RADON-NIKODYM DERIVATIVES OF GAUSSIAN MEASURES

By L. A. Shepp

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

I. SUMMARY

We give simple necessary and sufficient conditions on the mean and covariance for a Gaussian measure to be equivalent to Wiener measure. This was formerly an unsolved problem [26].

Another unsolved problem is to obtain the Radon-Nikodym derivative $d\mu/d\nu$ where μ and ν are equivalent Gaussian measures [28]. We solve this problem for many cases of μ and ν, by writing $d\mu/d\nu$ in terms of Fredholm determinants and resolvents. The problem is thereby reduced to the calculation of these classical quantities, and explicit formulas can often be given.

Our method uses Wiener measure μ_W as a catalyst; that is, we compute derivatives with respect to μ_W and then use the chain rule: $d\mu/d\nu = (d\mu/d\mu_W)/(d\nu/d\mu_W)$. Wiener measure is singled out because it has a simple distinctive property—the Wiener process has a random Fourier-type expansion in the integrals of any complete orthonormal system.

We show that any process equivalent to the Wiener process W can be realized by a linear transformation of W. This transformation necessarily involves stochastic integration and generalizes earlier nonstochastic transformations studied by Segal [21] and others [4], [27].

New variants of the Wiener process are introduced, both conditioned Wiener processes and free n-fold integrated Wiener processes. We give necessary and sufficient conditions for a Gaussian process to be equivalent to any one of the variants and also give the corresponding Radon-Nikodym (R-N) derivative.

Last, some novel uses of R-N derivatives are given. We calculate explicitly: (i) the probability that W cross a slanted line in a finite time, (ii) the first passage probability for the process $W(t + 1) - W(t)$, and (iii) a class of function space integrals. Using (iii) we prove a zero-one law for convergence of certain integrals on Wiener paths.

TABLE OF CONTENTS

Received 19 November 1965.

II. INTRODUCTION

All measures considered in this paper are *Gaussian* and are considered to be defined on the space of continuous functions $X = X(t), 0 \leqq t \leqq T$ with $T \leqq \infty$. Such a measure μ is determined by its mean m and covariance (cov) R

$$m(t) = \int X(t)\, d\mu(X),$$
$$R(s, t) = \int (X(s) - m(s))(X(t) - m(t))\, d\mu(X).$$

Wiener measure μ_W is the measure with mean zero and cov $=$ min (s, t). Two measures are equivalent (denoted $\sim$) when they have the same sets of measure zero.

1. A necessary and sufficient condition that $\mu \sim \mu_W$. We denote by L^2 and $\mathbf{L}^2$ the space of square-integrable functions on $[0, T]$ and $[0, T] \times [0, T]$ respectively; two functions are considered equal if they coincide almost everywhere.

Suppose μ is a measure with mean m and cov R.

THEOREM 1. *$\mu \sim \mu_W$ if and only if there exists a kernel $K \in \mathbf{L}^2$ for which*

$$R(s, t) = \min (s, t) - \int_0^s \int_0^t K(u, v)\, du\, dv \tag{1.1}$$

and

$$1 \notin \sigma(K) \tag{1.2}$$

and a function $k \in L^2$ for which

$$m(t) = \int_0^t k(u)\, du. \tag{1.3}$$

The kernel K is unique and symmetric and is given by $K(s, t) = -(\partial/\partial s)(\partial/\partial t) \cdot R(s, t)$ for almost every (s, t). The function k is unique and is given by $k(t) = m'(t)$ for almost every t.

As usual, $\sigma(K) = \{\lambda: K\varphi = \lambda\varphi, \varphi \neq 0\}$ is the spectrum, or the set of eigenvalues, of the Hilbert-Schmidt operator K. Here $K\varphi = \lambda\varphi$ means $\varphi \in L^2$ and

$$\int_0^T K(t, u)\varphi(u)\, du = \lambda\varphi(t).$$

Since K is symmetric and in $\mathbf{L}^2$ the eigenvalues $\lambda_1, \lambda_2, \cdots$ are real and $\sum \lambda_j^2 < \infty$. We shall show in Section 11 that if R is given by (1.1) with $K \in \mathbf{L}^2$, then R is nonnegative-definite if and only if $\lambda_j \leqq 1$ for all j. The condition (1.2) therefore says that $\lambda_j < 1$ for all j and should thus be interpreted as a statement of *strict* positive-definiteness for R. Note that λ_j may assume negative values.

The condition (1.1) has a simpler restatement in case $R_1(s, t) = (\partial/\partial s)R(s, t)$ is continuous for $s \neq t$. In this case (1.1) becomes (see Section 11 for the proof)

$$R_1(s, s+) - R_1(s, s-) \equiv 1, \qquad 0 < s < T. \tag{1.4}$$

This means that R must have a fixed discontinuity of unit size in its derivative along $s = t$, the same discontinuity that min (s, t) has in its derivative. A theorem of G. Baxter [2] implies the necessity of (1.4). We note that $R(0, \cdot) \equiv 0$ because $X(0) = 0$ under μ_W.

2. The R-N derivative $d\mu/d\mu_W$. Whenever $\mu \sim \mu_W$ the R-N derivative $d\mu/d\mu_W$ exists. We will show that $d\mu/d\mu_W$ can be written in terms of the Fredholm determinant and resolvent of the unique kernel K appearing in (1.1).

When $\sum |\lambda_j| < \infty$, K is said to be of *trace class* and the Fredholm determinant is

$$d(\lambda) = \prod_j (1 - \lambda\lambda_j). \tag{2.1}$$

For the general K, $\sum |\lambda_j|$ may not exist. The modified, or Carleman-Fredholm, determinant of K is

$$\delta(\lambda) = \prod_j (1 - \lambda\lambda_j)e^{\lambda\lambda_j} \tag{2.2}$$

which converges for all λ because $\sum \lambda_j^2 < \infty$. For each value of λ for which $\lambda^{-1} \notin \sigma(K)$ there is a unique kernel $H_\lambda \in \mathbf{L}^2$ called the Fredholm resolvent of K at λ. The resolvent equation

$$H_\lambda - K = \lambda H_\lambda K = \lambda K H_\lambda \tag{2.3}$$

determines H_λ uniquely. We denote H_1 by H for simplicity and note that H is defined because $1 \notin \sigma(K)$ by (1.2). The kernel H is symmetric and is continuous when K is continuous. There are known expansions of d, δ and H in powers of λ, cf. [7], pp. 1081–1086.

Let μ be a measure with mean m and cov R for which $\mu \sim \mu_W$. Let K be given by (1.1): $K(s, t) = -(\partial/\partial s)(\partial/\partial t)R(s, t)$ for almost every s and t.

THEOREM 2. *If K is continuous and of trace class then $d\mu/d\mu_W$ is given by*

$$\begin{aligned} d\mu/d\mu_W(X + m) &= [d(1)]^{-\frac{1}{2}} \\ &\cdot \exp\left[-\tfrac{1}{2}\int_0^T\int_0^T H(s, t)\, dX(s)\, dX(t) + \int_0^T k(u)\, dX(u) + \tfrac{1}{2}\int_0^T k^2(u)\, du\right] \end{aligned} \tag{2.4}$$

Here $k(t) = m'(t)$. When $m = 0$, (2.4) simplifies. The integral $\int_0^T k(u)\, dX(u)$

is the Wiener integral evaluated at the point X. It exists because $k \varepsilon L^2$. The integral $I(X) = \int_0^T \int_0^T H(s, t)\, dX(s)\, dX(t)$ is the double Wiener integral evaluated at X and is a (non-Gaussian) random variable with mean value $\int_0^T H(s, s)\, ds$. In our case the integral I can be defined because H is continuous. If H is also of bounded variation, we may integrate by parts to obtain an ordinary integral. Both H and the constant $d(1)$ may sometimes be obtainable in closed form even when $\sigma(K)$ is not so obtainable, as we shall see in Section 15.

In order to give a formula for $d\mu/d\mu_W$ valid for all K, the notion of a double Wiener integral must be extended slightly. We will use the *centered double Wiener integral* denoted

$$J(X) = \int_0^T c\int_0^T H(s, t)\, dX(s)\, dX(t).$$

J is introduced in Section 9. When the mean of I exists, J is the ordinary double Wiener integral I minus its mean. By this simple trick of subtracting off the mean, J can be defined for all $H \varepsilon \mathbf{L}^2$. K. Ito [10] was the first to consider the centered multiple Wiener integral. He obtained it in an equivalent way, by ignoring the values of H on the diagonal.

THEOREM 3. *If* $\mu \sim \mu_W$ *then*

$$(2.5) \quad (d\mu/d\mu_W)(X + m) = (\delta(1) \exp \operatorname{tr}(HK))^{-\frac{1}{2}} \cdot \exp\left[-\tfrac{1}{2}J(X) + \int_0^T k(u)\, dX(u) + \tfrac{1}{2}\int_0^T k^2(u)\, du\right].$$

The trace of a product always exists and we have

$$(2.6) \qquad \operatorname{tr}(HK) = \int_0^T \int_0^T H(s, t)K(s, t)\, ds\, dt.$$

When the hypothesis of Theorem 2 holds, (2.4) and (2.5) agree. Of course (2.4) is simpler. (2.5) has the advantage of being valid in general.

3. A representation for W. Let $\eta_1, \eta_2, \cdots$ be a sequence of independent standard normal variables (mean zero and variance one). Let $\varphi_1, \varphi_2, \cdots$ be an *arbitrary* complete orthonormal sequence in $L^2[0, T]$ and set

$$(3.1) \qquad \Phi_j(t) = \int_0^t \varphi_j(u)\, du, \qquad j = 1, 2, \cdots.$$

THEOREM 4. *For each* t, $0 \leqq t \leqq T$, *the series*

$$(3.2) \qquad \sum_{j=1}^{\infty} \eta_j \Phi_j(t) = W(t)$$

converges almost surely. The sum is the Wiener process on $[0, T]$.

This result appears to be new except for two special cases due to Wiener and Lévy. To prove the theorem we observe

$$(3.3) \qquad \Phi_j(t) = (\varphi_j, 1_t)$$

where 1_t is the indicator of the interval $[0, t]$ and $(f, g) = \int_0^T f(u)g(u)\, du$. The *completeness* implies

$$(3.4) \qquad \sum_j (\varphi_j, 1_s)(\varphi_j, 1_t) = (1_s, 1_t) = \min(s, t)$$

and (3.2) follows immediately from the 3-series theorem since the sum is Gaussian and has the required mean and covariance.

Wiener himself studied the special case of Fourier series:

$$\Phi_j(t) = 2^{\frac{1}{2}}(\sin (j - \tfrac{1}{2})\pi t)/(j - \tfrac{1}{2})\pi, \qquad 0 \leqq t \leqq 1. \tag{3.5}$$

In this case Φ as well as φ are orthogonal and this property characterizes the Fourier case. In the Fourier case, (3.2) converges uniformly in t with probability one. However, for some other choices of the sequence φ the convergence of (3.2) is even better. In fact, there does not seem to be any particular advantage to (3.5). To prove this, let us take φ to be the Haar sequence. Using double indices for convenience, set

$$\begin{aligned} \varphi_{-1} = 1, \qquad \varphi_{n,j}(t) &= 2^{n/2}, && 2^n t \,\varepsilon\, (j, j + \tfrac{1}{2}), \\ &= -2^{n/2}, && 2^n t \,\varepsilon\, (j + \tfrac{1}{2}, j + 1), \\ &= 0, && \text{otherwise} \end{aligned} \tag{3.6}$$

for $j = 0, 1, \cdots, 2^n - 1; n = 0, 1, 2, \cdots$. The φ's are complete and orthonormal. The representation (3.2) in this case is due to Lévy, cf. [11], p. 19, and takes the form

$$W(t) = \eta t + \textstyle\sum_{n=0}^{\infty} \sum_{j=0}^{2n-1} \eta_{n,j}\Delta_{n,j}(t) \tag{3.7}$$

where $\Delta_{n,j}(t) = \Phi_{n,j}(t) = \int_0^t \varphi_{n,j}$ is an isosceles triangle with base 2^{-n} centered at $(j + \frac{1}{2})2^{-n}$, and height, $\frac{1}{2}\cdot 2^{-n/2}$. This shows that the Wiener process is a random sum of triangles.

In Wiener's case, the series (3.2) fails to converge absolutely. By contrast (3.7) converges *absolutely* uniformly with probability one, as was pointed out by Z. Ciesielski. We follow [11], p. 19. Let

$$f_n(t) = \textstyle\sum_{j=0}^{2n-1} \eta_{n,j}\Delta_{n,j}(t).$$

We have

$$|f_n(t)| \leqq 2^{-n/2} \max_j |\eta_{n,j}|, \qquad 0 \leqq t \leqq 1, \tag{3.8}$$

because $\sum_j \Delta_{n,j}(t) \leqq 2^{-n/2}$ uniformly in t. Now $M_n = \max_j |\eta_{n,j}|$ is the maximum of 2^n independent standard normal variables and so

$$P\{M_n \leqq a_n\} = (\Phi(a_n) - \Phi(-a_n))^{2^n}. \tag{3.9}$$

Choosing $a_n = 2n^{\frac{1}{2}}$ it is easy to check that

$$\textstyle\sum_{n=1}^{\infty} P\{M_n > a_n\} < \infty. \tag{3.10}$$

By the Borel-Cantelli lemma we see that $M_n < a_n$ eventually and so, a.s.

$$\textstyle\sum_1^{\infty} 2^{-n/2} M_n < \infty. \tag{3.11}$$

It follows that (3.7) converges absolutely uniformly.

H. P. McKean has informed me that by using a recent theorem of J. Delporte

[30], p. 201, Corollary 6.4B, it can be shown that (3.2) converges *uniformly* a.s. We will not use this strong type of convergence and so we omit the proof.[1]

4. Simultaneous representation of μ and μ_W in terms of independent random variables. Let μ be a measure with mean m and cov R satisfying (1.1)–(1.3). Let $\varphi_1, \varphi_2, \cdots$ be the eigenfunctions of K; these are orthonormal and complete. The representation (3.2) gives

$$W(t) = \textstyle\sum_j \eta_j \Phi_j(t), \qquad 0 \leqq t \leqq T. \tag{4.1}$$

We shall define a Gaussian process Y *on the same space* as $\eta_1, \eta_2, \cdots$ with mean m and cov R. Write

$$k = m' = \textstyle\sum_j k_j \varphi_j, \qquad k_j = (k, \varphi_j), \tag{4.2}$$

and define

$$Y(t) = \textstyle\sum_j (\eta_j(1 - \lambda_j)^{\frac{1}{2}} + k_j)\Phi_j(t). \tag{4.3}$$

Y is clearly Gaussian and has mean

$$EY(t) = \textstyle\sum_j k_j(\varphi_j, 1_t) = (k, 1_t) = m(t) \tag{4.4}$$

and covariance

$$\textstyle\sum (1 - \lambda_j)\Phi_j(s)\Phi_j(t) = \min(s, t) - \int_0^s \int_0^t K(u, v)\, du\, dv = R(s, t). \tag{4.5}$$

In (4.5) we have used the $\mathbf{L}^2$ expansion of K

$$K(s, t) = \textstyle\sum \lambda_j \varphi_j(s)\varphi_j(t). \tag{4.6}$$

We have proved the following theorem.

THEOREM 5. *The processes* (4.1) *and* (4.3) *give a simultaneous representation of W and Y in terms of sums of independent variables.*

It is now possible to give a formal expression for $d\mu/d\mu_W(X)$. We expand a path X as

$$X(t) = \textstyle\sum_j X_j \Phi_j(t), \qquad X_j = \int_0^T \varphi_j(t)\, dX(t) \tag{4.7}$$

where $X_j = X_j(X)$ is the Wiener integral evaluated at X. $d\mu/d\mu_W(X)$ is the relative likelihood of X under μ and μ_W. Under μ_W, $X_j = \eta_j$, independent random variables. Under μ, $X_j = \eta_j(1 - \lambda_j)^{\frac{1}{2}} + k_j$ *also independent.* Because of independence, the probabilities multiply and we get

$$(d\mu/d\mu_W)(X) = \textstyle\prod_{j=1}^{\infty} (1 - \lambda_j)^{-\frac{1}{2}} \cdot \exp[-\frac{1}{2}(X_j - k_j)^2/(1 - \lambda_j)]/\exp[-\frac{1}{2} X_j^2]. \tag{4.8}$$

The product (4.8) always converges and represents $d\mu/d\mu_W$. A rigorous proof is given in Section 10. In Section 11 we show how (4.8) reduces to (2.4). The reduction to (2.4) is important because (2.4) is in terms of classical quantities and, in addition, does not explicitly involve the eigenvalues or eigenvectors, which are usually difficult to find.

[1] John Walsh has found a shorter proof, based on an abstract martingale convergence theorem.

5. Calculating $d\mu/d\nu$. Suppose that μ and ν are measures, both equivalent to μ_W. Then $\mu \sim \nu$ and by the chain rule

$$(5.1) \qquad (d\mu/d\nu)(X) = (d\mu/d\mu_W)(X)/(d\nu/d\mu_W)(X).$$

Applying (2.5) or (2.4) to obtain $d\mu/d\mu_W$ and $d\nu/d\mu_W$ we get an explicit formula for $d\mu/d\nu$, but only in the special case when $\mu \sim \mu_W$ and $\nu \sim \mu_W$. To get more general results we will study certain variants of W.

Let for $n = 0, 1, 2, \cdots$,

$$(5.2) \qquad W_n(t) = \int_0^t [(t-u)^n/n!]\, dW(u), \qquad 0 \leqq t \leqq T,$$

denote the n-fold integrated Wiener process. We have

$$(5.3) \qquad W_0(t) = W(t), \qquad W_n(t) = \int_0^t W_{n-1}(u)\, du, \quad n = 1, 2, \cdots.$$

The processes W_n satisfy $W_n^{(j)}(0) = 0, j = 0, 1, \cdots, n$.

Suppose μ is a measure with mean m and cov R whose sample paths Y are n-times differentiable a.s. Let $Y^{(n)}$ denote the nth derivative of Y. The process $Y^{(n)}$ is Gaussian with mean $m^{(n)}(t)$ and covariance

$$D_1^n D_2^n R(s, t) = (\partial^n/\partial s^n)(\partial^n/\partial t^n)R(s, t).$$

Let $\mu^{(n)}$ be the measure induced by $Y^{(n)}$; $\mu^{(n)}$ has the same mean and covariance as $Y^{(n)}$.

THEOREM 6. *Suppose $\mu \sim \mu_{W_n}$. Then $\mu^{(n)} \sim \mu_W$ and*

$$(5.4) \qquad (d\mu/d\mu_{W_n})(X) = (d\mu^{(n)}/d\mu_W)(X^{(n)}).$$

The mapping $X \to X^{(n)}$ is 1-1 on the set of n-times differentiable functions X for which $X^{(j)}(0) = 0, j = 0, 1, \cdots, n$ and standard arguments [9], pp. 163–164, give (5.4). The righthand side of (5.4) is given by (2.5).

Using (5.4), we can obtain $d\mu/d\nu$ explicitly whenever $\mu \sim \nu \sim \mu_{W_n}$ for some n. This generalizes (5.1) and, further, one may drop the assumption that n in (5.2) is an integer. However, the condition $\mu \sim \nu \sim \mu_{W_n}$ is still too restrictive. Excluded are *stationary measures* μ and ν because their sample paths do not vanish at zero. In order to remedy this lack we must unpin the process W_n at zero.

Let $\mathbf{W}_n$ be the *free* Wiener process

$$(5.5) \qquad \mathbf{W}_n(t) = \sum_{j=0}^n \xi_j t^j/j! + W_n(t), \qquad 0 \leqq t \leqq T,$$

where $\xi_0, \cdots, \xi_n$ are independent, standard normal variables. Suppose Y is a process for which $Y \sim \mathbf{W}_n$. The paths $Y(t)$ are then exactly n-times differentiable, and the derivatives of Y at $t = 0$ are *nonzero* random variables. The class of processes $Y \sim \mathbf{W}_n$ includes many processes of interest. We will see that stationary processes with rational spectral density are included as a special case. In the latter case, $d\mu/d\nu$ has already been found by Gelfand and Yaglom in a different way [28].

We shall give the conditions on m and R so that $Y \sim \mathbf{W}_n$. The condition on m is that $m^{(n+1)} \varepsilon L^2$, or

(5.6) $$m(t) = \sum_{j=0}^{n} [m^{(j)}(0)/j!]t^j + \int_0^t [(t-u)^n/n!]k(u)\,du,$$

where $k \varepsilon L^2$. In order to find the conditions on R, we first obtain a certain decomposition of an n-times differentiable covariance. The remainder of Section 5 will be needed only for Section 13 et seq.

An n-times differentiable process Y is called *nondegenerate at zero* when the random variables $Y(0)$, $Y'(0)$, $\cdots$, $Y^{(n)}(0)$ are linearly independent.

THEOREM 7. *The covariance R of an n-times differentiable process nondegenerate at zero may be written uniquely as*

(5.7) $$R(s,t) = \sum_{i=0}^{n} A_i(s)A_i(t) + R^*(s,t)$$

where

(5.8)
$$\begin{aligned}&\text{(i)} \quad R^* \text{ is a covariance,}\\ &\text{(ii)} \quad D_1^i D_2^i R^*(0,0) = 0, \qquad i = 0, 1, \cdots, n,\\ &\text{(iii)} \quad A_j^{(i)}(0) = 0, \quad i < j; \qquad A_i^{(i)}(0) > 0, \quad i = 0, \cdots, n.\end{aligned}$$

Condition (ii) means that if Z is a process with cov R^* then

(5.9) $$Z^{(j)}(0) = 0, \qquad j = 0, 1, \cdots, n.$$

We will have $Z \sim W_n$. The decomposition (5.7) is designed to reduce the problem to Theorem 6.

There are elegant formulas for $A_0, \cdots, A_n$ closely related to the decomposition formulas in Gauss's elimination method. To obtain them suppose R is a covariance that is n-times differentiable in each argument. Define

(5.10) $$R_{ij}(s,t) = D_1^i D_2^j R(s,t), \qquad R_{ij} = R_{ij}(0,0).$$

Let $\alpha_{-1} = 1$ and for $i \geqq 0$ let

(5.11) $$\alpha_i = \begin{vmatrix} R_{00} & \cdots & R_{0i} \\ \vdots & & \vdots \\ R_{i0} & \cdots & R_{ii} \end{vmatrix}.$$

If $R(s,t) = EY(s)Y(t)$ where Y is n-times differentiable then α_i is the Grammian of $Y(0), \cdots, Y^{(i)}(0)$ and α_i is strictly positive when Y is nondegenerate at zero. Whenever $\alpha_i > 0$, $i = 0, \cdots, n$ the functions $A_0, \cdots, A_n$ in (5.7) are unique and are given by

(5.12) $$A_i(t) = (\alpha_i\,\alpha_{i-1})^{-\frac{1}{2}} \begin{vmatrix} R_{00} & \cdots & R_{0i-1} & R_{00}(0,t) \\ \vdots & & R_{1i-1} & R_{10}(0,t) \\ & & \vdots & \vdots \\ R_{i0} & \cdots & R_{ii-1} & R_{i0}(0,t) \end{vmatrix},$$

$i = 0, 1, \cdots, n$. In particular, $A_0(t) = R(0,t)/(R(0,0))^{\frac{1}{2}}$. Note that A_i is independent of n for $n \geqq i$.

Define for $i = 0, \cdots, n$, $j = 0, \cdots, n$,

(5.13) $$A_{ji} = A_i^{(j)}(0) = (\alpha_i\,\alpha_{i-1})^{-\frac{1}{2}} \begin{vmatrix} R_{00} & \cdots & R_{0i-1} & R_{0j} \\ \vdots & & \vdots & \vdots \\ R_{i0} & \cdots & R_{ii-1} & R_{ij} \end{vmatrix}.$$

The $(n + 1) \times (n + 1)$ matrix $A = A_{ij}$ is lower semidiagonal and has an inverse. The inverse matrix is denoted by C and can be written explicitly (13.6).

The next theorem gives the conditions for a measure μ to be equivalent to μ_{W_n} as well as a formula for $d\mu/d\mu_{W_n}$ whenever it exists. It is clear that whenever $\mu \sim \mu_{W_n}$ the path functions Y must be n-times differentiable and nondegenerate at zero. In this case the covariance R of μ satisfies the hypothesis of Theorem 7, and R^* and $A_0, \cdots, A_n$ are uniquely defined.

THEOREM 8. *Let μ be a measure with mean m and* cov *R. $\mu \sim \mu_{W_n}$ if and only if: m satisfies* (5.6), *R has a unique decomposition* (5.7), *and*

$$D_1^n D_2^n R^*(s, t) = \min (s, t) - \int_0^s \int_0^t K(u, v)\, du\, dv \tag{5.14}$$

for a (unique, symmetric) kernel $K \varepsilon \mathbf{L}^2$ with $1 \not\varepsilon\, \sigma(K)$ and

$$A_i^{(n)}(t) = \int_0^t a_i(u)\, du + A_i^{(n)}(0) \tag{5.15}$$

for (unique) $a_0, \cdots, a_n$ in L^2.

When $\mu \sim \mu_{W_n}$ the R-N *derivative is given by* (5.16) *provided that K satisfies the conditions of Theorem* 2:

$$\begin{aligned}(d\mu/d\mu_{W_n})(X + m) = {} & [d(1)\alpha_n]^{-\frac{1}{2}} \exp [-\tfrac{1}{2}\textstyle\int_0^T \int_0^T H(s, t)\, dX^{(n)}(s)\, dX^{(n)}(t) \\ & + \textstyle\int_0^T k(u)\, dX^{(n)}(u) + \tfrac{1}{2}(k, k) \\ & - \tfrac{1}{2}\textstyle\sum_{j=0}^n (e_j^2 - (X^{(j)}(0) + m^{(j)}(0))^2) \\ & - \tfrac{1}{2}\textstyle\sum_{i=0}^n \sum_{j=0}^n e_i e_j ((I + H)a_i, a_j) \\ & + \textstyle\sum_{i=0}^n e_i \int_0^T ((I + H)a_i(t))\, dX^{(n)}(t)].\end{aligned} \tag{5.16}$$

Here I is the identity on L^2, H is the resolvent of K at $\lambda = 1$, $d(\cdot)$ is the determinant of K, and

$$e_j = \textstyle\sum_{i=0}^n C_{ji}(X^{(i)}(0) - m^{(i)}(0)), \qquad j = 0, 1, \cdots, n, \tag{5.17}$$

where $C = A^{-1}$, the inverse matrix of A in (5.13). *For general K we must replace the double stochastic integral in* (5.16) *by the centered integral and replace $d(1)$ by $\delta(1)$ exp tr (HK). This modification is completely analogous to that of Theorem* 3.

(5.16) is cumbersome. Its importance lies in its generality rather than in its simplicity; in many cases $d\mu/d\mu_{W_n}$ can be obtained more simply by other means.

6. Conditioned Wiener processes. We next consider sub-Wiener processes, obtained from W by linear conditioning. The Wiener integral

$$\eta = \eta(\psi) = \int_0^T \psi(t)\, dW(t) \tag{6.1}$$

is defined for $\psi \varepsilon L^2$; η is normal with mean zero and variance $(\psi, \psi) = \int_0^T \psi^2(u)\, du$.

Let V be any subspace of $L^2 = L^2[0, T]$ and let μ^V be the (Gaussian) measure obtained by conditioning μ_W so that $\eta(\psi) = 0$ for $\psi \varepsilon V$. Then μ^V has mean zero and cov

$$R^V(s, t) = \min (s, t) - \textstyle\sum_j \Psi_j(s)\Psi_j(t) \tag{6.2}$$

where

(6.3) $$\Psi_j(t) = \int_0^t \psi_j(u)\, du$$

and $\psi_1, \psi_2, \cdots$ is any orthonormal basis for V, contained in V.

We may realize a process with measure μ^V as follows:

(6.4) $$W^V(t) = W(t) - \sum_j \Psi_j(t) W_j, \qquad W_j = \eta(\psi_j).$$

It is easy to check that W^V has mean zero and cov R^V. Alternatively, let $\varphi_1, \varphi_2, \cdots$ be an orthonormal basis of the orthogonal complement of V in L^2. Of course, $\varphi_1, \varphi_2, \cdots$ are not complete. Let $\eta_1, \eta_2, \cdots$ be independent standard normal variables and define Φ_j as in (3.1). Then another realization of W^V is

(6.5) $$W^V(t) = \sum_j \eta_j \Phi_j(t).$$

The interest in (6.4) is that it is a realization on the same space as the original process W.

As an example, take V to be the space generated by $\psi = \psi_1 = 1$. Then with $T = 1, \Psi(t) = t$ and (6.2) is

(6.6) $$R^V(s, t) = \min(s, t) - st.$$

Now $W_1 = \eta(\psi) = \int_0^1 dW(t) = W(1)$ and (6.4) gives

(6.7) $$W^V(t) = W(t) - tW(1), \qquad 0 \leqq t \leqq 1.$$

The process (6.7) is called the pinned Wiener process [5].

The processes (6.4) are mutually singular for different subspaces V_1 and V_2. Indeed, if $\psi \in V_1$ but $\psi \notin V_2$ then $\eta(\psi) = 0$ for μ^{V_1} but is normal with nonzero variance for μ^{V_2} so that $\mu^{V_1} \perp \mu^{V_2}$. What is the condition on μ so that $\mu \sim \mu^V$? The answer is given by the next theorem.

THEOREM 9. *$\mu \sim \mu^V$ if and only if*

(6.8) $$R(s, t) = R^V(s, t) - \int_0^s \int_0^t K(u, v)\, du\, dv$$

where $K \in \mathbf{L}^2$ and in addition

(6.9) $$1 \notin \sigma(K) \quad \textit{and} \quad K\psi = 0 \quad \textit{for} \quad \psi \in V.$$

The mean must satisfy (1.3) *and in addition*

(6.10) $$(k, \psi) = 0, \qquad \psi \in V.$$

In case (6.8)–(6.10) *hold the* R-N *derivative is given by* (2.5) *where K is the unique kernel satisfying* (6.8). *When K satisfies the conditions of Theorem* 2, (2.4) *is also valid.*

7. Stochastic linear transformations of W. Interesting classes of processes can be obtained by various linear transformations of W. Segal and others consider some nonstochastic transformations and with them realize *some* processes equivalent to W [21], p. 464. By means of a transformation depending on a *stochastic integral* we will realize *any* process equivalent to W.

Suppose $M \varepsilon \mathbf{L}^2$. The stochastic integral

$$Z(s) = \int_0^T M(s, u)\, dW(u), \qquad 0 \leqq s \leqq T, \tag{7.1}$$

can be defined for each s in such a way that Z is a.s. measurable [6], p. 430. Assuming Z so defined, we set

$$Y(t) = W(t) - \int_0^t Z(s)\, ds \tag{7.2}$$

and call Y the *affine transformation of* W *with kernel* M. Of course when M is of bounded variation in the second argument, we may write Y as an ordinary non-stochastic transformation by integrating by parts. The process Y is Gaussian and has mean zero and covariance

$$R(s, t) = \min(s, t) - \int_0^s \int_0^t K(u, v)\, du\, dv \tag{7.3}$$

where

$$K = M + M^* - MM^*. \tag{7.4}$$

As usual, $M^*(u, v) = M(v, u)$ and, of course,

$$MM^*(u, v) = \int_0^T M(u, y)M(v, y)\, dy.$$

Let I denote the identity. Then by (7.4), $I - K = (I - M)(I - M^*)$. Since $\sigma(M) = \sigma(M^*)$,

$$1 \not\in \sigma(K) \Leftrightarrow 1 \not\in \sigma(M). \tag{7.5}$$

By Theorem 1 *we obtain:* $Y \sim W$ *if and only if*

$$1 \not\in \sigma(M). \tag{7.6}$$

Suppose R is any covariance satisfying (1.1) and (1.2). It is simple to show that there is an M satisfying (7.4) and (7.6). Indeed, the kernel K of R has the Mercer expansion in $\mathbf{L}^2$,

$$K(s, t) = \sum \lambda_j \varphi_j(s)\varphi_j(t) \tag{7.7}$$

where $\lambda_j < 1$ for all j. We may take M to be

$$M(s, t) = \sum_j [1 \pm (1 - \lambda_j)^{\frac{1}{2}}]\varphi_j(s)\varphi_j(t). \tag{7.8}$$

When all but a finite number of the signs in (7.8) are negative we have $M \varepsilon \mathbf{L}^2$ and it is easy to see that $M = M^*$ and (7.4) and (7.6) hold.

We have proved the following theorem. Let μ be Gaussian with mean m and cov R.

THEOREM 10. *Suppose* $\mu \sim \mu_W$. *There is an* M *for which the process* $Y + m$, *where* Y *is given by* (7.2), *is a realization of* μ. *In other words, the measure* μ_{Y+m} *induced by* $Y + m$ *satisfies*

$$\mu_{Y+m} = \mu. \tag{7.9}$$

M is not unique.

There are additional conditions one could put on M in order to make it unique in Theorem 8. However, none seems natural.

When M is Volterra

$$M(s, t) = 0 \quad \text{for} \quad s \leqq t \tag{7.10}$$

the process Y is *causal*: $Y(t)$ depends only on values $W(\tau)$ for $\tau \leqq t$. But we could not solve the problem of the existence and uniqueness of kernels M of Volterra type in (7.4).

8. A discussion of previous work. I. Segal considered a nonstochastic transformation S of W [20], p. 22; [21], p. 464. He defined

$$S(t) = W(t) + \int_0^T N(t, u)W(u)\, du, \qquad 0 \leqq t \leqq T. \tag{8.1}$$

N is assumed continuous and $N_t \varepsilon \mathbf{L}^2$. The transformation (8.1) generalizes an earlier one of Cameron and Martin [8]. It is easily seen that (8.1) is a special case of (7.2). Segal asserts that $S \sim W$ under the stated conditions on N. However, note that a spectral condition analogous to (7.6) is needed. Even if this correction is made, (8.1) is not "best possible" as Segal claims. The stochastic transformation (7.2) is better because it gives the most general transformation equivalent to W (Theorem 10).

In the case $\mu = \mu_{W+m}$, a translate of μ_W, the condition (1.3) as well as the formula for $d\mu/d\mu_W$ was known and is due to Cameron and Martin [6], and in the general case to Segal [21], p. 462.

D. E. Varberg [26] gave a formula for $d\mu/d\mu_W$ when R admits a certain factorization. His formula is based on a transformation of Woodward [27], which is similar to (8.1). Varberg requires many complicated additional assumptions. These complications arise because: (a) his approach is based on a transformation which is not general enough and (b) $d\mu/d\mu_W$ itself depends on the transformation only through the covariance of μ—many transformations give the same covariance as is shown by the manifold nonuniqueness of M in Theorem 10. For this reason it is better to work with the covariance directly. It has been brought to my attention by R. H. Cameron that D. E. Varberg in an unpublished manuscript has independently obtained an equivalent form of the sufficiency half of Theorem 1.

We should mention that a condition similar to (1.1) and (5.14) appears in work of Yu. A. Rozanov [19], p. 455.

We acknowledge with pleasure an informative private lecture given by J. Feldman and L. Gross and many profitable discussions with S. P. Lloyd.[3]

9. Double Wiener integrals. We will now define the *centered double Wiener integral*

$$J(X) \doteq \int_0^T c\int_0^T H(s, t)\, dX(s)\, dX(t) \tag{9.1}$$

for $H \varepsilon \mathbf{L}^2$ and for almost every function X.

A simple function H has the representation

$$H(s, t) = \sum_1^n \sum_1^n a_{jk}\chi_{jk}(s, t) \tag{9.2}$$

[3] Note added in proof: A. M. Yaglom communicated that I. M. Golosoy recently obtained results, announced in *Dokl. Akad. Nayk CCCP* **166** (1966) 263–266, which should be compared with ours. He obtains, among other things, an equivalent version of our Theorem 1

where

$$\chi_{jk}(s, t) = 1 \quad (s, t) \varepsilon (t_{j-1}, t_j) \times (t_{k-1}, t_k) \tag{9.3}$$
$$= 0 \quad \text{otherwise}$$

for some partition, $0 = t_0 < t_1 < \cdots < t_n = T$. When H is simple define

$$I(X) = \sum_1^n \sum_1^n a_{jk}(X(t_j) - X(t_{j-1}))(X(t_k) - X(t_{k-1})) \tag{9.4}$$

and

$$J(X) = I(X) - \int_0^T H(s, s)\, ds. \tag{9.5}$$

It is important and also easy to check that $J = J_H$ does not depend on the partition used to define it. J is not normally distributed, but has mean zero and variance

$$\int J^2\, d\mu_W = 2\int_0^T \int_0^T H^2(s, t)\, ds\, dt. \tag{9.6}$$

Now suppose that H is any element of $\mathbf{L}^2$ and that H_n is a sequence of simple functions for which $H_n \to H$ in $\mathbf{L}^2$. By (9.6), J_{H_n} is a Cauchy sequence in $L^2(\mu_W)$. We define $J_H = \lim J_{H_n}$. It is easy to check that J_H does not depend on the sequence H_n used to define it.

We call J the centered integral because

$$E_W J = \int J\, d\mu_W = 0. \tag{9.7}$$

The c between the integral signs in (9.1) calls attention to (9.7). For later use we point out that for $H(s, t) = \varphi(s)\varphi(t)$, a degenerate kernel, we get

$$\int_0^T c\int_0^T \varphi(s)\varphi(t)\, dX(s)\, dX(t) = \eta^2(\varphi) - (\varphi, \varphi) \tag{9.8}$$

where $\eta(\varphi)$ is the single Wiener integral

$$\eta(\varphi) = \int_0^T \varphi(t)\, dX(t). \tag{9.9}$$

The (*uncentered*) *double Wiener integral*

$$I(X) = \int_0^T \int_0^T H(s, t)\, dX(s)\, dX(t) \tag{9.10}$$

is now easy to define. We define $I = I_H$ for *continuous* H by

$$I(X) = J(X) + \int_0^T H(s, s)\, ds. \tag{9.11}$$

There is an important formula for I obtained by two integrations by parts when H is of bounded variation. We say H is of *bounded variation* when

$$\text{Var}\,(H) = \sup \sum_1^n \sum_1^n |H(t_j, t_k) - H(t_{j-1}, t_k) - H(t_j, t_{k-1}) + H(t_{j-1}, t_{k-1})| \tag{9.12}$$

is finite, the sup being taken over all partitions. In this case we have

$$I(X) = H(T, T)X^2(T) - X(T)\int_0^T X(s)\, d_sH(s, T) - X(T)\int_0^T X(t)\, d_tH(T, t) + \int_0^T \int_0^T X(s)X(t)\, d_s\, d_tH(s, t). \tag{9.13}$$

by a rather different method. He then applies the theorem to obtain conditions for equivalence and singularity for an arbitrary Gaussian measure and a Gauss-Markov measure.

One proves (9.13) first for simple functions and then for general H of bounded variation by passage to the limit. The advantage of (9.13) is that it does not involve stochastic integration and so is defined for *every* continuous function X.

III. PROOFS OF THE MAIN RESULTS

We first obtained Theorems 1–3 and 9 heuristically, using a direct evaluation of $d\mu/d\mu_W$ as a limit of finite-dimensional densities,

$$(d\mu/d\mu_W)(X) = \lim \,[p_\mu(x_1, \cdots, x_n)/p_{\mu_W}(x_1, \cdots, x_n)], \quad x_i = X(t_i),$$

where

$$p_\mu(\bar{x}) = (2\pi)^{-n/2}|R|^{-\frac{1}{2}} \exp\,(R^{-1}(\bar{x} - \bar{m}), (\bar{x} - \bar{m})).$$

The limit is taken as the partition $0 = t_0 < t_1 < \cdots < t_n = T$ becomes dense. While a rigorous proof along these lines is difficult and involves many details, it can be given when the kernel K of R is smooth. We proceed by the easier but indirect method, via (3.2) and the simultaneous representation.

The advantage of using μ_W and the related measures as the catalysts stems from the properties of *white noise*, formally $W'(t)$. Many have tried to calculate $d\mu/d\nu$ directly by getting a simultaneous representation of μ and ν in terms of independent variables. What would be needed is a set of simultaneous eigenfunctions. However, these *do not exist* in general (note that the claim in [28], p. 334, about the existence of generalized eigenfunctions is not correct). In case $\nu = \mu_W$ such eigenfunctions do exist and are in L^2 as we have seen. Basic is the fact that the covariance of white noise is formally the δ-function and δ can be expanded in any complete set

$$\delta(s, t) = \textstyle\sum_j \varphi_j(s)\varphi_j(t).$$

Of course, this expansion is only formal, but upon integration it becomes precise.

10. Proof of Theorem 1: Sufficiency. Suppose μ is a measure with mean m and cov R that satisfy (1.1)–(1.3). We will prove that $\mu \sim \mu_W$ by actually giving $d\mu/d\mu_W$. Let $\varphi_1, \varphi_2, \cdots$ denote the eigenfunctions of the kernel K and define the Wiener integral

(10.1) $$X_j = X_j(X) = \textstyle\int_0^T \varphi_j(t)\, dX(t), \qquad j = 1, 2, \cdots.$$

Define

(10.2) $$F_j(X) = (1 - \lambda_j)^{-\frac{1}{2}}\{\exp\,[-\tfrac{1}{2}(X_j - k_j)^2/(1 - \lambda_j)]/\exp\,[-\tfrac{1}{2}X_j^2]\}$$

where $k_j = (k, \varphi_j)$, $k = m'$.

Lemma 1. *The product*

(10.3) $$F(X) = \textstyle\prod_{j=1}^\infty F_j(X)$$

converges a.e. (μ_W) *and is integrable* μ_W.

Proof. First, $\prod_j (1 - \lambda_j)e^{\lambda_j}$ converges because $\sum \lambda_j^2 < \infty$. Using $\sum k_j^2 < \infty$ (by (1.3)) and the 3-series theorem, it is an easy exercise to prove that

$$\textstyle\sum_{j=1}^\infty (X_j - k_j)^2/(1 - \lambda_j) - X_j^2 - \lambda_j$$

converges for a.e. X. Thus (10.3) converges and F is a random variable.

To prove that $F \varepsilon L^1(\mu_W)$ we use an idea due to Kakutani [14]. Let

$$G_N = \prod_{j=1}^N (F_j)^{\frac{1}{2}}.$$

We already know that $G_N(X) \to (F(X))^{\frac{1}{2}}$ for a.e. X. We will show that G_N is a Cauchy sequence in $L^2(\mu_W)$. The limit of the Cauchy sequence must also be $(F)^{\frac{1}{2}}$. But L^2 is complete and so $(F)^{\frac{1}{2}} \varepsilon L^2(\mu_W)$, or $F \varepsilon L^1(\mu_W)$.

To prove that G_N is a Cauchy sequence observe that

$$E_W F_j = 1, \qquad E_W{}^2(F_j)^{\frac{1}{2}} = \beta_j^2 = \exp\,[-\tfrac{1}{2}k_j^2/(2 - \lambda_j)](1 - \lambda_j)^{\frac{1}{2}}/(1 - \lambda_j/2).$$

Now $X_1, X_2, \cdots$ are independent (μ_W) and hence so are $F_1, F_2, \cdots$. For $M < N$ we have by direct calculation

$$(10.4) \quad E_W(G_N - G_M)^2 = 2 - 2\prod_{M+1}^N E_W(F_j)^{\frac{1}{2}} = 2(1 - \prod_{M+1}^N \beta_j).$$

Again using $\sum \lambda_j^2 < \infty$ and $\sum k_j^2 < \infty$ it follows that $\prod \beta_j$ converges. The tail of the product tends to unity and so G_N is a Cauchy sequence. The lemma is proved.

LEMMA 2. *F is the* R-N *derivative*; $F = d\mu/d\mu_W$.

PROOF. What we must show is that

$$(10.5) \qquad \mu(A) = \int_A F(X)\, d\mu_W(X),$$

where A is any measurable set of functions X. We prove (10.5) first for sets A of the form

$$(10.6) \qquad A = \{X_1 < a_1, X_2 < a_2, \cdots, X_n < a_n\}.$$

Once again, $X_1, X_2, \cdots$ are independent (μ_W) and hence so are $F_1(X), \cdots, F_n(X)$. The expectation of a product of independent variables is the product of the expectations and by direct calculation

$$(10.7) \qquad \int_A F(X)\, d\mu_W(X) = \prod_{j=1}^n \Phi((a_j - k_j)/(1 - \lambda_j))^{\frac{1}{2}}$$

where Φ is the standard normal df.

Relative to μ the random variable X_j is Gaussian and has mean

$$(10.8) \qquad E_\mu \int_0^T \varphi_j(t)\, dX(t) = (\varphi_j, k) = k_j.$$

The covariance of $X_1, X_2, \cdots$ is

$$(10.9) \qquad E_\mu(X_i - k_i)(X_j - k_j) = \int_0^T \int_0^T \varphi_i(s)\varphi_j(t)\, d_s\, d_t R(s, t).$$

Applying (1.1), we obtain

$$(10.10) \quad \int_0^T \int_0^T \varphi_i(s)\varphi_j(t)\, d_s\, d_t R(s, t) = (\varphi_i, \varphi_j) - (K\varphi_i, \varphi_j) = (1 - \lambda_i)\delta_{ij}.$$

Thus $X_1, X_2, \cdots$ are also independent with respect to μ. It is now easy to check that $\mu(A)$ agrees with (10.7) and hence (10.5) follows, at least for sets of the form (10.6).

In order to prove (10.5) for any measurable set A we find by direct calculation that

$$(10.11) \qquad E_W(\textstyle\sum_j X_j \Phi_j(t) - X(t))^2 = 0$$

for each fixed t. It follows that almost surely (μ_W) we have

$$X(t) = \sum_j X_j\Phi_j(t) \tag{10.12}$$

and so the σ-field generated by $\{X_j\}$ and sets of probability zero (μ_W) includes the measurable sets (the σ-field generated by $X(t)$, for $0 \leqq t \leqq T$). We have already proved (10.5) for the σ-field generated by $\{X_j\}$ and Lemma 2 follows.

We have shown that when (1.1)–(1.3) hold, the R-N derivative $d\mu/d\mu_W$ exists. This derivative is a.s. positive because the infinite product converges and it follows that $\mu \sim \mu_W$. Without using the positivity, the equivalence would also follow from the dichotomy theorem of Feldman and Hájek. We turn to the other half of Theorem 1.

11. Proof of Theorem 1: Necessity. At this point we use an important theorem of Segal [21], p. 463. The following proof of the necessity was suggested by J. Feldman.

Let μ be a measure with mean m and cov R and suppose that $\mu \sim \mu_W$. Consider the Hilbert space $\mathbf{H} = L^2[0, T]$ and define the bilinear form $\mathbf{B}$: For $\varphi \in \mathbf{H}$ and $\psi \in \mathbf{H}$ define the random variables $\eta(\varphi)$ and $\eta(\psi)$ by the Wiener integral (9.9) and set

$$\mathbf{B}(\varphi, \psi) = \int \eta(\varphi)\eta(\psi)\, d\mu - \left(\int \eta(\varphi)\, d\mu\right)\left(\int \eta(\psi)\, d\mu\right). \tag{11.1}$$

The form $\mathbf{B}$ is positive, $\mathbf{B}(\varphi, \varphi) \geqq 0$, and is bounded,

$$\mathbf{B}(\varphi, \varphi) \leqq (\varphi, \varphi) \times \text{constant}$$

(see Lemma 1 of [15]). Consequently, there is a linear transformation B [18], p. 202, with

$$(B\varphi, \psi) = \mathbf{B}(\varphi, \psi). \tag{11.2}$$

Since B is a positive linear transformation it has a squareroot T [18], p. 265.

Segal's theorem says that if $\mu \sim \mu_W$ (in his terminology $n_T \sim n$) we must have

$$T^*T = B = I - K \tag{11.3}$$

where I is the identity and K is Hilbert-Schmidt, $K \in \mathbf{L}^2$. For all $\varphi \in \mathbf{H}$, $\psi \in \mathbf{H}$ we get by (11.3),

$$(B\varphi, \psi) = (\varphi, \psi) - (K\varphi, \psi) = \mathbf{B}(\varphi, \psi). \tag{11.4}$$

Now choose $\varphi = 1_s$, $\psi = 1_t$. We have $\eta(1_t) = \int_0^t dX(u) = X(t)$, since $X(0) = 0$ a.s. μ_W. We get by (11.1)

$$\mathbf{B}(1_s, 1_t) = R(s, t). \tag{11.5}$$

By (11.4)

$$R(s, t) = \min(s, t) - \int_0^s \int_0^t K(u, v)\, du\, dv. \tag{11.6}$$

This proves (1.1) necessary.

Next we prove the spectral condition on K, $\lambda_j < 1$ for all j. By (11.4) we get

$$(11.7) \qquad (\varphi, \varphi) \geqq (K\varphi, \varphi)$$

since $(B\varphi, \varphi) \geqq 0$. The largest eigenvalue of K is given by $\lambda_{\max} = \sup_\varphi (K\varphi, \varphi)/(\varphi, \varphi)$ and so $\lambda_j \leqq 1$. To show that $1 \not\in \sigma(K)$ we may argue as follows. Suppose $K\varphi = \varphi$, $(\varphi, \varphi) = 1$. We see that $\eta(\varphi)$ is an a.s. constant with respect to μ since it has variance $(B\varphi, \varphi) = 0$ by (11.4). On the other hand, with respect to μ_W, $\eta(\varphi)$ has variance one. This means $\mu \perp \mu_W$ and we get a contradiction. This proves (1.2) necessary.

(11.7) shows that if R has the form (1.1) *and R is a covariance* then $\sigma(K) \subseteq (-\infty, 1]$. Next we show that if R has the form (1.1) and $\text{sp}(K) \subseteq (-\infty, 1]$ then R is a covariance. This will prove the remark made below Theorem 1, R is nonnegative definite if and only if $\lambda_j \leqq 1$ for all j.

What we must show is that

$$(11.8) \qquad (R\varphi, \varphi) \geqq 0$$

for all $\varphi \in L^2$. Let

$$\Phi(t) = -\int_t^T \varphi(u)\, du.$$

Integration by parts applied twice gives

$$(11.9) \qquad (R\varphi, \varphi) = \int_0^T \Phi(s)\Phi(t)\, d_s\, d_t R(s, t) = (\Phi, \Phi) - (K\Phi, \Phi).$$

But $\sigma(K) \subseteq (-\infty, 1]$ and so $(\Phi, \Phi) - (K\Phi, \Phi) \geqq 0$. We get (11.8) immediately. We remark that (1.3) cannot be replaced by the condition: $(R\varphi, \varphi) > 0$ for nonzero $\varphi \in L^2$.

Next we prove (1.3). Let $\nu(R, m)$ denote the Gaussian measure with mean m and cov R so $\mu = \nu(R, m)$. We have

$$(11.10) \qquad \mu = \nu(R, m) \sim \nu(R_W, 0) = \mu_W$$

where $R_W(s, t) = \min(s, t)$. By a theorem of C. R. Rao and V. S. Varadarajan [17], p. 308, the measures must also be equivalent when the means are ignored:

$$(11.11) \qquad \nu(R, 0) \sim \nu(R_W, 0).$$

By considering the 1-1 path transformation $X(t) \to X(t) + m(t)$ we see that (11.11) gives

$$(11.12) \qquad \nu(R, m) \sim \nu(R_W, m) = \mu_{W+m}.$$

Comparing (11.10) and (11.12) and using the fact that $\sim$ is an equivalence relation we get

$$(11.13) \qquad \mu_W \sim \mu_{W+m}.$$

The condition on m in order that $\mu_W \sim \mu_{W+m}$ was found by Segal [21], p. 462. It is precisely (1.3). This completes the proof of Theorem 1.

In order to prove (1.4) we observe that whenever (1.1) holds K is given by

$$K(s, t) = -(\partial/\partial s)(\partial/\partial t)R(s, t), \qquad s \neq t. \tag{11.14}$$

Putting (11.14) back into (1.1) we get for $s < t$,

$$R(s, t) = s + \int_0^s \left(\int_0^u R_{12}(u, v)\, dv\right) du + \int_0^s \left(\int_u^t R_{12}(u, v)\, dv\right) du. \tag{11.15}$$

Integrating on v and using $R(0, t) = 0$ we get

$$R(s, t) = s + \int_0^s R_1(u, u-)\, du + \int_0^s (R_1(u, t) - R_1(u, u+))\, du. \tag{11.16}$$

This gives immediately

$$s = \int_0^s (R_1(u, u+) - R_1(u, u-))\, du \tag{11.17}$$

which is the integrated version of (1.4).

12. Proof of Theorems 2 and 3. We will now show that the product formula (10.3) for $F = d\mu/d\mu_W$ can be expressed in terms of the Fredholm quantities.

Let K denote, as usual, the kernel in (1.1) of the covariance of $\mu \sim \mu_W$. The eigenfunctions of the resolvent H of K are $\varphi_1, \varphi_2, \cdots$, the same as those of K, and the eigenvalues γ of H satisfy $\gamma_j = \lambda_j/(1 - \lambda_j)$, $j = 1, 2, \cdots$. We note that $\sum \gamma_j^2 < \infty$ and so H has the $\mathbf{L}^2$ expansion

$$H(s, t) = \textstyle\sum_j \gamma_j \varphi_j(s) \varphi_j(t). \tag{12.1}$$

With $X_j = \int_0^T \varphi_j(t)\, dX(t)$ as in (10.1) we get formally from (12.1),

$$\textstyle\sum_{j=1}^\infty X_j^2 \gamma_j = \int_0^T \int_0^T H(s, t)\, dX(s)\, dX(t). \tag{12.2}$$

Since $k = \sum k_j \varphi_j$ and $Hk = \sum k_j \gamma_j \varphi_j$ we get formally

$$\textstyle\sum_{j=1}^\infty X_j k_j/(1 - \lambda_j) = \int_0^T k(t)\, dX(t) + \int_0^T \int_0^T H(s, t) k(s)\, ds\, dX(t) \tag{12.3}$$

and

$$\textstyle\sum_{j=1}^\infty k_j^2/(1 - \lambda_j) = (k, k) + (Hk, k). \tag{12.4}$$

Substituting (12.2)–(12.4) into (10.3) we get

$$\begin{aligned} F(X) = \textstyle\prod_{j=1}^\infty F_j(X) = (d(1))^{-\frac{1}{2}} \exp\Big[& -\tfrac{1}{2} \int_0^T \int_0^T H(s,t)\, dX(s)\, dX(t) \\ & + \int_0^T k(u)\, dX(u) + \int_0^T Hk(t)\, dX(t) - \tfrac{1}{2}(k, k) - \tfrac{1}{2}(Hk, k)\Big]. \end{aligned} \tag{12.5}$$

Replacing X by $X + m$ gives (2.4).

To prove (12.2) rigorously we observe that by (9.8) the partial sum

$$\textstyle\sum_{j=1}^N X_j^2 \gamma_j = \sum_{j=1}^N \gamma_j + \int_0^T c \int_0^T H_N(s, t)\, dX(s)\, dX(t), \tag{12.6}$$

where

$$H_N(s, t) = \textstyle\sum_{j=1}^N \gamma_j \varphi_j(s) \varphi_j(t), \qquad N = 1, 2, \cdots.$$

Since $H_N \to H$ in $\mathbf{L}^2$ we have in $L^2(\mu_W)$

$$\textstyle\int_0^T c \int_0^T H_N(s, t)\, dX(s)\, dX(t) \to \int_0^T c \int_0^T H(s, t)\, dX(s)\, dX(t). \tag{12.7}$$

Now suppose that K satisfies the assumptions of Theorem 2 so that K is continuous and of trace class. It follows that the resolvent H is continuous and of trace class. H is also symmetric, of course, and it is *known* that these assumptions on H imply that

$$\text{(12.8)} \qquad \operatorname{tr}(H) = \sum_{j=1}^{\infty} \gamma_j = \int_0^T H(s, s)\, ds.$$

Although this fact is well known in the theory of integral equations, apparently no published reference exists. Passing to the limit in (12.6) we have

$$\text{(12.9)} \qquad \sum_{j=1}^{\infty} X_j^2 \gamma_j = \sum_{j=1}^{\infty} \gamma_j + \int_0^T c \int_0^T H(s, t)\, dX(s)\, dX(t).$$

Using (12.8) and

$$\int_0^T c \int_0^T H(s, t)\, dX(s)\, dX(t) = \int_0^T \int_0^T H(s, t)\, dX(s)\, dX(t) - \int_0^T H(s, s)\, ds,$$

we obtain (12.2).

To prove (12.3) rigorously we observe that

$$\text{(12.10)} \qquad \sum_{j=1}^{N} X_j k_j/(1 - \lambda_j) = \int_0^T (k_N(t) + Hk_N(t))\, dX(t),$$

where $k_N = \sum_{j=1}^{N} k_j \varphi_j$. As $N \to \infty$, $k_N \to k$ in L^2. By the convergence properties of Wiener integrals we get (12.3).

Since the rigorous justification of (12.4) is straightforward we have proved Theorem 2. Theorem 3 can be proved similarly. Instead of (12.2) we must use

$$\text{(12.11)} \qquad \sum_{j=1}^{\infty} (X_j^2 - 1)\gamma_j = \int_0^T c \int_0^T H(s, t)\, dX(s)\, dX(t),$$

which follows from (12.6), letting $N \to \infty$. Of course, now $\sum_j \gamma_j$ does not necessarily converge.

The trace of HK is

$$\text{(12.12)} \qquad \operatorname{tr}(HK) = \sum_j \gamma_j \lambda_j$$

and the proof of (2.5) is completed by a short calculation.

We omit the proof of Theorem 9 because it proceeds along the same lines as those of Theorems 1–3.

13. Proof of Theorem 7. Let R be the covariance of an n-times differentiable process Y nondegenerate at zero. Let $A_0, \cdots, A_n$ be defined by (5.12) and R^* be defined by (5.7). We will prove that R^* is a covariance by showing that it is the covariance of the process Z,

$$\text{(13.1)} \qquad Z(t) = Y(t) - \sum_{j=0}^{n} Y^{(j)}(0) B_j(t),$$

where

$$\text{(13.2)} \qquad B_j(t) = \sum_{i=0}^{n} B_{ij} R_{i0}(0, t), \qquad j = 0, \cdots, n,$$

and $B_{ij} = R_{ij}^{-1}$ is the inverse matrix of R in (5.10). The matrix $R = R_{ij}$ is positive-definite and so has a lower semidiagonal (lsd) square root which is unique and is given by Gauss's formula [8], p. 37. Comparing Gauss's formula with (5.13) we see that A is the lsd square root, that is

$$\text{(13.3)} \qquad R_{ij} = \sum_{k=0}^{n} A_{ik} A_{jk}, \qquad A_{ji} = A_i^{(j)}(0),$$

$i = 0, \cdots, n$ and $j = 0, \cdots, n$; in symbolic notation $R = AA^T$.

The covariance of Z is by direct calculation

$$(13.4) \qquad R_Z(s, t) = R(s, t) - \textstyle\sum_{j=0}^n \sum_{k=0}^n B_{jk} R_{j0}(0, s) R_{k0}(0, t).$$

We will now show that the second term on the right,

$$(13.5) \qquad \textstyle\sum_{j=0}^n \sum_{k=0}^n B_{jk} R_{j0}(0, s) R_{k0}(0, t) = \sum_{k=0}^n A_k(s) A_k(t),$$

which will prove that $R^*(s, t) = R_Z(s, t)$, the covariance of Z. Define the matrix $C = C_{ij}$, $i = 0, \cdots, n$ and $j = 0, \cdots, n$ by

$$(13.6) \qquad C_{ij} = (\alpha_i \alpha_{i-1})^{-\frac{1}{2}} \begin{vmatrix} R_{00} & \cdots & R_{0i-1} & 0 \\ & & \vdots & \vdots \\ \vdots & & R_{ji-1} & 1 \\ & & \vdots & \vdots \\ R_{i0} & \cdots & R_{ii-1} & 0 \end{vmatrix},$$

where the last column is zero except for the element in the jth row, $j \leqq i$, which is unity. We have $C_{ij} = 0$ if $i < j$ and so C is lsd. Multiplying (13.3) on the right by R_{jl}^{-1} and adding, $j = 0, \cdots, n$, we get

$$(13.7) \qquad \delta_{il} = \textstyle\sum_{k=0}^n A_{ik} \sum_{j=0}^n A_{jk} R_{jl}^{-1}.$$

But using the formula (5.13) for A_{jk} we get

$$(13.8) \qquad \sum_{j=0}^n A_{jk} R_{jl}^{-1} = (\alpha_k \alpha_{k-1})^{-\frac{1}{2}} \begin{vmatrix} R_{00} & \cdots & R_{0k-1} & \sum_{j=0}^n R_{0j} R_{jl}^{-1} \\ \vdots & & \vdots & \vdots \\ R_{k0} & & R_{kk-1} & \sum_{j=0}^n R_{kj} R_{jl}^{-1} \end{vmatrix} = C_{kl}.$$

Putting (13.8) into (13.7) we get $\delta_{il} = \sum_{k=0}^n A_{ik} C_{kl}$ and so $C = A^{-1}$.

We see that $R^{-1} = (AA^T)^{-1} = C^T C$ and so

$$(13.9) \qquad \sum_{j=0}^n \sum_{k=0}^n R_{jk}^{-1} u_j v_k = \sum_{l=0}^n (\alpha_l \alpha_{l-1})^{-1} \begin{vmatrix} R_{00} & \cdots & R_{0l-1} & u_0 \\ \vdots & & \vdots & \vdots \\ R_{l0} & \cdots & R_{ll-1} & u_l \end{vmatrix} \cdot \begin{vmatrix} R_{00} & \cdots & R_{0l-1} & v_0 \\ \vdots & & \vdots & \vdots \\ R_{l0} & \cdots & R_{ll-1} & v_l \end{vmatrix},$$

where $u_0, \cdots, u_n, v_0, \cdots, v_n$ are variables. Setting $u_j = R_{j0}(0, s)$, $j = 0, \cdots, n$ and $v_k = R_{k0}(0, t)$, $k = 0, \cdots, n$ we obtain (13.5).

We have proved that R^* is a covariance. In fact it is the covariance of Z. (5.8) (i) is proved. To prove (5.8) (ii) we observe that $D_1^i D_2^i R^*$ is the covariance of $Z^{(i)}$ and by (13.1)

$$(13.10) \qquad Z^{(i)}(0) = Y^{(i)}(0) - \textstyle\sum_{j=0}^n Y^{(j)}(0) B_j^{(i)}(0).$$

But $B_j^{(i)}(0) = \sum_{k=0}^n B_{kj}R_{ki} = \delta_{ij}$ by (13.2) since $B = R^{-1}$. Hence (5.9) holds and so (5.8) (ii) is proved. Now (5.8) (iii) follows immediately from the definition (5.12) of $A_0, \cdots, A_n$. Note that

$$(13.11)\qquad A_i^{(i)}(0) = (\alpha_i\, \alpha_{i-1})^{-\frac{1}{2}} \begin{vmatrix} R_{00} & \cdots & R_{0i-1} & R_{0i} \\ & & \vdots & \vdots \\ R_{i0} & \cdots & R_{ii-1} & R_{ii} \end{vmatrix} = (\alpha_i/\alpha_{i-1})^{\frac{1}{2}}.$$

It remains only to prove the uniqueness of (5.7). Suppose that (5.7) holds where $A_0, \cdots, A_n$ and R^* satisfy (5.8). Differentiating i times on s and j times on t and setting $s = t = 0$ we get for $i = 0, \cdots, n$ and $j = 0, \cdots, n$,

$$(13.12)\qquad R_{ij} = \sum_{k=0}^n A_k^{(i)}(0)A_k^{(j)}(0).$$

We have used the fact that $D_1^iD_2^jR^*(0, 0) = 0$ which follows from (i) and (ii) of (5.8). Using (5.8) (iii) we see that $A = A_k^{(i)}(0)$ is an lsd square root of R which we know to be unique. Now we differentiate (5.7) j times on s only and then set $s = 0$. Again, $D_1^jR^*(0, t) = 0$ and we get

$$(13.13)\qquad R_{j0}(0, t) = \sum_{i=0}^n A_i^{(j)}(0)A_i(t), \qquad j = 0, \cdots, n.$$

The matrix $A = A_i^{(j)}(0)$ is nonsingular because of the nondegeneracy at zero and so the linear equations (13.13) have a unique solution $A_0(t), \cdots, A_n(t)$ for each fixed t. We have proved Theorem 7.

14. Proof of Theorem 8. Let μ be a measure with mean m and cov R. The proof of Theorem 8 will be based on the decomposition (5.7). Given a process Z with covariance R^* we may realize a process Y with covariance R by setting

$$(14.1)\qquad Y(t) = \sum_{j=0}^n \xi_j A_j(t) + Z(t)$$

where $\xi_0, \cdots, \xi_n$ are standard normal variates independent of Z. Using (14.1) we will obtain the simultaneous representation of μ and $\mu_{\mathbf{W}_n}$.

Suppose that m and R satisfy the hypothesis of Theorem 8. Let $\varphi_1, \varphi_2, \cdots$ be the complete o.n. system of eigenfunctions of K in (5.14) and set

$$(14.2)\qquad \Phi_{j,n}(t) = \int_0^t [(t-u)^n/n!]\varphi_j(u)\, du.$$

The process $\mathbf{W}_n$ has the representation

$$(14.3)\qquad \mathbf{W}_n(t) = \sum_{j=0}^n \xi_j t^j/j! + \sum_{j=1}^\infty \eta_j \Phi_{j,n}(t),$$

where $\xi_0, \cdots, \xi_n, \eta_1, \cdots$ are independent standard normal variates on some space. The process Y defined on the same space as the ξ's and η's by

$$(14.4)\qquad Y(t) = \sum_{j=0}^n \xi_j A_j(t) + \sum_{j=1}^\infty [k_j + (1-\lambda_j)^{\frac{1}{2}}\eta_j]\Phi_{j,n}(t) + \sum_{j=0}^n m^{(j)}(0)t^j/j!$$

has mean

$$EY(t) = \sum_{k=1}^\infty k_j\Phi_{j,n}(t) + \sum_{j=0}^n m^{(j)}(0)t^j/j! = m(t)$$

by (5.6) and covariance R in (5.7) as a short calculation shows. Now it is easy to obtain a formal expression for $d\mu/d\mu_{W_n}$. We compute the formal likelihood ratio under the two measures as follows:

We expand (formally)

$$(14.5) \quad X(t) = \sum_{i=0}^{n} X^{(i)}(0)t^i/i! + \sum_{j=1}^{\infty} X_j \Phi_{j,n}(t), \qquad X_j = \int_0^T \varphi_j(u)\, dX^{(n)}(u)$$

and denote $X^i = X^{(i)}(0)$. Under μ_{W_n} we have $X^i = \xi_i$, $X_j = \eta_j$. Under μ we get with $m^i = m^i(0)$ from (14.4),

$$(14.6) \qquad X^i = m^i + \sum_{j=0}^{n} \xi_j A_j^{(i)}(0),$$

$$(14.7) \qquad X_j = k_j + \eta_j(1 - \lambda_j)^{\frac{1}{2}} + \sum_{i=0}^{n} \xi_i a_{ij},$$

where $a_{ij} = (a_i, \varphi_j)$. Inverting (14.6) with $C = A^{-1}$ given by (13.7), we get

$$(14.8) \qquad \xi_j = e_j = \sum_{i=0}^{n} C_{ji}(X^i - m^i).$$

Therefore,

$$(14.9) \qquad \eta_j = (X_j - k_j - u_j)/(1 - \lambda_j)^{\frac{1}{2}}, \qquad u_j = \sum_{i=0}^{n} e_i a_{ij}.$$

The likelihood ratio becomes

$$(14.10) \quad d\mu(X)/d\mu_{W_n}(X) = \exp\left[-\tfrac{1}{2} \sum_{j=0}^{n} (e_j^2 - (X^j)^2) - \tfrac{1}{2} \sum_{j=1}^{\infty} (\eta_j^2 - X_j^2)\right]|J|$$

where $|J|$ is the Jacobian of the transformation (14.6), (14.7) of $X^j, X_j \to \xi_j, \eta_j$. Now $|\partial\xi_j/\partial X^i| = |C_{ji}| = C_{00} \cdots C_{nn}$ because C is lsd. Now $C_{ii} = 1/A_{ii}$ and $A_{ii} = (\alpha_i/\alpha_{i-1})^{\frac{1}{2}}$ by (5.13). We get $C_{00} \cdots C_{nn} = (\alpha_n)^{-\frac{1}{2}}$. We have $|\partial\eta_i/\partial X_j| = [\prod_j (1 - \lambda_j)^{\frac{1}{2}}]^{-1} = [d(1)]^{-\frac{1}{2}}$ and so $|J| = (d(1)\alpha_n)^{-\frac{1}{2}}$. Now,

$$\sum_{j=1}^{\infty} (\eta_j^2 - X_j^2) = \sum_{j=1}^{\infty} [(X_j - k_j)^2/(1 - \lambda_j) - X_j^2] - 2 \sum (X_j - k_j)u_j/(1 - \lambda_j) + \sum u_j^2/(1 - \lambda_j).$$

For the first term on the right see (12.2)–(12.4). For the others we obtain

$$(14.11) \qquad \sum_{j=1}^{\infty} (X_j - k_j)u_j = \sum_{i=0}^{n} e_i \int_0^T a_i(t)\, d\bar{X}^{(n)}(t),$$

$$\sum_{j=1}^{\infty} (X_j - k_j)u_j\lambda_j/(1 - \lambda_j) = \sum_{i=0}^{n} e_i \int_0^T \int_0^T H(s, t)a_i(s)\, ds\, d\bar{X}^{(n)}(t),$$

where $\bar{X}(t) = X(t) - m(t)$ and also

$$(14.12) \qquad \sum_{j=1}^{\infty} u_j^2 = \sum_{i=0}^{n} \sum_{k=0}^{n} e_i e_k (a_i, a_k),$$

$$\sum_{j=1}^{\infty} u_j^2 \lambda_j/(1 - \lambda_j) = \sum_{i=0}^{n} \sum_{k=0}^{n} e_i e_k (Ha_i, a_k).$$

Now (14.11) and (14.12) give (5.16). The formal calculations can be made precise when the hypothesis—(5.6), (5.14) and (5.15)—of Theorem 8 holds. This procedure is an imitation of the sufficiency proof for Theorems 2 and 3 and is omitted.

The necessity of (5.14) is easy because if X is a process with $X(t) \sim W_n(t)$, it follows from Theorem 6 that $X^{(n)}(t) - X^{(n)}(0) \sim W(t)$, and so (5.14) follows from (1.1).

To prove the necessity of (5.6), suppose X is a process with mean m and $X \sim \mathbf{W}_n$. It follows, see Section 12, that $\mathbf{W}_n \sim \mathbf{W}_n + m$. Using the transformation $Y(t) \to Y^{(n)}(t) - Y^{(n)}(0)$ on both sides of the latter equivalence we obtain

$$(14.13)\quad \mathbf{W}_n^{(n)}(t) - \mathbf{W}_n^{(n)}(0) \sim \mathbf{W}_n^{(n)}(t) - \mathbf{W}_n^{(n)}(0) + m^{(n)}(t) - m^{(n)}(0).$$

But $W(t) = \mathbf{W}_n^{(n)}(t) - \mathbf{W}_n^{(n)}(0)$ and so $W \sim W + m^{(n)}(t) - m^{(n)}(0)$. Now (1.1) gives (5.6) immediately.

The necessity of (5.15) is formally obvious but apparently is difficult to prove. We proceed as follows: Suppose X is a process and $X \sim \mathbf{W}_n$. Using the simultaneous representation of X and $\mathbf{W}_n$, (14.3) and (14.4), we see that the sequences of random variables $S_j^1 = \eta_j$, $j = 1, 2, \cdots$, and $S_j^2 = (1 - \lambda_j)^{\frac{1}{2}}\eta_j + k_j + \sum_{i=0}^{n} a_{ij}\xi_i$, $j = 1, 2, \cdots$, are equivalent. Here $\xi_0, \cdots, \xi_n, \eta_1, \cdots$ are independent, standard normal variates. In order to prove (5.15) we must show that

$$(14.14)\qquad \sum_{j=1}^{\infty} a_{ij}^2 < \infty, \qquad i = 0, 1, \cdots, n.$$

Let μ_1 and μ_2 be the measures induced by S^1 and S^2 respectively on the space of infinite sequences. Then $\mu_1 \sim \mu_2$ and so the Hellinger integral [14]

$$(14.15)\qquad \int (d\mu_1\, d\mu_2)^{\frac{1}{2}} > 0.$$

The integral can be evaluated explicitly. Given $\xi_0, \cdots, \xi_n$ both S^1 and S^2 are sequences of independent random variables and so

$$(14.16)\quad \int (d\mu_1\, d\mu_2)^{\frac{1}{2}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\tfrac{1}{2}(\xi_0^2 + \cdots + \xi_n^2)\right) \prod_{j=1}^{\infty} r_j(\xi_0, \cdots, \xi_n)\, d\xi_0 \cdots d\xi_n,$$

where

$$(14.17)\quad r_j = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \exp\left(-\tfrac{1}{4}x^2\right) \exp\left(-\tfrac{1}{4}(x - b_j)^2/(1 - \lambda_j)\right) dx/(1 - \lambda_j)^{\frac{1}{4}},$$

and $b_j = k_j + \sum_{i=0}^{n} a_{ij}\xi_i$, $j = 1, 2, \cdots$. Evaluating the integral (14.17), (14.16) becomes

$$(14.18)\quad \Lambda \cdot \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(-\tfrac{1}{2}(\xi_0^2 + \cdots + \xi_n^2)\right) \cdot \prod_{j=1}^{\infty} \exp\left(-b_j^2/4(2 - \lambda_j)\right) d\xi_0 \cdots d\xi_n,$$

where

$$(14.19)\qquad \Lambda = \prod_{j=1}^{\infty} (1 - \lambda_j)^{\frac{1}{4}}/(1 - \lambda_j/2)^{\frac{1}{2}}.$$

The product (14.19) converges because $\sum \lambda_j^2 < \infty$. Now (14.18) is positive by (14.15) and it follows that

$$(14.20)\qquad \sum_{j=1}^{\infty} b_j^2 < \infty$$

for a set of $\xi_0, \cdots, \xi_n$ of positive measure in Euclidean $n + 1$ space. We have already proved the necessity of (5.6) and so $\sum k_j^2 < \infty$. Thus

$$(14.21)\qquad \sum_{j=1}^{\infty} \left(\sum_{i=0}^{n} a_{ij}\xi_i\right)^2 < \infty$$

for a set of $\xi_0, \cdots, \xi_n$ of positive measure. It follows readily that (14.14) must hold and so (5.15) is proved. This completes the proof of Theorem 8.

IV. EXAMPLES OF R-N DERIVATIVES

The examples given here will be applied in the next section in order to obtain new information about W and other processes. They will also serve to illustrate how Theorems 2 and 8 are applied.

15. Scale changes of W. Let h be an absolutely continuous and *increasing* function on $[0, T]$ with $h(0) = 0$. Define the process

$$Z(t) = [h'(t)]^{-\frac{1}{2}} W(h(t)), \qquad 0 \leq t \leq T. \tag{15.1}$$

As can be seen, the differential increments of Z have the same variance as those of W. Under very general conditions on h, which are given precisely by the following theorem, $Z \sim W$.

THEOREM 11. *$Z \sim W$ if and only if $h' = g^{-2}$ where g is absolutely continuous and $g' = \gamma \varepsilon L^2$.*

For smooth h, say $h \varepsilon C^3$ the R-N *derivative is*

$$(d\mu_Z/d\mu_W)(X) = (h'(T)/h'(0))^{\frac{1}{4}} \cdot \exp\{-[X^2(T)/4]h''(T)/h'(T) - \tfrac{1}{2}\textstyle\int_0^T f(t)X^2(t)\,dt\} \tag{15.2}$$

where $f = -\frac{1}{2}((h''/h')' - \frac{1}{2}(h''/h')^2)$, the Schwarzian derivative of h.

Doob [5] considered the more general scale change $Z(t) = v(t)W(u(t)/v(t))$ with covariance

$$R(s, t) = u(\min(s, t))v(\max(s, t)). \tag{15.3}$$

Such covariances were called *triangular* by Varberg [25], who calculated the R-N derivative for equivalent triangular processes by evaluating directly the limit of the finite dimensional densities. The densities can be written explicitly because Z is a Markov process.

In order to prove the theorem we find from (15.3) that

$$\begin{aligned} K(s, t) &= -u'(s)v'(t), \qquad s \leq t, \\ &= -u'(t)v'(s), \qquad t \leq s, \end{aligned} \tag{15.4}$$

where $u = h/(h')^{\frac{1}{2}}$, $v = (h')^{-\frac{1}{2}}$. Now if $\gamma \varepsilon L^2$ then $v' = \gamma$, $u' = h\gamma + |g|^{-1}$, are both in L^2 and so $K \varepsilon \mathrm{L}^2$. Conversely if $K \varepsilon \mathrm{L}^2$ then v' and hence $\gamma \varepsilon L^2$. We will see later that $1 \not\varepsilon \sigma(K)$ is automatically true and so by Theorem 1 we have proved that $\gamma \varepsilon L^2$ is necessary and sufficient.

Now assume that h is smooth, say $h \varepsilon C^3$. We make this assumption in order to use integration by parts at one step to express $d\mu_Z/d\mu_W$ in as simple a form as possible.

We will calculate the Fredholm determinant and resolvent of K at $\lambda = 1$. The eigenvalue problem $K\varphi = \lambda\varphi$ can be written

$$-v' \textstyle\int_0^t (u'\varphi) - u' \int_t^T (v'\varphi) = \lambda\varphi \tag{15.5}$$

where the arguments have been suppressed. Now it is easy to check that

$$(15.6) \qquad u'' = fu, \qquad v'' = fv,$$

$$(15.7) \qquad u(0) = 0, \qquad v'(T) = -\tfrac{1}{2}v(T)h''(T)/h'(T)$$

where

$$(15.8) \qquad f = -\tfrac{1}{2}[(h''/h')' - \tfrac{1}{2}(h''/h')^2]$$

and so R is in fact a Green function. Letting $y(t) = \int_0^t \varphi$, a direct calculation shows that the eigenvalues λ of (15.5) are those λ for which

$$(15.9) \quad y'' = (1 - \lambda^{-1})fy, \quad y(0) = 0, \quad y'(T)/y(T) = (1 - \lambda^{-1})v'(T)/v(T)$$

has a nontrivial solution y. Now it is clear that $\lambda = 1 \not\in \text{sp}\,(K)$ because if $K\varphi = \varphi$ then by (15.9) $y = 0$ and so $\varphi = 0$. To handle nonsmooth h, we would replace (15.6) and (15.9) by integrated versions, integral equations. The details are straightforward.

We want $d(1) = \prod (1 - \lambda_j)$ where λ_j satisfy (15.9). Define for any complex β the unique solution y_β to

$$(15.10) \qquad y_\beta'' = \beta f y_\beta, \qquad y_\beta(0) = 0, \qquad y_\beta'(0) = 1.$$

Set

$$(15.11) \qquad D(\beta) = y_\beta'(T) - \beta y_\beta(T)v'(T)/v(T).$$

It is known from the general theory of differential equations that $y_\beta(T)$ and $y_\beta'(T)$ depend analytically on β. Furthermore it is not difficult to show that $y_\beta(T)$ is entire and of order $\frac{1}{2}$. Also $y_\beta'(T)$ is entire and of order $\frac{1}{2}$. It follows that $D(\beta)$ is entire and of order $\frac{1}{2}$—hence $D(\beta)$ is its own canonical product:

$$(15.12) \qquad D(\beta) = \prod (1 - \beta/\beta_j),$$

where $\beta_1, \beta_2, \cdots$ are the zeros of D. Now, we have

$$(15.13) \qquad \beta_j = 1 - \lambda_j^{-1}$$

because $D(\beta) = 0$ if and only $\beta = 1 - \lambda^{-1}$ where λ satisfies (15.9) for some nonzero $y = y_\beta$ (note that y_β is nonzero because $y_\beta'(0) = 1$). Therefore

$$d(1) = \prod (1 - \lambda_j) = (\prod (1 - \beta_j^{-1}))^{-1} = (D(1))^{-1}.$$

What is $D(1)$? By (15.11),

$$(15.14) \qquad D(1) = [y_1'(T)v(T) - y_1(T)v'(T)]/v(T).$$

For $\beta = 1$ (and only for this value),

$$(15.15) \qquad y_\beta'(t)v(t) - \beta y_\beta(t)v'(t) = \text{constant}$$

being the Wronskian. Putting $t = 0$ to evaluate the constant we get

$$(15.16) \qquad D(1) = v(0)/v(T) = (h'(T)/h'(0))^{\frac{1}{2}}.$$

Finally,

$$d(1) = (h'(0)/h'(T))^{\frac{1}{2}}. \tag{15.17}$$

Next we calculate the resolvent kernel H. It will turn out that

$$H(s, t) = \theta(\max(s, t)) \tag{15.18}$$

where

$$\theta'(t) = -f(t), \qquad \theta(T) = \tfrac{1}{2}h''(T)/h'(T) \tag{15.19}$$

H is the unique continuous solution to

$$H - K = HK = KH. \tag{15.20}$$

It is straightforward to check that (15.18) satisfies (15.20) and so (15.18) is proved since H is unique in (15.20).

By (2.4) we have since $m = 0$,

$$(d\mu_Z/d\mu_W)(X) = [d(1)]^{-\frac{1}{2}} \exp [-\tfrac{1}{2} \textstyle\int_0^T \int_0^T H(s,t)\, dX(s)\, dX(t). \tag{15.21}$$

Applying (9.13) or proceeding formally (note that $X(0) = 0$) we have

$$\begin{aligned} &\textstyle\int_0^T \int_0^T H(s, t)\, dX(s)\, dX(t) \\ &\quad = 2 \textstyle\int_0^T (\int_0^t H(s, t)\, dX(s))\, dX(t) \\ &\quad = 2 \textstyle\int_0^T (\int_0^t \theta(t)\, dX(s))\, dX(t) = \int_0^T \theta(t)\, dX^2(t) \\ &\quad = X^2(T)\theta(T) + \textstyle\int_0^T X^2(u)f(u)\, du. \end{aligned} \tag{15.22}$$

We obtain (15.2) immediately and Theorem 11 is proved.

Let W^* be the pinned Wiener process,

$$W^*(t) = W(t) - tW(1), \qquad 0 \leqq t \leqq 1. \tag{15.23}$$

The scale change

$$Z(t) = [h'(t)]^{-\frac{1}{2}} W^*(h(t)) \tag{15.24}$$

will have $Z \sim W^*$ when h satisfies the hypotheses of Theorem 11 and in addition

$$h(1) = 1. \tag{15.25}$$

By identical techniques, this time using Theorem 9, one shows that the R-N derivative is

$$(d\mu_Z/d\mu_{W^*})(X) = (h'(0)h'(1))^{-\frac{1}{4}} \exp [-\tfrac{1}{2} \textstyle\int_0^1 f(t)X^2(t)\, dt] \tag{15.26}$$

where f is again the Schwarzian derivative of h. (15.26) will be applied in Section 18 in order to evaluate certain integrals.

16. The linear covariance. Consider the measure μ with $m = 0$ and covariance

$$R(s, t) = \tfrac{1}{2}(1 - |t - s|) \tag{16.1}$$

with $T \leqq 2$. Now $\mu \sim \mu_{W_0}$ and we will use the decomposition method of Theorems 7 and 8 in order to obtain $d\mu/d\mu_{W_0}$. Recall that $W_0(t) = \xi_0 + W(t)$.

For $n = 0$ and $m = 0$, (5.16) becomes

$$
\begin{aligned}
(d\mu/d\mu_{W_0})(X) &= [d(1)\, R(0,0)]^{-\frac{1}{2}} \exp\big[-\tfrac{1}{2} \textstyle\int_0^T\int_0^T H(s,t)\, dX(s)\, dX(t) \\
&\quad + \tfrac{1}{2}X^2(0)(1 - [R(0,0)]^{-1}((I+H)a_0, a_0)) \\
&\quad + X(0)[R(0,0)]^{-\frac{1}{2}} \textstyle\int_0^T (I+H)a_0(t)\, dX(t)\big].
\end{aligned} \tag{16.2}
$$

In the case (16.1) we have $A_0(t) = (1-t)/2^{\frac{1}{2}}$, $a_0(t) = -1/2^{\frac{1}{2}}$. We get in turn

$$
\begin{aligned}
R^*(s,t) &= s - st/2; \qquad s < t, \\
K(s,t) &= \tfrac{1}{2}, \\
H(s,t) &= 1/(2-T),
\end{aligned} \tag{16.3}
$$

so that K and H are constants. The spectrum of K is the single element $\lambda = T/2$. We must exclude $T = 2$ by (1.2) and indeed for $T = 2$ it is clear that $\mu \perp \mu_{W_0}$ since $Z(2) = -Z(0)$ a.s. for $\mu = \mu_Z$. We obtain $d(1) = 1 - T/2$.

We obtain further

$$
\begin{aligned}
(I+H)a_0(t) &= a_0(t) + \textstyle\int_0^T H(t,s)a_0(s)\, ds \\
&= -2^{\frac{1}{2}}/(2-T), \\
((I+H)a_0, a_0) &= T/(2-T), \\
\textstyle\int_0^T\int_0^T H(s,t)\, dX(s)\, dX(t) &= (X(t) - X(0))^2/(2-T).
\end{aligned} \tag{16.4}
$$

We put all the above ingredients into (16.2) and we get

$$
\begin{aligned}
&(d\mu/d\mu_{W_0})(X) \\
&\quad = [2/(2-T)^{\frac{1}{2}}] \exp[X^2(0)/2 - (X(0) + X(T))^2/2(2-T)].
\end{aligned} \tag{16.5}
$$

(16.5) can also be obtained by a direct passage to the limit from the finite dimensional densities of μ, which have been found explicitly by Slepian [23]. The results agree.

We draw attention to the normalization inherent in considering the Wiener process with *unit* diffusion constant. A process Z may be equivalent to σW, for some $\sigma \neq 1$, rather than to W itself. This is only a scale factor and gives no trouble. We get $Z/\sigma \sim W$ and

$$(d\mu_Z/d\mu_{\sigma W})(X) = (d\mu_{Z/\sigma}/d\mu_W)(\sigma^{-1}X). \tag{16.6}$$

Sometimes $\sigma = 1$ is not the most convenient normalization. Instead of (16.1) it is more usual to consider

$$\bar{R}(s,t) = 1 - |t-s|. \tag{16.7}$$

Let us denote by $\bar{\mu}$ the measure with cov $\bar{R}$ and mean 0. Let $\bar{\mu}_0$ denote the measure $\bar{\mu}$ conditioned so that $X(0) = x_0$ given. The mean of $\bar{\mu}_0$ is

$$\bar{m}_0(t) = x_0\bar{R}(0,t)/\bar{R}(0,0) \tag{16.8}$$

and the cov of $\bar{\mu}_0$ is

$$\bar{R}_0(s, t) = \bar{R}(s, t) - \bar{R}(0, s)\bar{R}(0, t)/\bar{R}(0, 0). \tag{16.9}$$

$\bar{\mu}_0$ is a Gaussian measure and is equivalent to the Wiener-type measure corresponding to the process

$$\omega(t) = x_0 + 2^{\frac{1}{2}}W(t), \qquad 0 \leqq t \leqq T. \tag{16.10}$$

The R-N derivative is easy to obtain and is another form of (16.5); it can be computed also from (2.4). We have

$$(d\bar{\mu}_0/d\mu_\omega)(X) = [2/(2 - T)]^{\frac{1}{2}} \exp\left(\tfrac{1}{2}x_0^2\right) \exp\left[-\tfrac{1}{2}(x_0 + X(T))^2/2(2 - T)\right]. \tag{16.11}$$

We will apply (16.11) in Section 17 to solve a certain first passage problem.

The example of this section was chosen for its simplicity. We can obtain with our methods all known examples of R-N derivatives in the Gaussian case as well as many new examples. The calculations, although straightforward, are sometimes quite tedious.

We have given illustrations by example of all the main theorems. The reader may find it instructive to obtain the R-N derivative in the Ornstein-Uhlenbeck case (Example 3 of [25]) by our methods.

V. SOME APPLICATIONS OF R-N DERIVATIVES

The R-N derivative has found its main application in statistics as the likelihood ratio. However, it can also be used as a theoretical tool, as Skorokhod [22], p. 408, pointed out. We will use it both to calculate probabilities and to evaluate function space integrals.

17. Calculating probabilities. We shall prove that if $b > 0$,

$$\Pr\{W(t) < at + b, 0 \leqq t \leqq T\} = \Phi((aT + b)/T^{\frac{1}{2}}) - e^{-2ab}\Phi((aT - b)/T^{\frac{1}{2}}). \tag{17.1}$$

We see that (17.1) is the probability that a gambler with income does not go broke in time $T \leqq \infty$ if b represents his initial capital, a his fixed rate of income (or outgo), and W his losses due to chance fluctuations. (17.1) was found by Doob [5] for $T = \infty$ and was given in general by Malmquist in a different form and by a different method [16].

To prove (17.1) we observe that by definition of $d\mu/d\mu_W$,

$$\mu(A) = \int_A (d\mu/d\mu_W)(X)\, d\mu_W \tag{17.2}$$

for any event A. Let $\mu = \mu_{W+m}$ where $m(t) = -at$ and take $A = \{X: X(t) < b, 0 \leqq t \leqq T\}$. By (2.4) we obtain since $k(t) = -a$,

$$F(X) = (d\mu/d\mu_W)(X) = e^{-aX(T)}e^{-\frac{1}{2}a^2T}. \tag{17.3}$$

Now by (17.2),

$$(17.4) \qquad \mu(A) = e^{-\frac{1}{2}a^2T}\int_A e^{-aX(T)}\, d\mu_W(X)$$
$$= e^{-\frac{1}{2}a^2T}\int_{-\infty}^{b} e^{-ax} \Pr\{M(T) < b;\, W(T) = x\}\, dx$$

where $M(T) = \max_{0 \leqq t \leqq T} W(t)$. But the reflection principle gives for $x < b$,

$$(17.5) \quad \Pr\{M(T) < b;\, W(T) = x\}$$
$$= \Pr\{W(T) = x\} - \Pr\{W(T) = 2b - x\}.$$

Using

$$\Pr\{W(T) = x\} = (2\pi T)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2/T}$$

and $\mu(A) = \Pr\{W(t) < at + b,\, 0 \leqq t \leqq T\}$ we obtain (17.1) after a simple calculation. The point of using (17.2) is that the reflection principle fails for slanted lines $at + b$, while (17.2) enables us to reduce the problem to the horizontal line.

As a second example of the same technique, we solve a first passage problem originally given by Slepian [23] in a different form. Let $S = S(t)$, $0 \leqq t \leqq T$ be the process with mean zero and covariance

$$(17.6) \qquad R(s, t) = 1 - |t - s|, \qquad |t - s| \leqq 1,$$
$$= 0, \qquad |t - s| > 1.$$

Suppose $T \leqq 1$. Let

$$(17.7) \qquad P_a^{+}(T \mid x_0) = \Pr\{S(t) < a,\, 0 \leqq t \leqq T \mid S(0) = x_0\}, \qquad x_0 < a,$$
$$P_a^{-}(T \mid x_0) = \Pr\{S(t) > a,\, 0 \leqq t \leqq T \mid S(0) = x_0\}, \qquad x_0 > a,$$

be the first passage probabilities.

We shall prove

$$(17.8) \quad P_a^{\pm}(T \mid x_0) = \Phi(\pm[a - x_0(1 - T)]/[T(2 - T)]^{\frac{1}{2}})$$
$$- e^{-\frac{1}{2}(a^2 - x_0^2)}\Phi(\pm[x_0 - a(1 - T)]/[T(2 - T)]^{\frac{1}{2}}).$$

The formulas break down for $T > 1$ and in this case the problem remains unsolved.

Let S_0 denote the process S conditioned to pass through x_0 at $t = 0$. We have seen that $S_0 \sim \omega = x_0 + 2^{\frac{1}{2}}W$ and the R-N derivative is given by (16.11). Let

$$A = \{X : X(t) \leqq a,\, 0 \leqq t \leqq T\}.$$

Applying (17.2) with $\mu = \mu_{S_0}$ and μ_W replaced by μ_ω we get

$$(17.9) \quad P_a^{+}(T \mid x_0)$$
$$= [2/(2 - T)]^{\frac{1}{2}} e^{\frac{1}{4}x_0^2}\int_A \exp\left[-\tfrac{1}{2}(x_0 + X(T))^2/2(2 - T)\right] d\mu_\omega(X).$$

Using the reflection principle to evaluate the integral, we obtain (17.8). P^{-} is obtained in a similar way. In comparing our form of the answer with Slepian's note that his $Q_a(T \mid x_0)$ is the derivative on T of our $P_a(T \mid x_0)$.

In principle one could use the method of this section to solve other first passage problems. However, these are the only cases in which we could evaluate the integrals explicitly.

18. Evaluating some special integrals. The integral of the R-N derivative over all space is unity. This fact will permit us to evaluate the Wiener integral of any functional that is the exponential of a quadratic form. Some special cases of this result are due to Kac and Siegert [12] who obtained them by other methods. Their results were later simplified by Anderson and Darling [1]. We will then apply the evaluations to prove a dichotomy theorem about Wiener paths.

We begin with a general identity. *Let L be any symmetric continuous kernel of trace class. Then*

$$(18.1) \qquad E_W \exp\left[\tfrac{1}{2}\lambda \int_0^T \int_0^T L(s, t)\, dX(s)\, dX(t)\right] = (D^L(\lambda))^{-\frac{1}{2}}$$

where E_W is the Wiener integral, d^L is the Fredholm determinant of L, and λ satisfies

$$(18.2) \qquad 1 - \lambda\lambda_j > 0, \qquad j = 1, 2, \cdots,$$

where $\lambda_1, \lambda_2, \cdots$ are the eigenvalues of L. When (18.2) fails the intearal is $+\infty$.

To prove (18.1) note that by (2.4) we have

$$(18.3) \qquad 1 = E_W[d(1)]^{-\frac{1}{2}} \exp\left[-\tfrac{1}{2}\int_0^T \int_0^T H(s, t)\, dX(s)\, dX(t)\right]$$

where $d(1) = d^K(1)$. Now if γ denotes eigenvalues of H then $\gamma - \lambda = \gamma\lambda$ and so

$$(18.4) \qquad d^K(1) = \prod_\lambda (1 - \lambda) = \prod_\gamma (1 + \gamma)^{-1} = d^H(-1) = d^{-H}(1).$$

Let now $L = -\lambda H$ and we obtain (18.1). When (18.2) fails the quadratic form

$$(18.5) \qquad (\varphi, \varphi) - \lambda(L\varphi, \varphi)$$

is not positive-definite and it is at least formally clear that (18.1) should diverge. The proof is omitted. It is amusing to expand both sides of (18.1) in powers of λ and check coefficients of λ^n. The general equality seems to be difficult to prove directly.

Next we evaluate

$$(18.6) \qquad A(f) = E_W \exp\left[-\tfrac{1}{2}\int_0^T f(t) X^2(t)\, dt\right]$$

for any, say continuous, f. *Solve the one-point problem*

$$(18.7) \qquad g'' = fg, \qquad g'(T) = 0, \qquad g > 0 \text{ on } [0, T).$$

Then

$$(18.8) \qquad A(f) = (g(T)/g(0))^{\frac{1}{2}}.$$

When (18.7) fails to have a solution positive on the half-open interval $[0, T)$ then

$A(F) = +\infty$. The result is also valid for $T = \infty$ provided $\lim g(x) = g(\infty) > 0$ and $\lim g'(x) = g'(\infty) = 0$.

Suppose now that the basic process is W^*, the pinned Wiener process, instead of W. *We have*

$$E_{W^*} \exp [-\tfrac{1}{2}\textstyle\int_0^1 f(t)X^2(t)\, dt] = (l(1))^{-\frac{1}{2}} \tag{18.9}$$

where l is the unique solution to

$$l'' = fl, \qquad l(0) = 0, \qquad l'(0) = 1, \tag{18.10}$$

provided $l > 0$ in $(0, 1]$. *When the solution l of* (18.10) *has a zero in* $(0, 1]$ *then* (18.9) *is* $+\infty$.

For positive f, (18.9) is finite and was obtained by Anderson and Darling [1] who simplified a previous formula due to Kac and Siegert [13].

To prove (18.9) we proceed as follows. By (15.26) we have for any h with $h(0) = 0$, $h(1) = 1$ and $h' > 0$,

$$E_{W^*} \exp [-\tfrac{1}{2}\textstyle\int_0^1 \bar{f}(t)X^2(t)\, dt] = (h'(0)h'(1))^{\frac{1}{2}} \tag{18.11}$$

where $\bar{f} = -\frac{1}{2}((h''/h')' - \frac{1}{2}(h''/h')^2)$. Now choose h to satisfy

$$h(0) = 0, \qquad h' = g^{-2}/\textstyle\int_0^1 g^{-2} \tag{18.12}$$

where g is any *positive* solution of

$$g'' = fg. \tag{18.13}$$

We have $h' > 0$, $h(1) = 1$ and the Schwarzian of h is f,

$$\bar{f} = -\tfrac{1}{2}((h''/h')^1 - \tfrac{1}{2}(h''/h')^2) = g''/g = f. \tag{18.14}$$

By (18.11) we have

$$E_{W^*} \exp [-\tfrac{1}{2}\textstyle\int_0^1 f(t)X^2(t)\, dt] = (g(0)g(1)\textstyle\int_0^1 g^{-2})^{-\frac{1}{2}}. \tag{18.15}$$

Let

$$l(t) = g(0)g(t)\textstyle\int_0^t g^{-2}. \tag{18.16}$$

It is easy to check that l satisfies (18.10) and this gives (18.9). Now it is known from Sturm-Liouville theory that a positive solution g exists to (18.13) if and only if the solution l of (18.10) is positive in $(0, 1]$. It is easy to see that when l has a zero then (18.9) is actually $+\infty$. A similar proof can be given for (18.8). The formula (18.9) for W^* is more symmetric than (18.8) for W as is evident by (18.15). This is because W^* is time-reversible.

We have assumed f to be continuous but this is not necessary. In general, one expresses the differential equation $g'' = fg$, $g'(T) = 0$ in the form of an integral equation

$$g(t) = g(T) + \textstyle\int_t^T (u - t)g(u)f(u)\, du. \tag{18.17}$$

One can now allow f to be simply measurable. Further, f can be replaced by a measure: $f du = dm$.

As mentioned above $A(f)$ has been calculated for continuous $f > 0$ by Kac and others by using various techniques. However, the most direct way to obtain $A(f)$ in this case was found by Gel'fand and Yaglom [29]:

By Riemann integration and bounded convergence we have

$$(18.18) \qquad A(f) = \lim_{n\to\infty} E_W \exp\left[-\tfrac{1}{2}\textstyle\sum_1^n f(jT/n)X^2(jT/n)T/n\right].$$

Now the expected value on the right is an n-dimensional integral that can be evaluated explicitly. Let $f_j = f(jT/n)$, $x_j = X(jT/n)$. The integral in (18.18) is

$$(18.19) \qquad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left[-\tfrac{1}{2}\textstyle\sum f_j X_j^2 T/n - \tfrac{1}{2}\sum (x_j - x_{j-1})^2 n/T\right] dx_1 \cdots dx_n/(2\pi T/n)^{n/2} = (|b_{ij}^{(n)}|_1^n)^{-\frac{1}{2}}$$

where

$$\begin{aligned} b_{ij}^{(n)} &= 2 + f_j T/n^2, & i &= j = 1, \cdots, n-1, \\ &= -1, & |i - j| &= 1, \\ &= 1 + f_n T^2/n^2, & i &= j = n, \\ &= 0, & |i - j| &> 1. \end{aligned}$$

Define

$$(18.20) \qquad g_k = g_k^{(n)} = |b_{ij}^{(n)}|_k^n, \qquad k = 1, 2, \cdots, n.$$

We see that $g_n = 1 + f_n T^2/n^2$, $g_n - g_{n-1} = O(T^2/n^2)$ and

$$(18.21) \qquad \Delta^2 g_k = f_{k+1} g_k, \qquad k = 1, \cdots, n-1.$$

We compare (18.21) with the differential equation

$$(18.22) \qquad g'' = fg, \qquad g(T) = 1, \qquad g'(T) = 0,$$

with $g(kT/n) \leftrightarrow g_k$. Standard techniques of comparison of difference and differential equations give

$$(18.23) \qquad A(f) = \lim (g_1^{(n)})^{-\frac{1}{2}} = (g(0))^{-\frac{1}{2}}$$

which agrees with (18.8) since $g(T) = 1$.

If μ and ν are Gaussian then $d\mu/d\nu(X)$ is always the exponential of a quadratic form in X. With $\nu = \mu_W$, *the quadratic form is diagonal,*

$$(18.24) \qquad \int_0^T X^2(u) f(u)\, du,$$

if and only if $\mu = \mu_Z$ *where* Z *is a scale change as in* (15.1). Indeed, as we have seen, (18.24) holds for $\mu = \mu_Z$. Conversely, we can obtain (18.24) for any f by means of a properly chosen scale change. Thus (18.24) characterizes scale change processes.

19. A dichotomy theorem for Wiener paths. Suppose $f \geq 0$ on $[0, T]$. *The integral*

$$(19.1) \qquad \int_0^T f(u) W^2(u)\, du$$

either converges a.s. *or diverges* a.s. *according as*

$$\int_0^T uf(u)\,du \tag{19.2}$$

converges or diverges. The convergence is trivial because if (19.2) is finite then

$$E_W\left(\int_0^T f(u)W^2(u)\,du\right) = \int_0^T uf(u)\,du < \infty \tag{19.3}$$

and so (19.1) must be finite a.s.

Now suppose $\int_0^T uf(u)\,du = \infty$. We will prove that

$$E_W \exp\left[-\int_0^T f(u)X^2(u)\,du\right] = 0 \tag{19.4}$$

and so $\int_0^T f(u)W^2(u) = \infty$ a.s. will follow. In order to avoid details we assume that f is continuous away from zero. By (18.8) it is enough to prove the following simple lemma.

LEMMA. *Suppose f is continuous on $(0, T]$ and $\int_0^T uf(u)\,du = \infty$. Then*

$$g'' = fg, \qquad g'(T) = 0, \qquad g(T) = 1, \tag{19.5}$$

has $g(0) = \infty$.

PROOF. Since g is convex and $g'(T) = 0$, g is monotonically decreasing. Suppose $g(0) = M < \infty$. Then $0 < -\epsilon g'(\epsilon) \leqq g(0) - g(\epsilon) \leqq M$ by convexity. We have, since $g \geqq 1$, for any $\epsilon > 0$

$$\begin{aligned}\int_\epsilon^T uf(u)\,du &< \int_\epsilon^T uf(u)g(u)\,du \\ &= \int_\epsilon^T ug''(u)\,du = Tg'(T) - \epsilon g'(\epsilon) - g(T) + g(\epsilon) \leqq 2M.\end{aligned} \tag{19.6}$$

This contradicts $\int_0^T uf(u)\,du = \infty$ and proves the lemma.

AN EXAMPLE. Let $f(u) = 2/(\epsilon + u)^2$, $0 < \epsilon$, $T < \infty$. The solution to (18.7) is

$$g(u) = A(\epsilon + u)^{-1} + B(\epsilon + u)^2, \qquad A = 2B(\epsilon + T)^3. \tag{19.7}$$

We obtain

$$E_W \exp\left[-\int_0^T (\epsilon + u)^{-2}W^2(u)\,du\right] = (3\epsilon/2(\epsilon + T)(1 + (\epsilon/\epsilon + T)^3/2))^{\frac{1}{2}}. \tag{19.8}$$

For other examples where (18.7) has an explicit solution see [1]. Letting $\epsilon \to 0$ we have by monotone convergence

$$E_W \exp\left[-\int_0^T u^{-2}W^2(u)\,du\right] = 0 \tag{19.9}$$

and so

$$\int_0^T u^{-2}W^2(u)\,du = \infty \tag{19.10}$$

for almost every path W. (19.10) also follows from a form of the iterated logarithm theorem recently found by V. Strassen [24]. However, it does not seem possible to obtain the general case of (19.2) in this way.

REFERENCES

[1] ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–212.

[2] BAXTER, G. (1956). A strong limit theorem for Gaussian processes. *Proc. Amer. Math. Soc.* **7** 522–527.

[3] CAMERON, R. H. and MARTIN, W. T. (1944). Transformations of Wiener integrals under translations. *Ann. Math.* **45** 386–396.

[4] CAMERON, R. H. and MARTIN, W. T. (1945). Transformations of Wiener integrals under a general class of linear transformations. *Trans. Amer. Math. Soc.* **58** 184–219.

[5] DOOB, J. L. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **20** 393–403.

[6] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[7] DUNFORD, N. and SCHWARZ, J. T. (1963). *Linear Operators Part II.* Interscience, New York.

[8] GANTMACHER, F. R. (1959). *The Theory of Matrices.* **1** Chelsea, New York.

[9] HALMOS, P. R. (1950). *Measure Theory.* Van Nostrand, New York.

[10] ITO, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 157–169.

[11] ITO, K. and MCKEAN, H. P. (1965). *Diffusion Processes and Their Sample Paths.* Academic Press, New York.

[12] KAC, M. (1951). On some connections between probability theory and differential and integral equations. *Proc. Second Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[13] KAC, M. and SIEGERT, A. J. F. (1947). On the theory of noise in radio receivers with square law detectors. *J. Appl. Phys.* **18** 383–397.

[14] KAKUTANI, S. (1948). On equivalence of infinite product measures. *Ann. Math.* **49** 214–224.

[15] KALLIANPUR, G. and OODAIRA, H. (1962). The equivalence and singularity of Gaussian measures, Chap. 19. *Proc. Symp. Time Series Analysis,* (M. Rosenblatt, ed.). Wiley, New York.

[16] MALMQUIST, S. (1954). On certain confidence contours for distribution functions. *Ann. Math. Statist.* **25** 523–533.

[17] RAO, C. R. and VARADARAJAN, V. S. (1963). Discrimination of Gaussian processes. *Sankhyā Ser. A* **25** 303–330.

[18] RIESZ, F. and SZ-NAGY, B. (1955). *Functional Analysis.* Ungar, New York.

[19] ROZANOV, YU. A. (1964). On probability measures in function spaces, corresponding to Gaussian stationary processes (in Russian). *Teor. Verojatnost. i Primenen.* **9** 448–465.

[20] SEGAL, I. E. (1958). Distributions in Hilbert space and canonical systems of operators. *Trans. Amer. Math. Soc.* **88** 12–41.

[21] SEGAL, I. E. (1965). Algebraic integration theory. *Bull. Amer. Math. Soc.* **71** 419–489.

[22] SKOROKHOD, A. V. (1957). On the differentiability of measures which correspond to stochastic processes. *Theor. Prob. Appl.* **2** 407–432.

[23] SLEPIAN, D. (1961). First passage time for a particular Gaussian process. *Ann. Math. Statist.* **32** 610–612.

[24] STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiets* **3** 211–226.

[25] VARBERG, D. E. (1961). On equivalence of Gaussian measures. *Pacific J. Math.* **11** 751–762.

[26] VARBERG, D. E. (1964). On Gaussian measures equivalent to Wiener measure. *Trans. Amer. Math. Soc.* **113** 262–273.

[27] WOODWARD, D. A. (1961). A general class of linear transformations of Wiener integrals. *Trans. Amer. Math. Soc.* **100** 459–480.

[28] YAGLOM, A. M. (1963). On the equivalence and perpendicularity of two Gaussian probability measures in function space, Chapter 22. *Proc. Symp. Time Series Analysis* (M. Rosenblatt, ed.). Wiley, New York.

[29] GEL'FAND, I. M. and YAGLOM, A. M. (1956). Integration in function spaces (in Russian). *Uspekhi Mat. Nauk.* **11** 77–114.

[30] DELPORTE, J. (1964). Fonctions aléatoires presque sûrement continues sur un intervalle fermé. *Ann. Inst. Henri Poincaré Sec. B* **1** 111–215.