiener process. Extensions to more general Markoff processes can be found in a recent book by Blanc-Lapierre and Fortet and (in a better form) in a Rand report by Darling and Siegert.

Although the differential equation pertinent to the calculation of (34) has not been solved, the very possibility of a reduction to a differential equation is noteworthy. No such reduction seems possible in two (or more) dimensions because we are then led to a Gaussian process $x(\sigma, \tau)$ with two-dimensional "time" (σ, τ) , and there is no adequate generalization of the notion of a Markoffian process to this case.

In closing Part I we must mention the simplest unsolved problem of the type discussed here.

In a square lattice consider all random-walk paths of n steps emanating from a fixed point and which do not have multiple points (each lattice point is

to be covered only once). Denoting by l_n their number, we are required to find

$$\lim_{n \rightarrow \infty} \frac{\log l_n}{n}.$$

PART II. NONEQUILIBRIUM PROBLEMS

5. Introductory remarks. As we have already mentioned in the Introduction, it is in nonequilibrium problems that the use of probability theory becomes of central and crucial importance.

The role of probability theory in classical physics is very peculiar. On the one hand the introduction of statistical reasoning proved to be enormously successful. On the other hand real understanding as to why we are allowed to use statistical methods is still largely lacking.

Boltzmann struggled with this problem, and so did Gibbs. Their efforts are fully and critically described in the celebrated article of P. and T. Ehrenfest on conceptual foundations of statistical mechanics, and we have progressed very little since the appearance of the article in 1912.

There are theories in statistical physics in which a probability mechanism is present from the very start. The best-known and most thoroughly discussed example of such a theory is the theory of Brownian motion. But already in the more basic theory of a gas not in thermal equilibrium there is considerable confusion (inherited from the early works of Boltzmann and due partly to the polemic nature of the subject at that time) as to the role and use of statistical assumptions.

Indeed, in reading over the older (as well as the more recent) works on the subject one is impressed by the desire of the authors not to use probability at all. It is only when faced with paradoxes (the reversibility paradox of Loschmitt and the recurrence paradox of Zermelo) that a belated appeal to probability is made.

We shall not touch upon the truly deep question of justification of probabilistic reasonings in kinetic theory. The question is far from settled and is not yet presentable in a form palatable to mathematicians.

What we shall do instead is to show that, in a special case, a consistent probabilistic approach is indeed possible.

Our approach may be said to lie halfway between Boltzmann's kinetic approach and the statistical approach of Gibbs, and may serve as an illustration of how probability is introduced into a problem of classical physics.¹

6. The Boltzmann equation and the "Stosszahlansatz." The basic equation of kinetic theory of dilute monoatomic gases is the nonlinear integrodifferential equation of Boltzmann.

¹ The full account may be found in a mimeographed report by the author, "Foundations of Kinetic Theory." This report was written under contract No. AF 18 (600)-685 with the Office of Scientific Research of the U.S. Air Force. The report will be published in the *Proceedings of the Third Berkeley Symposium on Probability and Statistics*.

In the case when the molecules are hard spheres, of diameter δ , which are allowed to exchange energy only through elastic collisions the equation assumes the form

$$(35) \quad \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f + \mathbf{X}(\mathbf{r}) \cdot \nabla_{\mathbf{v}} f = \frac{\delta^2}{2} \int d\mathbf{w} \int d\mathbf{l} \{ \tilde{f}\tilde{f}_1 - ff_1 \} |(\mathbf{w} - \mathbf{v}) \cdot \mathbf{l}|,$$

where $f d\mathbf{r} d\mathbf{v} \equiv f(\mathbf{r}, \mathbf{v}, t) d\mathbf{r} d\mathbf{v}$ is the number of molecules in $d\mathbf{r} d\mathbf{v}$ at $(\mathbf{r}, \mathbf{v})$, $\nabla_{\mathbf{r}}$ the gradient with respect to $\mathbf{r}$ and $\nabla_{\mathbf{v}}$ the gradient with respect to $\mathbf{v}$. Furthermore

$$\begin{aligned} f_1 &\equiv f(\mathbf{r}, \mathbf{w}, t), \\ \tilde{f} &\equiv f(\mathbf{r}, \mathbf{v} + (\mathbf{w} - \mathbf{v}) \cdot \mathbf{l}, t), \\ \tilde{f}_1 &\equiv f(\mathbf{r}, \mathbf{w} - (\mathbf{w} - \mathbf{v}) \cdot \mathbf{l}, t), \end{aligned}$$

$\mathbf{l}$ being a unit vector and $d\mathbf{l}$ the surface element of the unit sphere. Finally, $\mathbf{X}(\mathbf{r})$ is an outside force, *e.g.*, gravity, acting on a molecule at $\mathbf{r}$.

If $\mathbf{X}(\mathbf{r}) = 0$ and we set

$$(36) \quad f(\mathbf{r}, \mathbf{v}, t) = \frac{n}{V} f(\mathbf{v}, t)$$

($n \equiv$ total number of particles, $V \equiv$ the volume of the container), we note that (36) will be a solution of (35) if, and only if, $f(\mathbf{v}, t)$ is a solution of the reduced equation

$$(37) \quad \frac{\partial f}{\partial t} = \frac{n\delta^2}{2V} \int d\mathbf{w} \int d\mathbf{l} \{ \tilde{f}\tilde{f}_1 - ff_1 \} |(\mathbf{w} - \mathbf{v}) \cdot \mathbf{l}|.$$

Equation (37) governs the temporal evolution of the velocity distribution while the spatial distribution remains uniform.

It is from (37) that Boltzmann deduced his famous *H*-theorem, to the effect that

$$(38) \quad \frac{d}{dt} \int f \log f d\mathbf{v} \leq 0.$$

The derivation of (37) is based on the Stosszahlansatz (which goes back to Clausius) and which in its original (and most primitive) form was stated as follows:

The number of collisions (in time dt) between *A*-molecules (those whose velocities lie between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$) and *B*-molecules (those whose velocities lie between $\mathbf{w}$ and $\mathbf{w} + d\mathbf{w}$) which take place while the center line is in the direction $\mathbf{l}$ (within $d\mathbf{l}$) is

$$(39) \quad n_A n_B \delta^2 \frac{|(\mathbf{w} - \mathbf{v}) \cdot \mathbf{l}| - (\mathbf{w} - \mathbf{v}) \cdot \mathbf{l}}{2} d\mathbf{l} dt,$$

* The action of the walls of the container must be included in $\mathbf{X}(\mathbf{r})$, so that $\mathbf{X}(\mathbf{r}) = 0$ is incompatible with the fact that the gas is enclosed in a container. This is a minor difficulty which we shall disregard.

where n_A and n_B denote the numbers of A -molecules and B -molecules, respectively.

The derivation of (39) is intuitively very appealing (see almost any text on kinetic theory of gases), and it is only after one notices that the H -theorem contradicts the basic reversibility of the underlying model that one begins to be suspicious. It is certainly clear that the Stosszahlansatz (39) is not strictly derivable from the dynamics of the model and that we are dealing here with an assumption of statistical nature.

It is thus natural to interpret the Stosszahlansatz as an assumption concerning probability of a certain type of collision and to attempt to derive anew the Boltzmann equation (37). This procedure bypasses the deeper question of justification of such an assumption and avoids entirely the problem of how to "guess" a corresponding probability assumption in more general cases.

The more modest program (outlined above) is, however, executable, and we shall illustrate it by considering an instructive simplified model.

7. The model and the "master equation." Let

$$(40) \quad \mathbf{R} = (x_1, x_2, \dots, x_n)$$

be subject to the condition

$$(41) \quad \|\mathbf{R}\|^2 = x_1^2 + \dots + x_n^2 = n,$$

and let

$$(42) \quad A_{ij}(\theta)\mathbf{R} = (x_1, \dots, x_i \cos \theta + x_j \sin \theta, \dots, -x_i \sin \theta + x_j \cos \theta, \dots, x_n).$$

Let us assume that in time dt we have the probability

$$(43) \quad \frac{1}{n} g(\theta) d\theta dt, \quad (-\pi \leq \theta \leq \pi, g(\theta) \geq 0)$$

that $\mathbf{R}$ goes over into $A_{ij}(\theta)\mathbf{R}$ and the complementary probability

$$1 - \frac{n(n-1)}{2\pi} \int_{-\pi}^{\pi} g(\theta) d\theta dt$$

that nothing happens, that is, $\mathbf{R}$ remains unaltered.

The analogy of this model with the gas is obvious. The x_i are analogues of velocities of particles and θ the analogue of 1. Expression (43) is the analogue (probabilistically interpreted) of the Stosszahlansatz, and the great simplification is that the probability of the (i, j, θ) "collision" depends only on θ and not on x_i and x_j (this makes our model analogous to the so-called "Maxwell gas"). Formula (41) is the conservation-of-"energy" law (with one unit of "energy" per particle), and (unlike the case of a real gas) we have to abandon conservation of momentum in a collision.

If we start with a density $\varphi(\mathbf{R}, 0)$ of points on the sphere (41), the density $\varphi(\mathbf{R}, t)$ is easily seen to satisfy the "master equation"

$$(44) \quad \frac{\partial \varphi(\mathbf{R}, t)}{\partial t} = \frac{1}{n} \sum_{1 \leq i < j \leq n} \int_{-\pi}^{\pi} g(\theta) \{ \varphi(A_{ij}(\theta) \mathbf{R}, t) - \varphi(\mathbf{R}, t) \} d\theta.$$

8. Derivation of the Boltzmann equation. To get the Boltzmann equation we first note that $\varphi(\mathbf{R}, 0)$ must be assumed to be symmetric in the variables $x_1, \dots, x_n$ (particles are indistinguishable). It then follows easily that so is $\varphi(\mathbf{R}, t)$.

We define now a hierarchy of "contracted densities"

$$\begin{aligned} f_1^{(n)}(x, t) &= \int_{x_1^2 + \dots + x_{n-1}^2 = n - x^2} \varphi(\mathbf{R}, t) d\sigma_1, \\ f_2^{(n)}(x, y, t) &= \int_{x_1^2 + \dots + x_{n-2}^2 = n^2 - x^2 - y^2} \varphi(\mathbf{R}, t) d\sigma_2, \end{aligned}$$

etc., where $d\sigma_1, d\sigma_2, \dots$ are surface elements of spheres indicated under the integration signs.

Integrating (44) over all variables but one, one gets easily

$$(45) \quad \frac{\partial f_1^{(n)}(x, t)}{\partial t} = \frac{(n-1)}{n} \int_{-\sqrt{n-x^2}}^{\sqrt{n-x^2}} dy \int_{-\pi}^{\pi} g(\theta) \{ f_2^{(n)}(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, t) - f_2(x, y, t) \} d\theta,$$

which is strongly reminiscent of the Boltzmann equation. To complete the derivation one somehow has to justify that

$$f_2^{(n)}(x, y, t) \sim f_1^{(n)}(x, t) f_1^{(n)}(y, t)$$

for all x, y , and t .

This, of course, is not true in general, and one needs a new assumption on $\varphi(\mathbf{R}, 0)$.

We shall say that a sequence $\varphi_n(\mathbf{R}, 0)$ of probability densities defined on spheres $\|\mathbf{R}\|^2 = n$ has the "Boltzmann property" if

$$(46) \quad \lim_{n \rightarrow \infty} f_k^{(n)}(x_1, \dots, x_n, 0) = \prod_{j=1}^k \lim_{n \rightarrow \infty} f_1^{(n)}(x_j, 0).$$

It can then be shown that: (a) The Boltzmann property propagates in time, i.e., (46) holds with 0 replaced by t . (b) Setting

$$f(x, t) = \lim_{n \rightarrow \infty} f_1^{(n)}(x, t)$$

and assuming the Boltzmann property, one has

$$(47) \quad \frac{\partial f}{\partial t} = \int_{-\infty}^{\infty} dy \int_{-\pi}^{\pi} g(\theta) \{ f(x \cos \theta + y \sin \theta, t) f(-x \sin \theta + y \cos \theta, t) - f(x, t) f(y, t) \} d\theta,$$

which is the Boltzmann equation for our model.

Thus the nonlinear equation (47) appears as a consequence of the linear equation (44). The nonlinearity is, of course, due to the fact that we start with very special densities.

One encounters difficulties when one tries to extend our approach to the phase space and hence derive the full Boltzmann equation (35). The difficulties are again of conceptual nature, since one does not see how to introduce probabilistic assumptions. In the spatially homogeneous case we obtain the Stosszahlansatz and the "master equation" by averaging over all spatial configurations. In the general case there is nothing to average over, and we are back to the central problem of statistical mechanics of understanding the nature of the probabilistic method.

BIBLIOGRAPHY

1. M. Kac and J. C. Ward, *A combinatorial solution of the two-dimensional Ising model*, Phys. Rev. vol. 88 (1952) pp. 1332-1337.
2. T. H. Berlin and M. Kac, *The spherical model of a ferromagnet*, Phys. Rev. vol. 86 (1952) pp. 821-835.
3. L. van Hove, Physica vol. 15 (1949) p. 951.
4. For an excellent review of earlier work on condensation and related problems see B. Kahn, *On the theory of the equation of state*, Doctoral Dissertation, Amsterdam, 1938.
5. For an excellent review (as well as for an extensive list of references) of nonequilibrium problems see E. W. Montroll and M. S. Green, *Statistical mechanics of transport and nonequilibrium processes*, Ann. Rev. of Phys. Chem. vol. 5 (1954) pp. 449-476.

CORNELL UNIVERSITY,
ITHACA, N.Y.

INFINITE MODELS IN PHYSICS

BY

S. M. ULAM

The purpose of this paper is to present certain mathematical models of physical systems with the intention of indicating a whole class of interesting possibilities not normally apparent in the orthodox formulation. These models differ from the usual ones in two ways. First, they employ an actual infinity of points or objects present from the start; second, instead of using exclusively the Euclidean continuum (topologically) and the real-number system (algebraically), they will start with somewhat more general schemes. (The mathematical description of the physical world need not be specialized to the usually made assumptions that for a discrete description we use a finite system of mathematical points and for a continuum the familiar Euclidean n -dimensional space.) In the first part of this paper we shall discuss in a very special case the behavior of a system consisting of infinitely many discrete points and concentrate on some properties of such systems which cannot be obtained by a passage to the limit $n \rightarrow \infty$ of very large finite systems of n points. In the second part there will be a proposal of study of physical systems in which the number of points is of the power of the continuum and which are not Euclidean in character, but topologically these systems will be zero-dimensional, nowhere-dense perfect sets (Cantor sets).

PART I

Let us imagine the following "physical" system. On the real axis, in a one-dimensional space, we are given at time $t = 0$ an infinite collection of mass points distributed at random, on the points of the axis with integer coordinates. We assume that all the masses are equal and the random distribution is defined by the following procedure. For each integer value of x we have a priori probability $\frac{1}{2}$ that a mass point will occupy the position. Any such distribution of points may be denoted by a real number between zero and 1 written in the binary notation. We can, for example, reserve the even indices for the nonnegative integers, the odd ones for the negative ones, writing a 1 in the $2n$ th place if there is a point in the place $x = -n$, zero otherwise, 1 in the position $2n + 1$ if there is a mass point in the position $x = +n$, otherwise zero. In this fashion every initial distribution will correspond to a unique real number. The assumption of equal probability for zero and 1 leads to the ordinary Lebesgue measure on the interval.

We shall now assume that our mass points exert forces on each other. In the simplest case, we may consider attractions with the magnitude of the force proportional to the inverse square of the distance. Assuming an initial distribution as given at time $t = 0$, we want to study the properties of such a

system for subsequent times and establish some theorems about the behavior of such systems. It is important to notice that we introduce a probabilistic point of view from the beginning. A meaningful statement on properties of motion of such a system will concern "almost all" such initial distributions. An exceptional set of initial distributions, *i.e.*, a set such that the representative set of real numbers has measure zero in the sense of Lebesgue, will be of no concern to us. The statements made from now on about systems of this sort should be understood to be valid for almost all such systems. For simplicity we consider the objects of this sort located in a one-dimensional space. It is possible to make analogous constructions in three or n dimensions. Our mass points can then be thought of as located on the lattice points with all coordinates integer-valued in the (x,y,z) -space. (The representation of any such system of mass points could be, again, a point with all coordinates written in a binary system and, say, located in the unit cube.) Given such a system of points, we may also assume forces between any pair of these mass points. The first question that arises is whether the resultant total force acting on any point is well defined. In one dimension there is no problem since the sum of inverse squares of the distances converges absolutely. One can show, using rather elementary forms of the law of large numbers about sums of random vectors, that in two and even in three dimensions the resultant force due to attractions between points with an inverse-square law exists, with probability 1 for every point of our system, if this resultant is defined as a limit of resultant forces on a given point from all points located in spheres with the given point as a center and with radii tending to infinity. (In four dimensions this would no longer be true; it is the inverse square in the expression for force which ensures that convergence obtains in three dimensions.) The force on each point being well defined, we can write down an infinite system of total differential equations (Newton's equations of motion) defining the motion of each point of our system. One has to show, and this can be done, that the force is well defined not only for $t = 0$ but for positive values of time. We should notice here that a system of Newtonian equations is meaningful for such an infinite collection of points whereas some of the more general formulations of analytical mechanics, *e.g.*, a variational principle, can be defined only with difficulty. For example, a quantity like the total energy of a system is not defined since the total mass is infinite; the total kinetic energy may be infinite for $t > 0$, and so on. We hope to present a discussion of the question of how to formulate the more general principles in such infinite systems in another paper.

A convention is necessary to describe the process of collision between two mass points. Since we deal here, in our idealization, with finite masses located at mathematical points, the velocity of approach will become infinite before collision, and a special convention is necessary for describing the result of such a collision between a pair of points. These collisions are certain to happen in a one-dimensional system. Let us assume here that when two points collide, the momentum will be preserved. If masses m_1 and m_2 collide, a new mass

$m_1 + m_2$ is formed with a velocity given by the condition of preservation of momentum. The energy, then, will not be preserved, but we could imagine that it disappears in some internal degrees of freedom (of the point!), *e.g.*, as heat. It is clear that in a one-dimensional problem, collisions will lead to the appearance of ever-increasing masses. One is interested in the behavior of our system as a function of time. Some of the properties of (almost all) such systems can be established rather easily:

1. The masses appearing in the course of time will be unbounded. In other words, for almost every initial condition of our system there will exist for every M a time t such that a mass greater than M will appear after this time.

2. There will always exist single particles. In other words, for almost every system and for every t there will exist in the system a point with unit mass.

3. The asymptotic density of our system remains constant and equal to the original density. We define the asymptotic density as follows. Consider the totality of particles contained in an interval from $-N$ to $+N$, and denote by M_N the total mass of all particles in this interval. The $\lim_{N \rightarrow \infty} M_N/2N$ shall be

called the asymptotic density if it exists. With our initial masses equal to 1, and the random placing of these masses on integer points, this limit (from Bernoulli's theorem) is equal to $\frac{1}{2}$. It is easy to see that this limit will exist and be equal to $\frac{1}{2}$ for all t . This is simply due to the fact that, given any t , the displacement of each particle will be bounded. If we take a sufficiently large interval, the flux across its ends will constitute an arbitrarily small fraction of the total number of particles, and our assertion follows.

4. Arbitrarily large "holes" will appear in our system; *i.e.*, for almost every system and for all d there will exist a time t so that there will be infinitely many mass points separated by intervals larger than d . Moreover, for all greater times these long empty intervals will continue to exist.

All this is very easy to establish in one dimension. In two or more dimensions, we shall *not* have, in general, collision between point masses, and we should instead have to define captures, *i.e.*, formation of double or multiple systems. The corresponding theorems on the existence of stable or semistable captures are much harder to prove. Probably an easier way to deal with an analogue of our system would be to give to each point a finite size and then consider, say, certain collisions as leading to formations of larger masses. Properties 2 and 3 should be easy to establish.

More interesting would be a quantitative description of the behavior of such a system. For example, it is interesting to calculate in one dimension the average mass of a particle in our infinite system for any time t and to determine the shape of the distribution of masses as a function of time. Another set of data would concern the velocity distribution of our particles as a function of time. (To define a meaningful average velocity one would have to introduce a cutoff in the distance between two particles which approach each other just before a collision; otherwise, in the system consisting of mathematical points these velocities become arbitrarily large during the collision process.)

If we define a cutoff distance, then the average of our velocity will have a meaning for all t , and the question arises: What is the "temperature" of the system as a function of time?

In order to study these and other questions, a series of experiments was performed on a computing machine by John Pasta and the author. We tried to imitate our infinite system by a finite one composed of a great number of masses placed on points of a regular subdivision of an interval with a random decision made for having or not having points in successive positions. In order to "approximate" an infinite system somewhat realistically, one has to imagine the two end points of the interval on which the points are located as coinciding; *i.e.*, we have a finite system of points on the circumference of a circle. This attenuates the end effects. It is clear that a finite system will imitate an infinite one only for times less than a certain value. Given a finite system, it is certain that it will ultimately collapse to a single point, whereas in the infinite one, we know that the asymptotic density will remain constant for all times. Therefore, in interpreting the results of a calculation for a finite system, one has to carry them only up to a value of time T where the system still consists of "many" points. In order to make a rigorous analysis, for the study of any given functional of the distribution of distances, masses, velocities, etc., of our system, one would have to give a priori inequalities for T as a function of N and a $\delta > 0$; that is, given only a finite system of N points, T such that the functional under consideration computed up to time T will differ in value by less than δ from the value of this functional for an infinite system. Many finite systems were computed, *i.e.*, we started with many distributions of mass points given by our initial random procedure, and the subsequent motions of each point were calculated in each case with the hope of obtaining heuristic results on the behavior of such systems. The number of points initially taken was of the order of 1,000. Among the quantities studied were the distribution of masses at any time t , the distribution of distances, etc. The results of these experiments will be reported on in a separate paper. I shall indicate here very briefly some of the qualitative facts about them:

- a. The average mass of a particle appears to be increasing linearly with time.
- b. There is a suspicion, at least, that if one considers the distribution of masses of particles in the units of the average mass at time t , this distribution tends to a fixed function. In other words, a steady state may establish itself.
- c. A quantity which was called hierarchy was studied. This is defined by induction as follows: the original particles have by definition a hierarchy rank zero. When two particles of rank m and n collide, they form a particle whose hierarchy rank is equal to the greater of the two numbers m, n . In case $m = n$, the rank of the new particles is $m + 1$. This index gives an idea of the degree or order of conglomeration as distinguished from a mere increase of mass by accretion.
- d. The average kinetic energy was studied as a function of time. We used in the computation a cutoff in the distance of approach in order to eliminate

the arbitrarily high velocity just before collapse. The shape of this function cannot yet be discussed in detail, but it is obvious that this average energy increases with time and then starts decreasing again, which, of course, is due to the fact that our system ultimately will end up as one big particle at rest. Presumably, our finite systems imitate the infinite one only up to the time when this average energy stops increasing. Nothing as yet, therefore, can be stated about the behavior of "temperature" in time.

It should be pointed out here that the type of distribution which we have introduced in our numerical work (à la Bernoulli) could equally well be chosen differently. For example, we could assume that there is a fixed probability $\alpha \cdot dx$ of finding a unit mass in the interval dx ; that is, the initial randomness could be chosen à la Poisson. It is easy to see that the convergence of the force on each particle of the system would be equally valid for almost all initial conditions with this setup.

One could postulate a finite interval with infinitely many points of various masses distributed in this interval at time $t = 0$ by a given recipe and then discuss the ensuing motions. In this connection it is amusing to notice that there exists a sequence of numbers x_i , all between zero and 1, and a sequence

of masses m_i , such that $\sum_{i=1}^{\infty} m_i = 1$, with the following property. If we place

the mass m_i at the position x_i and let any two of these masses attract each other with a force inversely proportional to the square of their distance, the whole system will be in static equilibrium! Similarly with repulsive forces.

One could assume, of course, not only that the positions of our points on the line or in space are given at random but also that there exist, already at time $t = 0$, velocities of each of these points, distributed in a random fashion. It is proposed to study, heuristically, the behavior of such systems by performing computations on a large finite system.

PART II

We have briefly discussed an infinite but countable system of material points subject to forces acting on each of them. We shall now outline some models of mass distributions which combine such a discrete character with certain properties of a continuum. The most elementary schema of this sort would perhaps be the following.

Imagine a mass point with a mass equal to unity, located in the middle of the interval 0,1. This mass point can now, with a probability p_1 , remain forever in its original position or, with a probability $p_2 = 1 - p_1$, split into two parts, each of mass $\frac{1}{2}$, which will be located in the positions $\frac{1}{4}$ and $\frac{3}{4}$, respectively. If the latter eventuality has occurred, we shall assume again that either each of the two masses can stay where it is or each can independently split into two masses (equal to $\frac{1}{4}$ each) which will be located at $\frac{1}{8}$ and $\frac{3}{8}$ for the first point or $\frac{5}{8}$ and $\frac{7}{8}$ for the second point. This process is to continue indefinitely. We

imagine that each of the points can split into two equal ones, which will then be located to the left and right of it, with a probability p_2 , or stay "dead" forever. We have then a branching or multiplicative process which will define, at each stage, a possible distribution of masses on dyadically rational points of the interval. The process is defined by the two constants p_1 , p_2 . If p_2 should be equal to 1 and $p_1 = 0$, we would have a certainty of splitting every time, and the process would lead to all rational binaries, each having in the limit mass zero. The closure of our set would be the full continuum of real numbers between zero and 1. If p_2 is less than 1, we shall get as a result of our continuing branching process a countable set of points whose closure will be, *with probability 1*, a nowhere-dense set consisting of some isolated points and a perfect nowhere-dense set of the kind first defined by Cantor. There are three cases to distinguish, a subcritical, just critical, and supercritical system. In our simple setup, these correspond to the cases $p_2 < \frac{1}{2}$, $p_2 = \frac{1}{2}$, $p_2 > \frac{1}{2}$, respectively. In the last case there is a finite probability of the process never ending, and as a result, in addition to having a finite collection of points which are isolated, we shall obtain a perfect set on the interval as a closure of the "unending" part of the process. A detailed discussion of such branching can be found, *e.g.*, in [1]. We have to make here more precise the sense in which we speak of a result "of one such process." What is meant, of course, is that one considers all possible outcomes of such a branching process. There exists a measure in the space of all possible branching processes defined in a rather natural way [2]. When we speak of the process leading with probability 1 to a set with given properties, we mean that the subset of all outcomes with these properties has measure 1 in the space of all possible ones.

One can look upon the sets of points obtained by the above construction as describing "virtual positions of a physical object" or upon the space itself as being a collection of virtual symbols. These then will not, in general, form Euclidean continua. Neither will they consist of a discrete set of points. It is obvious that an analogous procedure can be effected in three dimensions. One could, for example, perform our branching independently in all three dimensions, or one could imagine the following single process. A particle with mass 1, located at the center of a unit sphere splits with probability p_2 into two particles, each with mass $\frac{1}{2}$, and located on the opposite ends of an interval with length α_1 . The direction of this interval can be obtained by a random process, say with isotropic distribution in space. Each of the ends with mass equal $\frac{1}{2}$ can again be subject to the same splitting possibility, say, with the same probability p_2 independently, and then split into two particles, each with mass equal to $\frac{1}{4}$, located at the ends of an isotropically chosen interval of length α_2 . If the process continues indefinitely, we shall obtain a three-dimensional analogue of the sets above. It might be of interest to add here that it is by no means easy to determine the topological character of the resulting set. It has been known (since Antoine) that some perfect nowhere-dense sets in 3-space are equivalent to sets located in one dimension under a homeomorphism of the

whole three-dimension space and others are not. The question which of the two is more likely under our process is not immediately answered.

This way of considering a space of symbols to correspond to a model of a physical situation is perhaps the simplest one of its kind and, so to say, applies to the "configuration space" of a particle. An analogous construction should be thought of as proceeding in the phase space. Not merely the positions but the momenta or velocities of a particle could be thought of as generated by a process like the above, leading to a Cantor-set structure for all possible values of the physical quantities in which we are interested. In this paper we shall restrict ourselves to the configuration space alone in order to bring out the features of the proposal more clearly. In addition, we shall confine ourselves here to very classical, *i.e.*, nonrelativistic and non-quantum-theoretical, features of such models. Any pretended attempt to take more seriously the implications of such models for physics would have to consider such constructions in Minkowski space.

We mentioned above some of the properties of sets obtained by our branching process; *e.g.*, the set will be, with probability 1, nowhere dense. To obtain quantitative data about this set seems quite difficult. Again exploratory numerical work seemed of some value, and a series of computations on an electronic machine was undertaken as follows. Starting with the original point in the middle of the interval, the process was continued by the use of random numbers. That is, a great number of finite sets was determined, the process being stopped each time after a certain number of "generations." Given the probabilities p_1 , p_2 , one may obtain a sufficient number N of sets produced by the splitting process for a statistical study. If the number of generations k for the splitting process is kept constant, we shall obtain a variable number of points in the set. One can then compute the average value of any functional of such a set. The integration in the space of all possible branching processes is then replaced by averaging on the N sets actually produced. To justify the averaging as an approximation to integration is very easy for many functionals. The quantities studied are as follows:

1. Given a set of mass points at generation time k , one can compute its moment of inertia I . (The center of gravity of each system, as is obvious from its definition, is located in the original position of the mass point since our splitting preserves the center of mass.) If we average the value of the moment of inertia over all N sets which we have manufactured, we obtain the value of the average moment of inertia of the infinite set with good approximation.

2. If we imagine that any two points of our set attract each other with a force proportional to the masses and inversely to the square of the distance, we can investigate the question of the value of the "gravitational" self-energy. Here again if we compute this quantity for each of our N sets and take the average, we shall get an approximation to the average or expected value of this quantity in our infinite process. This value exists in the three-dimensional case.

In addition to computing the average, one can get a fair idea of the distribution of the value of this self-energy. In order to ascertain it with any accuracy, a very great number N of sets has to be manufactured, of course.

3. One could ask about the attraction of two systems of the above sort, located on an interval of length 1, but apart from each other, if we assumed that any element of one set attracts any element of the other set with a force proportional to the product of the masses and inversely proportional to the square of the distance. In other words, the resultant of these forces can again be computed approximately by considering a number N^2 of such pairs. This resultant will, of course, be a function of the distance d between the centers of the two systems. As should be obvious, this resultant varies approximately as 1 divided by d^2 for large d but grows to infinity more rapidly than that when d is comparable to 1 and when the two systems overlap.

A detailed account of these experiments will be presented separately, together with the account of computations mentioned in Part I.

The motivation for a calculation of quantities like self-energy is, of course, that it is clear a priori that the values of these will be less than the values for corresponding Euclidean continuum models and may have finite values when these latter diverge. It is proposed to calculate in the near future a mass-splitting process more general than the one outlined above; *viz.*, the probabilities for splitting which we have assumed to be constant from generation to generation and from point to point could be assumed to depend on the existence of another object of the above sort in proximity to our given system. Indeed, if we assume that a function $V(x)$ is given and the probability of multiplication is proportional to this function, we obtain a distribution u as a result of our splitting process which will obey an equation similar to the Schrödinger equation

$$\Delta u + (E - V)u = 0.$$

In addition, the process, as it stands, has the unsatisfactory feature that it still leaves mathematical points with finite masses. If one wanted to insist on all mathematical points having masses zero, the following iteration of our procedure should be considered: let ω correspond to the first passage to the limit of our process as defined above. We shall now iterate it in this fashion. Each of the points which remained with finite mass will be again subject to a splitting procedure, say, with the same probabilities, into two equal masses but this time located on the opposite ends of an interval shorter than the one in the first process by a fixed ratio R . For example, let $R = \frac{1}{100}$; then if it happened that the first mass point located at $\frac{1}{2}$ has not split during the first process, it will have a probability p_2 of splitting into two masses equal to $\frac{1}{2}$, but located at positions $\frac{1}{2} - \frac{1}{100} \cdot \frac{1}{4}$ and at $\frac{1}{2} + \frac{1}{100} \cdot \frac{1}{4}$. Again, for example, if the point located at $\frac{3}{4}$ has not split in the first process, let it have probability p_2 of splitting into two masses each equal to $\frac{1}{4}$, and located at $\frac{3}{4} - \frac{1}{100} \cdot \frac{1}{8}$, and $\frac{3}{4} + \frac{1}{100} \cdot \frac{1}{8}$. If this continued, we would have a second

limiting set 2ω . If at the end of this procedure some points still have finite masses, we continue, assuming now another still smaller ratio R for the splitting distance and repeat this splitting. It is obvious that, with probability 1, a sequence of these processes will lead to a Cantor set of points so that at ordinal ω^2 all points will have mass zero, and we shall obtain a distribution of density without finite values at any point.

Another generalization seems strongly suggested. In the processes described, the objects which are obtained by our branching are still algebraically of the nature of real numbers. It would seem that this is too special. The formalism of the new quantum theory would suggest considering each of the "manufactured" points as having spinlike properties; *i.e.*, there should be several kinds $x_1, x_2, \dots, x_n$ of these points. This is, in fact, already the case for the branching process as we described it in three dimensions.

The aim in presenting this collection of problems about distributions of infinite sets of points endowed with masses and exerting "forces" on each other is twofold. In the first place, the modern theories of physics could very well use (in addition to their present point of view, oscillating as it does between the field and the discrete-particle description) something that combines from the beginning the two aspects and is not necessarily restricted to Euclidean topology. In the second place, it appears possible that a geometrization of physics, if it ever takes place, will not be effected through merely generalizing the Euclidean character of the distance in the large, which is the content of Riemannian geometry. Since physical phenomena change radically in nature in the small, it is clear that a differentiable metric will never bring out this feature. A more radical change, perhaps even that of local topology, will be necessary. Finally, the recent evidence for frequent transmutation in nature of what at each stage of the historical development was considered an elementary particle and the multiplicative character of phenomena in the small may make it appear amusing, at least for a mathematician, to consider models of this sort.

BIBLIOGRAPHY *

1. C. J. Everett and S. Ulam, *Multiplicative systems in several variables*, I, AECD-2164 (1948); II, AECD-2165 (1948); III AECD-2532 (1948), Technical Information Branch, Oak Ridge, Tenn.
2. ——— and ———, *Multiplicative systems*, I, Proc. Nat. Acad. Sci. U.S.A. vol. 34 (1948) pp. 403-405.

LOS ALAMOS SCIENTIFIC LABORATORY,
LOS ALAMOS, N. MEX.

QUANTUM THEORY AND THE FOUNDATIONS OF PROBABILITY

BY

B. O. KOOPMAN

1. Introduction. Ever since the advent of modern quantum mechanics in the late 1920's, the idea has been prevalent that the classical laws of probability cease, in some sense, to be valid in the new theory. More or less explicit statements to this effect have been made in large number and by many of the most eminent workers in the new physics [1]. Some authors have even gone farther and stated that the formal structure of logic must be altered to conform to the terms of reference of quantum physics [2].

Such a thesis is surprising, to say the least, to anyone holding more or less conventional views regarding the positions of logic, probability, and experimental science: many of us have been apt—perhaps too naively—to assume that experiments can lead to conclusions only when worked up by means of logic and probability, whose laws seem to be on a different level from those of physical science.

The primary object of this presentation is to show that the thesis in question is entirely without validity and is the product of a confused view of the laws of probability. The situation can be straightened out at a very elementary level: all that is needed is to make quite clear and explicit the concept of *event*. It will not be necessary either to adopt any particular position regarding the controversial matters at the foundations of probability or to commit oneself at all deeply on the level of physical laws.

2. Young's experiment. To put the issue in precise terms, let us consider one of the most-used arguments purporting to show that quantum physics is inconsistent with classical probability. It starts with the interference experiment of elementary physics, used historically by Young to establish the wave character of light. In this experiment, light emitted from the source S and collimated by passing through the slit C can reach any point x on a screen by passing through one or another or both of the slits A, B . When A or B is closed, there are no interference fringes on the screen; but when they are both open, such fringes are present, meaning that the intensity of light at a point x is not the sum of its intensities passing through A and B .

The account of this given by quantum physics is in terms of the statistics of photons. When such a corpuscule of light leaves C , it may reach x in any of the three cases: (i) with A open and B closed; (ii) with A closed and B open; (iii) with both open. And the intensity of light at x in each of these situations is proportional to the corresponding *probability* of the photon's reaching x . If the probabilities are denoted by $p_A, p_B, p_{A,B}$, respectively, we obviously do not have $p_{A,B} = p_A + p_B$, because of the interference. This is an alleged experimental violation of the law of total probability, according to the follow-

ing reasoning: "The event of the photon's reaching x when A and B are both open must occur in one of the following mutually exclusive ways: either by its passing through A (probability p_A) or through B (probability p_B); therefore the total probability should be $p_A + p_B$ and not the observed value, $p_{A,B}$."

It is easy to put one's finger on the fallacy contained in the above argument: the italicized words in the sentence "the photon going from C to x *passes through A*," while meaningful in classical mechanics, are without meaning in quantum physics. This fact is a special case of a general principle of the latter discipline, according to which the only meaning that can be ascribed to statements concerning the position, path, momentum, etc., of an elementary particle is in terms of the outcome of *experimental observations* concerning such position, path, momentum, etc. In the present instance, to say that the photon "passes through A " in case (iii) would imply that the passage through A

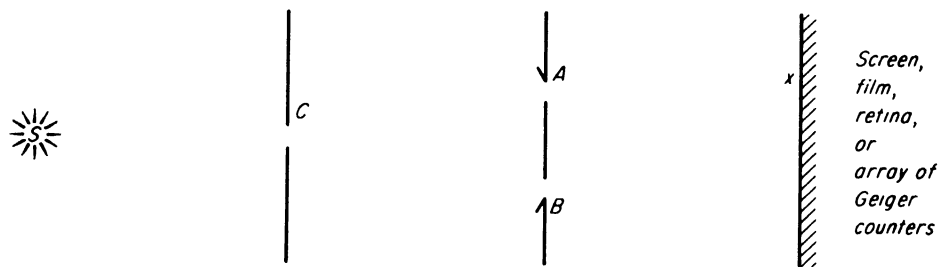


FIG. 1. Young's interference experiment.

could in principle be verified directly by the use of an appropriate mechanism of observation at A . But such information concerning the photon would be given only if it interacted with this mechanism. It is well known that such an interaction would change the wave function so that the probability $p_{A,B}$ would no longer apply [3]. The only correct characterization of p_A in the framework of quantum mechanics is as the probability not of the photon's reaching x passing through A but of its reaching x "by the agency of A alone being open." Similarly $p_{A,B}$ must be described as the probability of reaching x by the agency of both A and B being open.¹ Thus p_A and p_B are not the probabilities of two different outcomes of the same trial but of the same outcome of different types of trials: the principle of total probability does not even enter.

3. The laws of probability. It is safe to say that all schools of probability, however they may differ on questions of the definition of probability and the source of our knowledge of its laws, are in agreement on two points: (1) that it applies to some class of entities called "events" and (2) that it obeys the following three laws, in which α , β , γ , etc., denote events, O an impossible one, I an event certain to occur, while product and sum are to be understood in the logical sense: $\alpha\beta$ means the occurrence of both α and β , $\alpha + \beta$, that of at least

¹ Dirac suggests picturing this situation by thinking of one photon passing simultaneously through A and B in (iii). Of course no such picture is necessary to understand the facts.

one of them. Then we have for the probabilities $p(\alpha)$ of α and the conditional probability² $p(\alpha|\beta)$ of α given β , etc.:

- I. $0 \leq p(\alpha) \leq 1$; $p(O) = 0$; $p(I) = 1$.
- II. If $\alpha\beta = O$, then $p(\alpha + \beta) = p(\alpha) + p(\beta)$.
- III. $p(\alpha\beta) = p(\alpha)p(\beta|\alpha)$.

Here II is called the law of *total* probability, and III is the law of *compound* probability.³ From these three all the theorems of probability involving finite combinations follow. They need supplementing by continuitylike assumptions to obtain the more advanced part of the theory.

It should be clear that before any question of the validity or application of these laws can be asked, a precise meaning must be given to the concept of *event*. It would appear, indeed, from what we have already seen, that certain arguments for the invalidity of classical probability in quantum theory can be refuted by analyzing the nature of the things whose probability is being considered.

4. Events. A thorough examination of all the concrete applications of the theory of probability shows that the concept of *event* can always be interpreted as a *statement concerning the state of a material system on a specified occasion*. It is essential, furthermore, that the statement be *meaningful* according to the simplest interpretation of Bridgman's operational standards: in principle, capable of verification (true or false) by an observation. Examples: statements concerning the outcomes of games of chance on particular occasions; weather predictions; prediction of the outcome of a biological experiment; statements concerning the result of an observation of a quantum system. Counterexamples: generalities, such as the assertion that the law of conservation of energy is true; such unverifiable statements as concern the state of consciousness after death or whether the photon in (iii) "really" goes through *A* or *B*. "Events," in the present sense have been appropriately called "experimental propositions" [2]. They form the primitive building stones of the logical structure of experimental science.⁴

² Clearly $p(\alpha|\beta)$ is meaningless when $\beta = O$, that is, is impossible. Strictly, our equations should be generalized by being restated in terms of conditional probabilities. Then $p(I) = 1$ is replaced by *if β implies α , then $p(\alpha|\beta) = 1$* .

³ There is a frequent practice of regarding III as merely the definition of conditional probability $p(\alpha|\beta)$ (just as some authors regard Newton's second law as merely the definition of force). From the point of view of applied science this practice is objectionable since the concept defined (conditional probability, force, etc.) is already meaningful in the terms of the application. And from the point of view of a purely mathematical development it is equally objectionable, as ruling out $p(\alpha|\beta)$ when $p(\beta) = 0$, an important case. (Cf. the above definition of force when there is no motion: statics.)

⁴ The development of probability in terms of "sample space" X , in which events are replaced by subsets E of X , is reducible to the above by introducing $\alpha = "x \in E,"$ where x is the element of X representing the state of the system. Conversely, starting with events as experimental propositions, one can represent them as subsets E of X , by M. H. Stone's

When we take the next step of combining different events by the processes of logic, an issue is raised which proves to be fundamental: implicit in any event α is the nature of the experimental observation needed for its verification; we may denote it by O_α . If, therefore, we consider two events, α and β , we also have two types of observation O_α and O_β ; if these can be combined without alteration into a single composite observational process $O_{\alpha,\beta}$, which, when carried out, tells us whether α and β are true or not, then it is meaningful to speak of $\alpha\beta$ ("both are true") and $\alpha + \beta$ ("at least one is true"). But if the performance of O_α necessarily makes that of O_β impossible, then the expressions $\alpha\beta$ and $\alpha + \beta$ cease to have their usual meaning as logical combinations: under such conditions, we say that events α and β are *incompatible*, and similarly for O_α and O_β .

5. Compatibility and incompatibility. Examples of *compatible* events are easily given. If we toss a coin and throw a die, and α is the event of heads, β of ace, we can talk about $\alpha\beta$ and $\alpha + \beta$, since the combined toss and throw is a valid composite experiment $O_{\alpha,\beta}$, one aspect of which is O_α and the other, O_β . Similarly, if a single throw of a die is made and α is the statement that ace occurs, while β , that the number is even, there is evident compatibility, with $O_{\alpha,\beta} = O_\alpha = O_\beta$. Finally, as shown in quantum mechanics, if α and β are two statements about the position or two statements about the momentum of a particle, they are compatible. This is also true if α is a statement regarding one component of the position vector while β concerns an orthogonal component of the momentum vector.

The standard example of *incompatibility* is in quantum mechanics, where α and β are statements concerning the same components of position and momentum. But it is possible to give a much more obvious example. Suppose that, as a result of a gamma-ray mutation, an altogether different and unique creature were produced from rat parents, and suppose that its resistance to lethal doses of certain poisons A and B is of interest. If α is a statement concerning the number of days before it dies when injected solely with A , while β is a similar statement regarding B , then α and β are clearly incompatible, for the simple reason that you can kill your creature with A alone or else with B alone, but not "with both alone." Therefore no logical combinations such as $\alpha\beta$, $\alpha + \beta$ have any meaning: the situation assumed in the definition of α is inconsistent with that needed to make β meaningful. Nor is the situation improved by restating α and β as implications (α being interpreted as that the A experiment *implies* that the creature survives so many days, etc.), simply because, in order to become experimental propositions, a single experiment ("trial") must be capable of telling whether α is true ("succeeds") or false; and similarly for β —and we are back where we were.

representation theorem. This has the advantage of making the results of measure theory directly applicable to probability; but it has the drawback of replacing simple ideas by complicated constructs.

In the quantum example, which is less drastic than the biological one, there is a rather intriguing possibility. If α and β are the statements regarding x and p_x of the particle, we might denote by $O_\alpha O_\beta$ the combined operation of performing first O_α and then immediately O_β ; similarly $O_\beta O_\alpha =$ first O_β then O_α . Then $\alpha\beta$ could be used to denote the statement: "When $O_\alpha O_\beta$ is done, both α and β turn out to be true"; similarly, $\beta\alpha$ is defined on the basis of $O_\beta O_\alpha$. Now it is a fundamental principle of quantum mechanics (essentially, Heisenberg's) that $O_\alpha O_\beta$ gives different results in general than $O_\beta O_\alpha$, so that, in terms of truth values, $\alpha\beta \neq \beta\alpha$. This is not a violation of a law of logic; it is simply due to the fact that the symbol α has a different meaning (when its *complete* definition is envisaged) in each of the three expressions α , $\alpha\beta$, $\beta\alpha$; similarly for β .

Since the laws of probability II and III presuppose the definitions of the logical sum $\alpha + \beta$ and product $\alpha\beta$, their scope, in the nature of things, is restricted to cases where such definitions are possible, *viz.*, to compatibility. Moreover, the necessity of such a restriction was implicit in classical science, well before quantum physics. Of course no one would have blundered so badly as to have tried to apply II or III to the α , β defined in the biological example; but until Heisenberg emphasized that $O_\alpha O_\beta \neq O_\beta O_\alpha$, it did not seem like such a blunder to apply them in quantum mechanics. It is unfortunate that the "blame" was not placed on the procedural nature of O_α , O_β , $O_\alpha O_\beta$, etc., but on II and III: an evidence that sometimes physicists cling more tenaciously to their physical pictures than to general laws of thought.

6. Reformulation of probability. We return, in the light of these facts, to a more explicit statement of I, II, III. It must be acknowledged that every pair taken from the three experimental propositions (events) α , β , γ might be compatible without all three being so: just because $O_\alpha O_\beta = O_\beta O_\alpha = O_{\alpha\beta}$, and similarly for the two other pairs, it does not follow in principle that $O_\alpha O_\beta O_\gamma = \dots = O_{\alpha\beta\gamma}$. In other words, pairwise compatibility is not necessarily transitive. Consequently we are obliged to employ the notion of compatibility of *sets*, possibly infinite, of events, $S = \{\alpha, \beta, \dots\}$ meaning that a single observation O_S exists capable of telling which of its elements α , β , $\dots$ are true and which are false, and of which O_α , O_β , $\dots$ are subspects. Clearly if S is a compatible set, so is the enlarged set obtained by adding to S all logical combinations of its elements, such as $\alpha\beta$, $\alpha + \beta$, $\bar{\alpha}$ ("not α "), etc. (this could also be done with infinite sequences of elements of S). There is, accordingly, no restriction in assuming that the compatible sets are always *Boolean algebras* (and σ -algebras), *i.e.*, that they already contain all such combinations of their elements.

We may now formulate the laws of elementary probability as follows:

Let S be a compatible Boolean algebra (or σ -algebra) of events (experimental propositions). Then, for any α , β , γ , δ of S we have⁵

⁵ For a more general formulation, the following qualification should be added: " $\dots$ for which the indicated probabilities are defined."

- I. If $\gamma \neq O$, $0 \leq p(\alpha|\gamma) \leq 1$; and if in addition $\gamma \rightarrow \alpha$, then $p(\alpha|\gamma) = 1$.
- II. If $\gamma \neq O$ but $\alpha\beta\gamma = O$, then $p(\alpha + \beta|\gamma) = p(\alpha|\gamma) + p(\beta|\gamma)$.
- III. If $\alpha\gamma \neq O$, then $p(\alpha\beta|\gamma) = p(\alpha|\gamma)p(\beta|\alpha\gamma)$.

This formulation is as valid in quantum physics as in any other branch of science. It leads into certain fundamental concepts of the former: if S consists of all statements possible regarding the values of the quantum variables (observables) $X, Y, \dots, W$ of a certain system, then these latter are called *compatible variables*, the operation O_S fixing the values of all of them and being called a "maximal observation." For this to be the case, it is necessary and sufficient that $X, Y, \dots, W$ commute in all their combinations [4].

As noted in Sec. 3, these three laws have to be supplemented by certain laws of limits in order that probability have the properties of a totally additive measure. This involves nothing beyond our present compatibility assumptions, since our S is a σ -ring.

BIBLIOGRAPHY

1. References would be too numerous to give here, although we may cite as typical a recent expository note by R. P. Feynman, *The concept of probability in quantum mechanics* in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, University of California Press, Berkeley, Calif., 1951.
2. G. Birkhoff and J. von Neumann, *The logic of quantum mechanics*, Ann. of Math. vol. 37 (1936) pp. 823-843.
3. Cf. P. A. M. Dirac, *The principles of quantum mechanics*, Clarendon Press, Oxford, 1935, chap. I, sec. 2.
4. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932, chap. III.

COLUMBIA UNIVERSITY,
NEW YORK, N.Y.

INDEX

- Ambarzumian, V. A., 51
 Antoine, L., 92

 Bachelier-Wiener function, 2
 Batchelor, G. K., 47, 50
Berkeley Symposium, Third, 81n.
 Berlin, T. H., 76, 85
 Birkhoff, G., 102
 Blanc-Lapierre, A., 80
 Bok, B. J., 66
 Boltzmann, L., 81
 Boltzmann equation, 81, 84
 Boltzmann property, 84
 Boolean algebras, 101
 Boundary, adjoint, 39
 entrance, 39
 exit, 39
 natural, 39
 regular, 39
 Boundary points, 37
 Brelot, M., 32, 33
 Brownian function, 2
 Bruns-Charlier expansion, 47, 48

 Chandrasekhar, S., 66
 Characteristic functional, 42
 Choquet, G., 32, 33
 Clausius, R. J. E., 82
 Compatible events, 100
 Compatible sets, 101
 Compatible variables, 102

 Darling, D. A., 80
 Densities, asymptotic, 89
 contracted, 84
 Derman, C., 39
 Diffusion theory, 40
 Dirac, P. A. M., 98n., 102
 Dirichlet problem, 21ff.
 Doob, J. L., 32, 33

 Ehrenfest, P., 81
 Ehrenfest, T., 81

 Equation, Φ -, 43ff.
 Equilibrium problems, 73ff.
 Equivalent operators, 38
 Escape route, 37
 Events, 98ff.
 incompatible, 100
 Everett, C. J., 95
 Experimental propositions, 99
 Extremal elements, 36

 Feynman, R. P., 102
 Fortet, R., 80
 Fourier-Wiener series, 10
 Free energy per constituent, 74
 Function, individual, 3n.
 partition, 74ff.
 random, 3n.
 regular, 22
 resolutive, 24
 subregular, 22
 superregular, 22
 Functional derivatives, 42

 Gibbs, J. W., 81
 Gordeladse, S. G., 65, 66
 Green, M. S., 85

 Harris, T. E., 39
 Heisenberg, W., 48
 Heisenberg principle, 101
 Hierarchy, 90
 Hopf, E., 50

 Imperfect gas, problem of, 77
 Intrinsic first boundary-value problem, 30
 Ising models, 74

 Kac, M., 75, 85
 Kahn, B., 78, 85

- Kaufman, B., 75
 Khintchine, A., 4
 Kolmogoroff, A., 39, 40, 47, 48
 Kolmogoroff differential equations, 38, 39
 Krein, H., 35, 36

 Lee, T. D., 78
 Liouville's differential equation, 48, 49
 Loève, M., 13
 Loitsiansky's invariant, 72
 Loschmitt paradox, 81

 Markarian, B. E., 66
 Martin, M. H., 38
 Martingale, 27
 Master equation, 83-85
 Mathews, P. M., 66
 Maximal observation, 102
 Maxwell gas, 83
 Mayer, J. R. von, 78
 Mil'man, D., 35, 36.,
 Montroll, E. W., 85
 Münch, G., 66

 Navier-Stokes equations, 41, 49, 67ff.
 Neumann, J. von, 102
 Nonequilibrium problems, 81ff.

 Onsager, L., 75, 76

 Paley-Wiener formula, 10
 Pasta, J., 90
 Phase space, 41
 Φ -equation, 43ff.
 Point, irregular, 24
 regular, 24
 Positive cone, 35
 Probability, laws of, 98
 sojourn, 36
 Proudman, I., 50, 67n., 68, 70
 PWB method (Perron-Wiener-Brelot
 method), 23
 generalized, 29
 stochastic, 29

 Quantum mechanics, 97ff.

 Ramakrishnan, A., 66
 Regular average, 22
 Reid, W. H., 50
 Robbins, H., 39
 Rusakov, G. E., 66

 Sample space, 99n.
 Schrödinger equation, 94
 Semimartingale, 27
 Set, regular, 21
 sojourn, 36
 Siegert, A. J. F., 80
 Specific heat, 75
 Spectrum tensor, 67ff.
 Spherical models, 76
 Spitzer, L., 66
 Steepest descent, method of, 76
 Stochastic boundary function, 30
 Stone, M. H., 99n.
 Stosszahlansatz, 81-83, 85
 Stromgren, B., 66
 Support of a measure function, 22

 Tautz, G., 33
 Titt, E. W., 50
 Trajectories to the boundary, 26
 Trajectory, ω , 27

u-space, 41
 Uhlenbeck, G. E., 78
 Ulam, S., 95
 Ursell, F., 78

 van Hove, L., 78, 85
 von Neumann, J., 102

 Ward, J. C., 75, 85
 Wave-number space, 43
 Wiener process, 38

 Yang, C. N., 78
 Young, T., 97
 Young's experiment, 97

 Zermelo paradox, 81

