

INTEGRALS DEVISED FOR SPECIAL PURPOSES

E. J. MCSHANE

About forty-five years ago Professor T. H. Hildebrandt wrote a paper (Hildebrandt [1]) on integrals related to and extensions of the Lebesgue integral. At that time the Lebesgue integral, first invented in the preceding decade, had already displaced the Riemann integral from the monarchical position that it had occupied for over a generation. However, the subject was still new enough so that it was possible for Professor Hildebrandt to give an intelligible account of each one of the generalizations or extensions of the Lebesgue integral then extant, and to finish the job within a reasonable space of time. To do this today for the whole field of integration, covering all topics adequately, and still remaining within the bounds of a one hour talk, would be an impossible task. I have therefore chosen to confine myself to one particular line of integration theory. I shall pass by all developments of new types of integral for the sake of increasing the generality of the process of integration. Likewise, I shall pass by all research designed to increase our knowledge of already existent forms of integrals. The research that I intend to report on is extremely diverse in nature, but all the developments that I shall mention have one aspect in common. In each case a mathematician needed an integration process to attain some goal, found that none of the traditional types of integration did exactly what he wanted and proceeded to invent an integration method capable of producing the results that he wished.

A common feature of the integrals about which I wish to speak is that they are all what might be called "second growth" integrals. Probably none of them would have been thought of at all if the man who defined it had not been thoroughly familiar with an assortment of standard integration procedures, and a number of them in fact cannot even be defined without using concepts that themselves would never have arisen if it had not been for the development of the theory of the Lebesgue integral. This means that the one-time dominance of the Riemann integral has not been replaced by a similar dominance of any other integral. Competence in this field demands familiarity with a large assortment of concepts connected with procedures that might be called integration.

A compact example of the sort of theory that I am speaking of

Presidential address delivered before the Annual Meeting of the Society, January 25, 1963; received by the editors June 2, 1963.

occurs in a paper on conditional probability operators by Professor Robert Cogburn (Cogburn [1]). Let G denote the space of all bounded real-valued Borel functions on the Borel line. I now quote from Cogburn. "For functions ξ on R to G we introduce the order Riemann integral on $[a, b]$, denoted $\int_a^b \xi(x) dx$, and defined to be the order limit of the usual Riemann sums, provided this limit exists in G . It is easily seen that the integral does exist whenever ξ is order continuous in G on the interval $[a, b]$."

This example shows in a small space the peculiarities that I have mentioned. First, Cogburn needs an integral which is not one of the standard integrals. Nevertheless, anyone familiar with the concept of order limit and also familiar with ordinary Riemann integration procedure can put the two concepts together in the manner indicated by Cogburn, and thus produce a new integral. The relevant portion of the theory is very briskly developed, under the assumption that the reader is already thoroughly familiar with ordinary integration processes and can furnish the details without too much trouble; and the theory is carried just exactly as far as the author needs for the purposes of the research being carried out.

One technical difficulty that has given me some trouble in the preparation of this talk is that it is no longer perfectly clear just what an integral is. The classical integrals have a multitude of different properties. The integrals devised for special purposes frequently are designed so as to be stronger than the classical integration procedures with regard to some chosen property, but in emphasizing one property heavily it frequently happens that others are lost. We shall see this in several instances. However, the vagueness of what an integral is can be shown to you fairly easily by a theorem on averaging operators which I wish to present to you in its own right.

Let $(\Omega, \mathcal{A}, \mu)$ be a probability triple; that is, Ω is a set, \mathcal{A} is a σ -algebra of subsets of Ω , and μ is a countably additive (non-negative) measure on \mathcal{A} with $\mu\Omega=1$. Investigations in turbulence led to the study of operators with the following three properties.

(1) For some p in $[1, \infty]$, A is a linear operator on $L_p(\Omega, \mu)$ such that for all f in $L_p(\Omega, \mu)$

$$\int_{\Omega} |(Af)(\omega)|^p \mu(d\omega) \leq \int_{\Omega} |f(\omega)|^p \mu(d\omega).$$

(2) If $f \in L_p(\Omega, \mu)$ and g is essentially bounded and measurable on Ω , then

$$A(gAf) = (Ag) \cdot (Af)$$

(so that $(Ag) \cdot (Af) \in L_p(\Omega, \mu)$).

(3) If I is the function identically 1 on Ω , $AI = I$.

In 1960 Professor Gian-Carlo Rota proved (Rota [1]) that every averaging operator is necessarily a conditional probability operator; that is, there is a σ -subalgebra \mathfrak{B} of \mathfrak{A} such that whenever $f \in L_p(\Omega, \mu)$, Af is \mathfrak{B} -measurable and satisfies

$$\int_B (Af)(\omega) \mu(d\omega) = \int_B f(\omega) \mu(d\omega)$$

for every set B in \mathfrak{B} . Earlier, this had been proved by Mrs. Shu-Teh-Chen Moy (Moy [1]) under the stronger hypothesis that A satisfies (1) for $p=1$ and for $p=\infty$, hence for all p in $[1, \infty]$.

The averaging operators A do not at a glance look much like integrals. They possess stronger algebraic properties than ordinary integrals do, for in addition to linearity they have the multiplicative structure shown in (2). When exhibited as conditional probabilities they come closer to the usual types of integrals, for in some instances we can write conditional probabilities as Lebesgue-Stieltjes integrals, and when \mathfrak{B} consists of just two sets, Ω and the empty set, Af is a constant, the integral of f with respect to μ . I therefore do not feel completely certain as to whether or not a theorem about averaging operators can justly be called a theorem about integration. However, quite apart from classification, I feel that the theorems by Mrs. Moy and Professor Rota are interesting and important, and I have not the slightest hesitancy about presenting them.

One of the most conspicuous advantages of the Lebesgue integral, as contrasted with the Riemann integral, is the superiority of the convergence theorems associated with it. These theorems in turn rest on the countable additivity of Lebesgue measure, as contrasted with the finite additivity of Jordan content. It would not have been unreasonable to believe that with the invention of the Lebesgue integral, finitely additive measures would have lost all interest and importance. In spite of this, finitely additive measures have in recent years returned from limbo and proved useful in many cases. One of the simplest cases is given in full detail in the third chapter of Dunford and Schwartz (Dunford and Schwartz [1]). Although this is very well known, I shall give a brief sketch of it to refresh your memories and to provide a starting point for a later discussion. For simplicity, we shall suppose that \mathfrak{A} is a field of subsets of a set S , that m is a non-negative bounded finitely additive measure defined for all sets in \mathfrak{A} , and that all functions f , etc., mentioned are finite real-valued functions. A function f is simple if it assumes finitely many values,

each on a set belonging to \mathcal{G} . The integral of such a function is defined in the obvious way; it is obtained by multiplying each value of the function by the m -measure of the set on which it has that value, and adding the products. A sequence of functions f_1, f_2, \dots will be said to converge in measure to a function f_0 if for each positive ϵ there is an n_ϵ such that whenever n is greater than n_ϵ , there exists a set E_n with $m E_n < \epsilon$ such that

$$|f_n(x) - f_0(x)| < \epsilon \text{ for all } x \text{ in } S - E_n.$$

Now suppose that f_0 is a function for which it is possible to find a sequence f_1, f_2, \dots of simple functions with the following two properties:

1. f_n converges to f_0 in measure;
2. $\lim_{m, n \rightarrow \infty} \int_S |f_m(x) - f_n(x)| m(dx) = 0$.

From the second of these properties it is easily seen that the integrals of the f_n converge to some real number as limit, and in view of the first of these properties it is reasonable to assign this limit as the integral of f_0 with respect to the measure m .

The integral thus defined is in fact a generalization of the Lebesgue (or Lebesgue-Stieltjes) integral, since it reduces to that integral if the measure m happens to be countably additive. The reason for the generalization, however, is not merely a wish for greater generality in itself. When the Lebesgue integral is thus extended, the integral as just defined serves to express the most general form of linear functionals on spaces L_∞ (Dunford and Schwartz [1, p. 296]).

Our next integration procedure has its roots in the observation, in 1827, by the British biologist, Robert Brown, that a microscopic particle in a fluid underwent erratic changes of position. This incessant motion is now known as the Brownian motion and is recognized as being caused by the impact of molecules against the microscopic particle. A. Einstein and M. V. Smoluchowski provided a mathematical-physical model for this motion. Einstein (Einstein [1]) remarked that the microscopic particles constituted essentially a population of large molecules among the gas molecules, so that the theory of diffusion should be applicable to them. After a time interval Δt , the displacements, say in the direction of the x -axis, of a large aggregate of particles should be distributed as though the particles had all diffused for time Δt from an original concentration at $x=0$. This implies that the distribution is normal, with a variance proportional to Δt . (Einstein of course realized that this is applicable only when Δt is much larger than the mean time between molecular impacts.) In 1922 Norbert Wiener (Wiener [1], [2], [3]) provided the

mathematical model that was needed to support this physical theory. If we observe a particular particle starting at the origin at time $t=0$ its x coordinate at all future times will be given by a function $x(t)$ ($t \geq 0$). This function x is determined by chance; the details of the motion depend upon the unpredictable impact of molecules against the particle. Thus the functions $(x(t): 0 \leq t < \infty)$ may reasonably be looked upon as the points of a probability space; events will be sets of points (that is to say, sets of functions) in this space. Let Ω be a set of labels sufficient to distinguish the functions on $[0, \infty)$ from each other; it would be quite satisfactory to let the members of Ω be the names of the functions themselves. If ω is any point of Ω , the function labeled by it will be denoted by the symbol $x(\cdot, \omega)$, or $(x(t, \omega): 0 \leq t)$. It is the set of functions x , or alternatively the set Ω of labels for these functions, that should serve as the points of our probability space. Thus our problem is to introduce a probability measure into the space Ω in such a way as to retain the properties required of a physical model of the Brownian motion. That is, corresponding to each interval $[t_1, t_2]$ in the positive real axes, the increment $x(t_2, \cdot) - x(t_1, \cdot)$ should be a normally distributed random variable, and its mean should be 0 and its variance should be a constant multiple of $t_2 - t_1$. Also, the increments corresponding to nonoverlapping intervals should be statistically independent random variables. Wiener showed, not merely that this could be done, but that it could be done in such a way that the continuous functions had probability measure 1. That is, as a physical model would essentially demand, the probability of discontinuous motion was 0. In fact, he even showed that for every positive number $\alpha < 1/2$, the set of functions satisfying a Hölder condition with exponent α is a set of probability measure 1.

Instead of giving you any of the details of Wiener's original method of defining this measure in function space, I would like to give a sketch of a somewhat different definitional procedure recently worked out by S. H. Coleman (Coleman [1]). This is based on an extension (Stone [1], McShane [1]) of the Daniell integral (Daniell [1]). Instead of real-valued functions, let us consider a compact set X and the set X^T of all functions on the set $T = [0, \infty)$ to the set X . For the set Ω of labels to distinguish these functions we use X^T itself, and we topologize it with the usual product-space topology. If A is a subset of T , a function f on Ω will be said to be based on A if whenever ω and ω' are any two points of Ω such that the equation $x(t, \omega) = x(t, \omega')$ is satisfied for all points t in A , then the equation $f(\omega) = f(\omega')$ also is satisfied. Instead of assuming the existence of a joint distribution function for any finite set of increments, we shall equivalently assume that whenever A is a finite subset of T there is an integral I_A defined

on the class of all continuous functions based on A and having $I_A(1)=1$. We assume also that these integrals I_A satisfy a rather evident consistency condition, so that the subscript A may be dropped. Let us now say that a function u on Ω is a U -function if there is a set S , directed by \geq , of continuous finitely based functions such that $u = \bigvee S = (\sup_{f \in S} f(\omega) : \omega \in \Omega)$. We can then define $I(u)$ to be $\sup\{I(f) : f \in S\}$. The U -functions turn out to be exactly those functions that are lower semi-continuous and bounded below on Ω . From their integrals we can proceed to the integrals of a much larger class of functions, including all Borel functions, by the usual Daniell technique. There is however one useful by-product. It can be shown (McShane [1]) that if S is a set of U -functions directed by \geq and if I is bounded on S , then the supremum $\sup\{u(\cdot) : u \in S\}$ is itself an integrable U -function, and its integral is the supremum of the integrals $I(u)$ for all u in S . Herein there is no assumption of denumerability. This makes it possible to prove that certain important and self-suggesting sets of functions are in fact measurable sets. For instance, for each positive α the set of functions which satisfy a Hölder condition of exponent α is a measurable set. Similarly, the set of all functions of bounded variation is a measurable set.

Although the definition just given puts only extremely mild hypotheses on the behavior of the distributions of the random variables $x(t, \cdot)$ for various fixed t , the case that has received by far the greatest amount of attention is that of the Wiener integral, in which the increments are independent, normally distributed, and have variances proportional to the length of the t -interval. The theory of this integral has been studied in considerable detail by R. H. Cameron, W. T. Martin, and their students (Cameron [1]–[6], Cameron and Fagan [1], Cameron and Hatfield [1], [2], Cameron and Martin [1]–[5]). The importance of functional integrals in modern physics will be clearly demonstrated by a glance at a recent article by S. G. Brush (Brush [1]). Also, Wiener integrals can be used to write explicit formulas for the solutions of certain partial differential equations. The first use of the symbolism of functional integration to express the solution of the Schrödinger equation was in the Princeton dissertation of R. P. Feynman. This was developed extensively and rigorously by Kac, Feynman, Cameron and others (Cameron [4], Friedrichs, Schwartz et al. [1], Kac [1], [2], Montroll [1], Rosenblatt [1]). As an indication of the type of theorem found in these investigations we cite from Cameron [4]. Let R be the strip $\{(t, \xi) : 0 < t < t_0, -\infty < \xi < \infty\}$, where t_0 is either a positive real number or $+\infty$. Let $\sigma(\xi)$ be defined for all real ξ , let θ be defined on R ,

and let a be a positive number. Under certain conditions on σ and θ , which we shall not specify, the boundary value problem

$$\frac{\partial^2 G}{\partial \xi^2} - a \frac{\partial G}{\partial t} + \theta(t, \xi)G = 0, \quad (t, \xi) \in R,$$

$$\lim_{t \rightarrow 0+} G(t, \xi) = \sigma(\xi) \quad \text{for almost all } \xi$$

has a unique solution, and that solution is given by the Wiener integral

$$G(t, \xi) = \int \exp \left\{ \frac{t}{a} \int_0^1 \theta[2\sqrt{(t/a)x(s)} + \xi] ds \right\} \cdot \sigma[2\sqrt{(t/a)x(1)} + \xi] d_w x.$$

Wiener measure is one particularly important example of a measure in an infinite dimensional space. Another, and at first glance quite different, type of measure in function space is called for in the study of quantum field theory. To gain at least a vague idea of why this is so, let us go back to ordinary quantum mechanics. Given a system which in classical terms contains N particles, each of these particles will have three coordinates of momentum, so the whole system will have $3N$ coordinates of momentum, which we can number in some selected order as p_1, \dots, p_{3N} . In classical quantum mechanics, to each state of this system there will correspond a square integrable complex function $(\psi(p_1, \dots, p_{3N}): -\infty < p_j < \infty, j=1, \dots, 3N)$. To each physical observable there will correspond an operator, say A , which is self-adjoint on $L_2(R^{3N})$, such that if we measure this observable on the system in the state given by ψ , the expected value of the result will be the inner product $\langle A\psi, \psi \rangle$. For example, the expected value of the component p_1 of momentum will be

$$\int_{R^{3N}} p_1 |\psi(p_1, \dots, p_{3N})|^2 dp_1 \dots dp_{3N}.$$

The dimensionality of the space over which we integrate is three times the number of particles. But in quantum field theory, there is no finite upper bound on the number of particles that may be present. Even if we consider only a bounded part of space, the field in it is made up of superposed oscillations having infinitely many possible frequencies and correspondingly having infinitely many possible energies and momenta. The analog of the integral just above would be the integral over some space of infinitely many dimensions. This

is a clear indication of a need for study of integration over Banach spaces, in particular over Hilbert space.

Here, however, we meet with serious trouble. In the applications for which we wish to have such an integral, it is frequently the case that the nature of the situation being studied is such that we clearly wish to have an integral which is invariant under unitary transformations. Suppose that H is a Hilbert space of countably infinitely many dimensions, and that in it we have a countably additive measure defined, with the property that each point of the space is interior to measurable sets having arbitrarily small diameter and finite measure. If this measure is unitarily invariant, the measure of the set consisting of all of H except the origin can easily be shown to be zero. This means that if we wish to hold to the requirement of unitary invariance, we must abandon the requirement of countable additivity.

For the sake of simplicity let us restrict ourselves to integration in Hilbert spaces. A function f will be said to be based on a subspace A of the Hilbert space H if for all points x in H , $f(x) = f(P_A x)$, where P_A is the operation of projection on A . A function f based on a finite dimensional subspace of H will be called a cylinder function. If we introduce a measure in each finite dimensional subspace, we can define the integral of f over the whole Hilbert space to be the integral of $f(P_A x)$ over the finite dimensional subspace A , whenever f is based on A . (Of course we assume the obviously necessary consistency relations between these integrals.) The most important single case is that of the Gaussian normal distribution with a parameter c ; that is, if $\{x_1, \dots, x_k\}$ is an orthonormal basis for the subspace A , the measure in A is given by

$$\mu(dx) = (2\pi)^{-k/2} C^{-k} \exp\left[-x_1^2 - \dots - x_k^2\right]/2c dx_1 \dots dx_k.$$

This has the required consistency properties; if a function f can be regarded as based either on A or on B , its integral as a function based on A is the same as its integral considered as a function based on B . From this integral we can deduce a finitely additive measure, applying however only to sets whose indicator functions are finitely based cylinder functions. We could use this measure to define a finitely additive integral in the way described earlier (Dunford and Schwartz [1]). However, it turns out that this integral is not sufficiently powerful. The definition that we outlined in discussing the Wiener integral suggests that we might try to define U -functions as the limits of rising sequences of finitely based continuous functions, the integral of

the U -function being defined as the limit of the integrals of the finitely based functions. This would lead to a Daniell type of integral. This attempt fails completely. Everything has integral zero. However, J. Schwartz devised something somewhat related to this that does succeed in producing an interesting integral (Dunford and Schwartz [1, p. 402]). Let f be any real-valued function defined on the Hilbert space H , and let \mathcal{E}_1 be the set of all the bounded linear maps of H into itself that have finite-dimensional ranges. Then for all F in \mathcal{E}_1 , $f(F^*(\cdot))$ is a cylinder function based on the range of F . Let Φ be the class of all real functions f on H such that (i) $f(F^*(\cdot))$ is μ -measurable for all F in \mathcal{E}_1 (so that its absolute value has an integral, finite or infinite) and (ii) the infimum $\|f\|$, for all projections P in \mathcal{E}_1 and all positive ϵ , of the quantity

$$\sup \left\{ \int_H |f(F^*(x))| \mu(dx) : F \in \mathcal{E}_1 \text{ and } |P - FP| < \epsilon \right\}$$

is finite. It can be shown that this function $\|\cdot\|$ indeed possesses the properties of a norm. Also, if f happens to be a cylinder function and is integrable, $\|f\|$ is the same as the integral of the absolute value of f . So if g is any function which is the limit of a sequence f_1, f_2, \dots of integrable cylinder functions for which $\|f_m - f_n\|$ tends to zero, then necessarily the integrals $\int_H f_n(x) \mu(dx)$ tend to a limit, which we naturally define to be the integral of g . The integral thus defined is linear and can without too much trouble be shown to have most of the properties of integrals defined in terms of finitely additive measures. Moreover, it is quite easily seen to be unitarily invariant. It has the virtue that the definition of the integral involves only the original Hilbert space and functions defined on it without any extension of the Hilbert space to a larger space or extension of the concept of functional on the Hilbert space. However, in spite of these virtues, its theory has not thus far been very extensively developed.

Another method of introducing integration into Hilbert spaces is to imbed them first in some larger space over which some completely additive integral can be defined (Friedrichs and Shapiro [1]; Friedrichs, Schwartz et al. [1]). Suppose that H is a separable real Hilbert space. If e_1, e_2, \dots is any orthonormal sequence in H , every point x in H determines a sequence of real numbers, namely $\langle x, e_1 \rangle, \langle x, e_2 \rangle, \dots$, where $\langle x, y \rangle$ denotes the inner product of x with y . Thus the Hilbert space is mapped linearly into a subset of the space \hat{H} of all sequences of real numbers; the mapping, however, depends on the choice of orthonormal basis. The space \hat{H} can be regarded

as the cartesian product space R^J , where J is the set of positive integers. To introduce Gaussian measure on \hat{H} all that we have to do is to introduce Gaussian measure on each factor space (using the same variance parameter for each factor) and then form the product measure in the usual way. Since H may be regarded as a subset of \hat{H} this immediately gives us a countably additive measure in H . But this measure in H is quite useless, since it turns out at once that the measure of H is zero. To get any use out of the measure in \hat{H} , we must be more ingenious.

First, let us at least diminish the dependence of the theory on the choice of the orthonormal basis. Suppose that e_1, e_2, \dots and e'_1, e'_2, \dots are two orthonormal bases for H . Every point x in the Hilbert space H determines two sequences, namely its coordinates in the e_j -system and its coordinates in the e'_j -system, and thus provides a one-to-one mapping of a portion of \hat{H} onto itself. We extend this as follows. For every sequence ξ_1, ξ_2, \dots of real numbers we first form the finite sums $S_n = \xi_1 e_1 + \dots + \xi_n e_n$ ($n = 1, 2, \dots$). The sequence corresponding to ξ_1, ξ_2, \dots under the change of orthonormal basis could then reasonably be taken to be the sequence $\lim_{n \rightarrow \infty} \langle S_n, e_1 \rangle, \lim_{n \rightarrow \infty} \langle S_n, e_2 \rangle, \dots$, provided that all of these limits exist. Friedrichs and Shapiro showed (Friedrichs et al. [1, Chapter 12]) that this process indeed provides a one-to-one mapping for almost all of \hat{H} onto almost all of \hat{H} , the mapping being measure preserving. They did more than this; given any countable aggregate of orthonormal bases, they showed that there exists a subset of \hat{H} consisting of almost all of \hat{H} on which the mappings just described are all 1~1 and measure preserving. This is called the "corona" corresponding to the system of orthonormal bases used. Since we cannot successfully integrate functionals on H by contracting the measure from \hat{H} down into H , we work in the opposite direction. Given a functional f defined on H , we try to extend it to \hat{H} , or rather to the corona of H , and integrate the extension. In order to do this, the functional f must have sufficiently good behavior so that by some process of continuous extension of its domain from H to the corona we obtain an extended function uniquely determined at least almost everywhere on \hat{H} and integrable over \hat{H} . As one might reasonably expect, the cylinder functions are of this type. Determining larger classes of extendable functions is far from trivial. Gross succeeded in showing (Gross [1]) that if we introduce the topology H_2 , which is the weakest topology on the Hilbert space H such that all Hilbert-Schmidt operators on H are continuous from H with topology H_2 to H with the usual topology, and if f is a bounded complex-valued functional that is uniformly con-

tinuous near the origin in the topology H_2 , then this function f can be integrated over H by the procedure that we have just mentioned.

I. E. Segal has defined a different type of integration over Hilbert space, in which countable additivity is abandoned in favor of the preservation of unitary invariance. His definition applies to general Banach spaces and not merely to Hilbert space; also he considers a noncommutative extension of integration, which we intend to talk about later. For simplicity we shall still confine ourselves to separable Hilbert spaces and to real-valued functions on them. Whenever we are given a probability measure on H , each point x determines a random variable $(\langle x, \xi \rangle: \xi \in H)$; this is a linear mapping of H into the space of random variables on H , and the probability measure in H determines the joint distributions of all finite sets of such random variables. Now (as in Segal [5]) without assuming that any probability measure in H has been specified, let us suppose that we are given a linear mapping of H into some space of random variables. Any two such mappings will be called equivalent if for each finite set x_1, \dots, x_k of points of H , the joint distribution of the random variables corresponding to these x_j in one mapping is the same as the joint distribution of the corresponding random variables in the other mapping. By a *weak distribution* on H we shall mean an equivalence class of linear mappings of H into random variables. As a special case, the (isotropic) normal distribution is obtained by assigning a normally distributed random variable with variance $\|x\|^2$ to each vector x , orthogonal vectors corresponding to independent random variables. Now let C denote the set of all complex-valued cylinder functions u such that u is based on a finite-dimensional subspace of H and is bounded and Borelian on that subspace. Each such u can be integrated with respect to the joint distribution function of the random variables corresponding to a basis x_1, \dots, x_k of a finite-dimensional subspace on which u is based; its integral, or expected value, will be denoted by $E(u)$. If we denote the complex conjugate of a complex number a by a^* , we find readily that whenever u and v are in C , the function $E(uv^*)$ has all the properties of an inner product except that $E(uu^*)$ can be zero when u is not identically zero. Such functions u are called null-functions, and as usual we regard as equivalent any two functions that differ only by a null-function. The space of equivalence classes resulting is an inner-product or pre-Hilbert space, and can be completed in the usual way to form a Hilbert space H . Each u in C determines a transformation $T_u: v \rightarrow uv$ in C , hence in the completion H . This transformation is bounded (since u is bounded) and is self-adjoint for real u , hence the set of

all such T_u is a commutative algebra of bounded self-adjoint operators over H . We can choose a subset $\{T_{u_\alpha}: \alpha \in A\}$ whose algebraic combinations include all the T_u . Now we use the multiplicative representation of the operators (see, for example, McShane [5] or Segal [2]). There is a set $\{J_\alpha: \alpha \in A\}$ of bounded closed intervals of reals, a set B and a measure m on the cartesian product $\Pi: (\times_{\alpha \in A} J_\alpha) \times B$ such that H is isomorphic with $L_2(\Pi, m)$, each operator T_{u_α} ($\alpha \in A$) in H corresponding to the operation of multiplication by the coordinate z_α in $L_2(\Pi, m)$. Thus if $u \in C$, by hypothesis there is a polynomial p such that $T_u = p(T_{u_\alpha}, T_{u_{\alpha'}}, \dots)$, and this corresponds in $L_2(\Pi, m)$ to multiplication by $p(z_\alpha, z_{\alpha'}, \dots)$. C is thus mapped isomorphically onto the polynomial functions on Π . This latter set we close in $L_2(\Pi, m)$; the new elements are called "ideal functionals" (on H). They may fail to correspond to any true functional on H . The problem is to establish a correspondence between some set of functionals on H and a subset the "ideal functionals" in $L_2(\Pi, m)$ that will permit a useful interpretation of the integrals (with respect to m) of the "ideal functionals" as the integrals of the functionals over H .

The difficulty in establishing such a correspondence by limit processes arises from the poor convergence properties of the integral over H . It is possible for each positive ϵ to define a sequence u_1, u_2, \dots of cylinder functions such that $0 \leq u_1 \leq u_2 \leq \dots \rightarrow 1$ everywhere but the integrals of the u_n are all less than ϵ . If we are to establish a correspondence between a class of functionals on H and a subset of the class of ideal functionals just defined, in such a way as to permit us to interpret the integral of the ideal functional over $L_2(\Pi, m)$ as the integral of the true functional over H , we must restrict ourselves to functionals with strong continuity properties. A partial solution of this problem has been given by Gross (Gross [1]).

The rôle played by the Fourier transform in ordinary quantum mechanics is so important that we may feel confident that in order to be of any use in quantum field theory, the integrals over Hilbert space that we have been discussing must permit some sort of extension of the Fourier transform. To establish such a transform (following Segal [4]), we restrict our attention to Gaussian measures on Hilbert space, denoted by dN . If f is a cylinder function based on a finite dimensional subspace of H , and is a polynomial in the coordinates of points on that subspace, it is integrable with respect to Gaussian measure, and for each y in H the integral

$$F(y) = \int_H f(2^{1/2}x + iy)dN(x)$$

exists. This may not look much like a Fourier transform, but Gaussian measure has an exponential in it, and if (merely to keep the notation simpler) we take f to be based on a one-dimensional subspace, in which the one coordinate is x , we find that $dN(x) = \exp(-x^2/2c)dx$, and by means of an easy calculation we learn that $F(y) \exp(-y^2/4c)$ is the ordinary Fourier transform of $f(x) \exp(-x^2/4c)$. For every polynomial f , the mapping just defined is an isometry in the norm of $L_2(H, dN)$, and it extends to a unitary map of all of $L_2(H, dN)$ onto itself. This unitary map is called the Wiener transform. Its inverse is given by

$$f(y) = \int_H F(2^{1/2}x - iy)dN(x).$$

Although I have mentioned this Wiener transform first, it is not chronologically the first extension of the Fourier transform to a function space. Several years earlier, Cameron and Martin had defined a Fourier-Wiener transform (Cameron and Martin [3], [4]) of the function class $L_2(C, d_w x)$, where C is the set of all continuous functions (or, equivalently, of all real-valued functions) on $[0, 1]$, and $d_w x$ is Wiener measure. The formula given above for the Wiener transform bears a very close resemblance to that given by Cameron and Martin. This is of course not at all a coincidence. There is a transfer between the standard Wiener measure $d_w x$ and the Gaussian measure dN on a separable Hilbert space. Let H be the Hilbert space consisting of all absolutely continuous functions on the interval $[0, 1]$ that vanish at 0 and have square integrable derivatives; for each two such functions f, g we define the inner product to be the integral from 0 to 1 of $f'(g')^*$. This includes the piecewise linear continuous functions with finitely many corners, and we can choose an orthonormal basis e_1, e_2, \dots composed of such functions. By use of this orthonormal basis, we can extend H to a corona \hat{H} . But the extension from H to \hat{H} corresponds to the extension in function space from the set H to the set of real functions on $[0, 1]$ with Wiener measure. Moreover, this correspondence is such that for almost all the points of H , the corresponding function on $[0, 1]$ is a continuous function. We thus have a method of effecting a transition between the theory of Wiener measure and that of measure on Hilbert spaces. However, it does not follow that either one of them is thereby made obsolete. If for no other reason, we would be prevented from abandoning either one of these two integration processes because of the widely different nature of the generalizations of them.

As I mentioned in describing Segal's theory of integration on Hilbert spaces, Segal (Segal [3]) has also studied a process which he

calls a "noncommutative extension of abstract integration." Although similar extensions have also been considered by other people, I shall report only on what Segal has done, for a reason which I will mention later. In order to give some comprehension of this rather sophisticated mathematical theory, let us first start out with a space S , a σ -algebra \mathfrak{B} of subsets of S and a finite countably additive non-negative measure m on \mathfrak{B} . Define H to be $L_2(S, m)$. For each bounded measurable function f we define M_f to be that operator which carries h into $f \cdot h$ for all functions h in H . The operators thus defined thus constitute a commutative self-adjoint algebra \mathfrak{A} of bounded operators on H . If f is the indicator function of a set E in \mathfrak{B} (that is, $f(t) = 1$ if $t \in E$, $f(t) = 0$ if $t \notin E$) then M_f is a projection belonging to \mathfrak{A} . Moreover, every projection in \mathfrak{A} has this form. Thus to the measure m corresponds a real-valued function on the projections in \mathfrak{A} ; namely, if M_f is a projection in \mathfrak{A} , f being the indicator function of a set E , the number corresponding to the projection M_f is mE . On the other hand, suppose that we know the set S , the set of functions constituting $L_2(S, m)$, and the algebra \mathfrak{A} . Given a mapping from the projections in \mathfrak{A} into the non-negative reals that satisfies certain rather obvious additivity properties, this mapping will determine the measure m . Thus it will be possible to reword the theory of measure in such a way that it would be expressed in terms of the algebra \mathfrak{A} and the mapping m from projections in \mathfrak{A} onto reals. When this is done, it is possible to generalize the whole theory by abandoning the commutativity of \mathfrak{A} . We thus arrive at Segal's definition (Segal [3]):

"A gage space Γ is a system (H, \mathfrak{A}, m) composed of a Hilbert space H , a ring \mathfrak{A} of operators on H , and a non-negative valued function m on the projections in \mathfrak{A} , where m is completely additive, unitarily invariant, and such that every projection in \mathfrak{A} is the l.u.b. of projections on which m is finite. To be more explicit, m is said to be completely additive in case for any set S of mutually orthogonal projections in \mathfrak{A} with l.u.b. P , $m(P) = \sum m(Q) \{Q \in S\}$; and is *unitarily invariant* in case for every unitary operator $U \in \mathfrak{A}$ and projection $P \in \mathfrak{A}$, $m(U^*PU) = m(P)$. The function m associated with Γ is called a gage."

On this definition Segal bases a rather extensive theory of integration, including definitions of spaces L_p and analogs of the Riesz-Fisher theorem and the Radon-Nikodým theorem. However, it is manifestly impossible for me to give, in the few minutes at my disposal, even the rudiments of an idea of these developments.

These results of Segal did not spring full-panoplied from the brow of Jove. Their roots lie in the work of Neumann and Murray (Neu-

mann and Murray [1]); related ideas will be found in the work of Dixmier (Dixmier [1]); and Dye had generalized the Radon-Nikodým theorem to this noncommutative setting. The reason that I emphasize the work of Segal is that he was motivated in its development by a strong wish to put a sound mathematical foundation under quantum field theory. Thus he was not studying his noncommutative extension of abstract integration merely as a highly intriguing mathematical device, nor as a contribution to the theory of algebras of operators, but as an integration process invented for a special purpose, the advancement of the quantum field theory. It is for this reason that I regard Segal's contribution as coming so particularly within the scope of this talk.

Suppose that we are given some sort of device (enclosed of course in the canonical black box) with an input side and an output side. Some sort of stimulus is fed into the input side of the box, the total input between times t_1 and t_2 being $x(t_2) - x(t_1)$. We suppose that the mechanism inside the box is such that if an input of amount dx is fed into the box at time τ , the resulting contribution to output at a later time t will be $\Phi(t - \tau)dx$. We suppose also that the device is linear, so that the output at time t is the sum of all the contributions from all the inputs at earlier times; from our present completely nonrigorous point of view, the output is

$$\int_{-\infty}^t \Phi(t - \tau)dx(\tau).$$

There are in this world optimists who feel that any symbol that starts off with an integral sign must necessarily denote something that will have every property that they would like an integral to possess. This of course is quite annoying to us rigorous mathematicians; what is even more annoying is that by doing so they often come up with the right answer. If the function $x(\cdot)$ were of bounded variation, we would have of course an ordinary Stieltjes, or at worst Lebesgue-Stieltjes, integral before us. In the outstandingly important case of Gaussian input, the probability that $x(\cdot)$ will be of bounded variation is exactly zero. Consequently, the pseudo-definition that I have just presented for the thing that looks like an integral has probability 0 of meaning something. But the fact that this really undefined pseudo-integral still serves to give useful answers in physical and engineering situations, is enough to justify a strong suspicion that it is in fact a caricature of some rigorously definable mathematical entity. I now wish to sketch this definition and some of its generalizations.

Once again let us use the symbol Ω to mean a set of labels serving to distinguish all the real-valued functions on the real axis. We suppose that a probability measure P is given on some field \mathfrak{A} in Ω in such a way that for each interval $[t_1, t_2]$ the increment $x(t_2, \cdot) - x(t_1, \cdot)$ belongs to $L_2(\Omega, P)$, and that when $[t_1, t_2]$ and $[t_3, t_4]$ are nonoverlapping intervals the corresponding increments are orthogonal:

$$E([x(t_2) - x(t_1)][x(t_4) - x(t_3)]) = 0.$$

It follows at once that there is a function F defined on the real axis such that for every interval $[t_1, t_2]$ we have

$$E([x(t_2) - x(t_1)]^2) = F(t_2) - F(t_1).$$

Since F is monotonic nondecreasing, it can have only countably many discontinuities. Suppose that f is a step function on the real number system taking on the respective values c_1, \dots, c_n on the intervals $[a_0, a_1), [a_1, a_2), \dots, [a_{n-1}, a_n)$, where $a_0 < a_1 < \dots < a_n$ and all the a_j are points of continuity of F . It is then natural to define the integral of the step function f to be

$$\int_{-\infty}^{\infty} f(t) dx(t, \cdot) = \sum_{j=1}^n c_j [x(t_j, \cdot) - x(t_{j-1}, \cdot)].$$

Thus the integral is, as we should expect, a random variable, and it is easily seen to be defined and linear on the class of all step functions of the type just described. It is now rather easy to establish (Doob [1, p. 427]) the crucial equation

$$E\left(\left\{\int_{-\infty}^{\infty} f(t) dx(t, \cdot)\right\}^2\right) = \int_{-\infty}^{\infty} |f(t)|^2 dF(t).$$

Thus for the step functions of the type we have been considering, the stochastic integral defined by the finite sum just above maps step functions on the real number system into random variables in such a way as to preserve L_2 norm; that is, the stochastic integral is an isometry from part of $L_2(R, dF)$ to $L_2(\Omega, P)$. This gives us the opportunity to extend the range of definition of the integral from the class of step functions to a much larger class. Whenever f is in $L_2(R, dF)$ it is the limit of step functions of the type described. Each sequence of step functions of this type approaching f in the norm of $L_2(R, dF)$ is a Cauchy sequence, and hence maps onto a Cauchy sequence in $L_2(\Omega, P)$. The latter space being complete, the Cauchy sequence converges to a limit, which we define to be the integral of f with respect to $dx(t, \cdot)$. It is almost self-evident that this integral will possess the linearity and convergence properties that we would wish an

integral $\int f(t)dx(t)$ to have. It can also be shown to permit integration by parts,

$$\int_a^b h(t)dy(t, \omega) = h(b)y(b, \omega) - h(a)y(a, \omega) - \int_a^b y(t, \omega)h'(t)dt$$

for almost all ω , whenever h is absolutely continuous on the interval $[a, b]$ (Doob [1, p. 432]). Moreover, in gaining rigor we have not lost utility; this integral can be used to establish rigorously the theorems that the physicists wanted to use, such as Campbell's theorem (Doob [1, p. 433]).

There are however interesting and important problems for which the integral just defined turns out to be too restrictive because the integrands that we would like to use are themselves functions not only of t but also of ω . Suppose then that we are given a function $f(t, \omega)$ (t real, $\omega \in \Omega$), and we wish to define the integral of this f with respect to the stochastic process $x(t, \cdot)$ ($-\infty < t < \infty$). This can be done, but only under certain restrictions. One rather natural restriction is an analytic hypothesis which in effect asserts that for each value of t , the random variable $f(t, \cdot)$ ($-\infty < t < \infty$) is independent of the increments $x(t_2, \cdot) - x(t_1, \cdot)$ later than t . Besides this, it seems to be at least convenient and probably necessary to impose restrictions on the process $x(t)$. We assume it to be a Wiener process (Itô [2], [3], [6]) or at least a special kind of martingale (Doob [1, p. 436]). The method of defining this more general integral is conceptually quite close to that outlined above for the simpler integral of the functions of t alone, although of course the details are noticeably more complicated.

The stochastic integral just defined was in fact invented in order to permit the solution of certain stochastic differential equations, for example, the diffusion equation (Itô [2], [3], [6]; Doob [1, p. 277]). However, the subject can hardly be considered to be closed. There are other stochastic processes that occur in a quite natural manner and that lead to differential equations whose solution will call for new techniques. There seems to be room left for quite a bit of enjoyable research.

The stochastic integrals considered above are simple integrals. Multiple integrals can also be defined, and in fact were considered by Wiener as long ago as 1938 (Wiener [6]). The simple integral is, as we have seen, suggested to us by linear processes. Nonlinear processes lead just as naturally to multiple stochastic integrals, as Wiener has pointed out in a recent book (Wiener [7]). In order to get some idea of what is going on, let us once again abandon all pretense

at rigor. Once again we consider the black box with an input and an output side, the input being given by a random function x such that the amount of input during the time interval $[t_1, t_2]$ is $x(t_2) - x(t_1)$. In order to estimate the output at time t , we subdivide all times before t into short time intervals, which we can number from right to left as $\Delta t_1, \Delta t_2, \dots$. We shall assume that the output is adequately approximated by some infinitely differentiable (but not necessarily linear) function of all past increments $\Delta_j x$, where $\Delta_j x$ is the increment of the input x over the time interval Δt_j . The output is now represented by a function

$$\Phi(\Delta_1 x, \Delta_2 x, \dots).$$

This function we expand in a Taylor series:

$$\begin{aligned} \Phi(\Delta_1 x, \Delta_2 x, \dots) &= \Phi(0, 0, \dots) + \sum_j \Phi_j(0, 0, \dots) \Delta_j x \\ &\quad + \sum_{j,k} \Phi_{j,k}(0, 0, \dots) \Delta_j x \Delta_k x + \dots \end{aligned}$$

where the factors $\Phi_j(0, 0, \dots)$, etc. represent the partial derivatives of Φ with respect to the correspondingly numbered $\Delta_j x$. To be specific, let us look at the third term on the right. If t_j represents a point chosen in the interval Δt_j , we can change notation by replacing the symbol $\Phi_{j,k}(0, 0, \dots)$ by $\Psi(t_j, t_k)$. Now if the functions $x(t)$ are much better behaved than we have any reason to expect them to be, as the partitioning of the time axis becomes finer and finer, the term under consideration should approach a limit which could be reasonably indicated by the symbol

$$\int_{-\infty}^t \int_{-\infty}^t \Psi(t_1, t_2) dx(t_1) dx(t_2).$$

The other terms would with good luck have analogous limits. However, what we have already seen of the simple stochastic integral should convince us that a much more sophisticated approach than this is needed if we are to establish the existence of an integral with any sort of usable properties.

The subject of this talk being what it is, I shall confine my discussion of multiple stochastic processes to those in which stochastic integration plays an outstandingly prominent rôle; even these are too numerous for us even to sketch (Itô [5], [8], [11]; Wiener [6], [7]; McShane [2]), so we shall concentrate on two of them in which the underlying stochastic process is not required to be Gaussian. As before, we shall suppose that the set of functions on the real number system R is labeled by a set Ω of labels, and that we are given

a probability measure P in Ω . For simplicity we shall suppose that the process has no fixed points of discontinuity, and we shall suppose that the increments $x(t_2, \cdot) - x(t_1, \cdot)$ not only are independent, but have moments of every order, and that these moments define interval functions that are of bounded variation on every finite time interval. If N is any positive integer and I is an interval

$$I = (t_1, t'_1] \times \cdots \times (t_N, t'_N]$$

in R^N , and χ_I denotes the indicator function of the interval I , we define the stochastic integral of χ_I to be

$$\tilde{\chi}_I = [x(t'_1) - x(t_1)] \cdots [x(t'_N) - x(t_N)],$$

and we define a real-valued measure m_N by

$$m_N I = E(\tilde{\chi}_I).$$

By linearity we can define the stochastic integral of any step function on R^N , and we can define a corresponding real-valued integral of the step function by taking the expected value of the stochastic integral,

$$J_N f = E(f^-).$$

If f is a step function of N variables, we form from it a function of $2N$ variables defined by

$$f \otimes f = (f(t_1, \cdots, t_N) f(t_{N+1}, \cdots, t_{2N}) : t_1, \cdots, t_{2N} \in R).$$

This fails to be non-negative unless f happens to be of one sign. Nevertheless it can be shown that the equation

$$J_{2N}(f \otimes f) = E([f^-]^2)$$

is satisfied, so that the left member is non-negative. Furthermore, it is possible to show that the square root of the left member has all of the properties that we wish of a norm. It is this that takes over the rôle played by the L_2 norm in the definition of the simple stochastic integral. Since it is a real non-negative integral on a class of step functions, it can be extended by the Daniell process to an integral defined on a large class of functions and possessing the well-known linearity and closure properties. If f is any function in this class, and f_1, f_2, \cdots is a sequence of step functions converging to it according to the norm thus defined, namely

$$\lim_{n \rightarrow \infty} [J_{2N}([f_n - f] \otimes [f_n - f])]^{1/2} = 0,$$

then by isometry the stochastic integrals f_n^- of the step functions

form a Cauchy sequence in $L_2(\Omega, P)$. They therefore tend to a limit, and this limit is by definition the stochastic integral f^\sim .

Itô defines multiple stochastic integrals in quite another way, using as a tool the concept of random measure. If, for instance, $x(t, \cdot)$ ($0 \leq t \leq 1$) is the Brownian motion process, we have already defined

$$\int_0^1 f(t) dx(t, \cdot)$$

as a random variable for each f in $L_2([0, 1])$. If f is the indicator function of a measurable set E , the above integral can be denoted by $m(E, \cdot)$; for each E it is a random variable, and for disjoint sets E_1, E_2, \dots we have $m(E_j, \cdot)$ and $m(E_k, \cdot)$ independent if $j \neq k$ and also

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n m(E_j, \cdot) = m(\cup E_j, \cdot)$$

(convergence in probability). Random measures can be defined in any measure space.

Now let $X(t, \omega)$ ($-\infty < t < \infty, \omega \in \Omega$) define a separable process with stationary independent increments. In P. Lévy's formula for the characteristic function of ΔX there occurs a non-negative bounded measure $d\beta$; in the plane $\pi = R^2 = \{(t, u) : t \text{ real}, u \text{ real}\}$ we define two (real) measures μ, ν by

$$d\mu = (1 + u^2) dt d\beta(u), \quad d\nu = (1 + u^2) u^{-2} dt d\beta(u).$$

Now if E is any Borel set in π that is bounded and bounded away from the t -axis, for each ω we define $N(E, \omega)$ to be the number of points (t, u) in E for which $X(t+, \omega) - X(t-, \omega) = u$. This defines a random measure on the plane π , in terms of which the increments of X can be expressed:

$$\begin{aligned} X(b, \cdot) - X(a, \cdot) &= \gamma(b - a) + [\beta(0+) - \beta(0-)][B(b, \cdot) - B(a, \cdot)] \\ &= \lim_{n \rightarrow \infty} \int_a^b \int_{n^{-1} < |u| < n} [u dN(t, u) - u(1 + u)^{-1} d\nu(t, u)], \end{aligned}$$

where B is a Wiener process independent of the $N(E, \cdot)$. Next Itô defines a random measure $M(E)$ (E Borel, $E \subset \pi$, $\mu E < \infty$) by

$$\begin{aligned} M(E) &= \int_{B(0)} [\beta(0+) - \beta(0-)] dB(t) \\ &\quad + \lim_{n \rightarrow \infty} \int \int_{B(n)} [u dN(t, u) - u d\nu(t, u)], \end{aligned}$$

where $E(0)$ is the intersection of E with the t -axis and $E(n)$ is the set of points (t, u) of E with $n^{-1} < |u| < n$ ($n=1, 2, 3, \dots$).

If we define a "special rectangle-set" in R^{2p} (p a positive integer) to be a set of the form $E_1 \times \dots \times E_p$, where the E_j are disjoint Borel subsets of the plane π with $\mu(E_j) < \infty$, we can reasonably define the random measure of this set to be $M(E_1) \dots M(E_p)$. This leads immediately to a definition of the integrals of special simple functions, which are the linear combinations of indicator functions of special rectangle-sets. Now (trivially) the integral $I_p(f)$ of such a function f coincides with the integral of its symmetrization f^{sym} , obtained by permuting the arguments of f in all possible ways and averaging the results, and (not so trivially)

$$\begin{aligned} E(|I_p(f)|^2) &= E(|I_p(f^{\text{sym}})|^2) \\ &= \int_{\pi^p} |f(\xi_1, \dots, \xi_p)|^2 d\mu(\xi_1) \dots d\mu(\xi_p). \end{aligned}$$

Thus the domain of definition of the integral I_p can be extended to the closure in $L^2(\pi^p, (d\mu)^p)$ of the set of special simple functions; and under a mild continuity hypothesis on μ , this closure is all of that space.

This integral was defined for the purpose of analyzing the spectral structure of time-shift transformations (Itô [11]); for the Brownian motion this had been done previously by Kakutani (Kakutani [1]). Its relation to the integral described previously is not very transparent. At least they both specialize to Wiener's integral for Brownian-movement processes, and the overlap might be greater than is superficially evident.

In applications of stochastic integration the process $x(\cdot, \cdot)$ used is frequently either imperfectly known or intentionally idealized. It is therefore important to know that in some sense, "small" changes in the process will cause only "small" changes in the integral. Among stationary processes, one possible topology is the topology of convergence of all moments of $x(1, \cdot) - x(0, \cdot)$. If f, g are functions of m and n variables respectively, their stochastic integrals in the sense of McShane [2], denoted by \tilde{f}, \tilde{g} , will depend on the process. So will $E(\Phi(\tilde{f})\Psi(\tilde{g}))$, where Φ, Ψ are continuous functions of at most polynomial rate of growth. Subject to conditions on the magnitude of the moments of the increments $x(1, \cdot) - x(0, \cdot)$, this can be shown (McShane [2, p. 271]) to depend continuously on the process.

For simple stochastic integrals a more elementary topology is useful. Without going into details, this can be described as the topology

of the weak convergence of the joint distributions of $(x(t_1, \cdot), \dots, x(t_k, \cdot))$ for all finite sets $\{t_1, \dots, t_k\}$. With this topology, it can be shown (McShane [3], [4]) that under moderate requirements on f , the integral

$$y(t, \cdot) = \int_{-\infty}^t f(t - \tau) dx(\tau, \cdot)$$

defines a continuous mapping from stochastic processes $x(\cdot, \cdot)$ into stochastic processes $y(\cdot, \cdot)$. The continuity of solutions of stochastic differential equations has not yet been studied.

The connection between Wiener integrals and differential equations may be said to be seventeen years older than the Wiener integral itself, since Einstein deduced the distribution of the Brownian motion from the diffusion equation. But the principal development began with R. Feynman's dissertation of 1942 (see Feynman [1]). Let us first notice that if a functional f on $R^{[0,1]}$ depends only on the values of x at τ_1, \dots, τ_k , so that $f(x) = F(x(\tau_1), \dots, x(\tau_k))$, its Wiener integral is

$$\frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} F(x(\tau_1), \dots, x(\tau_k)) \cdot \exp \left(- \sum_{j=1}^k \frac{1}{2} \left[\frac{x(\tau_j) - x(\tau_{j-1})}{\tau_j - \tau_{j-1}} \right]^2 (\tau_j - \tau_{j-1}) \right) dx(\tau_1) \dots dx(\tau_k),$$

where the $x(\tau_j)$ are merely symbols for real numbers and $\tau_0 = 0$. This suggests denoting the Wiener measure symbolically by

$$d_w x = \frac{1}{N} \exp \left(- \frac{1}{2} \int_0^t \dot{x}(\tau)^2 d\tau \right) \Pi dx(\tau).$$

Now let us recall that the diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

led (Einstein [1]) to the Brownian motion and Wiener measure, which in turn can be used to write solutions of equations of the form

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} - U(x)u.$$

If we allow ourselves to replace t formally by $i\hbar t$ ($\hbar > 0$), we could hope that in some sense the corresponding "measure"

$$d_f x = \frac{1}{N} \exp \left(\frac{i}{2\hbar} \int_0^t \dot{x}(\tau)^2 d\tau \right) \Pi dx(\tau)$$

should permit us to write solutions of the equation

$$\frac{1}{i\hbar} \frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} - \frac{1}{\hbar^2} U(x)u,$$

which is the one-dimensional Schrödinger equation. The difficulty here is that the "measure" $d_f x$, and the "Feynman integral" based on it, are introduced merely as symbols, without any meaning assigned to them. This has not troubled some scientists, who used the non-existent "integral" in "proofs" that irritatingly wound up with the right results. In some cases it was remarked that if we replaced $i\hbar$ by $i\hbar + \delta$ ($\delta > 0$) the resulting convergence factor in the integrand would allow us to treat the integral in the same way as the Wiener integral, and the limit-passage $\delta \rightarrow 0+$ might be feasible after all other computations are finished. Cameron (Cameron [5]) pointed out that the statements about the integrals with $i\hbar + \delta$ in place of i were simply wrong; if we define the "measures" of finitely-based sets in the natural way, the resulting set function is not of bounded variation. If the "Feynman integral" is to be given any usable meaning, a more sophisticated approach is required. Within the past few months there has been considerable activity in this direction. (In fact, I have been favored with unpublished manuscripts by Itô, Feldman and Nelson since the January 1963 meeting, so the rest of this paper is in Congressional terminology an "extension of the remarks" made during that meeting.) These investigations fall into two classes. Those in one class proceed by defining some rigorous substitute for the invalid procedure of replacing $i\hbar$ by $i\hbar + \delta$ and then using analytic continuation and a passage to the limit analogous to $\delta \rightarrow 0$. Those in the other class (i.e., two papers by Itô) are essentially rigorous replacements for the (nonexistent) "rotation- and translation-invariant measure in Hilbert space."

Let us begin with Cameron's recent study (Cameron [6]) of the Feynman and related integrals. This begins with the observation that if we define $\lambda = \sigma^{-2}$ ("reciprocal variance parameter"), for the ordinary Wiener integral of a functional F we have

$$\int_{C[a,b]} F(x) d_{w,\lambda} x = \int_{C[a,b]} F(\lambda^{-1/2} x) d_{w,1} x$$

if the integrals exist; here $d_{w,\lambda} x$ is the Wiener measure corresponding to the Brownian-motion process with "reciprocal-variance parameter" λ , and $C[a, b]$ is the space of all real continuous functions on $[a, b]$ that vanish at a . Suppose now that F is a functional defined on $C[a, b]$ such that for some positive ρ_0 , $(F(\rho x) : \rho \in [0, \rho_0], x \in C[a, b])$

is integrable with respect to $d\rho d_{w,1}x$ measure on $[0, \rho_0] \times C[a, b]$. Then for almost all sufficiently large λ the integral

$$\int_{C[a,b]} F(\lambda^{-1/2}x) d_{w,1}x$$

will be finite. It is possible that the function of λ thus defined is, for large λ , the Laplace-Stieltjes transform of some function $(f(s): 0 \leq s < \infty)$; that is, f is left continuous and of bounded variation on $[0, h]$ for all finite h , and for sufficiently large λ the integral

$$\int_0^\infty e^{-s\lambda} |df(s)|$$

exists, and for almost all such λ

$$\int_0^\infty e^{-s\lambda} df(s) = \int_{C[a,b]} F(\lambda^{-1/2}x) d_{w,1}x.$$

Then $f(s)$ is the "Ilstow integral of F with parameter s ." Don't worry if you have never heard of anyone named Ilstow; it is merely a contraction of "inverse Laplace-Stieltjes transform of Wiener's."

So far λ has been positive. Now suppose that λ is complex (not 0) with non-negative real part. It may still happen (and will, if the real part of λ is large enough) that $|e^{-s\lambda}|$ has a finite integral over $[0, \infty]$ with respect to $|df(s)|$. In that case we take the preceding equation as the definition of its right member. In particular, if f is of bounded variation we can take $\lambda = -iq$ for any real nonvanishing number q , and the resulting integral is the "Feynman integral, with parameter q ," for which Cameron introduces the special notation

$$\int_{C[a,b]}^{f_q} F(x) dx = \int_0^\infty e^{isq} df(s).$$

However, the Feynman integral can be defined for a wider class of integrands by introducing convergence factors, that is by defining Wiener integrals with parameter λ in the right half-plane and letting λ tend to $-qi$. This time the trick works because the Wiener integrals involved have been truly defined. Cameron defines the "limiting Feynman integral with parameter q " to be the quantity

$$\int_{C[a,b]}^{-f_q} F(x) dx = \lim_{\eta \rightarrow 0; \operatorname{Re} \eta > 0} \int_0^\infty e^{s(iq-\eta)} df(s),$$

provided that the right member exists.

Since this whole treatment of complex-Wiener and Feynman inte-

grals rests upon the behavior of the "Itstow integrals" $f(s)$, the development of the theory requires extensive investigation of the latter. Cameron has carried out this investigation in great detail, which it would be absurd for me to try to outline. However, we can mention the last theorem. Under suitable boundedness and integrability conditions on the functions $(\theta(t, \xi): 0 \leq t \leq t_0, -\infty < \xi < \infty)$ and $(\psi(\xi): -\infty < \xi < \infty)$, the limiting Feynman integral

$$G(t, \xi) = \int_{C[0, t]}^{-\mathcal{I}_1} \exp \left\{ \int_0^t \theta(t-s, x(s) + \xi) ds \right\} \psi(x(t) + \xi) dx$$

exists for all (t, ξ) in $(0, t_0] \times (-\infty, \infty)$, and on that set it satisfies the partial differential equation

$$\frac{1}{i} \frac{\partial G}{\partial t} = \frac{1}{2} \frac{\partial^2 G}{\partial \xi^2} - i\theta(t, \xi)G$$

and the boundary condition

$$\lim_{t \rightarrow 0+} G(t, \xi) = \psi(\xi).$$

Thus the limiting Feynman integral serves to express the solution of the one-dimensional Schrödinger equation with suitably restricted potential function, and to this extent it does what the Feynman integral was meant to do.

We now turn to Itô's device for introducing a substitute for a uniform measure in Hilbert space. If we look back at the formula for the (nonexistent) Feynman measure $d_f x$, and replace x by y , we see that the paths x over which we integrate are uniquely specified by the functions y in $L_2[0, t]$, assuming that they all start at the same $x(0)$. Then $\int_0^t |\dot{x}(\tau)|^2 d\tau = \|y\|^2$, so the Feynman integral of a function $f(y)$ with respect to d_f should be of the form

$$N^{-1} \int_H \exp [i\|y\|^2/2h] f(y) \Pi dy(\tau),$$

with some suitable normalizing constant N . So our problem now is to assign a useful meaning to such integrals.

Consider first a finite-dimensional space R^n ; let h be a positive number. We shall define a "normalized uniform" measure $d_i y$ in R^n . (The subscript i stands for Itô; it is our notation, not his!) This will take the place of $N^{-1} \Pi dy(\tau)$ in the preceding formula. It will be a constant C times Lebesgue measure in R^n , where C depends only on h and n and will be specified shortly. If M is any bounded complex measure on R^n , the Fourier transform of M is the function FM

$= (FM(y) : y \in R^n)$ defined by

$$FM(y) = \int \exp i\langle y, z \rangle M(dz),$$

integration being over the whole space. If we write f for FM , an easy calculation shows

$$\begin{aligned} \int [\exp(i\|y\|^2/2h)] f(y) d_i y \\ = C \left\{ \sqrt{2h} \int_{-\infty}^{\infty} (\cos t^2 + i \sin t^2) dt \right\}^n \int \exp(h\|z\|^2/2i) M(dz). \end{aligned}$$

We choose C so that the coefficient of the integral in the right member is 1. Thus we find that the "normalized uniform" measure $d_i y$ in R^n possesses the following property: whenever f is the Fourier transform of a measure M , the equation

$$\int [\exp(i\|y\|^2/2h)] f(y) d_i y = \int [\exp(h\|z\|^2/2i)] M(dz)$$

holds.

This gives us the pattern of definition that Itô uses. No measure is defined in H ; but the integral in the left member of the last equation is defined for all functions f such that f is the Fourier transform of a bounded measure M , and its value is that given by the last equation. This is easily seen to be linear and to be invariant under rotation. If $f=1$, the value of the integral is 1. Also, an analogue of the Fubini theorem can be established.

From the heuristic discussion, the integral has some sort of spatial homogeneity. However, this can also be more precisely expressed in another way. Let V be any compact symmetric operator on H whose eigenvalues $\lambda_1, \lambda_2, \dots$ are positive and have finite sum. There is a unique probability measure G_V defined on the Borel sets in H whose Fourier transform is $\exp(-\langle Vy, y \rangle/2)$; this is called the Gaussian distribution with variance operator V . Then for every positive integer n , G_{nV} is a measure on H . Itô shows that if f is the Fourier transform of a measure,

$$\begin{aligned} \int_H \exp[i\|y\|^2/2h] f(y) d_i y \\ = \lim_{n \rightarrow \infty} \prod_{v=1}^{\infty} \left(1 + \frac{n\lambda_v}{hi} \right)^{1/2} \int_H \exp[i\|y\|^2/2h] f(y) G_{nV}(dy). \end{aligned}$$

Thus the left member is the limit of integrals with respect to measures which individually are not uniform over H , but which "spread out" over H as n increases. The measures tend to uniformity in the sense that if e_1, \dots, e_r are orthonormal in H and $\phi(\xi_1, \dots, \xi_r)$ is almost periodic on R^n ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_H \phi(\langle e_1, y \rangle, \dots, \langle e_r, y \rangle) G_n V(dy) \\ = \lim_{A \rightarrow \infty} (2A)^{-r} \int_{-A}^A \dots \int_{-A}^A \phi(\xi_1, \dots, \xi_r) d\xi_1 \dots d\xi_r. \end{aligned}$$

Itô shows that this integral furnishes solutions of the Schrödinger equation

$$\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = \frac{\hbar^2}{2} \frac{\partial^2 \psi}{\partial x^2} - q(x) \psi \quad (\psi(0+, x) = \phi(x))$$

(where $x \in R^1$) for a rather large class of potentials q , in the following way. If $\tau \in [0, t]$ and i_τ is the function which is 1 on $[0, \tau]$ and 0 on $(\tau, t]$, then (remembering that y was introduced as \hat{x}) $x + \langle i_\tau, y \rangle = x(t)$. We define

$$\begin{aligned} U_t \phi(x) = \int_H \exp[i \|y\|^2 / 2\hbar] \\ \cdot \exp \left[\hbar^{-1} i \int_0^t q(x + \langle i_\tau, y \rangle) d\tau \right] \phi(x + \langle i_t, y \rangle) d_i y. \end{aligned}$$

Itô shows that if q is the Fourier transform of a bounded complex measure μ in R' such that

$$\int_{R'} \xi^2 |\mu| (d\xi) < \infty,$$

the integral $U_t \phi(x)$ is well defined and satisfies the Schrödinger equation. This is true also if $q(x) = x$, even though this is not the Fourier transform of such a measure.

Finally, I wish to report briefly on three papers (Nelson [2], Feldman [2], Babbitt [1]) which are as yet unpublished and which I have not yet seen even in manuscript; however, I believe that two earlier publications (Nelson [1], Feldman [1]) indicate what two of these papers will contain. In each of these two the Schrödinger equation is solved in the form of a semi-group of transformations by starting with the heat equation and obtaining a semi-group capable of

analytic continuation to purely imaginary parameters. Nelson indicates that this technique also serves to define a Feynman integral. However, in these papers the integrals (other than the Wiener integral) play a secondary rôle, the chief goal being to solve the equations rather than to define an integral that will help in solving the equations; and therefore, in spite of their interesting contents, they are somewhat alien to our present subject. The abstracts of Babbitt's results indicate that analytic continuation of semi-groups plays an essential rôle in his results too.

BIBLIOGRAPHY

S. G. BRUSH

1. *Functional integrals and statistical physics*, Rev. Modern Phys. **33** (1961), 79–92.

R. H. CAMERON

1. *Some examples of Fourier-Wiener transforms of analytic functions*, Duke Math. J. **12** (1945), 485–488.
2. *A "Simpson's rule" for the numerical evaluation of Wiener integrals in function space*, Duke Math. J. **18** (1951), 111–130.
3. *The translation pathology of Wiener space*, Duke Math. J. **21** (1954), 623–627.
4. *The generalized heat flow equation and a corresponding Poisson formula*, Ann. of Math. (2) **59** (1954), 434–461.
5. *A family of integrals serving to connect the Wiener and Feynman integrals*, J. Math. and Phys. **39** (1960), 126–140.
6. *The Ilstow and Feynman integrals*, J. Analyse Math. **10** (1963), 287–361.

R. H. CAMERON and R. E. FAGAN

1. *Nonlinear transforms of Volterra type in Wiener space*, Trans. Amer. Math. Soc. **75** (1953), 552–575.

R. H. CAMERON and C. HATFIELD

1. *Summability of certain orthogonal developments of nonlinear functionals*, Bull. Amer. Math. Soc. **55** (1949), 130–145.
2. *Summability of certain series for unbounded nonlinear functionals*, Proc. Amer. Math. Soc. **4** (1953), 375–387.

R. H. CAMERON and W. T. MARTIN

1. *Transformation of Wiener integrals under translations*, Ann. of Math. (2) **45** (1944), 386–396.
2. *Transformation of Wiener integrals under a general class of linear transformations*, Trans. Amer. Math. Soc. **58** (1945), 184–219.
3. *Fourier-Wiener transforms of analytic functionals*, Duke Math. J. **12** (1945), 489–507.
4. *Fourier-Wiener transforms of functions belonging to L_2 over space C* , Duke Math. J. **14** (1947), 99–107.
5. *The transformations of Wiener integrals by nonlinear transformations*, Trans. Amer. Math. Soc. **66** (1949), 253–283.

ROBERT COGBURN

1. *Conditional probability operators*, Ann. Math. Statist. **33** (1962), 634–658 (in particular, p. 639).

S. H. COLEMAN

1. *Separability of stochastic processes*, Proc. Amer. Math. Soc. **14** (1963), 172–176.

P. J. DANIELL

1. *A general form of integral*, Ann. of Math. **19** (1918), 299–294.

JACQUES DIXMIER

1. *Formes linéaires sur un anneau d'opérateurs*, Bull. Soc. Math. France **81** (1953), 9–39.

N. DUNFORD and J. SCHWARTZ

1. *Linear operators*, Part I, Interscience, New York, 1958.

H. A. DYE

1. *The Radon-Nikodým theorems for finite rings of operators*, Trans. Amer. Math. Soc. **72** (1952), 243–280.

E. B. DYNKIN

1. *Markov processes and related problems of analysis*, Uspehi Mat. Nauk **15** (1960) No. 2, 3–24 (Russian); Translations of Russian Math. Surveys **15** (1960) No. 2, 1–21.

A. EINSTEIN

1. *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Ann. Physik (4) **17** (1905), 549–560.

JACOB FELDMAN

1. *On the Schrödinger and heat equations for bad potentials*, Mimeographed Notes, Department of Mathematics, Univ. California, Berkeley, July 1962.
2. *On the Schrödinger and heat equations for non-negative potentials*, Trans. Amer. Math. Soc. (to appear).

R. P. FEYNMAN

1. *Space-time approach to non-relativistic quantum mechanics*, Rev. Modern Phys. **20** (1948), 367–387.

K. O. FRIEDRICHS

1. *Mathematical aspects of the quantum theory of fields*, Interscience, New York, 1953.

K. O. FRIEDRICHS and H. N. SHAPIRO

1. *Integration over Hilbert space and outer extensions*, Proc. Nat. Acad. Sci. U.S.A. **43** (1957), 336–338.

K. O. FRIEDRICHS, J. SCHWARTZ, T. I. SEIDMAN, H. N. SHAPIRO and B. WENDROFF

1. Mimeographed notes on *Seminar on integration of functionals*, New York University, Institute of Mathematical Sciences, 1957.

I. M. GELFAND and A. M. YAGLOM

1. *Integration in function spaces and its application in quantum physics*, J. Math. Phys. **1** (1960), 48–69.

L. GROSS

1. *Integration and nonlinear transformations in Hilbert space*, Trans. Amer. Math. Soc. **94** (1960), 404–440.

T. H. HILDEBRANDT

1. *On integrals related to and extensions of the Lebesgue integral*, Bull. Amer. Math. Soc. **24** (1918), 113–177.

KIYOSI ITÔ

1. *Stochastic integral*, Proc. Imp. Acad. Tokyo **20** (1944), 519–524.
2. *On a stochastic integral equation*, Proc. Japan Acad. **22** (1946), 32–35.
3. *Stochastic differential equations in a differentiable manifold*, Nagoya Math. J. **1** (1950), 35–47.
4. *Brownian motions in a Lie group*, Proc. Japan Acad. **26** (1950), 4–10.
5. *Multiple Wiener integral*, J. Math. Soc. Japan **3** (1951), 157–169.
6. *On a formula concerning stochastic differentials*, Nagoya Math. J. **3** (1951), 55–64.

7. *Stochastic differential equations*, Mem. Amer. Math. Soc. No. 4 (1951).
 8. *Complex multiple Wiener integral*, Japan. J. Math. 22 (1952), 63–86.
 9. *Stationary random distributions*, Mem. Coll. Sci. Univ. Kyoto, Ser. A 28 (1953), 209–223.
 10. *Stochastic differential equations in a differentiable manifold. II*, Mem. Coll. Sci. Univ. Kyoto, Ser. A 28 (1953), 81–85.
 11. *Spectral type of the shift transformation of differential processes with stationary increments*, Trans. Amer. Math. Soc. 81 (1956), 253–263.
 12. *Wiener integral and Feynman integral*, Proc. of Fourth Berkeley Symposium on Math. Statistics and Probability, Vol. II (University of California Press, Berkeley and Los Angeles, 1961) pp. 228–238.
 13. *An improper integral on the Hilbert space with its application to the Feynman integral*, J. Math. Kyoto Univ. (to appear).
- M. KAC
1. *On distributions of certain Wiener functionals*, Trans. Amer. Math. Soc. 65 (1949), 1–13.
 2. *On some connections between probability theory and differential equations*, Proc. of Second Berkeley Symposium on Math. Statistics and Probability (1951), pp. 189–215.
- SHIZUO KAKUTANI
1. *Determination of the spectrum of the flow of Brownian motion*, Proc. Nat. Acad. Sci. U.S.A. 36 (1950), 319–323.
- A. N. KOLMOGOROV
1. *Foundations of the theory of probability*, Chelsea, New York, 1950.
- E. J. McSHANE
1. *Remark concerning integration*, Proc. Nat. Acad. Sci. U.S.A. 35 (1949), 46–49.
 2. *Stochastic integrals and non-linear processes*, J. Math. Mech. 11 (1962), 235–284.
 3. *Weak topologies for stochastic processes*, Proc. Nat. Acad. Sci. U.S.A. 48 (1962), 1148–1151.
 4. *A weak topology for stochastic processes*, Bol. Soc. Mat. São Paulo 14 (1962), 85–101.
 5. *Families of measures and representations of algebras of operators*, Trans. Amer. Math. Soc. 102 (1962), 328–345.
- E. W. MONTROLL
1. *Markoff chains, Wiener integrals and quantum theory*, Comm. Pure Appl. Math. 5 (1952), 415–453.
- SHU-TEH-CHEN MOY
1. *Characterizations of conditional expectation as a transformation on function spaces*, Pacific J. Math. 4 (1954), 47–63.
- EDWARD NELSON
1. *L'équation de Schrödinger et les intégrales de Feynman*, Mimeographed notes of a lecture delivered in Paris.
 2. Ann. of Math. (to appear).
- M. OWCHAR
1. *Wiener integrals of multiple variations*, Proc. Amer. Math. Soc. 3 (1951), 459–470.
- M. ROSENBLATT
1. *On a class of Markoff processes*, Trans. Amer. Math. Soc. 71 (1951), 120–135.
- GIAN-CARLO ROTA
1. *Une généralisation de l'espérance conditionnelle qui se présente dans la théorie statistique de la turbulence*, C. R. Acad. Sci. Paris 251 (1960), 624–626.

2. *On the representation of averaging operators*, Rend. Sem. Mat. Univ. Padova **30** (1960), 52–64.

I. E. SEGAL

1. *Decomposition of operator algebras*. I, II, Mem. Amer. Math. Soc. No. 9 (1951).
2. *Equivalence of measure spaces*, Amer. J. Math. **73** (1951), 275–313.
3. *A non-commutative extension of abstract integration*, Ann. of Math. (2) **57** (1953), 401–457. (See also Correction, Ann. of Math. (2) **58** (1953), 595–596.)
4. *Abstract probability space and a theorem of Kolmogorov*, Amer. J. Math. **76** (1954), 721–732.
5. *Tensor algebras over Hilbert space*. I, Trans. Amer. Math. Soc. **81** (1956), 106–134.
6. *Tensor algebras over Hilbert space*. II, Ann. Math. Statist. **63** (1956), 160–175.
7. *Distributions in Hilbert space and canonical system of operators*, Trans. Amer. Math. Soc. **88** (1958), 12–41.

T. I. SEIDMAN

1. *Linear transformations of a functional integral*, Comm. Pure Appl. Math. **12** (1959), 611–621.
2. *A survey of functional integration with applications of the Wiener integral*, Massachusetts Institute of Technology, Lincoln Laboratory Report 55-G-1, 16 October 1961.

M. H. STONE

1. *Notes on integration*. IV, Proc. Nat. Acad. Sci. U.S.A. **35** (1949), 50–58.

N. WIENER

1. *The average of an analytic functional*, Proc. Nat. Acad. Sci. U.S.A. **7** (1921), 253–260.
2. *The average of an analytic functional and the Brownian movement*, Proc. Nat. Acad. Sci. U.S.A. **7** (1921), 294–298.
3. *Differential space*, J. Math. Phys. Mass. Inst. Tech. **2** (1923), 131–174.
4. *The average value of a functional*, Proc. London Math. Soc. (2) **22** (1924), 454–467.
5. *Generalized harmonic analysis*, Acta Math. **55** (1930), 117–258.
6. *The homogeneous chaos*, Amer. J. Math. **60** (1938), 897–936.
7. *Non-linear problems in random theory*, Technology Press, Cambridge, 1958.

UNIVERSITY OF VIRGINIA