FUNCTIONAL INTEGRATION AND PARTIAL DIFFERENTIAL EQUATIONS

BY

MARK FREIDLIN

ANNALS OF MATHEMATICS STUDIES
PRINCETON UNIVERSITY PRESS

Annals of Mathematics Studies

Number 109

FUNCTIONAL INTEGRATION AND PARTIAL DIFFERENTIAL EQUATIONS

BY

MARK FREIDLIN

PRINCETON UNIVERSITY PRESS

PRINCETON, NEW JERSEY

1985

The Annals of Mathematics Studies are edited by
William Browder, Robert P. Langlands, John Milnor, and Elias M. Stein
Corresponding editors:
Stefan Hildebrandt, H. Blaine Lawson, Louis Nirenberg, and David Vogan

Clothbound editions of Princeton University Press books are printed on acid-free paper, and binding materials are chosen for strength and durability. Paperbacks, while satisfactory for personal collections, are not usually suitable for library rebinding

ISBN 0–691–08354–1 (cloth)
ISBN 0–691–08362–2 (paper)

Library of Congress Cataloging in Publication data will be found on the last printed page of this book

Printed in the United States of America
by Princeton University Press, 41 William Street
Princeton, New Jersey

CONTENTS

PREFACE

With every second-order elliptic differential operator L, one can associate a family of probability measures in the space of continuous functions on the half-line. This family of measures forms the Markov process corresponding to the operator L. If one knows some properties of the operator L, it is possible to draw conclusions about properties of the Markov process. And conversely, studying the Markov process one can obtain new information concerning the differential operator.

This book considers problems arising in the theory of differential equations. Markov processes (or the corresponding families of measures in the space of continuous functions) are here only a tool for examining differential equations. As a rule, the necessary results from the theory of Markov processes are given without proof in this book. We restrict ourselves to commentaries clarifying the meaning of these results. There are already excellent books where these results are set forth in detail, and we give references to these works.

The probabilistic approach makes many problems in the theory of differential equations very transparent; it enables one to carry out exact proofs and discover new effects. It is the latter—the possibility of seeing new effects—which seems to us the most significant merit of the probabilistic approach.

This book is intended not only for mathematicians specializing in the theory of differential equations or in probability theory but also for specialists in asymptotic methods and functional analysis. The book may also be of interest to physicists using functional integration in their research.

The two years I have spent writing this book were very hard, I would even say desperate, for me and my family. And I am glad to be able to thank my colleagues for their support. I have been happy to see convincing evidence of the high moral standards of many colleagues. I especially wish to express my gratitude to E. B. Dynkin for his constant attention and concern about all our problems.

Finally, I must say that this book would never be brought into the world without the enormous labor of my wife, Valeria Freidlin, in her editing, translating and retyping the manuscript. I feel even awkward about thanking her for this labor; in essence, she was my co-author.

MARK FREIDLIN

Functional Integration and Partial Differential Equations

INTRODUCTION

It was known long ago that there is a close relation between the theory of second-order differential equations and Markov processes with continuous trajectories. As far back as 1931, the parabolic equations for transition probabilities were written down in the article of Kolmogorov [1]. Still earlier, these equations on the theory of Brownian motion appeared in physics literature (Einstein [1]). It was also established that the mean values of some functionals of the trajectories of diffusion processes (as functions of an initial point) are the solutions of boundary value problems for the corresponding elliptic differential equations.

For a long time the connection between Markov processes and differential equations was used mainly in one direction: from the properties of the solutions of differential equations, some or other conclusions on Markov processes were made. Meanwhile, probabilistic arguments in problems of the theory of differential equations played at best the role of leading reasoning. This may be explained by lack of direct probabilistic methods for studying diffusion processes. Even the construction of such a process with given characteristics was carried out with the help of the existence theorems for the corresponding parabolic equations.

For the last quarter of a century the situation has changed in an essential way. The rapid development of direct probabilistic methods for examining Markov processes allowed one to construct and study them without turning to partial differential equations. Conversely, the construction and analysis of the trajectories of the corresponding diffusion process via direct probabilistic methods, enabled the solutions of differential equations to be constructed and the properties of these solutions to be examined.

It is not for the first time that such a situation arises in the theory of differential equations. For example, recall the mutual relations between differential equations and the calculus of variations. Originally, the differential equations served as the means of seeking solutions of extremal problems. With the development of the direct methods in the calculus of variations, the possibility appeared of constructing and studying the solutions of differential equations as the extremals of the corresponding functionals. Similar mutual relations have now been established between the theory of differential equations and that of diffusion processes.

Speaking somewhat inaccurately, one can say that, in the theory of second-order parabolic and elliptic differential equations, the trajectories of diffusion processes play the same part as characteristics do for first-order equations. Just as the theory of characteristics makes first-order equations geometrically descriptive, the probabilistic considerations make transparent many problems arising in the theory of second-order elliptic and parabolic equations.

Sometimes the probabilistic methods play the role of a tool for deriving delicate analytical results. Sometimes they are a basis for the extension of some analytical theory. However, in my view, the greatest value of such an approach consists in its visualization which turns this approach into an especially helpful instrument for discovering new effects, for a deeper qualitative understanding of the classical objects of mathematical analysis.

Among the tools of the direct probabilistic research of diffusion processes, one should, first of all, mention stochastic differential equations. The theory of such equations, originating in the works of Berstein, was basically founded by Ito and (independently) by Gihman, and then has been developed by a number of mathematicians. The stochastic integral introduced by Ito, Ito's formula, and the generalizations of these notions play the central part in the whole theory. The present state of the theory of stochastic differential equations is described in the monograph of Ikeda and Watanabe [2]; references to the original works can be found there too.

As another important factor permitting the direct study of diffusion processes, one should mention the convenient general concept of Markov process and Markov family introduced by Dynkin [1], [3] as well as the detailed analysis of the strong Markov property. The wide use of the theory of one-parameter semi-groups due to Feller is also worthwhile noting.

The theory of martingales serves as a highly suitable instrument for examining Markov processes (see Doob [1], Delacherie and Meyer [1]).

The transformations of Markov processes, in particular, those involving an absolutely continuous change of measure in the space of trajectories, are also very useful tools which enable one, in a transparent and explicit fashion, to understand the effects of potential terms and first order terms. This leads to an understanding of the affects of these terms on the behavior of the solution of the differential equation.

The last ten to fifteen years have seen a development of limit theorems for random processes—central limit theorem type results as well as theorems on the asymptotics of probabilities of large deviations. In particular, the counterpart of the asymptotic Laplace method for functional integrals pertains to the results of that kind. These results proved to be highly useful in a great number of problems in differential equations which have waited long to be solved.

The application of the probabilistic methods for examining differential equations is usually based on the representation of the solution of these equations as the mean value of some functional of the trajectories of a proper diffusion process. The mean value of a functional of the trajectories of a random process may be written down as the integral of the corresponding functional on the space of functions with respect to the measure in this space induced by the random process. This is why such representations of solutions are sometimes called the representations in the form of functional integrals.

The construction of the diffusion process corresponding to the differential operator

$$L = \frac{1}{2} \sum_{i,j=1}^{r} a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{r} f^i(x) \frac{\partial}{\partial x^i} \tag{1}$$

with the non-negative definite matrix $(a^{ij}(x))$, is carried out with the help of stochastic differential equations. The Wiener process W_t, the simplest of the non-trivial Markov processes serves as a starting point.

By a Wiener process (one-dimensional), we mean a random process $W_t = W_t(\omega)$, $t \geq 0$, having independent increments and continuous trajectories (with probability 1), and for which $EW_t = 0$, $EW_t^2 = t$ (E being the mathematical expectation sign).

It is established that such a process does exist and its finite-dimensional distributions are Gaussian. In particular, for every $t > 0$, the random variable $W_t(\omega)$ has the density function $(2\pi t)^{-1/2} \exp\left\{-\frac{x^2}{2t}\right\}$, $-\infty < x < \infty$. This process is connected, in the closest way, with the operator $\frac{1}{2}\frac{d^2}{dx^2}$ and with the simplest heat conduction equation. For instance, the solution of the Cauchy problem

$$\frac{\partial u(t,x)}{\partial t} = \frac{1}{2} \frac{\partial^2 u(t,x)}{\partial x^2}, \quad u(0,x) = g(x), \tag{2}$$

may be represented in the form

$$u(t,x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} g(x+y) e^{-\frac{y^2}{2t}} dy = Eg(x+W_t).$$

This assertion is checked by direct substitution into equation (2). Just as any random process, the Wiener process induces a measure in the space of functions. In the present case, it is a measure in the space of continuous functions on the half-line $t \geq 0$ with the values in R^1. This measure is referred to as the Wiener measure. It plays the principal role in all the questions to be considered in this book. The first construction of this measure was published by Wiener in 1923 [1]. Later on the Wiener process and the Wiener measure have been studied in detail.

An ordered collection of r independent Wiener processes $(W_t^1, \cdots, W_t^r) = W_t$ is termed an r-dimensional Wiener process. Such a process is connected with the Laplace operator in R^r. What process corresponds to the operator L in (1)? Let us assume for a moment that the coefficients of the operator are constant: $a^{ij}(x) = a^{ij}$, $b^i(x) = b^i$. Denote by $\sigma = (\sigma_j^i)$ a matrix such that $\sigma\sigma^* = (a^{ij})$ and consider the family of random processes

$$X_t^x = \sigma W_t + bt + x,\ x \in R^r, b = (b^1, \cdots, b^r), t \geq 0 . \tag{3}$$

It is not difficult to find the distribution function of the Gaussian process X_t^x and then to check that $u(t,x) = E\, g(X_t^x)$ is the solution of the Cauchy problem

$$\frac{\partial u}{\partial t} = L\, u(t,x), \quad u(0,x) = g(x) , \tag{4}$$

for any continuous bounded function $g(x)$. Therefore, the random process (3) is associated with the operator L with constant coefficients.

It is natural to expect that, in the vicinity of every point $x \in R^r$, the process corresponding to the operator L with variable (sufficiently smooth) coefficients, must behave just as the process corresponding to the operator with the constant coefficients frozen at this point x. On the basis of this reasoning, for the family of the processes X_t^x corresponding to the operator L with variable coefficients, we obtain the differential equation

$$d\, X_t^x = \sigma(X_t^x)\, d\, W_t + b(X_t^x)\, dt,\ X_0^x = x , \tag{5}$$

where the matrix $\sigma(x)$ is such that $\sigma(x)\sigma^*(x) = (a^{ij}(x))$, $b(x) = (b^1(x), \cdots, b^r(x))$.

If the trajectories of the Wiener process were differentiable functions or at least had bounded variation, then equation (5) could be treated within the framework of the usual theory of ordinary differential equations. But, with probability 1, the trajectories of the Wiener process have infinite variation on every non-zero time interval. Therefore, equation (5) should

be given a meaning. Ito's construction is most convenient for this. This construction is given in the beginning of Chapter I.

One can demonstrate that, under mild assumptions on the coefficients, equation (5) has a unique solution X_t^x. The random functions X_t^x, $x \in R^r$, together with the corresponding probability measure, form a Markov family connected with the operator L. A solution of Cauchy's problem (4) may be written in the form $u(t,x) = E\, g(X_t^x)$.

A solution of Dirichlet's problem for the operator L may also be represented in the form of the mathematical expectation of some functional of the process X_t^x. For example, if D is a bounded domain in R^r with a smooth boundary ∂D and the operator L does not degenerate for $x \in D \cup \partial D$, then the solution of the Dirichlet problem

$$L\, u(x) = 0, x \in D;\ u(x)|_{\partial D} = \psi(x), \tag{6}$$

where $\psi(x)$ is a continuous function on ∂D, may be written as follows

$$u(x) = E\, \psi(X_{\tau^x}^x). \tag{7}$$

Here $\tau^x = \inf\{t : X_t^x \notin D\}$ is the first exit time of the process X_t^x from the domain D.

If the term with a potential v is added to the operator L, then the solutions of various problems for the operator $L + v$ may also be represented in terms of the trajectories of the process X. For example, the solution of the Cauchy problem

$$\frac{\partial u(t,x)}{\partial t} = L\, u(t,x) + v(x)\, u(t,x),\ u(0,x) = g(x) \tag{8}$$

is given by the Feynman-Kac formula

$$u(t,x) = E\, g(X_t^x) \exp\left\{\int_0^t v(X_s^x)\, ds\right\}. \tag{9}$$

Notice that equation (5) may be looked upon as the mapping of the space $C_{0,\infty}(R^r)$ of continuous functions on the half-line with values in

R^r, into itself: $I: W_\cdot \to X^x_\cdot$. This mapping is defined a.e. with respect to the Wiener measure in $C_{0,\infty}(R^r)$. The value of X^x_t at time t is defined as a functional of the Wiener trajectory in the interval $[0,t]$ which depends on x as a parameter: $X^x_t = I^x(W_s, 0 \le s \le t)(t)$. This mapping allows formulae (7) and (9) to be rewritten in the form of integrals with respect to the Wiener measure.

Chapter I describes the construction and properties of the Wiener process. The necessary information on stochastic integrals, stochastic differential equations, and Markov processes and their transformations is given here. Some limit theorems for random processes are included as well. In particular, we provide the definition and properties of the action functional related to the Laplace type asymptotics for functional integrals. In short, Chapter I introduces those notions and methods which are necessary for the direct probabilistic analysis of processes (measures in the space of functions) connected with differential operators.

Today there are a number of monographs presenting these results in detail. Also, in this book, random processes are a tool rather than an object of research themselves. For this reason the results of Chapter I are, as a rule, cited without proof. We restrict ourselves to short comments and references.

In Chapter II, the formulas representing the solutions of differential equations in the form of functional integrals (in the form of the mean values of the functionals of the trajectories of the corresponding processes) are studied. Besides formulas (7) and (9), this chapter gives representations for the solutions of the second boundary value problem as well as some other problems. The behavior of random processes as $t \to \infty$ is a traditional subject of probability theory. This is closely related to problems concerning the stabilization, as $t \to \infty$, of the solutions of Cauchy's problem as well as of mixed problems. It is also related to the statement of boundary valued problems in unbounded domains. These questions are also considered in Chapter II. Speaking somewhat inaccurately, one can say that a solution of the first boundary

value problem is unique if and only if the trajectories of the corresponding diffusion process leave the domain D with probability 1. Hence the question of the correct statement of the problem in an unbounded domain is closely related to the behavior of the trajectories as $t \to \infty$. If, with positive probability, the trajectories go to infinity without hitting the boundary of the domain, then supplementary conditions at infinity are required to single out the unique solution. For example, the Wiener process in R^2 does not run to infinity, and so the solution of the exterior Dirichlet problem for the Laplace operator in R^2 is unique in the class of bounded functions. Meanwhile, the Weiner process in R^r, for $r \geq 3$, goes to infinity with positive probability, and hence, when considering the exterior Dirichlet problem for the Laplace equation in these spaces, one must in addition define the value of the limit of the solution at infinity. In the case of equations of a more general form, "the boundary at infinity" may have a more complicated structure. Everything depends on the final (i.e. as $t \to \infty$) behavior of the trajectories of the corresponding diffusion process.

Probabilistic methods have proved to be greatly effective in examining degenerate elliptic and parabolic equations. Chapter III is devoted to these questions. If the coefficients are Lipschitz continuous, then existence and uniqueness theorems are valid for equation (5) regardless of any degeneration of the diffusion matrix $(a^{ij}(x))$. This enables one to examine the peculiarities of the statement of boundary value problems for degenerate equations. In particular, the behavior of the corresponding process near the boundary points is in exact agreement with where and how the boundary conditions will be taken. After the process corresponding to the operator has been constructed, it is not difficult to prove the existence theorem and to clarify uniqueness conditions. The generalized solution is described in the form of functional integral (7). This allows one to examine its local properties. Under broad assumptions, the generalized solution turns out to be Hölder continuous. In order to ensure Lipschitz continuity or smoothness, one should make some special

assumptions. Chapter III clarifies the conditions under which the generalized solution is smooth and gives an example illustrating the importance of these conditions. Roughly speaking, the smoothness of the generalized solution is due to the relation between the rate of scattering of the trajectories of system (5) starting from close points and the first eigenvalue (generalized) of the boundary value problem. The rate of scattering of the trajectories is defined by a number which is expressed in terms of the Lipschitz constant of the coefficients of equation (5).

The results of Chapter III, besides being interesting on their own, serve as a basis for Chapter IV where elliptic equations with small parameter in higher derivatives are dealt with. The analysis of how the solutions of boundary value problems depend on these parameters reduces to the following two questions: first, to analyzing the dependence of the trajectories of "ordinary" equation (5) on these parameters, and then to examining the dependence of the functional integral on the parameters contained in the integrand. Here the dependence on the parameters may be understood in a rather broad sense. This may be the dependence on the initial point—in this way Chapter III studies the modulus of continuity and the smoothness of the generalized solutions. This may also be the dependence on various parameters involved in the operator $L + v(x)$. Here, for example, belongs the problem on the behavior of the solutions of the equations with fast oscillating coefficients and various versions of the averaging principle.

The fact that equations (5) are not sensitive to degenerations makes the probabilistic approach especially suitable in problems with small parameter in higher derivatives. Consider the Dirichlet problem in a bounded domain D:

$$L^{\varepsilon}u^{\varepsilon}(x) = (L_0 + \varepsilon L_1)\,u^{\varepsilon}(x) = \frac{1}{2}\sum_{i,j=1}^{r} A^{ij}(x)\,\frac{\partial^2 u^{\varepsilon}}{\partial x^i \partial x^j} + \sum_{i=1}^{r} B^{i}(x)\,\frac{\partial u^{\varepsilon}}{\partial x^i} + \tag{10}$$

$$+ \frac{\varepsilon}{2}\sum_{i,j=1}^{r} a^{ij}(x)\,\frac{\partial^2 u^{\varepsilon}}{\partial x^i \partial x^j} = 0,\ x \in D;\ u^{\varepsilon}(x)|_{\partial D} = \psi(x)\ ,$$

where $\psi(x)$ is a continuous function on ∂D. We admit that the small parameter may precede not all the second-order derivatives and thus the operator L_0 also may, generally speaking, involve terms with second-order derivatives.

The random process corresponding to the operator L^ε may be constructed with the help of the stochastic differential equations

$$\begin{aligned} d X_t^{\varepsilon,x} &= \sigma(X_t^{\varepsilon,x})\, d W_t + B(X_t^{\varepsilon,x})\, dt + \\ &+ \sqrt{\varepsilon}\, \tilde{\sigma}(X_t^{\varepsilon,x})\, d \tilde{W}_t, X_0^{\varepsilon,x} = x \in R^r, \varepsilon \geq 0 , \end{aligned} \tag{11}$$

where $\sigma(x)\, \sigma^*(x) = (A^{ij}(x))$, $\tilde{\sigma}(x)\, \tilde{\sigma}^*(x) = (a^{ij}(x))$, and W_t and $\tilde{W}_t$ are independent Wiener processes. For $\varepsilon = 0$, equation (11) defines the random functions $X_t^{0,x}$, $x \in R^r$, $t \geq 0$, corresponding to the operator L_0.

From equation (11), one can easily deduce that

$$\lim_{\varepsilon \downarrow 0} P\{ \sup_{0 \leq t \leq T} |X_t^{\varepsilon,x} - X_t^{0,x}| > \delta \} = 0 \tag{12}$$

for any $T > 0$, $\delta > 0$.

Denote by $\tau^{\varepsilon,x} = \inf\{t : X_t^{\varepsilon,x} \notin D\}$, $\varepsilon > 0$, the first exit time of the process $X_t^{\varepsilon,x}$ from the domain. The behavior of $\tau^{\varepsilon,x}$ as $\varepsilon \downarrow 0$ is an important characteristic of problem (10). If one supposes that, with probability 1, the trajectories of the degenerate process $X_t^{0,x}$, $x \in D$, leave the domain D in a finite time and, moreover, behave in a rather regular way near the boundary, then it is not difficult to conclude from (12) that $\tau^{\varepsilon,x}$ has a finite limit as $\varepsilon \downarrow 0$, $\lim_{\varepsilon \downarrow 0} u^\varepsilon(x) = u_0(x)$ exists, does not depend on the perturbating operator L_1 and is a unique solution of the equation $L_0 u_0(x) = 0$, $x \in D$, with the corresponding boundary conditions. This is the simplest case. If there are no second-order derivatives in the operator L_0, then we have the known result due to Levinson [1].

If $\tau^{\varepsilon,x}$ grows like ε^{-1} or faster (as $\varepsilon \downarrow 0$), then the limit behavior of $u^\varepsilon(x)$ already depends, generally speaking, on perturbations. For

example, §4.3 considers the case when the operator L_0, in a sense, does not help, but does not hinder the trajectories $X_t^{\varepsilon,x}$ from hitting ∂D either. Here, under some extra conditions, hitting the boundary, and thereby $\lim_{\varepsilon \downarrow 0} u^\varepsilon(x)$ are controlled by a certain operator which is obtained from L_1 by means of averaging with respect to a measure specified by the operator L_0. If this averaged operator vanishes, then $\lim_{\varepsilon \downarrow 0} u^\varepsilon(x)$ is defined by the subsequent approximation which is of the central limit theorem nature.

Next, Chapter IV discusses the case when the operator L_0, in a sense, hinders the process $X_t^{\varepsilon,x}$ from leaving D. In these problems, $\tau^{\varepsilon,x}$ grows very fast as $\varepsilon \downarrow 0$, approximately like $\exp\{\text{const.}\, \varepsilon^{-1}\}$. The case is typical when there are no second-order derivatives in L_0 and the field $B(x) = (B^1(x), \cdots, B^r(x))$ is such that its integral curves everywhere cross the boundary ∂D of the domain D from the outside toward the interior. Here the exit from the domain is defined by the large deviations of the process $X_t^{\varepsilon,x}$ from $X_t^{0,x}$, and the result is formulated and established via the action functional.

In the last section of Chapter IV, a problem is treated where the small parameter precedes the terms of first order, but due to the presence of degenerations, these terms become the main ones. This section sets forth results of the averaging principle type and of large deviations type.

The last three chapters are devoted to the analysis of quasi-linear equations. Chapter V goes into the question of the existence "in the large" (that is for all $t \geq 0$) of a continuous solution of Cauchy's problem and of some mixed problems. The results of this chapter are based on transformations of Markov processes leading to an absolutely continuous change of measure. The last section of Chapter V is devoted to the analysis, as $t \to \infty$, of the solutions of one class of quasi-linear systems admitting a simple probabilistic interpretation.

Chapters VI and VII consider various generalizations of the Kolmogorov-Petrovskii-Piskunov equation [1]

$$\frac{\partial u(t,x)}{\partial t} = \frac{D}{2}\frac{\partial^2 u}{\partial x^2} + f(u),\ u(0,x) = \begin{cases} 1, & x \le 0 \\ 0, & x > 0 \end{cases}. \qquad (13)$$

As the function $f(u)$, one can, for example, take $u(1-u)$ or $u(u-\mu)(1-u)$. It is known that, for large t, the solution of problem (13) behaves as a wave with some profile $V(\xi)$, $-\infty < \xi < \infty$, travelling with some velocity c^* from left to right: $u(t,x) \approx V(x-c^*t)$, $t \to \infty$. A small parameter may be introduced in the problem so that the wave (considered in a rough preliminary approximation) may become a step travelling with the velocity c^*. Such a consideration enables one to generalize the problem widely and examine a number of new effects, such as the appearance of "new sources" in space non-homogeneous media, the wave propagation at the expense of boundary conditions, and some others. This chapter also discusses the question of going over from the description of the wave propagation via equations of type (13) to the axiomatic theory of excitable media. The last section considers the problem of wave propagation in some systems of differential equations. As the basic apparatus in this and in the next chapters, we use asymptotic bounds of the Laplace type for functional integrals.

Chapter VII, the last one, examines the behavior, as $t \to \infty$, of the solution of equation (13) type in which the non-linear term has the form $f = f(x,u)$. As $f(x,u)$, we take either a function periodic in x, or a random field homogeneous in x. In either case the notion of wave propagation velocity is introduced (strictly speaking, the wave itself does not exist, though). This velocity is expressed in terms of some spectral characteristics of the operators.

Today a number of reviews and monographs are available where there are some applications of function integration and probabilistic methods in analysis, differential equations, and physics (e.g. Kac [1], [2], Feynman and Hibbs [1], Freidlin [7], Dynkin and Yushkevich [1], McKean [1], Friedman [2], Wentzell and Freidlin [2], Simon [1]). The overlap of this book with the above monographs is not large. However, to make this book

self-contained we had to include some results which are contained in the foregoing works.

We note that in this book there are no results on potential theory. Special monographs are already available on this subject (Blumenthal and Getoor [1], Meyer [1]). The problems connected with analyzing the spectra for second-order operators are also barely mentioned in this book. The materials on these questions may be found in the works of Kac [2, 3], Simon [1], Wentzell and Freidlin [2].

In conclusion we will explain how formulas and theorems are numbered. For example, Theorem 3.2.1 is the first theorem in the second section in Chapter III. Inside Chapter III, it is written as Theorem 2.1 only. Formulas are numbered in a similar fashion. Figures are numbered consecutively within each chapter.

Chapter I

STOCHASTIC DIFFERENTIAL EQUATIONS AND RELATED TOPICS

§1.1 *Preliminaries*

As is customary in mathematical probability theory, we start with a *probability space* $(\Omega, \mathcal{F}, P)$. Here Ω is an arbitrary set which is interpreted as the space of elementary events. The second component $\mathcal{F}$ is a *σ-field* of subsets of the space Ω, i.e. the system of subsets of the space Ω containing Ω itself and being closed with respect to unions and intersections in finite or countable numbers, as well as with respect to the operation of taking the complement. The elements of the *σ-field* $\mathcal{F}$ are called *events*. The third component of the probability space, P, is a *probability measure* on the σ-field $\mathcal{F}$, i.e. it is a non-negative, countably additive function defined on $\mathcal{F}$ and such that $P(\Omega) = 1$.

A function $\xi(\omega)$ on Ω with real values, for which $\{\omega : \xi(\omega) < x\} \in \mathcal{F}$ for any $x \in (-\infty, \infty)$, is called a *random variable*.

Given a set X with some σ-field $\mathcal{B}$ of its subsets (a *measurable space* $(X, \mathcal{B})$), one can define a random variable $\xi(\omega)$ with values in X. Indeed, it is required that the function $\xi(\omega)$ be $(\mathcal{F}, \mathcal{B})$-measurable: $\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$ for any $B \in \mathcal{B}$.

A probability measure $\mu(D) = P\{\xi(\omega) \in D\}$, $D \in \mathcal{B}$ is termed the *distribution* of the random variable $\xi(\omega)$.

If a space X is equipped with a topology, then the minimal σ-field containing all open sets, is called a *Borel σ-field* of the topological space X. The Borel σ-field in the Euclidean space R^r is denoted by $\mathcal{B}^r$; $(\mathcal{F}, \mathcal{B}^r)$-measurable (or briefly, $\mathcal{F}$-measurable) functions on Ω with values in R^r, are termed r-dimensional random variables. The mathematical

expectation of a random variable $\xi(\omega)$ will be denoted by $E\xi$:

$$E\xi = \int_{\Omega} \xi(\omega)\, P(d\omega) .$$

To every r-dimensional random variable ξ, there is a corresponding *characteristic function* $f_{\xi}(\lambda)$: $f_{\xi}(\lambda) = E \exp\{i(\lambda, \xi)\}$, $\lambda \in R^r$.

If a characteristic function $f_{\xi}(\lambda)$ has the form

$$f_{\xi}(\lambda) = \exp\left\{-\frac{1}{2}\,(Q(\lambda-m), \lambda-m)\right\}, \ \lambda \in R^r,$$

for some $m \in R^r$ and $Q = (q^{k\ell})$, $k, \ell = 1,2, \cdots, r$, then the random variable ξ is called *Gaussian*. Here $m = (m^1, \cdots, m^r)$ is the mathematical expectation of $\xi = (\xi^1, \cdots, \xi^r)$, and $(q^{k\ell}) = Q$ is the matrix of covariances: $q^{k\ell} = E(\xi^k - m^k)(\xi^\ell - m^\ell)$, $m^k = E\xi^k$. Here and henceforth we denote by $(.\,,.)$ the Euclidean scalar product. We remind that $q^{k\ell} = 0$ if and only if the components ξ^k and ξ^ℓ of a Gaussian random variable are independent. The class of Gaussian random variables is closed with respect to linear transformations and with respect to limit passage.

Suppose that in the space Ω, there is a σ-subfield $\mathcal{Y}$ of the underlying σ-field $\mathcal{F}: \mathcal{Y} \subseteq \mathcal{F}$. By a *conditional expectation* $E(\xi\,|\mathcal{Y})$ of a one-dimensional random variable ξ, we mean a $\mathcal{Y}$-measurable function on Ω for which the equality

$$\int_{\Lambda} E(\xi\,|\mathcal{Y})\, P(d\omega) = \int_{\Lambda} \xi(\omega)\, P(d\omega)$$

holds for any $\Lambda \in \mathcal{Y}$. These conditions define the conditional expectation in a unique way up to a set of zero measure. The equalities between conditional expectations are all fulfilled almost surely (P-a.s. or P-a.e.).

We shall list some basic properties of conditional expectations:

1. $\mathrm{E}(\xi|\mathcal{Y}) \geq 0$, if $\xi \geq 0$.

2. $\mathrm{E}(\xi+\eta|\mathcal{Y}) = \mathrm{E}(\xi|\mathcal{Y}) + \mathrm{E}(\eta|\mathcal{Y})$

if the summands on the right-hand side exist.

3. $\mathrm{E}(\xi\eta|\mathcal{Y}) = \xi\mathrm{E}(\eta|\mathcal{Y})$ if $\mathrm{E}\xi\eta$ and $\mathrm{E}\eta$ are defined and ξ is $\mathcal{Y}$-measurable.

4. Let $\mathcal{Y}_1$ and $\mathcal{Y}_2$ be two σ-fields such that $\mathcal{Y}_1 \subseteq \mathcal{Y}_2 \subseteq \mathcal{F}$. Then $\mathrm{E}(\xi|\mathcal{Y}_1) = \mathrm{E}(\mathrm{E}(\xi|\mathcal{Y}_2)|\mathcal{Y}_1)$.

5. Suppose that a random variable ξ does not depend on a σ-field $\mathcal{Y}$, that is $\mathrm{P}\{(\xi \in D) \cap \mathcal{A}\} = \mathrm{P}\{\xi \in D\}\mathrm{P}(\mathcal{A})$ for any Borel set $D \in R^1$ and $\mathcal{A} \in \mathcal{Y}$. Then $\mathrm{E}(\xi|\mathcal{Y}) = \mathrm{E}\xi$ provided $\mathrm{E}\xi$ exists.

Let $\chi_{\mathcal{A}}(\omega)$ be the indicator of a set $\mathcal{A} \in \mathcal{F}$, i.e. $\chi_{\mathcal{A}}(\omega) = 1$ for $\omega \in \mathcal{A}$ and $\chi_{\mathcal{A}}(\omega) = 0$ for $\omega \in \Omega \setminus \mathcal{A}$. Then the random variable $\mathrm{E}(\chi_{\mathcal{A}}(\omega)|\mathcal{Y})$ is said to be the conditional probability of the event $\mathcal{A}$ with respect to the σ-field $\mathcal{Y}$ and is denoted by $\mathrm{P}(\mathcal{A}|\mathcal{Y})$.

A family of r-dimensional random variables $\xi_t(\omega)$ depending on a real parameter $t \in T \subset (-\infty, \infty)$ is called a *random process*. Thus, for every fixed $t = t_0$, we obtain a random variable $\xi_{t_0}(\omega)$. For a fixed $\omega = \omega_0$ we obtain a function of t which is called a trajectory or a sample function of the process $\xi_t(\omega)$.

The totality of distributions $\mu_{t_1,\ldots,t_n}$ of random variables $(\xi_{t_1},\ldots,\xi_{t_n})$ for various $n = 1,2,3,\cdots$ and $t_1,\cdots,t_n \in T$, is termed the *family* of *finite-dimensional distributions* of the process $\xi_t(\omega)$. In this book, we shall, as a rule, consider random processes having parameter t which varies over the half-line $[0,\infty)$ or over the interval $[0,T_1]$ and having trajectories which are continuous with probability 1. Such processes are, in fact, defined in a unique way by their finite-dimensional distributions.

If the finite-dimensional distributions of a process $\xi_t(\omega)$, $t \in [0,\infty)$, are all Gaussian ones, then the process is called a *Gaussian random process*. The finite-dimensional distributions of such a process are all defined in a unique way by two functions—by the mathematical expectation $m(t) = \mathrm{E}\xi_t$ and by the correlation function $R(s,t) = \mathrm{E}(\xi_s - m(s))(\xi_t - m(t))$.

If $\xi_t = (\xi_t^1, \cdots, \xi_t^r)$ is an r-dimensional random process, then $m(t) = (E\xi_t^1, \cdots, E\xi_t^r)$; $R(s,t) = (R^{ij}(s,t))$, $R^{ij}(s,t) = E(\xi_s^i - m^i(s))(\xi_t^j - m^j(t))$.

With every random process $\xi_t(\omega)$, $t \in T$, one can associate an increasing system of σ-fields $\mathcal{F}_{\leq t} = \mathcal{F}^{\xi}_{\leq t} = \sigma(\xi_s, s \leq t)$ where $\sigma(\xi_s, s \leq t)$ is the minimal σ-field with respect to which the random variables ξ_s are all measurable for $s \leq t$, $s \in T$. Sometimes we shall also consider the σ-fields $\mathcal{F}_{\geq t} = \mathcal{F}^{\xi}_{\geq t} = \sigma(\xi_s, s \geq t)$. We shall use the notation $E(\eta|\xi_s, s \in S)$ for the conditional expectation of the random variable η with respect to the σ-field generated by the random variables ξ_s for $s \in S$ (i.e. the minimal σ-field with respect to which the random variables ξ_s are all measurable for $s \in S$).

All the objects introduced here and their properties, as well as other elementary information from probability theory, which is assumed to be known to the reader, are considered in detail in many courses in the theory of stochastic processes (see, e.g. Doob [1], Gihman and Skorohod [1], Wentzell [1]).

§1.2 *The Wiener measure*

In this section, the Wiener measure in the space of continuous functions will be constructed. This measure is connected with the Laplace operator and with the simplest heat operator. Solutions of some problems for such operators admit a representation in the form of integrals of appropriate functionals with respect to this measure. In many respects, the measures connected with general second-order elliptic operators are similar to the Wiener measure, and it is convenient to construct such measures proceeding from the Wiener measure.

Because the Wiener measure is so important, we will describe a few constructions for it and list its basic properties. As a rule, we shall drop proofs restricting ourselves to brief comments on what needs to be proved. The detailed proofs of these properties are available in many manuals and monographs, so the reader is referred to the corresponding literature.

Let $C_{0,T} = C_{0,T}(R^1)$ be the space of continuous functions on $[0,T]$, $T < \infty$, with values on the real line R^1; $C_{0,\infty} = C_{0,\infty}(R^1)$ being the space of continuous functions on $[0,\infty)$. We put $C^0_{0,T} = \{\phi \in C_{0,T} : \phi_0 = 0\}$, $0 < T \leq \infty$. The spaces C_{0T} and C^0_{0T} will be thought of as equipped with the uniform convergence topology (uniform convergence on every finite interval whenever $T = \infty$).

As the σ-field in the spaces $C^0_{0,T}$, $0 < T \leq \infty$, one can take the Borel σ-field $\mathcal{B}_{0,T}$, that is the minimal σ-field containing all open sets of the space $C^0_{0,T}$.

Suppose we are given a measurable space $(X, \mathcal{B})$, i.e. there is a set X, with a σ-field of subsets $\mathcal{B}$ being defined in it. How can one put a measure on this space? There are several ways of assigning a measure. Presumably, it is the easiest way to define a measure with the aid of a density function with respect to some standard measure defined on the same σ-field $\mathcal{B}$. For example, let $X = R^1$ be the real line equipped with the σ-field of Borel sets. Then, as the standard measure, it is sometimes natural to take the Lebesgue measure; that is, a measure on the line, unique up to scale factor and invariant with respect to translations. For instance, Gaussian measure on the line is specified by the density function

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}$$

with respect to the Lebesgue measure, where a and σ are real parameters.

So far we have had no measure in the space $C^0_{0,T}$. It should be noted that it is not easy to define some non-trivial measure in this space, and there is no measure which is invariant with respect to translations on $C^0_{0,T}$. Therefore, this way does not fit as a starting point, but we shall bear it in mind for later constructions.

Another way of defining a measure consists of prescribing it with the help of some mapping. Let a measurable space $(Y, \mathcal{B}')$ and a probability measure $P'(A)$, $A \in \mathcal{B}'$ be already available. Suppose we are given a

measurable mapping $f: (Y, \mathcal{B}') \to (X, \mathcal{B})$; that is, a mapping for which $f^{-1}(B) \in \mathcal{B}'$ for $B \in \mathcal{B}$. Then this mapping induces the probability measure $P(B) = P'(f^{-1}(B))$ on the σ-field $\mathcal{B}$ in X. And this way is already suitable for constructing the Wiener measure in $C^0_{0,T}$. We shall provide this construction later on in this section.

The third way of defining a measure is with the help of passage to the limit. Namely, it is possible to construct a sequence of measures μ_n in $C^0_{0,T}$ being described in a comparatively simple manner, and then to consider the limit of μ_n as $n \to \infty$. For example, limit in the sense of weak convergence of measures is convenient. Recall that a sequence of measures μ_n in $C^0_{0,T}$ converges weakly to a measure μ if

$$\int_{C^0_{0,T}} f(x)\,\mu_n(dx) \to \int_{C^0_{0,T}} f(x)\,\mu(dx), \quad n \to \infty,$$

for any continuous bounded functional $f(x)$ on the space $C^0_{0,T}$.

In this way, the Wiener measure may be constructed in $C^0_{0,T}$ as well.

Finally, we recall another approach widely used in probability theory—Kolmogorov's extension of measure. According to this method, a measure for some collection of relatively simple sets must be defined, and then it must be continued, by the countable additivity property, onto the smallest σ-field containing the original collection of sets. Of course, in doing so one must demonstrate that such an extension will not face obstacles and will give a measure, countably additive on this minimal σ-field. If one wants to get a measure on the σ-field which has been set beforehand, then, in addition, one should make certain that this σ-field is contained in the minimal σ-field generated by the simple sets.

We begin by outlining the construction of the Wiener measure via the last procedure.

So, first of all, a collection of "simple" sets in the space of continuous functions must be indicated. Let $0 < t_1 < t_2 < \cdots < t_n$, where n is

any positive integer. Moreover, let $\Gamma_1, \Gamma_2, \cdots, \Gamma_n$ be intervals of the real line (not necessarily different). We will denote by $\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n}$ the following set in the space $C^0_{0,\infty}$ (Fig. 1):

$$\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n} = \{\phi \in C^0_{0,\infty} : \phi_{t_1} \in \Gamma_1, \cdots, \phi_{t_n} \in \Gamma_n\}.$$

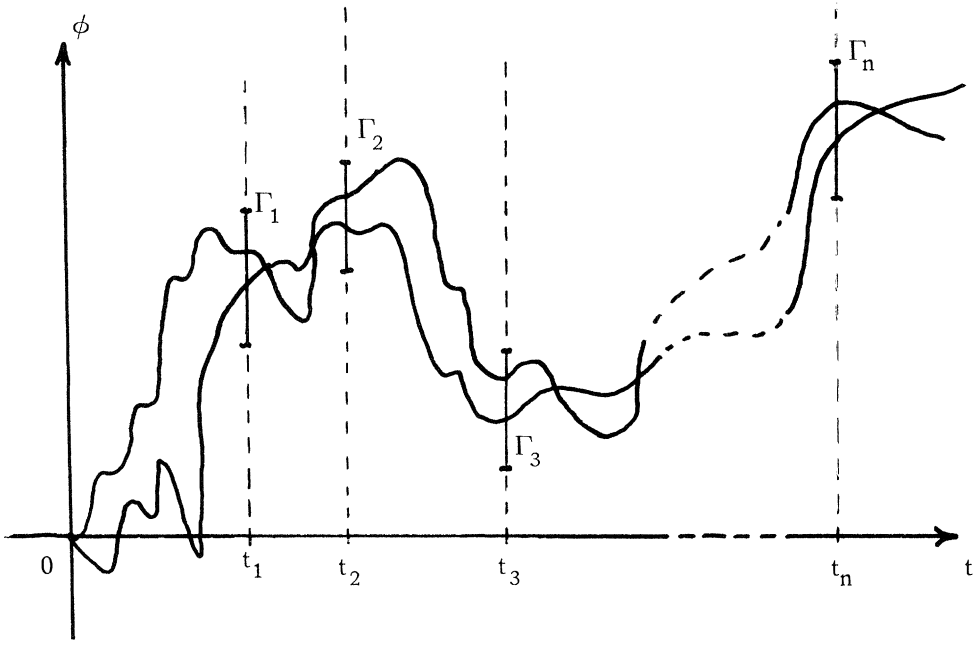

Fig. 1

These sets belong to the class of so-called cylinder sets. By *cylinder sets* in the space $C^0_{0,\infty}$, we mean the sets of the form $\{\phi \in C^0_{0,\infty}: (\phi_{t_1}, \cdots, \phi_{t_n}) \in B\}$, where B belongs to the Borel σ-field $\mathcal{B}^n$ in R^n. In the case of the sets $\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n}$ we choose $B = \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_n$.

Now let us introduce a measure on the sets $\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n}$. We designate

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}, x,y \in R^1, t > 0, \tag{2.1}$$

and put

$$\mu(\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n}) = \tag{2.2}$$

$$= \int_{\Gamma_1} dy_1 \cdots \int_{\Gamma_n} dy_n p(t_1,0,y_1)\ p(t_2-t_1,y_1,y_2) \cdots p(t_n-t_{n-1},y_{n-1},y_n)\ .$$

Notice that if, for some i, Γ_i coincides with the whole space R^1, then

$$\Pi^{t_1,\ldots,t_{i-1},t_i,t_{i+1},\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_{i-1},R^1,\Gamma_{i+1},\ldots,\Gamma_n} = \Pi^{t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_{i-1},\Gamma_{i+1},\ldots,\Gamma_n}\ .$$

In other words, the superscript t_i and the corresponding interval $\Gamma_i = R^1$ may be omitted. Hence, one and the same set in the space of functions may be written in the form $\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n}$ with different number of indices.

Formula (2.2) for the measure of this set can also be written in various ways. Thus, for our definition of the measure of "simple" sets to be correct, it is necessary that, if $\Pi^{t_1,\ldots,t_n}_{\Gamma_1,\ldots,\Gamma_n} = \Pi^{t'_1,\ldots,t'_m}_{\Gamma'_1,\ldots,\Gamma'_m}$ then the corresponding values of the measure defined by formula (2.2) coincide. This property is referred to as the compatibility of a family of distributions. It is possible to check that formula (2.2) does define a compatible family of distributions. This compatibility comes from the Kolmogorov-Chapman equation

$$p(s+t,x,y) = \int_{-\infty}^{\infty} p(s,x,z)\, p(t,z,y)\, dz \tag{2.3}$$

which is fulfilled for function (2.1).

It is easily seen that the function $p(t,x,y)$ defined by (2.1) also has the following properties:

$$p(t,x,y) = p(t,0,y-x), \quad \int_{-\infty}^{\infty} x^{\alpha} p(h,0,x)\,dx = \text{const} \times h^{\alpha/2}, \qquad (2.4)$$

where α is an arbitrary positive number.

By using the compatibility of the distributions given by formula (2.2) and properties (2.4), it is now possible to prove that the function defined for simple sets by formula (2.2) may be extended to a measure on the minimal σ-field of subsets of the space $C^0_{0,\infty}$ which contains all possible $\Pi^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n}$ (see e.g. Ito and McKean [1], Wentzell [1]). This minimal σ-field will be denoted by $\mathfrak{N}^{\infty}_0$, $\mathfrak{N}^t_0$ will designate the σ-subfield of the σ-field $\mathfrak{N}^{\infty}_0$ generated by the sets $\Pi^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n}$ for arbitrary natural numbers n and $0 < t_1 < t_2 < \cdots < t_n \leq t$.

What kind of sets belong to the σ-fields $\mathfrak{N}^t_0$ and $\mathfrak{N}^{\infty}_0$? Generally speaking, the σ-fields $\mathfrak{N}^t_0$ and $\mathfrak{N}^{\infty}_0$ are rather extensive. For example, let us consider the set:

$$\mathfrak{A} = \{\phi \in C^0_{0,\infty} : f_s < \phi_s < g_s \text{ for } s \in [0,t]\},$$

where f_s, g_s, $s \in [0,t]$ are arbitrary continuous functions on the segment $[0,t]$, $f_s < g_s$, $f_0 < 0 < g_0$. Since we deal with continuous functions,

$$\mathfrak{A} = \bigcup_{n=1}^{\infty} \bigcap_{s_i \in \Lambda_{0,t}} \left\{ f_{s_i} + \frac{1}{n} < \phi_{s_i} < g_{s_i} - \frac{1}{n} \right\},$$

where $\Lambda_{0,t}$ denotes the set of rational numbers in the segment $[0,t]$. Therefore, the set $\mathfrak{A}$ may be represented in the form of a union of the intersections of a countable number of simple sets, and thus $\mathfrak{A} \in \mathfrak{N}^t_0$.

Hence it appears clear that the σ-field $\mathfrak{N}^{\infty}_0$ contains the Borel σ-field $\mathfrak{B}_{0,\infty}$. It is possible to make sure that the opposite inclusion is also valid, and therefore, $\mathfrak{B}_{0,\infty} = \mathfrak{N}^{\infty}_0$. For $0 < t_1 < \cdots < t_n \leq T$, the sets $\Pi^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n}$ may be projected into the space $C^0_{0,T}$. The σ-field generated by these sets in $C^0_{0,T}$ will also be called $\mathfrak{N}^T_0$. It is readily checked that this σ-field in $C^0_{0,T}$ coincides with the Borel σ-field $\mathfrak{B}_{0,T}$.

We proceed to present some other examples of sets from $\mathfrak{N}^T_0$ and $\mathfrak{N}^{\infty}_0$.

Let $x > 0$, $\tau_x(\phi) = \min\{t : \phi_t = x\}$, i.e. $\tau_x(\phi)$ is the first moment when $\phi \in C^0_{0,\infty}$ hits the level x; if the set $\{t : \phi_t = x\}$ is empty (the function ϕ never attains the level x), then we set $\tau_x(\phi) = +\infty$. Consider the set $\mathcal{D} = \{\phi \in C^0_{0,\infty} : \tau_x(\phi) \le t\}$. It is clear that $\mathcal{D} = \bigcap_{n=1}^{\infty} \bigcup_{s_i \in \Lambda_{0,t}} \left\{\phi \in C^0_{0,t} : \phi_{s_i} > x - \frac{1}{n}\right\}$ which implies that $\mathcal{D} \in \mathcal{N}_0^t$. It is not difficult to verify that the set $\mathcal{E} = \{\phi \in C^0_{0,\infty} : \tau_x(\phi) < \tau_x(-\phi)\}$ of the trajectories reaching the point x before the point $-x$, belongs to the σ-field $\mathcal{N}_0^\infty$. Throughout this book, we will often consider the first hitting times of closed sets.

So, in the space $C^0_{0,T}$ there is a σ-field $\mathcal{N}_0^T$ which coincides with the Borel σ-field $\mathcal{B}_{0,T}$ of this space (if $T = \infty$, then $\mathcal{N}_0^\infty = \mathcal{B}_{0,\infty}$).

The measure on the σ-field $\mathcal{N}_0^T$, $0 < T \le \infty$, defined by equality (2.2) for the "simple" sets $\Pi^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n}$, is called the *Wiener measure*.

The first mathematically correct construction of such a measure and the analysis of its basic properties was brought about by N. Wiener [1] in 1923.

REMARK. The statement that the function given by (2.2) for simple sets may be extended to a measure on the σ-field $\mathcal{N}_0^T$ in the space $C^0_{0,T}$, is based on two different results. The first of them is Kolmogorov's general theorem on the extension of a measure in the space $H_{0,T}$ of all functions on $[0,T]$ with values in R^1. The second result is concerned with the conditions under which this measure is concentrated on the space of continuous functions. In order to clarify the meaning of these results, let us consider the "simple" sets in $H_{0,T}$:

$$\widetilde{\Pi}^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n} = \{\phi \in H_{0,T} : \phi_{t_1} \in \Gamma_1, \cdots, \phi_{t_n} \in \Gamma_n\}.$$

Here n is any natural number, $\Gamma_1, \cdots, \Gamma_n$ are intervals on the line, $0 < t_1 < \cdots < t_n \le T$. For the simple sets $\widetilde{\Pi}^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n} \subset H_{0,T}$ we will define the function

$$\tilde{\mu}(\tilde{\Pi}^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n}) = \tag{2.5}$$

$$= \int_{\Gamma_1} dy_1 \cdots \int_{\Gamma_n} dy_n p(t_1,0,y_1)\, p(t_2-t_1,y_1,y_2) \cdots p(t_n-t_{n-1},y_{n-1},y_n)\,,$$

where $p(t,x,y)$ is specified by equality (2.1).

Just as in the case of the simple sets in $C^0_{0,T}$, one and the same simple set in $H_{0,T}$ may be described in different ways, for example, with a different number of indices. However, the function $\tilde{\mu}$ depends only on the set $\tilde{\Pi}^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n} \subset H_{0,T}$ and not on how it is written down. This follows from the Kolmogorov-Chapman equation (2.3). By the theorem on the extension of measure, the function $\tilde{\mu}$, originally defined only on simple sets in $H_{0,T}$, may be extended to a measure on the minimal σ-field $\tilde{\mathfrak{N}}^T_0$ in $H_{0,T}$ which contains all simple sets $\tilde{\Pi}^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n}$, $0 < t_1 < t_2 < \cdots < t_n < T$.

The explicit form of the function $p(t,x,y)$ is not essential for the extension of $\tilde{\mu}$ to a measure on $\tilde{\mathfrak{N}}^T_0$ to exist. For example, the function $p_1(t,x,y) = \dfrac{t}{\pi[t^2 + (y-x)^2]}$ also satisfies the Kolmogorov-Chapman equation. If in equalities (2.5), $p(t,x,y)$ is replaced by $p_1(t,x,y)$, then we have some new function $\tilde{\mu}_1$ on the simple sets in $H_{0,T}$. This function may be extended to a measure on the σ-field $\tilde{\mathfrak{N}}^T_0$ in $H_{0,T}$ as well.

Suppose now that in equalities (2.2), p_1 is substituted for p. If these modified equalities are used to define the function $\mu_1(\Pi^{t_1,\dots,t_n}_{\Gamma_1,\dots,\Gamma_n})$ for simple sets in $C^0_{0,T}$, then this function may no longer be extended to a measure on $\mathfrak{N}^T_0$: In order that one could perform the extension to a measure on the σ-field $\mathfrak{N}^T_0$ in the space $C^0_{0,T}$, it is sufficient to check that the outer measure of the set $C^0_{0,T}$ in $H_{0,T}$ is equal to 1 ($C^0_{0,T}$ does not belong to $\tilde{\mathfrak{N}}^T_0$, thus the measure $\tilde{\mu}(C^0_{0,T})$ is not defined):

$$\inf\{\tilde{\mu}(\mathfrak{A}) : C^0_{0,T} \subset \mathfrak{A} \in \tilde{\mathfrak{N}}^T_0\} = 1 .$$

It turns out that, for the last equality to be fulfilled, simple sufficient conditions may be given. Namely, for the outer measure of the set $C^0_{0,T}$ to be 1, and therefore, in order that one could extend μ from simple sets in $C^0_{0,T}$ to a measure on $\mathfrak{N}^T_0$, it suffices that, for some $\alpha, \beta, c > 0$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x_1 - x_2|^{\alpha} \mu_{t_1,t_2}(dx_1 \times dx_2) \le c|t_1 - t_2|^{1+\beta} , \tag{2.6}$$

where $\mu_{t_1,t_2}(\Gamma_1 \times \Gamma_2) = \mu(\Pi^{t_1,t_2}_{\Gamma_1,\Gamma_2})$. The above assertion is the Kolmogorov theorem on the existence of a continuous modification. In the case of the function $p(t,x,y)$ such constants α, β, c exist. This follows, for example, from (2.4) with $\alpha = 4$. Therefore, one can make the extension of the measure defined by (2.2) for simple sets in $C^0_{0,T}$ to a measure on the σ-field $\mathfrak{N}^T_0$. For μ_1, introduced previously, such an extension in $C^0_{0,T}$ is impossible.

If the compatibility condition and bound (2.6) holds, then one can arrange the proof of the existence of an extended measure immediately in the space $C^0_{0,T}$ (see Ito and McKean [1]), without dividing it into these two stages.

The notion of the Wiener measure is closely related to that of Wiener random process. A random process $W_t(\omega)$, $t \ge 0$, on a probability space $(\Omega, \mathcal{F}, P)$ is called a *Wiener process* if its trajectories are continuous with probability 1, $P\{W_0 = 0\} = 1$, and the finite-dimensional distributions are given by

$$P\{W_{t_1} \in \Gamma_1, \cdots, W_{t_n} \in \Gamma_n\} = \tag{2.7}$$

$$= \int_{\Gamma_1} dy_1 \cdots \int_{\Gamma_n} dy_n \, p(t_1,0,y_1)\, p(t_2 - t_1, y_1, y_2) \cdots p(t_n - t_{n-1}, y_{n-1}, y_n)$$

where $p(t,x,y)$ is defined by equality (2.1).

Formula (2.7) implies that W_t is a Gaussian process. By using (2.7) it is not difficult to calculate that $EW_t = 0$, and the correlation function $R(s,t)$ of the process W_t has the form $R(s,t) = EW_sW_t = s \wedge t$. Note that the correlation function and the expectation of a Gaussian process determine, in a unique way, its finite-dimensional distribution functions. Thus, the Wiener process could be defined as a mean zero Gaussian process with continuous trajectories having the correlation function $R(s,t) = s \wedge t$.

Every one-dimensional random process $X_t(\omega)$, $t \geq 0$, $\omega \in \Omega$ whose trajectories are continuous with probability 1, induces a mapping $\Omega \to C_{0,\infty} : \omega \to X_\cdot(\omega)$. This mapping induces a probability measure on the σ-field[1] $\mathcal{N}_0^\infty$ in $C_{0,\infty}$. Comparing (2.7) with (2.2) we draw the conclusion that the process $W_t(\omega)$ induces the Wiener measure in the space $C_{0,\infty}^0$.

On the other hand, given the space $C_{0,\infty}^0$ with the Wiener measure μ_W on the σ-field $\mathcal{N}_0^\infty$, one can take $(C_{0,\infty}, \mathcal{N}_0^\infty, \mu_W)$ as a probability space, and define the random process $W_t(\omega) = \phi_t$, $t \in [0,\infty)$, for $\omega = \phi$. By virtue of (2.2) this process is a Wiener process. Therefore, the trajectories of the Wiener process $W_t(\omega)$ are simply elements of the space $C_{0,\infty}^0$. The elements of the σ-field $\mathcal{N}_0^t$ are events defined by the motion of the Wiener process W_s for $s \in [0,t]$.

The construction of Wiener measure and the examination of its properties is in essence equivalent to the construction and examination of the properties of Wiener process.

Let us consider other constructions of the Wiener measure. We will construct the Wiener measure in $C_{0,1}^0$. Afterwards, it is not hard to define the Wiener measure in $C_{0,T}^0$ for any $T > 0$.

---

[1] We preserve the notation $\mathcal{N}_0^t$, $0 \leq t \leq \infty$, for the σ-field of all cylinder sets of $C_{0,t}$ (not only of $C_{0,t}^0$), i.e. $\mathcal{N}_0^t$ is the σ-field generated by the sets $\{\phi \in C_{0,t} : (\phi_{t_1},\dots,\phi_{t_n}) \in B\}$ for any natural n, any $0 \leq t_1 < t_2 < \cdots < t_n \leq t$, and $B \in \mathcal{B}^n$.

Suppose we are given a probability space $(\Omega, \mathcal{F}, P)$ and a sequence of independent mean zero Gaussian random variables on it having variance $1: \xi_0, \xi_1, \xi_2, \cdots$. Such a probability space may be constructed, for example, by taking an infinite product of lines equipped with Borel σ-fields and the standard Gaussian measure. We will define the mapping $(\Omega, \mathcal{F}) \to (C^0_{0,1}, \mathcal{R}^1_0)$ in such a fashion that in $C^0_{0,1}$ the Wiener measure will be induced.

Let us consider the series

$$\phi_s(\omega) = \xi_0 s + \frac{\sqrt{2}}{\pi} \sum_{n=1}^{\infty} \sum_{k=2^{n-1}}^{2^n - 1} \xi_k \frac{\sin k\pi s}{k}, \quad 0 \le s \le 1 . \tag{2.8}$$

This series converges uniformly on $[0,1]$ with probability 1 (see, e.g. Ito and McKean [1]).

Thus, formula (2.8) gives the mapping $f: \Omega \to C^0_{0,1}$. This mapping is readily checked to be measurable, i.e. $f^{-1}(B) \in \mathcal{F}$ for $B \in \mathcal{R}^1_0$. Since the sum of independent Gaussian random variables is also a Gaussian random variable, and since the limit of Gaussian random variables is also a Gaussian random variable, we conclude that the process $\phi_s(\omega)$ is Gaussian. Now let us verify that it is a Wiener process. As was said above, for this it suffices to check that $E\phi_s = 0$, $E\phi_s\phi_t = s \wedge t$. The first of these equalities comes from the fact that series (2.8) may be integrated with respect to the measure $P(d\omega)$ on Ω termwise. Noting that $E\xi_i\xi_j = \delta_{ij}$ we conclude from (2.8) that

$$E\phi_s\phi_t = st + \sum_{k=1}^{\infty} \frac{2}{\pi^2} \frac{\sin k\pi s \sin k\pi t}{k^2}, \quad s,t \in [0,1] .$$

We leave to the reader to make certain that the expression on the right-hand side is equal to $s \wedge t$. Hence, $\phi_s(\omega)$ is a Wiener process. The mapping $f: \Omega \to C^0_{0,1}$ induces the Wiener measure in $C^0_{0,1}$.

Finally we mention another way of constructing the Wiener measure ensuring a transparent image of the nature of the set of functions this measure is concentrated on.

We shall consider a symmetric random walk over a lattice on the line. Let a particle (being at time t at a point kh) jump to one of the neighboring points $(k-1)h$ or $(k+1)h$ with equal probabilities at time $t+\Delta$. Here $h > 0$ is the lattice spacing, $\Delta > 0$ is a time interval between sequential jumps, $k = 0,1,2,\cdots$. Let us denote by $X^{\Delta,h}_{k\Delta} = X^{\Delta,h}_{k\Delta}(\omega)$, $k = 0,1,2,\cdots$, the trajectory of this particle starting from zero at time $t = 0$. Clearly, $X^{\Delta,h}_{k\Delta}$ is a random sequence. Let $(\Omega, \mathcal{F}, P)$ be the probability space this sequence is defined on. We will introduce the random broken lines $X^{\Delta,h}_s$ $s \in [0,1]$, consisting of the segments which connect sequential points $(k\Delta, X^{\Delta,h}_{k\Delta}(\omega))$ and $((k+1)\Delta, X^{\Delta,h}_{(k+1)\Delta})$, $k = 0,1,\cdots,\left[\frac{1}{\Delta}\right], \left[\frac{1}{\Delta}\right]+1$. The random broken lines $X^{\Delta,h}_s$ induce in $C^0_{0,1}$ certain measure $\mu^{\Delta,h}$ concentrated on the broken lines with vertices at the points $(k\Delta, \ell h)$, $k = 0,1,2,\cdots$; $\ell = 0, \pm 1, \pm 2, \cdots$. For any integers $k_1 < k_2 < \cdots < k_n$ and any intervals $\Gamma_1, \Gamma_2, \cdots, \Gamma_n \subset R^1$, the value of the measure $\mu^{\Delta,h}$ of the simple set

$$\{X^{\Delta,h}_\cdot \in C^0_{0,1} : X^{\Delta,h}_{k_1\Delta} \in \Gamma_1, X^{\Delta,h}_{k_2\Delta} \in \Gamma_2, \cdots, X^{\Delta,h}_{k_n\Delta} \in \Gamma_n\}$$

may be written down explicitly via binomial probabilities.

Now let $\Delta, h \downarrow 0$ in such a way that $h^2\Delta^{-1} = 1$. Then the family of measures $\mu^{\Delta,h}$ proves to converge weakly to the Wiener measure in the space $C^0_{0,1}$. The proof of this statement may be decomposed into two stages. First, one must prove that the measures $\mu^{\Delta,h}$ of simple sets in $C^0_{0,1}$ converge to the Wiener measure of these sets. This convergence is an implication of a version of the central limit theorem. Secondly, one must verify that the family of measures $\mu^{\Delta,h}$, $h^2 = \Delta$, $\Delta \downarrow 0$ is weakly compact. This implies that the measures $\mu^{\Delta,h}$ converge weakly to the Wiener measure. For a detailed proof, see Donsker [1], Ito and McKean [1].

A more strong statement is also true (see Knight [1], Ito and McKean [1]): one can find a probability space $(\Omega, \mathcal{F}, P)$ and a family of random walks $X_{k\Delta}^{\Delta,h}(\omega)$, $\omega \in \Omega$, such that the broken lines $X_s^{\Delta,h}(\omega)$ drawn by these walks, with probability 1, converge uniformly on $[0,1]$ to the continuous functions $X_s = X_s(\omega)$. The random process $X_s(\omega)$ is a Wiener process.

So, the functions the Wiener measure is concentrated on, may be imagined as the limits of random broken lines $X_s^{\Delta,h}$ for $\Delta = h^2$ as $\Delta \downarrow 0$. Whence, via the properties of symmetric random walk over a lattice on the line, we can obtain the properties of the Wiener process. For example, the fact that, with probability 1, random walk returns to zero an infinite number of times, implies that, for almost all trajectories of the Wiener process $W_s(\omega)$, the set $\{s \in [0,1], W_s(\omega) = 0\}$ is a perfect one (i.e. a closed set every point of which is a limit point for this set).

So, suppose we are given a Wiener process $W_t(\omega)$, $t \geq 0$, on the probability space $(\Omega, \mathcal{F}, P)$. We will list some important properties of the Wiener process (and, thereby, of the Wiener measure as well) in addition to those mentioned previously.

We will show that the Wiener process has independent increments, i.e. for any $0 < t_1 < t_2 < \cdots < t_n$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \cdots, W_{t_n} - W_{t_{n-1}}$ are independent. As W_t is a Gaussian process, all the n differences are jointly Gaussian distributed. Hence, to prove the independence of these random variables, it is sufficient to check that they are uncorrelated, i.e. that $E(W_{t_{k+1}} - W_{t_k})(W_{t_{i+1}} - W_{t_i}) = 0$ for $i \neq k$. The last equality may be easily checked by remembering that $EW_s W_t = s \wedge t$. It is possible to demonstrate that the Wiener process is the only random process having continuous, with probability 1, trajectories and independent increments for which $EW_t = 0$, $EW_t^2 = t$.

From the independence of increments, we can readily deduce the so-called *Markov property* of the Wiener process which may be stated as follows: for any $T > 0$, the random process $\widetilde{W}_t = W_{T+t}(\omega) - W_T(\omega)$ does

not depend on the σ-field $\mathcal{F}^W_{\leq T} = \sigma(W_s, s \in [0,T])$[2] and is a Wiener process.

The proof of this statement is immediate from the independence of increments and from the following observation: for any $0 \leq t_1 < t_2 < \cdots < t_n < T$ the σ-fields $\sigma(W_{t_1}, \cdots, W_{t_n})$ and $\sigma(W_{t_1}, W_{t_2} - W_{t_1}, \cdots, W_{t_n} - W_{t_{n-1}})$ coincide.

May we substitute a random variable for T? Will the process $\widetilde{W}_t = W_{t+\tau} - W_\tau$ be a Wiener process if the time $\tau = \tau(\omega)$ itself depends on the trajectory? One can show that one cannot take an arbitrary random variable. For example, let $\theta = \sup\{t \in [0,1] : W_t = 0\}$. By the definition of θ, the process $\widetilde{W}_t = W_{t+\theta} - W_\theta$ does not change the sign in the interval $(0, 1-\theta)$ with probability 1. For a Wiener process, such a behavior (no zeroes in some interval $(0, 1-\theta)$) has probability 0. For instance, this follows from the foregoing Knight's result asserting that the Wiener trajectories may be represented in the form of the limit of symmetric walks over a lattice, if one takes into account that symmetric random walk returns to zero infinitely many times. Therefore, $\widetilde{W}_t = W_{t+\theta} - W_\theta$ is not a Wiener process. However, an important class of random variables $\tau(\omega)$ does exist for which the process $W_{t+\tau} - W_\tau$ is a Wiener process. These are the so-called Markov times.

A random variable $\tau(\omega)$ which is allowed to take non-negative values and the value $+\infty$ is said to be a *Markov time*,[3] whenever for every $t \geq 0$ the event $\{\tau(\omega) \leq t\}$ belongs to $\mathcal{F}^W_{\leq t} = \sigma(W_s, 0 \leq s \leq t)$.

In other words, a Markov time is one whose occurrence may be known without any information about what will happen after this time. This property is emphasized by another term for the Markov time—a random variable independent of future. In particular, the random variable θ introduced above (the last time, before 1, when the Wiener trajectory

[2] We will remind that if $\mathfrak{A}$ is a collection of random variables on a space $(\Omega, \mathcal{F}, P)$ then $\sigma(\mathfrak{A})$ is the minimal σ-field with respect to which the variables of the class $\mathfrak{A}$ are all measurable.

[3] To be more exact, a Markov time with respect to the expanding system of the σ-fields $\mathcal{F}^W_{\leq t}$. Later on we shall introduce a more general notion of Markov time.

visits zero) is not a Markov time. On the other hand, $\tau(\omega) = \inf\{t : W_t = a\}$ –(the first hitting time to the point a)–is a Markov time.

It is easy to ascertain the elementary properties of the Markov times. For instance, if τ_1 and τ_2 are Markov times, then $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$ are also Markov times. The variables $\tau_1 + 1$ and $2\tau_1$, for example, are Markov times as well; but the variables $\frac{\tau_1}{2}$ and $\tau_2 - 1$ are not, in general, Markov times.

A σ-field $\mathcal{F}^W_{\leq\tau} \subset \mathcal{F}$ is associated with every Markov time $\tau(\omega)$. Namely, an event $\mathcal{A} \in \mathcal{F}$ belongs to $\mathcal{F}^W_{\leq\tau}$ if and only if, for every $t \geq 0$, the inclusion $\mathcal{A} \cap \{\tau \leq t\} \in \mathcal{F}^W_{\leq t}$ holds. It is not difficult to check that the collection of such events actually forms a σ-field. Intuitively, this σ-field is the collection of events which are defined by the behavior of the Wiener process W_s up to the time $\tau(\omega)$.

It is possible to prove that if $\tau(\omega)$ is a Markov time and $P\{\tau(\omega) < \infty\} = 1$, then the random process $W_t = W_{t+\tau} - W_\tau$ does not depend on the σ-field $\mathcal{F}^W_{\leq\tau}$ and is a Wiener process (see Dynkin [1], Hunt [1], Ito and McKean [1]). This property is called the *strong Markov property* of a Wiener process.

Up to now we have been dealing with the Wiener trajectories starting solely from zero. In what follows it will be suitable for us to consider a family of Wiener processes starting from various points $a \in R^1 : W^a_t = a + W_t$. Then the Markov property may be interpreted like this: the process $\widetilde{W}^{W_T}_t = W_{T+t}$ is a new Wiener process starting from the point W_T, at time $t = 0$ and conditionally independent of the behavior of the process W_t for $t \leq T$, given W_T.

The strong Markov property may be applied for deducing the helpful relation (see, e.g. Ito and McKean [1]):

$$P\{\max_{0\leq s\leq t} W_s > a\} = 2P\{W_t > a\}.$$

In fact, by denoting $\tau_a = \min\{t : W_t = a\}$ we arrive at

$$P\{\max_{0\leq s\leq t} W_s > a\} = P\{\tau_a \leq t\} = P\{\tau_a \leq t, W_t < a\} + \\ + P\{\tau_a \leq t, W_t \geq a\}.$$

The terms on the right-hand side of this equality are equal. This is immediate from the strong Markov property and from the symmetry of the distribution of W_s with respect to zero. Noting this we derive

$$P\{\max_{0\le s\le t} W_s > a\} = P\{\tau_a \le t\} = 2P\{\tau_a \le t, W_t \ge a\} = \tag{2.9}$$
$$= 2P\{W_t \ge a\} = \frac{2}{\sqrt{2\pi t}} \int_a^\infty e^{-\frac{x^2}{2t}} dx = \sqrt{\frac{2}{\pi}} \int_{\frac{a}{\sqrt{t}}}^\infty e^{-\frac{y^2}{2}} dy .$$

From (2.9) we conclude that $P\{\tau_a < \infty\} = \lim_{t\to\infty} P\{\tau_a < t\} = 1$, but $E\tau_a = \infty$.

The joint distribution of W_t and $\max_{0\le s\le t} W_s = \zeta_t$ may also be computed (see Ito and McKean [1]). The density function $p_{W,\zeta}(x,y)$ of this two-dimensional random variable is the following

$$p_{W,\zeta}(x,y) = \left(\frac{2}{\pi t^3}\right)^{1/2} (2y-x) e^{-\frac{(2y-x)^2}{2t}}, \quad t \ge 0, y \ge 0, x \le y .$$

Now we dwell briefly on the local properties of Wiener trajectories. By remembering that the increments of Wiener process are independent, it is not difficult to prove that, for every fixed t, almost all Wiener trajectories W_t are not differentiable. Paley, Wiener, and Zygmund [1] established a stronger result: Wiener trajectories are nowhere differentiable with probability 1.

With probability one, the trajectories of Wiener process are Hölder continuous with exponent $\frac{1}{2} - \varepsilon$ for every $\varepsilon > 0$, and do not satisfy the Hölder condition with the exponent $\frac{1}{2}$. Wiener trajectories have infinite variation in every non-empty interval. For the proof of all these assertions, the reader is referred to Ito and McKean [1], Chapter 1.

The properties of symmetry and self similarity of the Wiener process play a highly significant role.

It is obvious that if $W_t(\omega)$ is a Wiener process, $W_0(\omega) = 0$, then the process $\widetilde{W}_t = -W_t(\omega)$ is also a Wiener process. No matter what the positive number α, the process $\widetilde{\widetilde{W}}_t = \alpha W_{t/\alpha^2}$ is a Wiener process as well. Indeed, the process $\widetilde{\widetilde{W}}_t$ has continuous trajectories and is a Gaussian process; $E\widetilde{\widetilde{W}}_t = \alpha E W_{t/\alpha^2} = 0$. The correlation function of the process $\widetilde{\widetilde{W}}_t$ has the form: $E\widetilde{\widetilde{W}}_s\widetilde{\widetilde{W}}_t = \alpha^2 E W_{s/\alpha^2} W_{t/\alpha^2} = \alpha^2\left(\frac{s}{\alpha^2} \wedge \frac{t}{\alpha^2}\right) = s \wedge t$. As was remarked above, these properties characterize the Wiener process.

There is another interesting transformation which preserves Wiener processes. Let W_t be a Wiener process. The process

$$\hat{W}_t = \begin{cases} 0, & t = 0, \\ t W_{1/t}, & t > 0 \end{cases} \tag{2.10}$$

is a Wiener process too. In order to prove this, we again make use of the fact that a mean zero Gaussian process with continuous trajectories and the correlation function $R(s,t) = s \wedge t$, is a Wiener process. The continuity of the function $\hat{W}_t$ may be broken at zero only, but one can see that for any $\varepsilon > 0$

$$\lim_{t \downarrow 0} \frac{|\hat{W}_t|}{t^{(1/2)-\varepsilon}} = 0 \quad \text{P-a.s.}$$

This can be checked, for example, with the law of iterated logarithm for $t \to \infty$ (see below). Thus, the functions $\hat{W}_t$ are continuous at zero. Taking into consideration that $E\hat{W}_t = 0$, $E\hat{W}_s\hat{W}_t = st\left(\frac{1}{s} \wedge \frac{1}{t}\right) = s \wedge t$, we deduce that $\hat{W}_t$ is a Wiener process. This transformation enables us to determine the properties of a Wiener process as $t \to 0$, by studying its properties as $t \to \infty$, and vice versa.

There is an interesting property, the so-called law of iterated logarithm which holds for almost all the Wiener trajectories: namely, with probability 1.

$$\overline{\lim_{t\to\infty}} \frac{W_t}{\sqrt{2t \ln \ln t}} = 1 , \quad \underline{\lim_{t\to\infty}} \frac{W_t}{\sqrt{2t \ln \ln t}} = -1 .$$

The proof of this statement may be found in Ito and McKean [1]. Using transformation (2.10), the law of the iterated logarithm becomes the local law of the iterated logarithm:

$$\overline{\lim_{t\downarrow 0}} \frac{W_t}{\sqrt{2t \ln \ln t^{-1}}} = 1 , \quad \underline{\lim_{t\downarrow 0}} \frac{W_t}{\sqrt{2t \ln \ln t^{-1}}} = -1$$

almost surely. The local law of the iterated logarithm, in particular, implies that a Wiener process starting from zero, returns to the point 0 with probability 1 in an arbitrarily small interval $(0,h)$, $h > 0$. We have used this property when stating that the variable $\theta = \sup\{t \in [0,1], W_t = 0\}$ is not a Markov time. Furthermore the law of the iterated logarithm implies that the set of zeroes of a Wiener trajectory is unbounded with probability 1 as $t \to \infty$.

Notice that the Lebesgue measure of the set $\{t : W_t = 0\}$ is zero with probability 1. Really, the time $\Lambda_\Gamma(t)$, a trajectory spends in the set $\Gamma \subset R^1$ up to time t, may be written in the form $\Lambda_\Gamma(t) = \int_0^t \chi_\Gamma(W_s)ds$ where $\chi_\Gamma(x)$ is the indicator of the set Γ. From this we conclude that $E\Lambda_\Gamma(t) = \int_0^t E\chi_\Gamma(W_s)ds = \int_0^t \int_\Gamma p(s,0,y)dy\,ds$. In particular, if Γ has zero measure, then $E\Lambda_\Gamma(t) = 0$ and therefore $\Lambda_\Gamma(t) = 0$ with probability 1. Many interesting properties of the Wiener process are available in the books of Levy [1] and Ito and McKean [1].

A collection of r independent Wiener processes $(W_t^1, \cdots, W_t^r) = W_t$, $t \geq 0$, defined on a probability space $(\Omega, \mathcal{F}, P)$, is said to be an *r-dimensional Wiener process*. The measure in the space $C_{0,\infty}(R^r)$ of continuous functions on $[0, \infty)$ with values in R^r, induced by the process $W_t(\omega)$ in $C_{0,\infty}(R^r)$, is called a *Wiener measure in* $C_{0,\infty}(R^r)$.

A family of σ-fields $\mathcal{F}^W_{\leq t} = \sigma(W_s^1, \cdots, W_s^r ; s \in [0,t])$ is associated with the process W_t. Corresponding to this family of σ-fields (as to any

increasing family of σ-fields) there is a collection of Markov times: namely, non-negative random variables $\tau(\omega)$ for which $\{\tau(\omega) \leq t\} \in \mathcal{F}^W_{\leq t}$, $t \geq 0$. This concept of a Markov time, if considered in connection with the one-dimensional process W_t^1, is broader than that introduced before, because the σ-fields $\mathcal{F}^W_{\leq t}$ are broader than $\mathcal{F}^{W^1}_{\leq t}$. For these new Markov times the strong Markov property is fulfilled as well: $W_{\tau+t} - W_\tau = \widetilde{W}_t$ is an r-dimensional Wiener process independent of events belonging to the σ-field $\mathcal{F}^W_{\leq \tau} = \{\mathfrak{A} \in \mathcal{F} : \mathfrak{A} \cap \{\tau \leq t\} \in \mathcal{F}^W_{\leq t}$ for any $t \geq 0\}$.

We will note some differences in the behavior of r-dimensional Wiener processes as we vary the dimension r (these distinctions follow from the properties of the one-dimensional Wiener process and from the independence of the components $(W_t^1, \cdots, W_t^r)$). As we observed, after any time t, a Wiener process in R^1 will hit zero and, hence, will hit any other point of the line as well. For $r \geq 2$, however, with probability 1, the Wiener trajectory will never hit any fixed point of R^r. But for $r = 2$, with probability 1, the Wiener trajectory does hit any open set after any fixed time. For $r > 2$, this is not the case. What is more, one can prove that, for $r \geq 3$, $\lim_{t\to\infty} |W_t| = \infty$ a.s..

Finally, we remark that the Wiener process in R^r is invariant with respect to rotations: if W_t is an r-dimensional Wiener process, then $\widetilde{W}_t = Q W_t$ is also a Wiener process for any orthogonal matrix Q. This assertion follows from the definition of W_t and from the properties of the Gaussian distribution. The family of Wiener processes $W_t^x = x + W_t$ in R^r is invariant with respect to the group of all rigid motions of the space.

Concluding this section it is worthwhile to draw the reader's attention to the close connection between the Wiener measure (or process) and the Laplace operator[4] Δ.

---

[4]More precisely, this is the operator $\frac{1}{2}\Delta$ rather than Δ. This is seen in considering Cauchy's problem. In the case of Dirichlet's problem for the homogeneous equation $\Delta u = 0$, this difference is of course imperceptible.

The simplest example of this connection is the following: the mean value $E\,g(W_t^x) = u(t,x)$ (here $g(x)$ is a continuous bounded function in R) is a solution of Cauchy's problem for the simplest heat conduction equation:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u\,, \quad \lim_{t \downarrow 0} u(t,x) = g(x)\,.$$

This is straightforward since according to the definition of W_t^x

$$u(t,x) = E\,g(W_t^x) = \int_{R^r} g(y)\,\frac{1}{(\sqrt{2\pi t})^r}\,e^{-\frac{|x-y|^2}{2t}}\,dy\,.$$

Now we shall show how the Wiener process is linked with Dirichlet's problem for the Laplace operator.

Suppose we are given a bounded domain $D \subset R^r$ with a boundary ∂D, the domain D, for simplicity, being assumed convex. Consider the Dirichlet problem in the domain D:

$$\Delta u(x) = 0\,,\ x \in D\,, \quad \lim_{x \to x_0 \in \partial D} u(x) = \psi(x_0)\,, \tag{2.11}$$

where $\psi(x)$ is a continuous function defined on ∂D. We will show how the function $u(x)$ which is a solution of problem (2.11) can be represented in the form of the average value of a proper functional of the Wiener trajectory, or (what is the same) in the form of an integral with respect to the Wiener measure.

Let W_t be a Wiener process in R^r. We will introduce the Markov time $\tau_D^x = \tau_D^x(\omega) = \inf\{t : x + W_t \in \partial D\}$ —the first hitting time to the boundary of the domain D starting from $x \in D \cup \partial D$. Let us show that $P\{\tau_D^x < \infty\} = 1$ for arbitrary $x \in D$. In fact, let $W_t^{1,x}$ be the first component of $W_t^x = W_t + x$ and let a number a be such that $D \subset \{x \in R^r; x^1 < a\}$. We shall denote by $\tau_a^{1,x}$ the first hitting time of the one-dimensional process $W_t^{1,x}$ to the point a: $\tau_a^{1,x} = \inf\{t : W_t^{1,x} = a\}$. It is clear that $\tau_D^x < \tau_a^{1,x}$ so the finiteness of τ_D^x comes from the fact that $P\{\tau_a^{1,x} < \infty\} = 1$.

The last equality results from (2.9). At the time τ_D^x, the point $W^x_{\tau_D^x}$ belongs to ∂D, thus one can consider $\psi(W^x_{\tau_D^x})$ and the function

$$V(x) = E\,\psi(W^x_{\tau_D^x}) . \tag{2.12}$$

Let us show that the function $V(x)$ is a solution of Dirichlet's problem (2.11).

First we will make sure that the function $V(x)$ defined by (2.12) satisfies the boundary conditions, i.e. $\lim_{x \to x_0 \in \partial D} V(x) = \psi(x_0)$. To that end, let us draw the line of support Γ (for the sake of visualization, we shall speak of the two-dimensional case) through the point $x_0 \in \partial D$ (see Fig. 2). Since the Wiener process is invariant with respect to the group of movements of the plane, the point x_0 may be considered as coinciding with the origin and the line Γ –with the x^2-axis. Let $\tau_0^{1,x}$ be the first hitting time of the first component $W_t^{1,x}$ of the Wiener process W_t^x to the point 0. Then noting equality (2.9) we obtain

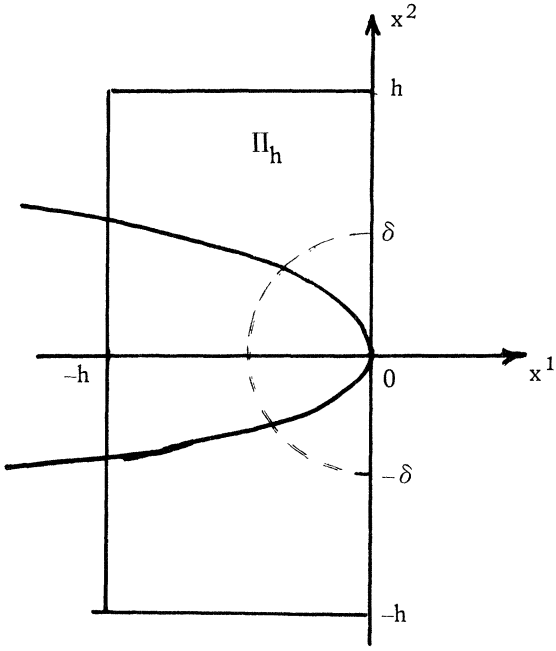

Fig. 2

$$P\{\tau_D^x > t\} < P\{\tau_0^{1,x} > t\} = \frac{2}{\sqrt{2\pi}} \int_0^{|x^1|t^{-1/2}} e^{-\frac{y^2}{2}} dy \to 0 \qquad (2.13)$$

as $x^1 \to 0$. Formula (2.9) also readily implies that, for any fixed $h > 0$, the probability of leaving the rectangle Π_h shown in Fig. 2 before time t across the horizontal sides or across the side lying on the line $x^1 = -h$, tends to zero as $t \downarrow 0$ uniformly in the set of the initial points $x \in D \cap U_\delta(0)$, where $U_\delta(0) = \{x : |x| < \delta\}$. This remark along with (2.13) implies that $P\{W^x_{\tau_D^x} \in \gamma_h\} \to 1$ as $x \to 0$, where $\gamma_h = \partial D \cap \Pi_h$. Now it is easy to verify that $V(x)$ takes on the prescribed boundary values:

$$|V(x) - \psi(0)| = |E(\psi(W^x_{\tau_D^x}) - \psi(0))| \le E|\psi(W^x_{\tau_D^x}) - \psi(0)|\chi_{\gamma_h} +$$

$$+ E|\psi(W^x_{\tau_D^x}) - \psi(0)|(1 - \chi_{\gamma_h}) \le \sup_{x \in \partial D, |x-0| \le 2h} |\psi(x) - \psi(0)| +$$

$$+ 2 \max_{y \in \partial D} |\psi(y)| \times P\{W^x_{\tau_D^x} \notin \gamma_h\}. \qquad (2.14)$$

Here $\chi_{\gamma_h}(\omega) = 1$ for $\omega \in \{W^x_{\tau_D^x} \in \gamma_h\}$ and $\chi_{\gamma_h}(\omega) = 0$ on the rest of Ω. The first summand on the right-hand side of (2.14) can be made arbitrarily small by making h small. The second summand, as it has been shown above, tends to zero for a fixed h as $x \to 0 \in \partial D$. Hence,

$$\lim_{x \to x_0 \in \partial D} V(x) = \psi(x_0).$$

Now we will show that the function $V(x)$ is continuous for $x \in D$. First let $\bar{\tau}^x$ be a Markov time with respect to the family of σ-fields $\mathcal{F}^W_{\le t}$, which satisfies $P\{\bar{\tau}^x \le \tau_D^x\} = 1$. Let $\eta = \eta(\omega) = W^x_{\bar{\tau}^x}$. Using the strong Markov property one can write down the following chain of equalities:

$$V(x) = E\,\psi(W^x_{\tau^x_D}) = E(E(\psi(\tilde{W}^\eta_{\tau^\eta_D})|\eta)) = \tag{2.15}$$

$$= EV(\eta) = EV(W^x_{\hat{\tau}^x}) ,$$

where $\tilde{W}_t$ is a Wiener process independent of W_t. Let us introduce the Markov time $\hat{\tau} = \hat{\tau}^{x,y} = \tau^x_D \wedge \tau^y_D$, $x, y \in D$. Using (2.15) leads to

$$V(x) - V(y) = E(V(W^x_{\hat{\tau}}) - V(W^y_{\hat{\tau}})) . \tag{2.16}$$

One of the points $W^x_{\hat{\tau}}$ or $W^y_{\hat{\tau}}$ belongs to ∂D, and the distance of the other point from D is bounded by $|x-y|$, because the trajectory $W^y_{\hat{\tau}}$ is a translation of $W^x_{\hat{\tau}}$ by the vector $(y-x)$. Hence, on account of the foregoing bounds near the boundary we have:

$$|V(W^x_{\hat{\tau}}) - V(W^y_{\hat{\tau}})| \to 0 \quad \text{as} \quad |x-y| \to 0 \quad \text{a.s.}$$

Noting that the function $V(x)$ is bounded, we deduce from (2.16) that $|V(x)-V(y)| \to 0$ as $|x-y| \to 0$.

Let us show that the function $V(x)$ given by (2.12) has the mean-value property: for any circle K having its center at a point $x \in D$ and lying entirely in $D \cup \partial D$, the function $V(x)$ satisfies

$$V(x) = \int_K V(y)\, m(dy) \tag{2.17}$$

where $m(dy)$ is a uniform distribution on the circle with the condition $m(K) = 1$. Indeed, let us put $\tau^x_K = \min\{t : W^x_t \in K\}$; that is, the first hitting time of the circle K starting from $x \in D$. The random variable τ^x_K is a Markov time satisfying $P\{\tau^x_K < \tau^x_D\} = 1$, and by (2.15)

$$V(x) = EV(W^x_{\tau^x_K}) .$$

Further, remembering that the Wiener process is invariant with respect to rotations one can conclude that when starting from the center of the circle K, the distribution at the first exit time from K is uniform: that is, $P\{W^x_{\tau^x_K} \in \gamma\} = m(\gamma)$ for every arc $\gamma \subset K$. Therefore,

$$V(x) = EV(W^x_{\tau^x_K}) = \int_K V(y)\, m(dy) .$$

So, we have shown that $V(x)$ is a continuous function which takes on the prescribed boundary values and has the mean-value property. As is known in the theory of differential equations, a solution $u(x)$ of Dirichlet's problem (2.11) possesses these properties too. Consequently, to prove that $u(x) = V(x)$, it remains to note that there is but one continuous function taking the value $\psi(x)$ on ∂D and satisfying the mean-value property. Indeed, suppose on the contrary, that there are two such functions and that $a(x)$ is their difference. For $x \in \partial D$, the function $a(x)$ vanishes identically and satisfies the mean-value property as well. Assume that $a(x)$ attains its maximum for $x = x_0 \in D$. Let us draw a circle with center at x and with radius equal to the distance from x to ∂D (Fig. 3).

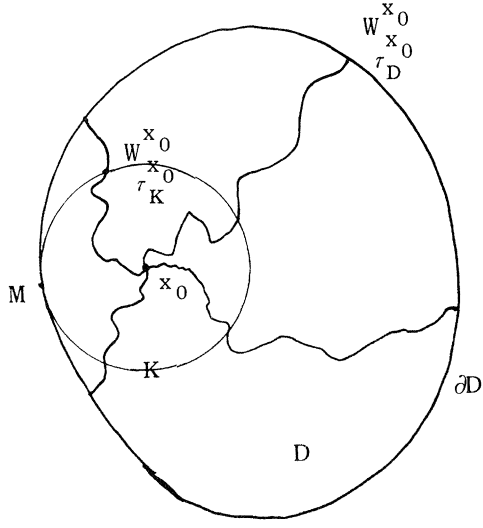

Fig. 3

Using the continuity of $a(x)$ and the mean-value theorem yields that $a(x_0) = a(M) = 0$ (M is shown in Fig. 3). Thus, $a(x) \equiv 0$ for $x \in D$ and hence $u(x) \equiv V(x)$.

So we have established that the solution of Dirichlet's problem can be represented in the form

$$u(x) = E\,\psi(W^x_{\tau^x_D}) . \tag{2.18}$$

Intuitively this means that, for the value of $u(x)$ at any point $x \in D$ to be determined, we are to let out Wiener trajectories from the point x, then to watch where they first hit ∂D and what the boundary values are at the points of the first hitting ∂D, and then to average these values over all the trajectories $W^x_t(\omega)$.

§1.3 *Stochastic differential equations*

Our next goal now is to construct a family of measures in $C_{0,\infty}(R^r)$ (or, equivalently, a family of random processes with continuous trajectories) which corresponds to an elliptic, possibly degenerate, second-order differential operator of general form. As we have seen in the previous section, a family of Wiener processes W^x_t is associated with the operator $\frac{1}{2}\Delta$. Now suppose that we are given a second-order operator

$$\tilde{L} = \frac{1}{2}\sum_{i,j=1}^{r} a^{ij}\frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{r} b^i \frac{\partial}{\partial x^i}$$

with constant coefficients and non-negative characteristic form: $\sum a^{ij}\lambda_i\lambda_j \geq 0$, $a^{ij} = a^{ji}$. Consider the family of random processes

$$\tilde{X}^x_t = x + \sigma W_t + bt \tag{3.1}$$

where $x \in R^r$ is an initial point, and σ is a real matrix[5] such that

[5] The existence of such a matrix follows from the fact that the matrix (a^{ij}) is non-negative.

$\sigma \cdot \sigma^* = (a^{ij})$, $b = (b^1, \cdots, b^r)$. Knowing the density function of the random variable W_t it is easy to evaluate the distribution of $\tilde{X}_t^x$ and to make sure that the function $u(t,x) = E\, g(\tilde{X}_t^x)$ is a solution of Cauchy's problem

$$\frac{\partial u}{\partial t} = \tilde{L}u, \quad \lim_{t \downarrow 0} u(t,x) = g(x),$$

where $g(x)$, $x \in R^r$, is a bounded continuous function. Therefore, the Gaussian process $\tilde{X}_t^x$ and the corresponding measure in $C_{0,\infty}(R^r)$ are related in a natural way to the operator $\tilde{L}$.

Now consider the operator of general form

$$L = \frac{1}{2} \sum_{i,j=1}^{r} a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{i=1}^{r} b^i(x) \frac{\partial}{\partial x^i}, \quad \sum_{i,j=1}^{r} a^{ij}(x) \lambda_i \lambda_j \geq 0,$$

with variable coefficients. The question suggests itself: what stochastic process X_t^x is related to this operator L? If the coefficients of the operator possess some continuity properties, then it is natural to expect that, near the point x, such a process X_t^x is close to the process $\tilde{X}_t^x$ defined in (3.1) with σ and b held fixed to their values at point x. In other words, the desired process must satisfy the differential equation

$$dX_t^x = \sigma(X_t^x)\, dW_t + b(X_t^x)dt, \quad X_0^x = x, \tag{3.2}$$

where $b(x) = (b^1(x), \cdots, b^r(x))$, $\sigma(x)\sigma^*(x) = (a^{ij}(x))$.

Later on we will see that actually such is the case, but first we must make sense of equation (3.2). To make the meaning of equation (3.2) more precise, we will follow the ideas of K. Ito. Namely, we integrate it from 0 to t, taking into account the initial condition:

$$X_t^x - x = \int_0^t \sigma(X_s^x)\, dW_s + \int_0^t b(X_s^x)\, ds. \tag{3.3}$$

If the matrix σ does not depend on x, then the existence and uniqueness of a solution of such an equation for a fixed trajectory W_s, follow from the corresponding results for ordinary differential equations (non-stochastic), whenever one assumes, for example, that the function $b(x)$ satisfies a Lipschitz condition. In the general case, the situation is more complicated: here the first integral in (3.3) must be given a meaning. But this is not easy to do, because the Wiener trajectories W_s have with probability one infinite variation on every interval.

We proceed now to construct Ito's stochastic integral and describe its properties (see, e.g. Gihman and Skorohod [1], [2], McKean [1], Wentzell [1]). The general outline is as follows (now W_s is supposed to be a one-dimensional Wiener process). To begin with, we define $\int_a^b f(s,\omega)\,dW_s$ for the "simple" real-valued functions $f(s,\omega)$, $s \in [a,b]$. Here it turns out that $\int_a^b f(s,\omega)\,dW_s = \eta(\omega)$ is a random variable such that $E|\eta(\omega)|^2 = \int_a^b E|f(s,\omega)|^2\,ds$. Therefore, to every simple function $f(s,\omega)$, the stochastic integral is a random variable $\eta(\omega)$. We introduce the Hilbert space $H^2([a,b]\times\Omega)$ of functions with the norm $\|f\|_{H^2} = \int_a^b E|f(s,\omega)|^2\,ds$ and the Hilbert space $L^2(\Omega)$ of the random variables $\eta(\omega)$, $\|\eta\|_{L^2} = E|\eta(\omega)|^2$. Then integration becomes a linear isometric mapping of the set $\tilde{H}^2_{a,b}$ of simple functions from $H^2([a,b]\times\Omega)$ into $L^2(\Omega)$. This mapping can be extended in a continuous way to the closure of the set $\tilde{H}^2_{a,b} \subset H^2([a,b]\times\Omega)$ with the isometry being preserved. Such a continuation defines an integral on the closure of the set of simple functions which turns out to be a sufficiently extensive set. Now, we proceed with the details.

Suppose we are given a one-dimensional Wiener process W_t, $t \geq 0$, and an increasing family of σ-fields N_t, $t \geq 0$. We assume that $\mathcal{F}^W_{\leq t} \subset N_t$ and that the increments $W_s - W_t$ do not depend on the σ-field N_t for $s > t$. Such a family of σ-fields N_t will be called a *family adapted* to the Wiener process W_t.