Journal of Statistical Physics, Vol. 1, No. 3, 1969

# New Applications of Wiener Integrals to Engineering and Physics[1]

Michael Schilder[2]

Received June 24, 1969

It is shown that the solutions to quite general problems in nonequilibrium statistical physics and engineering can be expressed as Wiener integrals. A new way is also given for numerically evaluating these Wiener integrals.

**KEY WORDS:** Wiener integrals; diffusion processes; function space integrals; stochastic optimal control; Kalman filtering; Langevin equation; stochastic Hamilton–Jacobi equation; Fokker–Planck equation.

## 1. INTRODUCTION

**1.1.** A new method will be given for obtaining numerical answers to quite general problems in nonlinear statistical control theory and in nonequilibrium statistical physics. It will also be shown that the solution to the nonlinear Kalman filtering problem can be expressed as the ratio of the solutions of a partial differential equation.

We consider stochastic equations of the type

$$\ddot{x}(t) = f\big(t, x(t), \dot{x}(t)\big) + \lambda\dot{z}(t), \qquad x(S) = X, \qquad \dot{x}(S) = Y \tag{1}$$

where $\dot{z}(t)$ is the derivative of Brownian motion, or white noise, $x(t)$ is an $n$-dimensional vector, and $f(t, x(t), \dot{x}(t))$ is a fixed vector-valued function of $n$ arguments. The initial values of the Ito stochastic equation (1) can be either fixed or random. Boundary conditions can also be included.

It will be shown that the expected values of functionals with respect to the $x(t)$ distribution defined by (1) can be expressed as Wiener integrals. Thus, the variance, or the mean of $x(t)$ for any $t \geqslant S$, can be expressed as a Wiener integral; or the probability that $x(t)$ is in any (measurable) set of Euclidean space at any time $t \geqslant S$.

[1] This is a rewritten version of Research Report RC 68-1, General Precision Systems, Inc., Little Falls, New Jersey.

[2] Department of Mathematics, Newark College of Engineering, Newark, New Jersey.

A new method will then be given, involving a power series in $\lambda$, for numerically evaluating Wiener integrals. The coefficients of this power series can be calculated by quadratures if the solution to (1) with $\lambda = 0$ is known and the solutions to a linear and a ricatti equation are known. This method essentially provides a rigorous justification for "linearizing" (1) and gives the series of best approximation.

It will also be shown that the representation of the expected values of functionals of solutions of (1) as Wiener integrals gives a new, easy way of deriving some partial differential equations of control theory and physics.

Finally, it will be shown that Wiener integrals can be used to choose the $f$ of (1) in such a way that the expected values of functionals are minimized, much in the same way that the calculus of variations is used to minimize nonrandom integrals.

The remainder of this paper is organized as follows: Section 2 defines the Wiener integral and gives some of the basic theory needed for the other sections. In Section 3, three partial differential equations are derived. Section 4 considers specific applications to statistical control theory, and Section 5 considers specific applications to statistical physics. Section 6 gives an approximation method.

## 2. THE WIENER INTEGRAL

**2.1.** The Wiener integral, named after Norbert Wiener, who first devised it (see Kac[24] for an excellent history of the Wiener integral), is the integral associated with Brownian motion. Brownian motion is, of course, the motion of a small particle suspended in a fluid which is due to the impacts of the molecules of the fluid. Einstein was the first to axiomitize this motion. His axioms can be stated as follows:

($\gamma_1$) Given a particle at position $A$ at time $S$, its motion after time $S$ is independent of the motions of the particle before time $S$.

($\gamma_2$) The distribution, at time $t_1 > S$, of the position of a particle which was at position $A$ at time $S$ is normal (Gaussian) with mean $A$ and variance $t_1 - S$.

We now make two assumptions which will hold throughout the remainder of this paper. The mass of the particles considered is unity, and all particles move in one dimension only. All the results presented in this paper are invariant with respect to the dimension of the space in which the particle moves, so these assumptions will lead to considerable notational convenience. All results can be considered as vector results, if the proper transpositions are made.

From axiom ($\gamma_2$), the probability that the position of the particle at time $t_1$ is less than $A$, given it was at $A$ at time $S$, which we will denote as

$$\operatorname{prob}\{z(t_1) \leqslant A_1 \mid z(S) = A\},$$

is

$$\frac{1}{[2\pi(t_1 - S)]^{1/2}} \int_{-\infty}^{A_1} \exp\left\{-\frac{[z(t_1) - A]^2}{2(t_1 - S)}\right\} dz(t_1) \tag{2}$$

Using axioms ($\gamma_1$) and ($\gamma_2$) and some elementary probability theory, we find

$$\begin{aligned}&\operatorname{prob}\{z(t_1) \leqslant A_1, z(t_2) \leqslant A_2, \ldots, z(t_n) \leqslant A_n \mid z(S) = A\} \\ &= \frac{1}{[(2\pi)^n (t_n - t_{n-1}) \cdots (t_1 - t_0)]^{1/2}} \int_{-\infty}^{A_1} \cdots \int_{-\infty}^{A_n} \exp\left\{-\frac{1}{2} \sum \frac{[z(\tau_i) - z(\tau_{i-1})]^2}{t_i - t_{i-1}}\right\} \\ &\quad \times dz(t_1) \cdots dz(t_n) \end{aligned} \tag{3}$$

where $t_0 = S$, and $z(t_0) = A$. The expression (3) gives the probability that the Brownian-motion path $z(t)$ is in the set (see Fig. 1)

$$\{z(S) = A, z(t_1) \leqslant A_1, \cdots, z(t_n) \leqslant A_n\} \tag{4}$$

We note that, if, for example, we want the probability that $z(t) \leqslant 10, S \leqslant t \leqslant T$, then we must extend the number of integrations from $n$ to infinity. The integral on the right-hand side of (3) with $n$ finite or infinite is called a Wiener integral. For a more detailed exposition of the Wiener integral, see Gelfand and Yaglom,[15] Kac,[23] or Wiener *et al.*[44] The Wiener integral of some function of the paths

$$(z(S) = A, S \leqslant t \leqslant T)$$

will be denoted by $E_z^w\{F(z)|z(S) = A\}$.

The Wiener integral can be used to study Brownian motion and the processes closely related to Brownian motion, such as the process (1). For example, the autocorrelation coefficient of a stochastic process, $\rho(t_1, t_2)$, is defined as

$$\rho(t_1, t_2) = \left\langle [z(t_1) - \langle z(t_1)\rangle_{av}][z(t_2) - \langle z(t_2)\rangle_{av}] \right\rangle_{av}$$

where the argular brackets denote averaging.

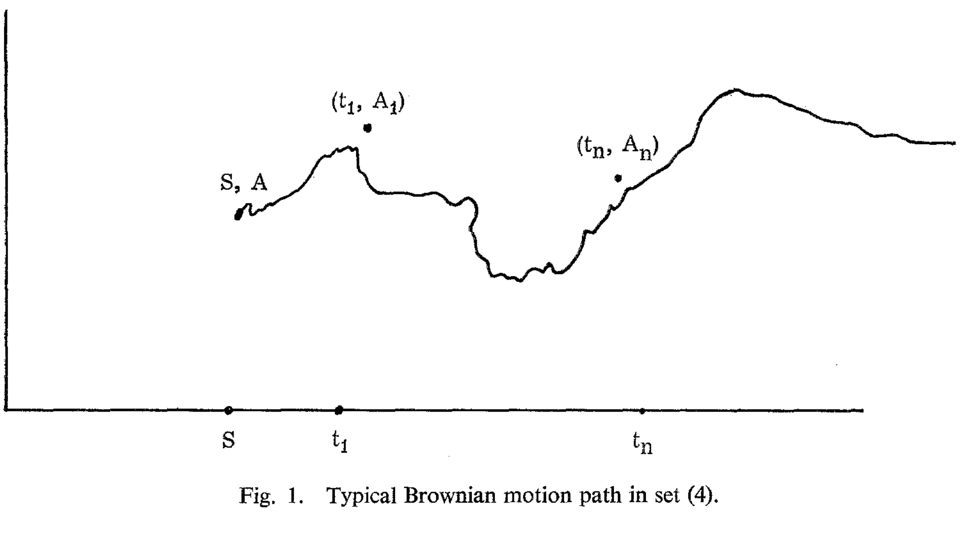


Fig. 1. Typical Brownian motion path in set (4).

In the present case, since, by definition [or by the same reasoning that leads to (2)],

$$E_z^w\{z(t_1)|\, z(S) = A\} = E_z^w\{z(t_2)|\, z(S) = A\} = A$$

We have, if $t_2 \geqslant t_1$,

$$\begin{aligned} \rho(t_1, t_2) &= E_z^w\{[z(t_1) - A][z(t_2) - A]|\, z(S) = A\} \\ &= \frac{1}{[(2\pi)^2 (t_2 - t_1)(t_1 - S)]^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [z(t_1) - A][z(t_2) - A] \\ &\quad \times \exp\left\{-\frac{1}{2}\left[\frac{[z(t_1) - A]^2}{t_1 - S} + \frac{[z(t_2) - z(t_1)]^2}{t_2 - t_1}\right]\right\} dz(t_1)\, dz(t_2) \end{aligned}$$

Some calculation shows this to be $t_1 - S$. If $t_1 \geqslant t_2$, then in the same way, $\rho(t_1, t_2) = t_2 - S$. Therefore, we have used the Wiener integral to prove that the correlation coefficient of Brownian motion is

$$\rho(t_1, t_2) = \min(t_1 - S, t_2 - S) \tag{5}$$

In the same way, it follows

$$\begin{aligned} &E_z^w\left\{\frac{z(t + \Delta) - z(t)}{\Delta} \,\Big|\, z(S) = A\right\} \\ &= \frac{1}{[(2\pi)^2 (\Delta)(t - S)]^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{z(t + \Delta) - z(t)}{\Delta} \\ &\quad \times \exp\left\{-\frac{1}{2}\left[\frac{[z(t + \Delta) - z(t)]^2}{\Delta} + \frac{[z(t) - A]^2}{t - S}\right]\right\} dz(t)\, dz(t + \Delta) \\ &= 0 \qquad \text{if} \quad \Delta \neq 0 \end{aligned} \tag{6a}$$

and

$$\begin{aligned} &E_z^w\left\{\left[\frac{z(t + \Delta) - z(t)}{\Delta}\right]^2 \,\Big|\, z(S) = A\right\} \\ &= \frac{1}{[(2\pi)^2 (\Delta)(t - S)]^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\frac{z(t + \Delta) - z(t)}{\Delta}\right]^2 \\ &\quad \times \exp\left\{-\frac{1}{2}\left[\frac{[z(t + \Delta) - z(t)]^2}{\Delta} + \frac{[z(t) - A]^2}{t - S}\right]\right\} dz(t)\, dz(t + \Delta t) \\ &= \frac{1}{\Delta} \qquad \text{if} \quad \Delta \neq 0 \end{aligned} \tag{6b}$$

Therefore, while the mean of $\dot{z}(t)$ can be defined as

$$\lim_{\Delta \to 0} E_z^w\{[z(t + \Delta) - z(t)/\Delta]|\, z(S) = A\} = 0$$

the variance of $\dot{z}(t)$, if it exists, is clearly infinite, by (6).

**2.2.** One might therefore equestion exactly what Eq. (1) means. Ito[21] was the first to show rigorously how stochastic equations of the type

$$\dot{x}(t) = f(t, x(t)) + \lambda\dot{z}(t)$$

could be defined (however, this type of equation was considered before Ito by physicists as the Langevin equation; see Chandrasekhar[6]).

It will now be shown how stochastic equations of the type (1) can be rigorously defined using a generalization of Ito's method. We suppose $\dot{x}(t)$ is defined as

$$\dot{x}(t) = \lambda y(t) \tag{7}$$

Since $x(S) = X$ and $\dot{x}(S) = Y$, it follows that

$$x(t) = \lambda \int_S^t y(\alpha)\, d\alpha + X \tag{8}$$

$$\lambda y(S) = \dot{x}(S) = Y \tag{9}$$

If we substitute (7)–(9) into (1), the result is

$$\lambda\dot{y}(t) = f\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) + \lambda\dot{z}(t), \qquad \lambda y(S) = Y \tag{10}$$

Integrating both sides of (10) from $S$ to $t$ and dividing by $\lambda$, we get

$$y(t) = (1/\lambda) \int_S^t f\left(\alpha, \lambda \int_S^\alpha y(\beta)\, d\beta + X, \lambda y(\alpha)\right) d\alpha + z(t) - A + Y/\lambda \tag{11}$$

where the substitutions $z(S) = A$ and $y(S) = Y/\lambda$ have been made.

Since it can be proved that Wiener measure can be defined so that almost all $z$'s are in $C[S, A, T]$, where $C[S, A, T]$ is the set of all continuous functions on the interval $[S, T]$ whose value at $S$ is $A$, it follows that solutions to the integral equation (11) can be defined in the usual manner. If $f(\cdot, \cdot, \cdot)$ is not suitably restricted, of course, (11) may have none or more than one solution for some $z$'s in $C[S, A, T]$. If, however, certain conditions are put on $f$ (see, for example, Doob[8] or Dynkin[9]), then (11) has exactly one solution, $y(t)$, for each $z \in C[S, A, T]$.

This correspondence can be written explicitly as

$$y(t) = T(t, z), \tag{12}$$

since, as stated above, for each $z(t)$, $S \leqslant t \leqslant T$, there is exactly one solution, $y(t)$, of (11). Equation (12) just writes out this correspondence. It is, of course, a nonlinear transformation of $C[S, A, T]$ into $C[S, Y/\lambda, T)$, where $Y/\lambda$ is, of course, the initial position of the $y(t)$ process. We can now define a $\sigma$ field and a probability distribution on the $y$ space using (12). The distribution on the $y$ space is called the Ito distribution, because it is induced by the Ito equation (1). If $F(y)$ is some measurable, integrable function defined on the $y$ space, we denote its integral as

$$E_y^I\{F(y) \mid y(S) = Y/\lambda\} \tag{13}$$

We *define* its integral to be [using (12)]

$$E_y^I\{F(y(\cdot))|\, y(S) = Y/\lambda\} = E_z^w\{F(T(\cdot\,, z))|\, z(S) = A\} \tag{14}$$

For an exposition of the theory of measure-preserving transformations, see Halmos.[19] Since the integral (13) is now defined, it is possible to define *any* parameter of the $y$ distribution. The variance of $y$ at time $T$ is, for example,

$$E_y^I\{ y^2(T)|\, y(S) = Y/\lambda\} = E_z^w\{[T(T, z)]^2|\, z(S) = A\}$$

As another example, let $B$ be a measurable set in $C[S, Y/\lambda, T]$ and let $\chi_B(y)$ be the following function:

$$\begin{aligned} \chi_B(y) &= 1 \qquad \text{if} \quad y \in B \\ &= 0 \qquad \text{if} \quad y \notin B \end{aligned} \tag{15}$$

Then the probability $y$ is in $B$ ($B$ might be, for example, $\{y \in C[S, Y/\lambda, T]\; y(t) \leqslant 10,$ $S \leqslant t \leqslant T\}$)is, by definition $E_y^I\{\chi_B(y)|\, y(S) = Y/\lambda\}$, which is, by (14), $E_z^w\{\chi_B(T(\cdot,z))|$ $z(S) = A\}$.

It will be shown (Theorem 2.1), by changing variables, that the Wiener integral on the right-hand side of (14) can be expressed in a manner which does not require solution of the nonlinear integral equation (11) for all $z \in C[S, A, T]$, which is, of course, in general, a hopeless task. Now that the $y$ integral and distribution have been defined, it is possible, in the same way, using the 1–1 set of transformation equations (7)–(9), to define the $x$ distribution of Eq. (1), which we again call the Ito distribution. Suppose $F(x(\cdot), \dot{x}(\cdot))$ is a function which is measurable and integrable with respect to the $x$ distribution to be defined. Then, by definition,

$$\begin{aligned} &E_x^I\{F(x(\cdot), \dot{x}(\cdot))|\, x(S) = X, \dot{x}(S) = Y\} \\ &= E_y^I \left\{F\left(\lambda \int_S^{(\cdot)} y(\alpha)\, d\alpha + X, \lambda y(\cdot)\right)\middle|\, y(S) = Y/\lambda\right\} \\ &= E_z^w \left\{F\left(\lambda \int_S^{(\cdot)} T(\alpha, z)\, d\alpha + X, \lambda T(\cdot\,, z)\right)\middle|\, z(S) = A\right\} \end{aligned} \tag{16}$$

That this definition is, in fact, the same as the classical definition will be seen when it is shown that the distribution function of $x(t)$ and $\dot{x}(t)$ satisfies the classical backwards equation.

The case when Eq. (11) does not have unique solutions, while not considered in this paper, is interesting and physically relevant. See Cameron[4] for some details on this question.

**2.3.** Now let $G(t, z(\cdot))$ be a measurable and (possibly) nonlinear function space operator which takes continuous functions on $[S, T]$ into continous functions on $[S, T]$ and is such that $G(t, z)$ does not depend on $z$ past time $t$. The $G(t, z)$ might be, for example, $G(t, z(\cdot)) = \int_S^t z(\alpha)\, d\alpha + X$, or the transformation (12). However,

$G(t, z(\cdot))$ could not be $G(t, z(\cdot)) = \int_S^T z(\alpha)\, d\alpha$, $t < T$, since this $G(t, z(\cdot))$ depends on $z$ past time $t$. Ito[21] has defined an integral

$$\int_S^T f\Big(t, G(t, z(\cdot)), z(t)\Big)\, dz(t)$$

which makes sense for $f$'s that are necessarily continuous. (See also Doob,[8] Dynkin,[9] and Nelson.[32] Nelson calls this integral the Wiener integral, and gives no name to what we and most other authors call the Wiener integral).

It can be shown that (see Dynkin[9]), for rather weak conditions on $f$ and $G$,

$$E_z^w \left\{ \int_S^t f\Big(\alpha, G(\alpha, z(\cdot)), z(\alpha)\Big)\, dz(\alpha) \Big| \, z(S) = A \right\} = 0 \tag{17}$$

and

$$E_z^w \left\{ \Big( \int_S^t f(\alpha, G(\alpha, z(\cdot)), z(\alpha))] d\alpha \Big)^2 \Big| \, z(S) = A \right\}$$
$$= E_z^w \left\{ \int_S^t f\Big(\alpha, G(\alpha, z(\cdot)), z(\alpha)\Big)^2 d\alpha \Big| \, z(S) = A \right\} \tag{18}$$

**2.4.** Three lemmas will now be proved which will be needed later. They are a generalization of a result given in Gelfand and Yaglom.[15]

**Lemma 2.1.** If $f_1(\alpha),\ldots,f_n(\alpha)$ are in $L^2[S, T]$, then

$$E_z^w \left\{ \int_S^t f_1(\alpha_1)\, dz(\alpha_1) \cdots \int_S^t f_m(\alpha)\, dz(\alpha) \Big| \, z(S) = A \right\} = 0 \qquad \text{if} \quad m \quad \text{is odd,}$$
$$\int_S^t f_1(\alpha) f_2(\alpha)\, d\alpha \int_S^t f_3(\alpha) f_4(\alpha)\, d\alpha \cdots \int_S^t f_{m-1}(\alpha) f_m(\alpha)\, d\alpha$$
$$+ \int_S^t f_1(\alpha) f_3(\alpha)\, d\alpha \cdots \int_S^t f_{n-2}(\alpha) f_n(\alpha)\, d\alpha \tag{19}$$
$$+ \cdots + \int_S^t f_1(\alpha) f_m(\alpha)\, d\alpha + \cdots + \int_S^t f_{(m/2)-1}(\alpha) f_{(m/2)+1}(\alpha)\, d\alpha \qquad \text{if} \quad m \quad \text{is even}$$

*Proof.* The integral $\int_S^t f_i(\alpha)\, dz(\alpha)$ is defined to be the $L^2$ limit (with respect to Wiener measure) of step functions of the type $\sum f_i(t_{i_k})\, \Delta z_{i_k}$, where

$$\Delta z_{i_k} = z(t_{i_{k+1}}) - z(t_{i_k}).$$

Let $\{t_j\}$ be a partition of the interval $[S, T]$ which is a refinement of the partitions generated by the $t_{i_k}$. Let $\Delta z_j = z(t_{j+1}) - z(t_j)$. We note, by axiom $(\gamma_1)$ or by the definition of the Wiener integral, that the $\Delta z$'s are independent, and, therefore,

$$E_z^w\{(\Delta z_{j_1})^{n_1} (\Delta z_{j_2})^{n_2} \cdots (\Delta z_{j_l})^{n_l} \mid z(S) = A\}$$
$$= \big(E_z^w\{(\Delta z_{j_1})^{n_1} \mid z(S) = A\}\big) \cdots \big(E_z^w\{(\Delta z_{j_l})^{n_l} \mid z(S) = A\}\big)$$

It can also be seen that

$$E_z^w\{(\Delta z_j)^n \mid z(S) = A\}$$
$$= \frac{1}{[(2\pi)^2 (t_{j+1} - t_j)(t_j - S)]^{1/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [z(t_{j+1}) - z(t_j)]^n$$
$$\times \exp\left\{-\frac{1}{2}\left[\frac{[z(t_{j+1}) - z(t_j)]^2}{t_{j+1} - t_j} + \frac{[z(t_j) - A]^2}{t_j - S}\right]\right\} dz(t_j)\, dz(t_{j+1})$$

let

$$w_1 = z(t_{j+1}) - z(t_j)/(t_{j+1} - t_j)^{1/2}, \qquad w_0 = z(t_j) - A/(t_j - S)^{1/2}.$$

Therefore,

$$E_z^w\{(\Delta z_j)^n \mid z(S) = A\}$$
$$= (t_{j+1} - t_j)^{n/2} \frac{1}{2\pi} \int_{-\infty}^{\infty} (w_1)^n \exp\left[-\frac{1}{2}(w_1^2 + w_0)^2\right] dw_1\, dw_0$$
$$= (t_{j+1} - t_j)^{n/2} \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} (w_1)^n \exp\left[\frac{-w_1^2}{2}\right] dw_1$$

From tables, we find this to be zero if $n$ is odd and $(t_{j+1} - t_j)^{n/2}\,[1 \cdot 3 \cdots (n-1)]$ if $n$ is even. Some rather tedious combinatorial analysis shows that

$$E_z^w\left\{\left[\sum_k f_1(t_{1_k})\, \Delta z_{1_k}\right] \cdots \left[\sum_k f_m(t_{m_k})\, \Delta z_{m_k}\right] \middle|\, z(S) = A\right\} = 0 \tag{20}$$

if $m$ is odd, since, if one interchanges the above finite sums and the integral, there will always be an odd number of factors in each integral. If $m$ is even, then (19) is

$$\sum_j f_1(t_j) f_2(t_j)(t_{j+1} - t_j) \sum_j f_2(t_j) f_3(t_j)(t_{j+1} - t_j) \cdots \sum f_{m-1}(t_j) f_m(t_j)(t_j - t_{j-1})$$
$$+ \cdots + \sum_j f_1(t_j) f_m(t_j)(t_{j+1} - t_j) \cdots \sum_j f_{(m/2)-1}(t_j) f_{(m/2)+1}(t_j)(t_{j+1} - t_j) \tag{21}$$

Since the functions $f_i(t)$ are in $L^2[S, T]$, it follows that the sequence of step functions converges properly, and the expression (21) converges to the right-hand side of (19), which proves the lemma.

**Lemma 2.2.** If $f_1(t),\ldots,f_m(t)$ are in $L^2(S, T)$, then

$$E_z^w\left\{\int_S^{t_1} f_1(\alpha)\, dz(\alpha) \int_S^{t_2} f_2(\alpha)\, dz(\alpha) \cdots \int_S^{t_m} f_m(\alpha)\, dz(\alpha) \middle|\, z(S) = A\right\} = 0$$

if $m$ is odd, and is

$$\int_S^{\min(t_1,t_2)} f_1(\alpha) f_2(\alpha)\, d\alpha \int_S^{\min(t_3,t_4)} f_3(\alpha) f_4(\alpha)\, d\alpha \cdots \int_S^{\min(t_{n-1},t_n)} f_{n-1}(\alpha) f_n(\alpha)\, d\alpha$$
$$+ \cdots + \int_S^{\min(t_1,t_m)} f_1(\alpha) f_m(\alpha)\, d\alpha \cdots \int_S^{\min(t_{(m/2)-1},t_{(m/2)+1})} f_{(m/2)-1}(\alpha) f_{(m/2)+1}(\alpha)\, d\alpha$$

if $m$ is even.

*Proof.* The proof follows immediately from lemma 2.1 by letting

$$t = \max(t_1, ..., t_m)$$

and by extending the $f_i(\alpha)$ to the interval $[S, T]$ as $f_i(\alpha) = 0$, $t_i \leqslant \alpha \leqslant T$.

**2.5.** We now quote a result on the Ito integral known as Ito's integration by-parts formula.

**Lemma 2.3.** Suppose $F(t, X, Y)$ is defined for all real $X$ and $Y$ and $S \leqslant t \leqslant T$. Suppose $F_t$, $F_X$, and $F_{YY}$ exist and are continuous there, and suppose $z(t)$ is Brownian motion with $z(S) = A$.
Then

$$\begin{aligned} F\left(T, \int_S^T z(\alpha)\, d\alpha, z(T)\right) &- F(S, 0, A) \\ = \int_S^T F_t \left(t, \int_S^t z(\alpha)\, d\alpha, z(t)\right) dt &+ \int_S^T F_X \left(t, \int_S^t z(\alpha)\, d\alpha, z(t)\right) z(t)\, dt \\ + \frac{1}{2} \int_S^T F_{YY} \left(t, \int_S^t z(\alpha)\, d\alpha, z(t)\right) dt &+ \int_S^T F_Y \left(t, \int_S^t z(\alpha)\, d\alpha, z(t)\right) dz(t) \end{aligned} \tag{22}$$

The last integral on the right-hand side of (22) is an Ito integral.

*Proof.* The proof follows directly from Dynkin's[9] theorem 7.2 and an application of the rule for differentiating composite functions.

We note that, if $z(t)$ were an ordinary differentiable function of $t$, (22) would follow directly from the rule for differentiating composite functions, if it were not for the term $\frac{1}{2}\int_S^T F_{YY}(t, \int_S^t z(\alpha)\, d\alpha, z(t))\, dt$. The reason for the appearance of this term is given by Eq. (6); $\lim_{\Delta\to 0}(\Delta/1 \{[z(t + \Delta t) - z(t)]/\Delta\}^2$ is not zero (as it is for ordinary functions), but 1.

**2.5.** The Wiener integral of $F(z)$, $E_z^w \{F(z) | z(S) = A\}$ is usually denoted in physics works as

$$\int F(z) \exp \left\{-\frac{1}{2} \int_S^T [\dot{z}(t)]^2\, dt\right\} dz$$

(see Feynman[11] or Donsker,[7] for example). This is so because the Wiener integral of the function $F(z)$ can be defined as [see Eq. (3)]

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{[(2\pi)^n (t_n - t_{n-1}) \cdots (t_1 - t_0)]^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F(z(t_1), ..., z(t_n)) \\ \times \exp \left\{-\frac{1}{2} \sum_i \frac{[z(t_{i+1}) - z(t_i)]^2}{t_{i+1} - t_i}\right\} dz(t_1) \cdots dz(t_n) \end{aligned} \tag{23}$$

where $t_0 = S$ and $z(t_0) = A$, and where the sequences $(t_j)$ are becoming dense in

$[S, T]$. Multiplying numerator and denominator of the fractions in the exponent by $(t_{i+1} - t_i)$, we get

$$\lim_{n\to\infty} \frac{1}{[(2\pi)^n (t_n - t_{n-1}) \cdots (t_1 - t_0)]^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F(z(t_1),..., z(t_n))$$

$$\times \exp \left\{ -\frac{1}{2} \sum \left[ \frac{z(t_{i+1}) - z(t_i)}{t_{i+1} - t_i} \right]^2 (t_{i+1} - t_i) \right\} dz(t_1) \cdots dz(t_2) \tag{24}$$

If $\dot{z}(t)$ existed and was in $L^2[S, T]$, then the term inside the exponent would become $-\frac{1}{2} \int_S^T [\dot{z}(t)]^2 \, dt$ as $n \to \infty$, and then we could write, using (19),

$$E_z^w \left\{ F(z) \middle| z(S) = A \right\} = \int F(z) \exp \left[ -\frac{1}{2} \int_S^T [\dot{z}(t)]^2 \, dt \right] \delta z \tag{25}$$

where $\int (\cdots) \, \delta z$ means

$$\lim_{n\to\infty} \frac{1}{[(2\pi)^n (t_n - t_{n-1}) \cdots (t_1 - S)]^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\cdots) \, dz(t_1) \cdots dz(t_n) \tag{26}$$

Un fortunately, however, the integral (26) does not exist (see Gross[18]), and by (6), neither does $(\dot{z}(t))^2$. Yet the representation (25) is useful for understanding certain theorems about the Wiener integral.

**2.6.** We now suppose the $f(t, X, Y)$ of Eq. (1) to have the following properties:

($\alpha$) $f(t, X, Y)$ is jointly measurable in each of its three variables.

($\beta$) The nonlinear integral equation (11) has a unique solution for almost every $z \in C[S, A, T]$.

($\gamma$) There exists a nondecreasing function of a real variable $f_0$ such that $|f(t, \int_S^t z(\alpha) \, d\alpha, z(t))| < f_0(\sup |z(t)|)$. Conditions ($\alpha$) and ($\gamma$) are easily verified for a given $f$. Sufficient conditions for ($\beta$) to hold are given by Doob[8] and by Dynkin,[9] Chapter 11.

One of the important theorems of this paper will now be stated (see Dynkin,[9] Girsanov,[17] or Schilder[37] for other hypothesis to this theorem)

**Theorem 2.1.** Suppose $f(t, X, Y)$, satisfies conditions ($\alpha$)–($\gamma$), and $F(x, \dot{x})$ is any integrable function with respect to the Ito distribution defined by (16); then

$$\begin{aligned}
&E_x^I\{F(x(\cdot), \dot{x}(\cdot)) | \, x(S) = X, \quad \dot{x}(S) = Y\} \\
&= E_z^w \left\{ F\left(\lambda \int_S^{(\cdot)} T(t, z) \, dt + X, \lambda T(\cdot\,, z)\right) \middle| z(S) = A \right\} \\
&= E_y^w \Big\{ F\left(\lambda \int_S^{(\cdot)} y(\alpha) \, d\alpha + X, \lambda y(\cdot)\right) \\
&\quad \times \exp \left[ (-1/2\lambda^2) \int_S^T f^2 \left(t, \lambda \int_S^t y(\alpha) \, d\alpha + X, \lambda y(t)\right) dt \right. \\
&\quad \left. + (1/\lambda) \int_S^T f\left(t, \lambda \int_S^t y(\alpha) \, d\alpha + X, \lambda y(T)\right) dy(t) \right] \Big| \, y(S) = Y/\lambda \Big\}
\end{aligned} \tag{27}$$

The last integral in the exponent of the Wiener integral is an Ito integral.

*Proof.* The proof follows directly from Girsanov's[17] theorem 2. To see why formula (27) is true, we note, from (16) (by definition),

$$E_x^{I}\{F(x, \dot{x})|\, x(S) = X, \dot{x}(S) = Y\}$$
$$= E_z^{w}\left\{F\left(\lambda \int_S^{(\cdot)} T(\alpha, z)\, d\alpha + X, \lambda T(\cdot\,, z)\right)\middle|\, z(S) = A\right\} \tag{28}$$

By (25), the last expression is, heuristically,

$$\int F\left(\lambda \int_S^{(\cdot)} T(\alpha, z)\, d\alpha + X, \lambda T(\cdot, z)\right) \exp\left\{-\frac{1}{2}\int_S^T [\dot{z}(t)]^2\, dt\right\} \delta z \tag{29}$$

By (12), $T(t, z) = y(t)$, and, by (10), $\dot{z}(t) = \dot{y}(t) - (1/\lambda) f(t, \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t))$. Substituting these two expressions into (29) (we are, of course, actually changing variables), we get

$$\int F\left(\lambda \int_S^{(\cdot)} y(t)\, dt + X, \lambda y(\cdot)\right)$$
$$\times \exp\left\{-\frac{1}{2}\int_S^T \left[\dot{y}(t) - (1/\lambda) f\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right)\right]^2 dt\right\} \delta y$$

This integral becomes, upon squaring out the term in the exponent,

$$\int F\left(\lambda \int_S^{(\cdot)} y(\alpha)\, d\alpha + X, \lambda y(\cdot)\right) \exp\left\{(-1/2\lambda^2) \int_S^T f^2\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) dt\right.$$
$$\left.(1/\lambda) \int_S^T f\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) \dot{y}(t)\, dt - \frac{1}{2}\int_S^T [\dot{y}(t)]^2\, dt\right\} \delta y$$

By (25) again, this is

$$E_y^{w}\left\{F\left(\lambda \int_S^{(\cdot)} y(\alpha)\, d\alpha + X, \lambda y(\cdot)\right)\right.$$
$$\times \exp\left[(-1/2\lambda^2) \int_S^T f^2\left(t, \lambda \int_S^T y(\alpha)\, d\alpha + X, \lambda y(t)\right) dt\right.$$
$$\left.\left. + (1/\lambda) \int_S^T f\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) dt\right]\middle|\, y(S) = Y/\lambda\right\}$$

which "proves" Theorem 2.1. The reason this is not a good proof is that the integral defined by (25) does not really exist.

Equation (16) shows that the expected values of functionals with respect to the $x$ distribution defined by (1) may be expressed as Wiener integrals. This representation is, however, of little value, since it requires, in general, a solution of a nonlinear integral equation. Theorem 2.1 shows that expected values of functionals of the $x$ distribution may be expressed as Wiener integrals; one simply has to multiply them by the factor $\exp[(-1/2\lambda^2) \int_s^T f^2\, dt + (1/\lambda) \int_s^T f\, dy(t)]$. This factor is, of course, a Radon–Nikodym derivative. The notation $f = f(t, \lambda \int_s^t y(\alpha)\, d\alpha + X, \lambda y(t))$ will be used for the remainder of the paper if the meaning is clear.

**Corollary 2.1.** Suppose $F(t, X, Y)$ satisfies the conditions of Theorem 2.1; then

$$E_z^w \left\{ \exp \left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dz(t) \right] \Big| \, z(S) = Y/\lambda \right\} = 1$$

*Proof.* We let the $F(x(\cdot), \dot{x}(\cdot))$ of Theorem 2.1 be identically equal to one. Then the left hand side of Eq. (27) is one, since $E_x^I\{ \cdots \mid x(S) = X, \dot{x}(S) = Y\}$ is the integral of a probability distribution. By Theorem 2.1, the right hand side of (27) is one.

**2.7.** The formula for conditional probabilities is

$$\mathrm{Prob}\{B \mid C\} = \mathrm{Prob}\{B \cap C\} / \mathrm{Prob}\{C\}.$$

This formula can be generalized to integrals as follows. By definition, for $B$ and $C$ measurable sets in function space,

$$\begin{aligned} &E_x^I\{F(x(\cdot), \dot{x}(\cdot)) \mid x(S) = X, \quad \dot{x}(S) = Y, \quad x \in B, \quad \dot{x} \in C\} \\ &= \frac{E_x^I\{\chi_B(x)\, \chi_C(\dot{x})\, F(x(\cdot), \dot{x}(\cdot)) \mid x(S) = X, \dot{x}(S) = Y\}}{E_x^I\{\chi_B(x)\, \chi_C(\dot{x}) \mid x(S) = X, \dot{x}(S) = Y\}} \end{aligned} \tag{30}$$

The notation on the left hand side of (30) reads, "The expected value of the function $F(x(\cdot), \dot{x}(\cdot))$ with respect to the Ito distribution given that $x$ is in $B$ and $\dot{x}$ is in $C$ is ...".

A lemma is now given that will be used in Section 5.

**Lemma 2.4.** Suppose the integral of $F(z)$ with respect to Wiener measure exists; then

$$\begin{aligned} &E_z^w\{F(z) \mid z(S) = A\} \\ &= \int_{-\infty}^{\infty} E_z^w \left\{ F(z) \Big| \, z(S) = A, \quad \int_S^T z(\alpha)\, d\alpha = R \right\} d\,\mathrm{Prob} \left\{ \int_S^T z(\alpha)\, d\alpha = R \right\} dR \end{aligned} \tag{31}$$

$d\,\mathrm{Prob}\{\int_S^T z(\alpha)\, d\alpha = R\}$ is the density function of the random variable $\int_S^T z(\alpha)\, d\alpha$.

*Proof.* If $C_i$, $i = 1, 2, 3, \ldots$, is a collection of disjoint sets whose union is the whole set, then a well known formula from elementary probability theory states $\mathrm{Prob}\{B\} = (\mathrm{Prob}\{B \mid C_i\})(\mathrm{Prob}\{C_i\})$. The analog of the formula for integrals is

$$\begin{aligned} E_z^w\{F(z) \mid z(S) = A\} &= \sum_i E_z^w \left\{ F(z) \Big| \, z(S) = A, \int_S^T z(\alpha)\, d\alpha \in C_i \right\} \\ &\times \mathrm{Prob} \left\{ \int_S^T z(\alpha)\, d\alpha \in C_i \, \Big| \, z(S) = A \right\} \end{aligned}$$

Let $C_i$ be a collection of nonoverlapping intervals on the real line whose union is the entire real line, and let $|C_i|$ be the length of $C_i$.

From the above,

$$
\begin{aligned}
&E_z{}^w\{F(z)|\, z(S) = A\} \\
&\quad = \sum_i E_z{}^w \left\{F(z)\middle|\, z(S) = A,\ \int_S^T z(\alpha)\, d\alpha \in C_i\right\} \\
&\quad \times \text{Prob}\left\{\int_S^T z(\alpha)\, d\alpha \in C_i \,\middle|\, z(S) = A\right\} |\, C_i \,|/|\, C_i \,|
\end{aligned}
$$

If the $C_i$ now shrink down to points, then $\text{Prob}\{\int_S^T z(\alpha)\, d\alpha \in C_i \mid z(S) = A\}/C_i$ becomes $d\,\text{Prob}\{\int_S^T z(\alpha)\, d\alpha = R \mid z(S) = A\}$, $C_i$ becomes $dR$, and $\sum_i$ becomes an integral, which proves Lemma 2.4.

## 3. PARTIAL DIFFERENTIAL EQUATIONS ASSOCIATED WITH WIENER INTEGRALS

**3.1.** It is well known (see Doob,[8] for example) that the probability distribution function associated with Eq. (1) satisfies a parabolic partial differential equation called the backwards equation. From the discussion of the previous section and Theorem 2.1, it follows that this probability distribution function may be expressed as a Wiener integral, and therefore that the resulting Wiener integral satisfies the backwards equation. This fact will now be shown directly. First, however, a lemma is needed.

**Lemma 3.1.** Suppose $g(t, X, Y)$ is jointly continuous and defined for $S \leqslant t \leqslant T$, with $X$ and $Y$ real. Suppose $f(t, X, Y)$ satisfies the conditions of Theorem 2.1; then

$$
\begin{aligned}
&E_y{}^w \left\{\int_S^T g\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) dy(t)\right. \\
&\quad - (1/\lambda) \int_S^T g\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) f\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(T)\right) dt \\
&\quad \times \exp\left[(-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dy(t)\right]\left|\, y(S) = Y/\lambda\right\} = 0 \qquad (32)
\end{aligned}
$$

*Proof.* Let $T(t, z)$ be defined by (12). Dynkin[9] shows the Ito integrals

$$\int_S^T g\left(t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t)\right) dy(t) \text{ and } \int_S^T g\left(t, \lambda \int_S^t T(\alpha, z)\, d\alpha + X, \lambda T(t, z)\right) dz(t)$$

are well defined. By Theorem 2.1,

$$E_z{}^w \left\{\int_S^T g\left(t, \lambda \int_S^t T(\alpha, z)\, d\alpha + X, \lambda T(t, z)\right) dz(t) \middle|\, z(S) = A\right\} \qquad (33)$$

equals the left-hand side of (33) [again, this is really only the change of variables $\dot{z} = \dot{y} - (1/\lambda) f$]. But (33) is zero, by (17).

**Theorem 3.1.** Suppose $f(t, X, Y)$ satisfies the conditions of Theorem 2.1. Suppose $Q(t, X, Y)$ satisfies, for $S \leqslant t \leqslant T$, with $X$ and $Y$ real,

$$Q_t(t, X, Y) + f(t, X, Y)\, Q_Y + YQ_X + L(t, X, Y) + \tfrac{1}{2}\lambda^2 Q_{YY} = 0 \tag{34}$$

with

$$Q(T, X, Y) = \phi(X, Y).$$

It is assumed that, in the region considered $L$, $Q_t$, $Q_X$, $Q_Y$, and $Q_{XY}$ exist and are continuous. Then

$$Q(S, X, Y) = E_x^I \left\{ \phi(x(T), \dot{x}(T))^2 + \int_S^T L(t, x(t), \dot{x}(t))\, dt \,\middle|\, x(S) = X, \dot{x}(S) = Y \right\} \tag{35}$$

*Proof.* By Theorem 2.1, the right-hand side of (35) is

$$\begin{aligned} E_y^w \Big\{ &\phi \left( \lambda \int_S^T y(\alpha)\, d\alpha + X, \lambda y(T) \right) + \int_S^T L \left( t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(T) \right) dt \\ &\times \exp \left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dy(t) \right] \Big|\, y(S) = Y/\lambda \Big\} \end{aligned} \tag{36}$$

By hypothesis,

$$\phi \left( \lambda \int_S^T y(\alpha)\, d\alpha + X, \lambda y(T) \right) = Q\left( T, \lambda \int_S^T y(\alpha)\, d\alpha + X, \lambda y(T) \right)$$

Substituting this into (36) and using the Ito integration by-parts formula (22), (36) becomes

$$\begin{aligned} E_y^w \Big\{ &Q(S, X, \lambda y(S)) + \int_S^T Q_t \left( t, \lambda \int_S^t y(\alpha)\, d\alpha + X, \lambda y(t) \right) dt \\ &\times \lambda \int_S^T Q_X\, y(t)\, dt + 1/2\lambda^2 \int_S^T Q_{YY}\, dt + \lambda \int_S^T Q_Y\, dy(t) + \int_S^T L\, dt \\ &\times \exp \left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dy(t) \right] \Big|\, y(S) = Y/\lambda \Big\} \end{aligned} \tag{37}$$

The partial differential equation (34) is now substituted into (31), which is equal to

$$\begin{aligned} E_y^w \Big\{ &Q(S, X, \lambda y(S)) + \lambda \int_S^T Q_Y\, dy(T) - \int_S^T Q_Y f\, dt \\ &\times \exp \left[ (1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dy(t) \right] \Big|\, y(S) = Y/\lambda \Big\} \end{aligned}$$

which equals

$$\begin{aligned} E_y^w \Big\{ &Q(S, X, \lambda y(S)) \exp \left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dy(t) \right] \Big|\, y(S) = Y/\lambda \Big\} \\ &\times \lambda E_y^w \Big\{ \left( \int_S^T Q_Y\, dy(t) - (1/\lambda) \int_S^T Q_Y f\, dt \right) \\ &\times \exp \left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dy(t) \right] \Big|\, y(S) = Y/\lambda \Big\} \end{aligned}$$

The second integral is zero, by Lemma 3.1; the first is $Q(S, X, Y)$, by corollary 2.1. This proves the theorem.

If, in (34), we let $L = 0$ and $\phi(X, Y) = 1$ if $X \leqslant X_1$ and $Y \leqslant Y_1$ and zero otherwise, then (34) becomes the backwards equation. Also, the integral (35) becomes exactly the probability that $x(T) \leqslant X_1$ and $y(T) \leqslant Y_1$, as it should.

**3.2.** Consider now the problem of minimizing

$$R(S, X, Y, u(\cdot\,, \cdot\,, \cdot))$$
$$= E_x{}^I \left\{ \int_S^T L(t, x(t), \dot{x}(t), u(t, x(t), \dot{x}(t)))\, dt + \phi(x(T), \dot{x}(T)) \Big|\, x(S) = X, \dot{x}(S) = Y \right\} \tag{38}$$

where $x$ has the Ito distribution defined by

$$\ddot{x} = f\big(t, \dot{x}(t), x(t), u(t, x(t), \dot{x}(t))\big) + \lambda \dot{z}(t), \qquad x(S) = X, \qquad \dot{x}(S) = Y \tag{39}$$

over a set of "control functions" $u$, where $u(t, X, Y)$ is contrained to lie in some set $U$.

We will derive an equation for the minimal $R$, using Wiener integrals, and show that, for an important special case, the solution of this equation can be expressed in terms of a Wiener integral. A $u$ which minimizes (if one exists) the $R$ of (38) will be called $u^*(t, X, Y)$. Let $R^*(S, X, Y) = R(S, X, Y, u^*(\cdot, \cdot, \cdot))$, and $f^*(t, x, \dot{x}) = f(t, x, \dot{x}, u^*(t, x, \dot{x}))$.

Also let

$$\bar{H}(t, X, Y, R) = \inf_{u \in U} \{ f(t, X, Y, u(t, X, Y))\, R_Y + Y R_X + L(t, X, Y, u(t, X, Y))\} \tag{40}$$

Definition (40) means that, for each *fixed* $t$, $X$, $Y$, and $R$ minimize the right-hand side for $u \in U$. Call a $u$ so obtained $\bar{u}$.

**Theorem 3.2.** Suppose $f(t, x, \dot{x}, u(t, x, \dot{x}))$ satisfies the assumptions of Theorem 2.1 for $u$ in some control class $U$. Suppose the function $L(t, X, Y, u(t, X, Y))$ is defined and continuous, $S \leqslant t \leqslant T$, with $X$ and $Y$ real or each $u \in U$. Suppose that the differential equation (the stochastic Hamilton–Jacobi equation)

$$\bar{R}_t + \bar{H}(t, X, Y, \bar{R}) + \tfrac{1}{2}\lambda^2 \bar{R}_{YY} = 0 \tag{41}$$

has a continuously differentiable solution $\bar{R}$. Then $\bar{R} = R^*$ (the minimal value of $R$) and the $u$ defined by (40) is a $u^*$.

*Proof.* From (22), if $z(t)$ is Brownian motion

$$\bar{R}\left(T, \lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(T)\right) - \bar{R}(S, X, \lambda z(S))$$
$$= \int_S^T [\bar{R}_t + \tfrac{1}{2}\lambda^2 \bar{R}_{YY} + \lambda \bar{R}_X z(t)]\, dt + \lambda \int_S^T \bar{R}_Y\, dz(t) \tag{42}$$

From (40) and (41), it can be seen that the integral with the brackets is

$$\int_S^T -\bar{f}\left(t, \lambda \int_S^t z(\alpha)\, d\alpha + X, \lambda z(t)\right) \bar{R}_Y\left(t, \lambda \int_S^t z(\alpha)\, d\alpha + X, \lambda z(t)\right) - \bar{L}\left(t, \lambda \int_S^t z(\alpha)\, d\alpha + X, \lambda z(t)\right) dt \tag{43}$$

where

$$\bar{f}(t, X, Y) = f(t, X, Y, \bar{u}(t, X, Y)) \qquad \text{and} \qquad \bar{L}(t, X, Y) = L(t, X, Y, \bar{u}(t, X, Y))$$

Since

$$\bar{R}(S, X, Y) = E_z^w \left\{ \bar{R}(S, X, \lambda z(S)) \exp\left[ -(2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t) \right] \middle| \, z(S) = Y\lambda^{-1} \right\}$$

where $f(t, X, Y) = f(t, X, Y, u(t, X, Y))$, and $u$ is any $u \in U$, it follows by substituting (42) and (43) into the last equation that

$$\begin{aligned}
\bar{R}(S, X, Y) &= E_z^w \left\{ \left[ \int_S^T (\bar{f}\bar{R}_Y + \bar{L})\, dt - \lambda \int_S^T \bar{R}_Y\, dz(t) + \bar{R}\left(T, \lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(T)\right) \right] \right. \\
&\quad \left. \times \exp\left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dz(t) \right] \middle| \, z(S) = Y/\lambda \right\} \\
&= E_z^w \left\{ \int_S^T (\bar{f}\bar{R}_Y + \bar{L})\, dt + \int_S^T (\bar{R}_Y f - \bar{R}_Y f)\, dt \right. \\
&\quad - \lambda \int_S^T \bar{R}_Y\, dz(t) + \bar{R}\left(T, \lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(T)\right) \\
&\quad \left. \times \exp\left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dz(t) \right] \middle| \, z(S) = Y/\lambda \right\}
\end{aligned}$$

From Lemma 3.1, it follows that

$$E_z^w \left\{ \left[ \int_S^T \bar{R}_Y f\, dt - \lambda \int_S^T \bar{R}_Y\, dz(t) \right] \times \exp\left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dz(t) \right] \middle| \, z(S) = Y/\lambda \right\} = 0$$

Thus,

$$\begin{aligned}
\bar{R}(S, X, Y) &= E_z^w \left\{ \left[ \int_S^T (\bar{f}\bar{R}_Y + \bar{L} - \bar{R}_Y f)\, dt + \bar{R}\left(T, \lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(T)\right) \right] \right. \\
&\quad \left. \times \exp\left[ (-1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dz(t) \right] \middle| \, z(S) = Y/\lambda \right\}
\end{aligned}$$

For any $u \in U$, it follows that, since $\bar{R}(T, X, Y) = \phi(X, Y)$ from the definition (38) and Theorem 2.1,

$$R(S, X, Y) = E_z^w \left\{ \left[ \int_S^T L\,dt + \bar{R}\left(T, \lambda \int_S^T z(\alpha)\,d\alpha + X, \lambda z(T)\right) \right. \right.$$
$$\left. \left. \times \exp\left[(-1/2\lambda^2) \int_S^T f^2\,dt + (1/\lambda) \int_S^T f\,dz(t)\right] \right| z(S) = Y/\lambda \right\}$$

Subtracting $\bar{R}$ from $R$, we have

$$R(S, X, Y) - \bar{R}(S, X, Y) = E_z^w \left\{ \int_S^T (f\bar{R}_Y + L) - (\bar{R}_Y \bar{f} + \bar{L})\,dt \right.$$
$$\left. \times \exp\left[-(1/2\lambda^2) \int_S^T f^2\,dt + (1/\lambda) \int_S^T f\,dz(t)\right] \right| z(S) = Y/\lambda \right\}$$

but, by (40) $f\bar{R}_Y + L \geqslant \bar{f}\bar{R}_Y + \bar{L}$ for all $u \in U$ and all $X$, $Y$, and $t$. Thus, $R(S, X, Y) \geqslant \bar{R}(S, X, Y)$. QED.

The proof is practically a literal translation into the terminology of this paper of the one given in Wonham.[46] Wonham specifically states that his proof is not rigorous. The proof of Theorem 3.2 is, however, perfectly rigorous—due to the assumption that Eq. (41) has a continuously differentiable solution.

**3.3.** In this section, it will be shown that, if the $f$ and $L$ of Theorem 3.2 are of a special form, then the solution of Eq. (41) can be given in terms of Wiener integrals. Let $L$ be of the form

$$L(t, X, Y, u(t, X, Y)) = L(t, X, Y) + \tfrac{1}{2}u^2(t, X, Y) \tag{44}$$

and $f$ be of the form

$$f(t, X, Y, u(t, X, Y)) = f(t, X, Y) + u(t, X, Y) \tag{45}$$

and let the control class $U$ be the set of all functions such that the $f$ of (39) satisfies the assumptions of Theorem 2.1.

It is easily seen that [see (40)], for this case,

$$\begin{aligned} &\bar{H}(t, X, Y, R) \\ &\quad = \min\{(f(t, X, Y) + u(t, X, Y))\, R_Y + YR_X + L(t, X, Y) + \tfrac{1}{2}u^2(t, X, Y)\} \\ &\quad = f(t, X, Y)\, R_Y + YR_X + L(t, X, Y) - \tfrac{1}{2}R^2{}_Y \end{aligned}$$

since the $u$ which minimizes $\bar{H}$ in this case is $u = -R_Y$, or $u(t, X, Y) = -R_Y(t, X, Y)$. Thus, in this case, Eq. (41) becomes

$$\bar{R}_t + f\bar{R}_Y + Y\bar{R}_X + \bar{L} - \tfrac{1}{2}(\bar{R}_Y)^2 + \tfrac{1}{2}\lambda^2\bar{R}_{YY} = 0, \quad \bar{R}(T, X, Y) = \phi(X, Y) \tag{46}$$

where $\phi$ can be any continuous function.

In this case, the Ito distribution of $x$ for the optimal $u$ by Theorem 3.2 is given by

$$\ddot{x} = f(t, x(t), \dot{x}(t)) - R_Y(t, x(t), \dot{x}(t)) + \lambda \dot{z}(t), \qquad x(S) = X, \quad \dot{x}(S) = Y \tag{47}$$

It will now be shown that the expected values of functions with respect to this distribution, henceforth called the optimal distribution, can be expressed in terms of Wiener integrals, without solving (46), and that the solution to (46) can be expressed as a Wiener integral. It is shown in Section 6 how to obtain series expansions for $\bar{R}$ and $\bar{u}$.

**Theorem 3.3.** Suppose the functions $f$ and $L$ satisfy the hypothesis of Theorem 3.2 and that they are in the form given by (44) and (45). Suppose, also, that (46) has a well-defined solution; then

$$\begin{aligned}
&E_x^0\{F(x, \dot{x}) \mid x(S) = X, \dot{x}(S) = Y\}\\
&\quad = E_z^w \Big\{F\Big(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X, \lambda z(\cdot)\Big)\\
&\qquad \times \exp\Big[-\lambda^{-2} \int_S^T \Big(L + \frac{1}{2} f^2\Big)\, dt + \lambda^{-1} \int_S^T f\, dz(t)\\
&\qquad - \lambda^{-2}\phi\Big(\lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(T)\Big)\Big] \Big| z(S) = Y/\lambda\Big\}\\
&\quad \div E_z^w \Big\{\exp\Big[\lambda^{-2} \int_S^T \Big(L + \frac{1}{2} f^2\Big)\, dt + \lambda^{-1} \int_S^T f\, dz(t)\\
&\qquad - \lambda^{-2}\phi\Big(\lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(t)\Big)\Big] \Big| z(S) = Y/\lambda\Big\}
\end{aligned}$$

where $E_x^0\{\cdots\}$ is the distribution defined by (46) and (47).

*Proof.* From Theorem 2.1 and (47),

$$\begin{aligned}
E_x^0\{F(x, \dot{x}) \mid x(S) = X, \dot{x}(S) = Y\} &= E_z^w \Big\{F\Big(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X, \lambda z(\cdot)\Big)\\
&\quad \times \exp\Big[(-1/2\lambda^2) \int_S^T (f - R_Y)^2\, dt\\
&\quad + (1/\lambda) \int_S^T f\, dz(t) - (1/\lambda) \int_S^T \bar{R}_Y\, dz(t)\Big] \Big| z(S) = Y/\lambda\Big\}
\end{aligned}$$

From (22), with $\bar{R}$ substituted for $F$,

$$\begin{aligned}
&\bar{R}\Big(T, \lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(t)\Big) - \bar{R}(S, X, \lambda z(S))\\
&\quad = \int_S^T \bar{R}_t\, dt + \lambda \int_S^T \bar{R}_Y\, dz(t) + \lambda \int_S^T \bar{R}_X z(t)\, dt + \frac{1}{2}\lambda^2 \int_S^T \bar{R}_{YY}\, dt
\end{aligned}$$

Solving this expression for $(1/\lambda)\int_S^T \bar{R}_Y\,dz(t)$, and squaring out the

$$-(1/2\lambda^2)\int_S^T (f-\bar{R}_Y)^2\,dt$$

term, we get

$$\begin{aligned}
E_x^{\,0}\{F(x,\dot{x})\,|\,x(S)=X,\dot{x}(S)=Y\} &= E_z^{\,w}\Big\{F\Big(\lambda\int_S^{(\cdot)} z(\alpha)\,d\alpha+X,\lambda z(\cdot)\Big)\\
&\times \exp\Big\{\lambda^{-2}\int_S^T -\frac{1}{2}f^2+f\bar{R}_Y\\
&-\frac{\bar{R}_Y{}^2}{2}+\lambda^{-2}\bar{R}_t+\lambda^{-1}\bar{R}_X z(t)+\frac{1}{2}\bar{R}_{YY}\Big)\,dt\\
&+1/\lambda^{-1}\int_S^T f\,dz(t)+\Big[\bar{R}(S,X,\lambda z(S))\\
&-\bar{R}\Big(T,\lambda\int_S^T z(\alpha)\,d\alpha+X,\lambda z(T)\Big)\Big]\lambda^2\Big\}\Big|\,z(S)=Y/\lambda\Big\}
\end{aligned}$$

From the partial differential equation (46), it follows, by substitution into the integral of the exponent, that

$$\begin{aligned}
&E_x^{\,0}\{F(x,\dot{x})\,|\,x(S)=X,\dot{x}(S)=Y\}\\
&\quad= E_z^{\,w}\Big\{F\Big(\lambda\int_S^{(\cdot)} z(\alpha)\,d\alpha+X,\lambda z(\cdot)\Big)\\
&\quad\times\exp\Big[(-1/\lambda^2)\int_S^T\Big(L+\frac{1}{2}f^2\Big)\,dt+(1/\lambda)\int_S^T f\,dz(t)\\
&\quad-\phi\Big(\lambda\int_S^T z(\alpha)\,d\alpha+X,\lambda z(T)\Big)\Big/\lambda^2\Big]\Big|\,z(S)=Y/\lambda\Big\}\exp[\bar{R}(S,X,Y)/\lambda^2] \qquad (48)
\end{aligned}$$

If, in (48), we let $F(x,\dot{x})\equiv 1$, then, from Corollary 2.1, it follows that

$$E_x^{\,0}\{1\,|\,x(S)=X,\dot{x}(S)=Y\}=1$$

and, therefore, that

$$\begin{aligned}
1 &= E_z^{\,w}\Big\{\exp\Big[\frac{-1}{\lambda^2}\int_S^T\Big(L+\frac{1}{2}f^2\Big)\,dt\\
&+\frac{1}{\lambda}\int_S^T f\,dz(t)-\phi\Big(\lambda\int_S^T z(\alpha)\,d\alpha+X,\lambda z(T)\Big)\Big/\lambda^2\Big]\Big|\,z(S)=\frac{Y}{\lambda}\Big\}\exp\frac{\bar{R}(S,X,Y)}{\lambda^2} \qquad (49)
\end{aligned}$$

From (48) and (49), this time with a general $F(x,\dot{x})$ in (48) it follows, by solving (49) for $\exp[\bar{R}(S,X,Y)/\lambda^2]$, that

$$\begin{aligned}
&E_x^{\,0}\{F(x,\dot{x})\,|\,x(S)=X,\dot{x}(S)=Y\}\\
&\quad= E_z^{\,w}\Big\{F\Big(\lambda\int_S^{(\cdot)} z(\alpha)\,d\alpha+X,\lambda z(\cdot)\Big)
\end{aligned}$$

$$\times \exp\left[(-1/\lambda^2)\int_S^T\left(L+\frac{1}{2}f^2\right)dt+(1/\lambda)\int_S^T f\,dz(t)\right.$$

$$\left.-\phi\left(\lambda\int_S^T z(\alpha)\,d\alpha+X,\lambda z(T)\right)\Big/\lambda^2\right]\Big|\,z(S)=Y/\lambda\Big\}$$

$$\div E_z^w\Big\{\exp\left[(-1/\lambda^2)\int_S^T\left(L+\frac{1}{2}f^2\right)dt\right.$$

$$\left.+(1/\lambda)\int_S^T f\,dz(t)-\phi\left(\lambda\int_S^T z(\alpha)\,d\alpha+X,\lambda z(T)\right)\Big/\lambda^2\right]\Big|\,z(S)=Y/\lambda\Big\}$$

which was to be proved.

The following corollary follows immediately from (49) by solving it for $\bar{R}(S, X, Y)$.

**Corollary 3.1.** If $f$, $L$, and $\bar{R}$ satisfy the conditions of Theorem 3.3, it follows that

$$\bar{R}(S,X,Y)=-\lambda^2\ln\Big(E_z^w\Big\{\exp\left[(-1/\lambda^2)\int_S^T\left(L+\frac{1}{2}f^2\right)dt+(1/\lambda)\int_S^T f\,dz(t)\right.$$

$$\left.-\phi\left(\lambda\int_S^T z(\alpha)\,d\alpha+X,\lambda z(T)\right)\Big/\lambda^2\right]\Big|\,z(S)=Y/\lambda\Big\}\Big) \tag{50}$$

**Corollary 3.2.** If $f$, $L$, and $R$ satisfy the conditions of Theorem 3.3, then the solution to (46) is given by (50).

*Proof.* The proof is immediate.

**Lemma 3.2.** Continuously differentiable solutions to the nonlinear partial differential equation (46) are equivalent to positive, continuously differentiable solutions of the linear equation

$$W_t+fW_Y+YW_X-(LW/\lambda^2)+\tfrac{1}{2}\lambda^2W_{YY}=0,\qquad W(T,X,Y)=\exp[-\phi(X,Y)/\lambda^2] \tag{51}$$

under the transformation

$$W(t,X,Y)=\exp[-(1/\lambda^2)\,\bar{R}(t,X,Y)] \tag{52}$$

*Proof.* Let $W$ be defined by (52); then

$$W_t=-WR_t/\lambda^2,\qquad W_X=-WR_X/\lambda^2,\qquad W_Y=-WR_Y/\lambda^2,$$
$$W_{YY}=(-WR_{YY}/\lambda^2)+(WR_{YY}/\lambda^4) \tag{53}$$

Upon multiplying (46) through by $-(1/\lambda^2)\,W$ and using relations (53), the result follows. The terminal condition is obvious.

**Theorem 3.4.** Suppose equation (51) has a positive, continuously differentiable solution $W(t, X, Y)$, and suppose

$$g(t,X,Y)=f(t,X,Y)+\lambda^2(\ln W)_Y$$

satisfies conditions ($\alpha$)–($\gamma$) of Section 2; then

(i)

$$W(S, X, Y) = E_z^w \Big\{ \exp \Big[ (-1/\lambda^2) \int_S^T \Big( \frac{1}{2} f^2 + L \Big) dt + (1/\lambda) \int_S^T f \, dz(t) - \phi \Big( \lambda \int_S^T z(\alpha) \, d\alpha + X, \lambda z(T) \Big) \Big/ \lambda^2 \Big] \Big| \, z(S) = Y/\lambda \Big\}$$

(ii) The hypotheses of Theorem 3.3 are satisfied by $g$ and $L$, and therefore Theorem 3.3 and Corollary 3.1 hold for this $g$ and $L$.

*Proof.* The proof follows immediately from Lemma 3.2, Theorem 3.3, and Corollary 3.1.

**3.3.** Dynkin[9] has proved Theorems 2.1, 3.1, and 3.4 under different sets of hypotheses. In particular, he shows all these theorems are true, if suitably modified, under the hypothesis that the particle stays in a certain region of space, or equivalently, that the solutions to the partial differential equations (34) and (51) vanish outside a given region.

Unfortunately, he usually does not consider the time-dependent case. He shows existence and uniqueness for differential equations (34) and (51), and these results makes Theorem 3.3 applicable in a rather general sense. Due to lack of space, his results cannot be listed here. (See also Skorokhod.[41])

The nonlinear transformation (52) was first used by Hopf.[20] It was first used for Wiener integrals by Donsker. Varadhan[40] was the first to show that the Wiener integral could be used to solve an equation similar to (46). The author[38] was apparently the first to notice that Wiener integrals could be used to solve stochastic extremal problems. A hint of how Theorem 3.3 might be generalized to solve more-general stochastic Hamilton–Jacobi equations might be in Varadhan's work.[40]

A more general version of (1) is

$$\dot{x} = f\big(t, x(t)\big) + B(t) \dot{z}(t) \tag{54}$$

where $x(t)$, $f\big(t, x(t)\big)$, and $z(t)$ are possibly vectors, $\dot{z}(t)$ is Brownian motion, and $B(t)$ is a possibly singular matrix. In Schilder,[38] it is shown how this more general type of equation can be transformed into (1).

## 4. APPLICATIONS TO STOCHASTIC CONTROL THEORY

In this section, we give some applications to the theory of modern stochastic control theory. For a more detailed discussion of the problems and procedures of modern control theory, see Friedland *et al.*[14] or Wonham.[45,46]

A fairly general problem in control theory is to describe the probabilistic structure of the path of an airplane, satellite, or rocket whose dynamics are controlled by an equation of the form

$$\ddot{x} = f\big(t, x(t), \dot{x}(t)\big) + \lambda \dot{z}(t), \qquad x(S) = X, \quad \dot{x}(S) = Y$$

This problem, of course, has been discussed previously in this paper. It follows from Theorem 2.1 that, for any function $F(x, \dot{x})$ measurable and integrable on the Ito process defined by (1),

$$\begin{aligned} &E_x^{I}\{F(x, \dot{x}) | \, x(S) = X, \quad \dot{x}(S) = Y\} \\ &\quad = E_z^{w} \Big\{ F\Big(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X, \lambda z(\cdot)\Big) \\ &\qquad \times \exp\Big[-(1/2\lambda^2) \int_S^T f^2\, dt + (1/\lambda) \int_S^T f\, dz(t)\Big] \Big| \, z(S) = Y/\lambda \Big\} \end{aligned} \tag{55}$$

If we let $F(x, \dot{x}) = [x(T)]^n$, then (55) is a formula for expressing the value of the $n$th moment of the position of the vehicle at any time $T > S$.

If we let $F(x, \dot{x}) = [\dot{x}(T)]^n$, then (55) is a formula for expressing the value of the $n$th moment of the velocity at a time $T > S$.

If we let $F(x, \dot{x}) = 1$ when $x \in A$, and $F(x, \dot{x}) = 0$ when $x \notin A$, then (55) is a formula for the probability that the vehicle stays in a given region $A$, which might be, for example, a vicinity of the moon.

The control problem was considered in Sections 3.2 and 3.3. The problem is now, not only the description of the path of the vehicle, but also the necessity to make the vehicle move in such a way that the quantity $R$ of (38) is minimized. The quantity $L$ of (38) might be, for example, the amount of fuel used by the vehicle, and $\phi(x, \dot{x})$ might be the sum of the distance the vehicle is from a given spot (the moon) and the square of its velocity. Clearly, in designing a control $u$ for the vehicle, one wants to minimize the total amount of fuel used, the distance from the terminal place at the terminal time, and the velocity at the terminal time (if one wants to stop at the terminal time).

It was shown (Section 3.3) that, for a special case, this problem can be explicitly solved in terms of Wiener integrals. This special case could cover the above example, if the force is additive. More research will undoubtedly show how to express the solutions to more general control problems in terms of Wiener integrals or integrals closely enough related to Wiener integrals so that the approximation techniques of Section 6 can be used.

We consider now a slightly different problem. In most control problems, one has, at best, only a foggy notion of the exact position or velocity of the vehicle. The control problem described in Section 3.2 implicitly assumed that the position and velocity were exactly known to the control $u$—since $u$ was assumed to be a function of $x$ and $\dot{x}$. Thus, a closer approximation to reality can be obtained if it is assumed that one can know $x$ and $\dot{x}$ only up to some noise factor.

The general solution to the control problem when there is noise in the force factor and when there is noise in the control system has not yet been obtained—not by Wiener integrals, nor by any other method, except in the linear case.

However, Kalman and Bucy[25,26] showed how an optimal estimate of the position and the velocity of the vehicle could be obtained given certain noisy observations of position and velocity. Kalman also showed that his method had close connection to the Wiener technique of analyzing stationary time series. Kalman, however, assumed that the force equation (1) was a linear equation.

It will now be shown that a general solution to this problem, henceforth called the Kalman problem, can be obtained for nonlinear equations in terms of Wiener integrals.[3]

Consider now the system of equations

$$
\begin{aligned}
\ddot{x}_1 &= f_1(t, x_1(t), \dot{x}_1(t)) + \lambda \dot{z}_1(t) \\
\ddot{x}_2 &= f_2(t, x_2(t), x_1(t), \dot{x}_1(t)) + \lambda \dot{z}_2(t) \\
x_1(S) &= X_1, \qquad \dot{x}_1(S) = Y_1
\end{aligned} \tag{56}
$$

where $z_1(t)$ and $z_2(t)$ are two independent Brownian motions and $x_1$ and $x_2$ are two Ito processes.

The Kalman problem now is, given that we know $x_2$, what is $x_1$? Or, more (less) exactly, since $x_1$ is a random quantity, what is the distribution of $x_1$ given that we know $x_2(t)$, $S \leqslant t \leqslant T$?

The distribution of $x_1$ given $x_2$ will, of course, following the main theme of this paper, be given by giving the expected value of every well-defined function $F(x_1, \dot{x}_1)$ with respect to distribution of $x_1$ given $x_2$.

The formula for conditional expectations is, from Eq. 30, for $B$ a set of positive measure,

$$
\begin{aligned}
&E_x{}^I\{F(x_1, \dot{x}_1) | \, x_1(S) = X_1, \dot{x}_1(S) = Y_1, x_2 \in B\} \\
&\qquad = \frac{E^I_{x_1,x_2}\{F(x_1, \dot{x}_1)\, \chi_B(x_2) | \, x_1(S) = X_1, \dot{x}_1(S) = Y_1\}}{E^I_{x_1,x_2}\{\chi_B(x_2) | \, x_1(S) = X_1, \dot{x}_1(S) = Y_1\}}
\end{aligned} \tag{57}
$$

From Theorem 2.1 (for two-dimensional processes) this is equal to

$$
\begin{aligned}
&E^w_{z_1,z_2} \Big\{ F\left(\lambda \int_S^{(\cdot)} z_1(\alpha)\, d\alpha + X_1, \lambda z_1(\cdot)\right) \chi_B \left(\lambda \int_S^{(\cdot)} z_2(\alpha)\, d\alpha + X_2\right) \\
&\qquad \times \exp\Big[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt \\
&\qquad + \lambda^{-1} \left(\int_S^T f_1\, dz_1(t) + \int_S^T f_2\, dz_2(t)\right)\Big] \Big| \, z_1(S) = Y_1, z_2(S) = Y_2 \Big\} \\
&\div E^w_{z_1,z_2} \Big\{ \chi_B \left(\lambda \int_S^{(\cdot)} z_2(\alpha)\, d\alpha + X_2\right) \\
&\qquad \times \exp\Big[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt \\
&\qquad + \lambda^{-1} \left(\int_S^T f_1\, dz_1(t) + \int_S^T f_2\, dz_2(t)\right)\Big] \Big| \, z_1(S) = Y_1, z_2(S) = Y_2 \Big\}
\end{aligned}
$$

where $x_2(S) = X_2$, $\dot{x}_2(S) = Y_2$.

We now let $B$ be a measurable set containing the observed $x_2$, and multiply and divide the above fraction by $E_z{}^w\{\psi_B(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X_2) | \, z(S) = Y_2/\lambda\}$.

---

[3] Other solutions to the nonlinear case have been obtained by Friedland and Bernstein,[13] Kushner,[29] Bryson and Frazier,[1] and others.

We now let $B$ shrink down to the observed $x_2$. It may be shown that the integration on the $z_2$ variable disappears and that, wherever the $z_2$ variable appeared, it can be replaced by $\dot{x}_2/\lambda$. Therefore, in the limit as $B \to x_2$, (57) becomes

$$\begin{aligned}
&E_x^I\{F(x_1, \dot{x}_1) \mid x_1(S) = X, \dot{x}_1(S) = Y, X_2\} \\
&= E_z^w \Big\{F\Big(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X, \lambda z(\cdot)\Big) \\
&\quad \times \exp\Big[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt \\
&\quad + \lambda^{-1} \int_S^T f_1\, dz(t) + \lambda^{-2} \int_S^T f_2\, d\dot{x}_2(t)\Big] \Big| z(S) = Y_1\lambda^{-1}\Big\} \\
&\div E_z^w \Big\{\exp\Big[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt \\
&\quad + \lambda^{-1} \int_S^T f_1\, dz(t) + \lambda^{-2} \int_S^T f_2\, d\dot{x}_2(t)\Big] \Big| z(S) = Y_1\lambda^{-1}\Big\} \qquad (58)
\end{aligned}$$

We state the above result as a theorem.

**Theorem 4.1.** Suppose the two-dimensional vector-valued function $f_1(t, x_1, \dot{x}_1), f_2(t, x_2, x_1, \dot{x}_1)$ satisfies the conditions of Theorem 2.1 in vector form, and that the second derivative of the observed function, $x_2(t)$, is absolutely continuous. Suppose further that $f_{2t}, f_{2Y_1}, f_{2X_2}, f_{2X_1}$, and $f_{2Y_1Y_1}$ exist and are continuous.[4] Then

$$\begin{aligned}
&E_{x_1}^I\{F(x_1(\cdot), \dot{x}(\cdot)) \mid x_1(S) = X, \dot{x}_1(S) = Y_1, x_2(\cdot)\} \\
&= E_z^w \Big\{F\Big(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X, \lambda z(\cdot)\Big) \\
&\quad \times \exp\Big[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt \\
&\quad + \lambda^{-1} \int_S^T f_1\, dz(t) + \lambda^{-2} \int_S^T f_2\, d\dot{x}_2(t)\Big] \Big| z(S) = Y_1\lambda^{-1}\Big\} \\
&\div E_z^w \Big\{\exp\Big[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt \\
&\quad + \lambda^{-1} \int_S^T f_1\, dz(t) + \lambda^{-2} \int_S^T f_2\, d\dot{x}_2(t)\Big] \Big| z(S) = Y_1\lambda^{-1}\Big\} \qquad (59)
\end{aligned}$$

Theorem 4.1 gives a complete solution to the Kalman problem, since the expected value of any function of the $x_1$ process can be expressed given $x_2$ and $\dot{x}_2$. It is, of course, not in a convenient computational form. In Section 6, it will be shown how to evaluate ratios of Wiener integrals. Schilder[38] shows that, in the case where $f_1$ and $f_2$ are linear, the result given here agrees with the original Kalman–Bucy result.

[4] These last conditions are necessary to ensure that $\int_S^T f_2\, dz_2(t) \to \int_S^T f_2\, d\dot{x}_2(t)$ as $B \to x_2$.

It will be shown now that, in the case that $F(x, \dot{x})$ happens to be of the form $F(x, \dot{x}) = \phi(x(T), \dot{x}(T))$, (59) can be evaluated using partial differential equations.

**Corollary 4.1.** Suppose that $f_1$ and $f_2$ satisfy the conditions of Theorem 4.1 and that the partial differential equation

$$W_t(t, X, Y) + f_1 W_Y - Y W_X + (1/\lambda^2)\,[f_2(t, x_2(t), X, Y)\,\ddot{x}_2(t) - \tfrac{1}{2} f_2^2(t, X, Y)]\,W + \tfrac{1}{2}\lambda^2 W_{YY} = 0 \tag{60}$$

has a well-defined solution for each of the terminal conditions

$$W(T, X, Y) = \phi(X, Y) > 0 \tag{61}$$

$$W(T, X, Y) \equiv 1 \tag{62}$$

Then

$$W_1(S, X, Y) = E_z^w \left\{ \phi\left(\lambda \int_S^T z(\alpha)\, d\alpha + X_1, \lambda z(T)\right) \exp\left[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt + \lambda^{-1} \int_S^T f_1\, dz(t) + \lambda^{-2} \int_S^T f_2 \ddot{x}_2(t)\, dt\right] \Big|\, z(S) = Y\lambda^{-1} \right\}$$

which is the numerator of the right-hand side of (59), solves (60) with terminal conditions (61), and

$$W_2(S, X, Y) = E_z^w \left\{ \exp\left[-(2\lambda^2)^{-1} \int_S^T (f_1^2 + f_2^2)\, dt + \lambda^{-1} \int_S^T f_1\, dz(t) + \lambda^{-2} \int_S^T f_2 \ddot{x}(t)\, dt\right] \Big|\, z(S) = Y\lambda^{-1} \right\}$$

which is the denominator of the right-hand side of (59), solves (60) with terminal conditions (62). Both $x_2(t)$ and $\dot{x}_2(t)$ are assumed known.

*Proof.* The proof is immediate from Theorem 3.4.

We note that, if $f_2(t, x_2(t), x_1(t), \dot{x}_1(t))$ does not depend on $x_1(t), \dot{x}_1(t)$, then knowledge of $x_2$ gives no knowledge of $x_1$, and $f_2$ is just a function of $t$ if $x_2$ is given. It is a well-known result that, in this case, solutions of (60) can be written in the form

$$W(t, X, Y) = W_0(t, X, Y) \times \exp\left\{(1/\lambda^2) \int_t^T [f_2 \ddot{x}_2(t) - \tfrac{1}{2} f_2^2]\, dt\right\},$$

where $W_0(t, X, Y)$ satisfies (60) with $f_2 = 0$. If $W_0(t, X, Y)$ is given terminal conditions (61), then it can be seen that $W_1(t, X, Y)$, defined above, is simply

$$W_0(t, X, Y) \exp\left\{(1/\lambda^2) \int_t^T \left[f_2 \ddot{x}_2(t) - \frac{1}{2} f_2^2\right] dt\right\}$$

while by Corollary 2.1, $W_2(t, X, Y)$ is $\exp\{(1/\lambda^2)\int_t^T [f_2\ddot{x}_2(t) - \frac{1}{2}f_2^2]\,dt\}$. Their ratio is the solution to the Kalman problem, and is just $W_0(t, X, Y)$. By definition, $W_0(t, X, Y)$ satisfies the ordinary backwards equation, as it should, since, in this case, there is no additional knowledge of $x_1$. Other authors who considered the nonlinear Kalman problem apparently did not make this check.

For other applications of Wiener integrals to engineering, see Wiener,[43] Wiener *et al.*,[44] MacDonald,[30] and Schilder.[36]

## 5. APPLICATIONS OF WIENER INTEGRALS TO PHYSICS

**5.1.** As stated in Section 2, Brownian motion was first devised to describe the motion of a particle in a fluid. While a number of characteristics of this model fit very well to the physical situation, by (6), the variance of the velocity of the particle is infinite. Classical kinetic theory holds[22] that the kinetic energy is proportional to the variance of the velocity, so that, clearly, some modification of the original hypotheses must be made.

The equation

$$\ddot{x}(t) = f(t, x(t)) + \beta\dot{x}(t) + \lambda\dot{z}(t), \qquad x(S) = X \sim \rho(S, X), \quad \dot{x}(S) = u(S, X) \tag{63}$$

has been proposed to describe the motions of a particle in a solution whose position is $x(t)$ at time $t$, which is acted upon by an external force $f(t, x(t))$, by a drag force $\beta\dot{x}(t)$, and by a random force $\lambda\dot{z}(t)$ corresponding to the particle being bumped by molecules of the solution. The initial position of the particle, $X$, is now assumed to be random with a distribution equal to the normalized density of the particles at initial time $S$, and the initial velocity is assumed to be a function of the particle's initial position, $u(S, x(S))$. For more details, history, and applications of this model, see Chandrasekhar[6] and Nelson.[32] Equation (63) is, of course, simply a statement of Newton's second law with a random force. Now $\dot{x}(t)$ has the dimension of $z(t)$, so that it has finite variance. While the model (63) was first proposed just to describe the motions of heavy particles in a relatively lighter solution, Kirkwood[27] and Kirkwood *et al.*[28] show, for "liquids and other condensed systems," that the Ito equation (63) is actually a consequence of the (nonrandom) equations of motion of all the particles of the fluid; that is, for liquids and other condensed systems, $x(t)$ can denote the position of a molecule of the system. Other authors[30,33] have proposed that (63) can be used as a model for studying departures from equilibrium.

We note that (63) is a special case of (1) if the initial conditions of (1) are randomized. If, now, $F(x, \dot{x})$ is some function of the position of the paths and their velocity, then it can be seen that

$$\begin{aligned} &E_x^I\{F(x, \dot{x}) \mid x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X)\} \\ &\quad = \int_{-\infty}^{\infty} E_x^I\{F(x, \dot{x}) \mid x(S) = X, \dot{x}(S) = u(S, X)\}\, \rho(S, X)\, dX \end{aligned} \tag{64}$$

where again the notation $X \sim \rho(S, X)$ means that $X$ has the distribution $\rho(S, X)$.

If $f(t, x(t)) + \beta\dot{x}(t)$ and $F(x, \dot{x})$ satisfy the conditions of Theorem 2.1, it follows from Theorem 2.1 and Eq. (64) that

$$\begin{aligned} &E_x^I\{F(x(\cdot), \dot{x}(\cdot)) |\, x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X)\} \\ &= \int_{-\infty}^{\infty} E_z^w \Big\{F\Big(\lambda \int_S^{(\cdot)} z(\alpha)\, d\alpha + X, \lambda z(\cdot)\Big) \\ &\quad \times \exp \Big\{-(2\lambda^2)^{-1} \int_S^T \Big[f\Big(t, \lambda \int_S^t z(\alpha)\, d\alpha + X\Big) + \beta\lambda z(t)\Big]^2 dt \\ &\quad + \lambda^{-1} \int_S^T f\Big(t, \lambda \int_S^t z(\alpha)\, d\alpha + X\Big)\, dz(t) \\ &\quad + \beta\lambda^{-1} \int_S^T z(t)\, dz(t)\Big\}\Big|\, z(S) = \lambda^{-1} u(S, X)\Big\}\, \rho(S, X)\, dX \end{aligned} \tag{65}$$

From (65), it follows immediately that *any* thermodynamic function of a fluid or collection of particles whose equations of motion are described by (63) can be expressed as a Wiener integral. The Wiener integral thus becomes a kind of partition function for nonequilibrium problems. Two of the important parameters of fluid dynamics are the density at a point $P$ at time $T$, $\rho(T, P)$, and the average velocity $u$ of all particles at point $P$ at time $T$. They can be defined as

$$\rho(T, P) = \lim_{B \to P} \{E_x^I[\chi_B(x(T)) |\, x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X)]/|\, B\,|\} \tag{66}$$

(where $B$ is an interval containing $P$, and $|\, B\,|$ is the length of $B$), and

$$u(T, P) = E_x^I\{\dot{x}(T) |\, x(S) = X \sim \rho(S, X), \quad \dot{x}(S) = u(S, X), \quad X(T) = P\} \tag{67}$$

They can be expressed as Wiener integrals as follows:

**Theorem 5.1.**[5] Suppose $f(t, x(t), \dot{x}(t))$ satisfies the conditions of Theorem 2.1; then the density $\rho(T, P)$ at point $P$ and time $T$ is

$$\begin{aligned} \rho(T, P) = E_z^w \Big\{&\rho\Big(S, P - \lambda \int_S^T z(\alpha)\, d\alpha\Big) \\ &\times \exp \Big[-(2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t)\Big] \\ \Big|\, &z(S) = \lambda^{-1} u\Big(S, P - \lambda \int_S^T z(\alpha)\, d\alpha\Big)\Big\} \end{aligned} \tag{68}$$

and the flow (average) velocity at $T$, $P$ is

$$\begin{aligned} u(T, P) = [1/\rho(T, P)]\, E_z^w \Big\{&\lambda z(T)\, \rho\Big(S, P - \lambda \int_S^T z(\alpha)\, d\alpha\Big) \\ &\times \exp \Big[-(2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t)\Big]\Big|\, z(S) \\ &= \lambda^{-1} u\Big(S, P - \lambda \int_S^T z(\alpha)\, d\alpha\Big)\Big\} \end{aligned} \tag{69}$$

[5] Since no extra difficulty is involved, this theorem is stated and proved for a general $f(t, x(t), \dot{x}(t))$ and not one of the form $f(t, x(t)) + \beta\dot{x}(\tau)$.

*Proof.* Let $B(P, \epsilon)$ be the interval $[P - \frac{1}{2}\epsilon, P + \frac{1}{2}\epsilon]$. Then $| B |$ (the length of $B$) is $\epsilon$. Let, as before,

$$\begin{aligned} \chi_{B(P,\epsilon)}(X) &= 1 \qquad \text{if} \quad P - \tfrac{1}{2}\epsilon \leqslant X \leqslant P + \tfrac{1}{2}\epsilon \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

By (65) and (66),

$$\begin{aligned} \rho(T, P) &= \lim_{B \to P} | B |^{-1} E_x{}^{I}\{\chi_B(x(T)) | \, x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X)\} \\ &= \lim_{\epsilon \to 0} \epsilon^{-1} \int_{-\infty}^{\infty} E_z{}^{w} \Big\{ \chi_{B(P,\epsilon)} \Big(\lambda \int_S^T z(\alpha)\, d\alpha + X\Big) \\ &\quad \times \exp\Big[-(2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t)\Big] \Big| \, z(S) = \lambda^{-1} u(S, X)\Big\} \rho(S, X)\, dX \end{aligned}$$

Multiply and divide (under the one-dimensional integral) by

$$E_z{}^{w} \Big\{ \chi_{B(P,\epsilon)} \Big(\lambda \int_S^T z(\alpha)\, d\alpha + X\Big) \Big| \, z(S) = (1/\lambda)\, u(S, X)\Big\}$$

Thus,

$$\begin{aligned} \rho(t, P) = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} & \frac{\begin{array}{c} E_z{}^{w}\{\chi_{B(P,\epsilon)} [\lambda \int_S^T z(\alpha)\, d\alpha + X] \\ \times \exp[(-2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t)] | \, z(S) = \lambda^{-1} u(S, X)\} \end{array}}{E_z{}^{w}\{\chi_{B(P,\epsilon)}(\lambda \int_S^T z(\alpha)\, d\alpha + X) | \, z(S) = (\lambda^{-1})\, u(S, X)\}} \\ & \times \epsilon^{-1} E_z{}^{w} \Big\{ \chi_{B(P,\epsilon)} \Big(\lambda \int_S^T z(\alpha)\, d\alpha + X\Big) \Big| \, z(S) = \lambda^{-1} u(S, X)\Big\} \rho(S, X)\, dX \end{aligned}$$

By (30), the first ratio of Wiener integrals is

$$\begin{aligned} E_z{}^{w} & \Big\{ \exp\Big[-(2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t)\Big] \Big| \, z(S) \\ &= \lambda^{-1} u(S, X), \lambda \int_S^T z(\alpha)\, d\alpha + X \in B(P, \epsilon)\Big\} \end{aligned}$$

By definition, $\epsilon^{-1} E_z{}^{w}\{\chi_{B(P,\epsilon)}(\lambda \int_S^T z(\alpha)\, d\alpha + X) | \, z(S) = \lambda^{-1} u(S, X)\}$ is

$$\epsilon^{-1} \operatorname{Prob} \Big\{ P - \frac{1}{2}\epsilon \leqslant \lambda \int_S^T z(\alpha)\, d\alpha + X \leqslant P + \frac{1}{2}\epsilon \,\Big|\, z(S) = \lambda^{-1} u(S, X)\Big\}$$

Therefore,

$$\begin{aligned} \rho(T, P) = \int_{-\infty}^{\infty} & E_z{}^{w} \Big\{ \exp\Big[-(2\lambda^2)^{-1} \int_S^T f^2\, dt + \lambda^{-1} \int_S^T f\, dz(t)\Big] \Big| \, z(S) \\ &= \lambda^{-1} u(S, X), \lambda \int_S^T z(\alpha)\, d\alpha + X = P\Big\} \\ & \times d \operatorname{Prob} \Big\{ \lambda \int_S^T z(\alpha)\, d\alpha + X = P \,\Big|\, z(S) = \lambda^{-1} u(S, X)\Big\} \rho(S, X)\, dX \end{aligned}$$

where

$$d\,\text{Prob}\left\{\lambda \int_S^T z(\alpha)\,d\alpha + X = P \,\middle|\, z(S) = \lambda^{-1} u(S, X)\right\}$$

is the density function of the random variable $\lambda \int_S^T z(\alpha)\,d\alpha + X$. The function $\rho(S, X)$ is now moved under the Wiener integral sign. Since $\lambda \int_S^T z(\alpha)\,d\alpha + X = P$ in this integral, it follows that $X = P - \lambda \int_S^T z(\alpha)\,d\alpha$. Making this substitution whenever $X$ appears,

$$\begin{aligned}
\rho(T, P) = \int_{-\infty}^{\infty} E_z{}^w &\Big\{\rho\left(S, P - \lambda \int_S^T z(\alpha)\,d\alpha\right) \\
&\times \exp\left[-(2\lambda^2)^{-1} \int_S^T f^2 + \lambda^{-1} \int_S^T f\,dz(t)\right]\Big|\, z(S) \\
&= \lambda^{-1} u\left(S, P - \lambda \int_S^T z(\alpha)\,d\alpha\right), \lambda \int_S^T z(\alpha)\,d\alpha = P - X\Big\} \\
&\times d\,\text{Prob}\left\{\lambda \int_S^T z(\alpha)\,d\alpha = P - X \,\middle|\, z(S) = u\left(S, P - \lambda \int_S^T z(\alpha)\,d\alpha\right)\right\}\,dX
\end{aligned}$$

We now change variables in the one-dimensional integral, letting $P - X = R$, and the result follows from Lemma 2.4 (read from right to left).

By definitions (67) and (30)

$$u(T, P) = \lim_{B \to P} \frac{E_x{}^I\{\dot{x}(T)\,\chi_B(x(T)) \mid x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X)\}}{E_x{}^I\{\chi_B(x(T)) \mid x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X)\}}$$

This expression can be rewritten, in the same way as $\rho(T, X)$, as

$$u(T, P) = \lim_{\epsilon \to 0} \frac{\int_{-\infty}^{\infty} E_z{}^w\{\lambda z(T)\,\zeta\}\,\rho(S, X)\,dX}{\int_{-\infty}^{\infty} E_z{}^w\{\zeta\}\,\rho(S, X)\,dX}$$

with

$$\begin{aligned}
\zeta \equiv \chi_{B(P,\epsilon)} &\left(\lambda \int_S^T z(\alpha)\,d\alpha + X\right) \\
&\times \exp\left[-(2\lambda^2)^{-1} \int_S^T f^2\,dt + \lambda^{-1} \int_S^T f\,dz(t)\right]\Big|\, z(S) = \lambda^{-1} u(S, X)
\end{aligned}$$

Multiplying and dividing this expression under each one-dimensional integral sign by $E_z{}^w\{\chi_{B(P,\epsilon)}(\lambda \int_S^T z(\alpha)\,d\alpha + X) \mid z(S) = \lambda^{-1} u(S, X)\}$ and using (30), we have

$$u(T, P)$$

$$= \lim_{\epsilon \to \infty} \frac{\begin{array}{r}\int_{-\infty}^{\infty} E_z{}^w\{\lambda z(T)\,\eta\} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \\ \times \int_{-\infty}^{\infty} E_z{}^w\{\chi_{B(P,\epsilon)}(\lambda \int_S^T z(\alpha)\,d\alpha + X) \mid z(S) = \lambda^{-1} u(S, X)\}\,\rho(S, X)\,dX\end{array}}{\int_{-\infty}^{\infty} E_z{}^w\{\eta\} \int_{-\infty}^{\infty} E_z{}^w\{\chi_{B(P,\epsilon)}(\lambda \int_S^T z(\alpha)\,d\alpha + X) \mid z(S) = \lambda^{-1} u(S, X)\}\,\rho(S, X)\,dX}$$

with

$$\begin{aligned}
\eta &\equiv \exp\left[-(2\lambda^2)^{-1} \int_S^T f^2\,dt + \lambda^{-1} \int_S^T f\,dz(t)\right]\Big|\, z(S) \\
&= \lambda^{-1} u(S, X), \lambda \int_S^T z(\alpha)\,d\alpha + X \in B(P, \epsilon)
\end{aligned}$$

Upon multiplying and dividing this expression by $\epsilon$ and taking the limit in the same way as with $\rho(t, P)$, we obtain the desired result.

Jeans[22] defines the kinetic energy of a nonequilibrium fluid at a point to be proportional to the mass of the molecules concerned times the variance of their velocity at the given point. Thus, if the units are picked properly, kinetic energy can be defined in the present case at time $T$ and point $P$ as

$$\text{K.E. } (T, P) = E_x^I\{[\dot{x}(T) - u(T, P)]^2 \mid x(S) = X \sim \rho(S, X), \dot{x}(S) = u(S, X), x(T) = P\} \tag{70}$$

Jeans also shows that the pressure at a given point is proportional to the kinetic energy of the fluid at the point times the density of the fluid at the point. Thus, if the units are picked properly,

$$\text{Pressure}(T, P) = \text{K.E. } (T, P)\, \rho(T, P) \tag{71}$$

where K.E. $(T, P)$ is defined by (70) and $\rho(T, P)$ is defined by (66). Of course, both kinetic energy and pressure have representations as Wiener integrals in the same way as the density and average velocity do.

The proof of the following theorem is given in Schilder.[39] It is too long to be given here. It shows that $\rho$, $u$, and Pressure satisfy a system of nonlinear partial differential equations similar to the Navier–Stokes equations.

**Theorem 5.2.** Suppose $f(t, x(t), \dot{x}(t))$ satisfies the conditions of Theorem 2.1. Then

$$\lim_{T\to S} \rho(T, P) = \rho(S, P), \qquad \lim_{T\to S} u(T, P) = u(S, P) \tag{72}$$

$$\rho_t + (u\rho)_P = 0 \tag{73}$$

$$\begin{aligned} u_t + uu_P = (-Q_P/\rho) + E_x^I\{f(T, x(T), \dot{x}(T)) | \, x(S) \\ = X \sim \rho(S, X), \quad \dot{x}(S) = u(S, X), \quad x(T) = P\} \end{aligned} \tag{74}$$

with $\rho$ and $u$ definedfby (66) and (67), and $Q(T, P) = \text{Pressure}(T, P)$. In the case that $f(t, x(t), \dot{x}(t))$ is of the form $f(t, x(t)) + \beta\dot{x}(t)$, the last term of (74) can be written as $f(T, P) + \beta u(T, P)$.

For other applications of Wiener integrals to physics, see Feynman and Hibbs,[11] Kac,[23] Wiener,[43] Wiener *et al.*,[44] MacDonald,[30] Nelson,[32] Martin and Segal,[31] Gimibre,[16] and Gelfand and Yaglom.[15]

## 6. EVALUATION OF WIENER INTEGRALS

**6.1.** In this section, a power-series expansion in even powers of $\lambda$ for Wiener integrals will be given. This series first appeared in Schilder,[35] although the coefficients were not evaluated there. First, a lemma is needed. It shows how to evaluate "Gaussian Wiener integrals" (see Feynman and Hibbs[11]).

**Lemma 6.1.** Suppose that the functions $a_1(t),..., a_6(t)$ are continuous in the interval $S \leqslant t \leqslant T$, and that the ricatti equation

$$\begin{aligned}\dot{C}_4(t) &= [a_4(t) - C_4(t)]^2 + 2a_1(t) - 2C_5(t)\\ \dot{C}_5(t) &= [a_4(t) - C_4(t)][a_5(t) - C_5(t)] + a_2(t) - C_6(t)\\ \dot{C}_6(t) &= [a_5(t) - C_5(t)]^2 + 2a_3(t)\end{aligned} \tag{75}$$

with $C_4(T) = -2a_6$, $C_5(T) = -a_7$, and $C_6(T) = -2a_8$, has a solution.

Let $B(t)$ be the matrix

$$B_{11}(t) = a_4(t) - C_4(t), \qquad B_{12} = a_5(t) - C_5(t), \qquad B_{21}(t) = 1, \qquad B_{22}(t) = 0$$

Suppose that the matrix $\Phi_{hj}(h, j = 1, 2)$ is an invertable matrix solution[6] to the equation

$$\frac{d}{dt}\,\Phi_{hj}(t) = \sum_i^2 B_{hi}\Phi_{ij}(t) \tag{76}$$

Then[7]

$$\begin{aligned} E_z{}^w \Big\{&F(z) \exp\Big(\int_S^T \Big\{a_1(t)\, z^2(t) + a_2(t)\, z(t) \int_S^t z(\alpha)\, d\alpha + a_3(t)\Big[\int_S^t z(\alpha)\, d\alpha\Big]^2\Big\}\, dt\\ &+ \int_S^T \Big[a_4(t)\, z(t) + a_5(t) \int_S^t z(\alpha)\, d\alpha\Big]\, dz(t)\\ &+ a_6 z^2(T) + a_7 z(T) \int_S^T z(\alpha)\, d\alpha + a_8 \Big[\Big[\int_S^T z(\alpha)\, d\alpha\Big]^2\Big)\Big|\, z(S) = 0\Big\}\\ = E_x{}^w &\Big\{F\Big(\sum_i^2 \Phi_{1i}(\cdot) \int_S^{(\cdot)} \Phi_{i1}^{-1}(\alpha)\, dx(\alpha)\Big)\Big|\, x(S) = 0\Big\} \exp\Big[-\int_S^T C_4(t)\, dt\Big] \end{aligned} \tag{77}$$

*Proof.* Let the $a_4(t)$ and $a_5(t)$ in the Ito integral in the Wiener integral on the left-hand side of (77) be

$$a_4(t) = b_4(t) + C_4(t), \qquad a_5(t) = b_5(t) + C_5(t) \tag{78}$$

where $C_4(t)$ and $C_5(t)$ are defined by the hypothesis of this theorem [i.e., (75)].

Consider now the following function:

$$F(t, X, Y) = \tfrac{1}{2}C_4(t)\, Y^2 + C_5(t)\, YX + \tfrac{1}{2}C_6(t)\, X^2$$

Substituting this $F$ into (22), we have that

$$\begin{aligned} \frac{1}{2}\, C_4(T)\, z^2(T) &+ C_5(T)\, z(T) \int_S^T z(\alpha)\, d\alpha + \frac{1}{2}\, C_6(T)\Big[\int_S^T z(\alpha)\, d\alpha\Big]^2 - C_4(S)\, z^2(S)\\ = \int_S^T &\Big\{\frac{1}{2}\, \dot{C}_4(t)\, z^2(t) + \frac{1}{2}\, \dot{C}_5(t)\, z(t) \int_S^t z(\alpha)\, d\alpha + \frac{1}{2}\, \dot{C}_6(t)\Big[\int_S^t z(\alpha)\, d\alpha\Big]^2\Big\}\, dt\\ &+ \int_S^T \Big[C_5(t)\, z(t) + C_6(t) \int_S^t z(\alpha)\, d\alpha\Big]\, z(t)\, dt\\ &+ \int_S^T \Big[C_4(t)\, z(t) + C_5(t) \int_S^t z(\alpha)\, d\alpha\Big]\, dz(t) + \frac{1}{2} \int_S^T C_4(t)\, dt \end{aligned} \tag{79}$$

[6] It is shown in books on differential equations that such a $\Phi$ always exists.

[7] This lemma is a generalization of Cameron and Martin's work.[2]

In the Ito integral in (79), substitute the values of $C_4$ and $C_5$ given by (78) and then solve the equation for the Ito integral involving the $a$'s. Thus,

$$\int_S^T \Big[a_4(t)\,z(t) + a_5(t)\int_S^t z(\alpha)\,d\alpha\Big]\,dz(t) = \int_S^T \Big[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\Big]\,dz(t)$$
$$- \int_S^T \Big\{\frac{1}{2}\,\dot{C}_4(t)\,z^2(t) + \dot{C}_5(t)\,z(t)\int_S^t z(\alpha)\,d\alpha + \frac{1}{2}\,\dot{C}_6(t)\Big[\int_S^t z(\alpha)\,d\alpha\Big]^2\Big\}\,dt$$
$$- \int_S^T \Big[C_5(t)\,z(t) + C_6(t)\int_S^t z(\alpha)\,d\alpha\Big]\,z(t)\,dt - \frac{1}{2}\int_S^T C_4(t)\,dt + \frac{1}{2}\,C_4(T)\,z^2(T)$$
$$+ \frac{1}{2}\,C_5(T)\,z(T)\int_S^T z(\alpha)\,d\alpha + \frac{1}{2}\,C_6(T)\Big[\int_S^T z(\alpha)\,d\alpha\Big]^2 - \frac{1}{2}\,C_4(S)\,z^2(S)$$

Substituting this result into the Wiener integral on the left-hand side of (77) and collecting terms with like coefficients, $z$, $z^2$, and $dz$, we find that the left-hand side of (77) becomes

$$E_z^w\,\Big\{F(z)\exp\Big(\int_S^T \Big[a_1(t) - \frac{1}{2}\,\dot{C}_4(t) - C_5(t)\Big]\,z^2(t)\,dt$$
$$+ \int_S^T \Big[a_2(t) - \dot{C}_5(t) - C_6(t)\Big]\,z(t)\Big[\int_S^t z(\alpha)\,d\alpha\Big]\,dt$$
$$+ \int_S^T \Big[a_3(t) - \frac{1}{2}\,\dot{C}_6(t)\Big]\Big[\int_S^t z(\alpha)\,d\alpha\Big]^2 + \Big[a_6 + \frac{1}{2}\,C_4(t)\Big]\,z^2(T)$$
$$+ [a_7 + C_5(T)]\,z(T)\int_S^T z(\alpha)\,d\alpha + \Big[a_8 + \frac{1}{2}\,C_6(T)\Big]\Big[\int_S^T z(\alpha)\,d\alpha\Big]^2$$
$$+ \int_S^T \Big[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\Big]\,dz(t)$$
$$- \frac{1}{2}\int_S^T C_4(t)\,dt - \frac{1}{2}\,C_4(S)\,z^2(S)\Big)\,\Big|\,z(S) = 0\Big\} \tag{80}$$

Using the differential equations (75) with the terminal conditions and remembering that $b_4(t) = a_4(t) - C_4(t)$ and $b_5(t) = a_5(t) - C_5(t)$, we find that (80) becomes

$$E_z^w\,\Big\{F(z)\exp\Big(-\frac{1}{2}\int_S^T b_4^2(t)\,z^2(t)\,dt - \Big[\int_S^T b_4(t)\,b_5(t)\,z(t)\int_S^t z(\alpha)\,d\alpha\Big]\,dt$$
$$- \frac{1}{2}\int_S^T b_5^2(t)\Big[\int_S^t z(\alpha)\,d\alpha\Big]^2\,dt + \int_S^T \Big[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\Big]\,dz(t)$$
$$- \frac{1}{2}\int_S^T C_4(t)\,dt - \frac{1}{2}\,C_4(S)\,z^2(S)\Big)\,\Big|\,z(S) = 0\Big\}$$

This can be rewritten as

$$E_z^w\,\Big\{F(z)\exp\Big(-\frac{1}{2}\int_S^T \Big[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\Big]^2\,dt$$
$$+ \int_S^T \Big[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\Big]\,dz(t)\Big)\,\Big|\,z(S) = 0\Big\}\,\exp\Big[-\frac{1}{2}\int_S^T C_4(t)\,dt\Big] \tag{81}$$

which is a more familiar form [see (27)]. We can write (81) in the flat-integral form as

$$\int F(z) \exp\left(-\frac{1}{2}\int_S^T \left[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha - \dot{z}(t)\right]^2 dt\right)\Big|\, z(S) = 0\Big\}$$
$$\times\ \delta z \exp\left[-\frac{1}{2}\int_S^T C_4(t)\,dt\right] \tag{82}$$

which shows that we have just completed the square in the Wiener integral on the left-hand side of (77). To complete the analysis now, we let

$$x(t) = z(t) - \int_S^t \left[b_4(\alpha)\,z(\alpha) + b_5(\alpha)\int_S^\alpha z(\beta)\,d\beta\right] d\alpha \tag{83}$$

and substitute this expression into the right-hand side of (77). Using the transformation theorem, we get for the right-hand side of (77)

$$E_z^w \Big\{F\Big(\sum_i^2 \Phi_{1i}(\cdot)\int_S^{(\cdot)} \Phi_{i1}^{-1}(\alpha)\,dz(\alpha) - \sum_i^2 \Phi_{1i}(\cdot)\int_S^{(\cdot)} \Phi_{i1}^{-1}(\alpha)\,b_4(\alpha)\,z(\alpha)\,d\alpha$$
$$-\sum_i^2 \Phi_{1i}(\cdot)\int_S^{(\cdot)} \left[\Phi_{i1}^{-1}(\alpha)\,b_5(\alpha)\int_S^\alpha z(\beta)\,d\beta\right] d\alpha\Big)$$
$$\times \exp\Big(-\frac{1}{2}\int_S^T \left[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\right]^2 dt$$
$$+\int_S^T \left[b_4(t)\,z(t) + b_5(t)\int_S^t z(\alpha)\,d\alpha\right] dz(t)\Big)\Big|\, z(S) = 0\Big\} \exp\left[-\frac{1}{2}\int_S^T C_4(\alpha)\,d\alpha\right]$$

To complete the proof of this corollary, it is now sufficient to show that [see (81)]

$$\sum_i^2 \Phi_{1i}(t)\int_S^t \Phi_{i1}^{-1}(\alpha)\,dz(\alpha) - \sum_i^2 \Phi_{1i}(t)\int_S^t \Phi_{i1}^{-1}(\alpha)\,b_4(\alpha)\,z(\alpha)\,d\alpha$$
$$-\sum_i^2 \Phi_{1i}(t)\int_S^t \left[\Phi_{i1}^{-1}(\alpha)\,b_5(\alpha)\int_S^\alpha z(\beta)\,d\beta\right] d\alpha = z(t) \qquad (S \leqslant t \leqslant T) \tag{84}$$

We note, that if we integrate by parts[8] the first term on the left-hand side of (84), we get

$$\sum_i^2 \Phi_{1i}(t)\int_S^t \Phi_{i1}^{-1}(\alpha)\,dz(\alpha) = \sum_i^2 \Phi_{1i}(t)\Big\{\Phi_{i1}^{-1}(t)\,z(t) - \int_S^t \frac{d}{d\alpha}[\Phi_{i1}^{-1}(\alpha)]\,z(\alpha)\,d\alpha\Big\} \tag{85}$$

Since $\sum_j \Phi_{hj}(t)\,\Phi_{jk}^{-1}(t) = \delta_{hk}$,[9] it follows that

$$\sum_j^2 \frac{d}{dt}[\Phi_{hj}(t)]\,\Phi_{jk}^{-1}(t) = -\sum_j^2 \Phi_{hj}(t)\,\frac{d}{dt}[\Phi_{jk}^{-1}(t)]$$

---

[8] We again use (22), of course, with $F_i(T, \int_S^T z(\alpha)\,d\alpha, z(T)) = \Phi_{i1}^{-1}(T)z(T)$. Since the $F_i$ are just linear functions of $z$, their second derivatives with respect to $z$ vanish, and we have the ordinary integration-by-parts formula.

[9] Here, $\delta_{hk}$ is the identity matrix.

From (76) and the last equation, it follows that

$$\sum_i^2 \sum_j^2 B_{hi}(t)\, \Phi_{ij}(t)\, \Phi_{jk}^{-1}(t) = -\sum_j \Phi_{hj}(t) \frac{d}{dt}\, [\Phi_{jk}^{-1}(t)]$$

which is clearly equivalent to

$$B_{hk}(t) = -\sum_j^2 \Phi_{hj}(t) \frac{d}{dt}\, [\Phi_{jk}^{-1}(t)]$$

Multiplying through by $-\Phi^{-1}$, we get

$$-\sum_h \Phi_{ih}^{-1}(t)\, B_{hk}(t) = \frac{d}{dt}\, [\Phi_{ik}^{-1}(t)] \tag{86}$$

Substituting this into (85) and multiplying it out, it follows that

$$\sum_i^2 \Phi_{1i}(t) \int_S^t \Phi_{i1}^{-1}(\alpha)\, dz(\alpha) = z(t) + \sum_i \sum_h \Phi_{1i}(t) \int_S^t \Phi_{ih}^{-1}(\alpha)\, B_{h1}(\alpha)\, z(\alpha)\, d\alpha \tag{87}$$

Writing out the $h$ summation on the right-hand side of (87), we get

$$z(t) + \sum_i \Phi_{1i}(t) \int_S^t \Phi_{i1}^{-1}(\alpha)\, B_{11}(\alpha)\, z(\alpha)\, d\alpha + \sum_i \Phi_{1i}(t) \int_S^t \Phi_{i2}^{-1}(\alpha)\, B_{21}(\alpha)\, z(\alpha)\, d\alpha \tag{88}$$

From (84), (85), (88), and the definition of the $B$ matrix, it follows that it is sufficient to show that

$$\sum_i \Phi_{1i}(t) \int_S^t \Phi_{i2}^{-1}(\alpha)\, B_{21}(\alpha)\, z(\alpha)\, d\alpha = \sum_i^2 \Phi_{1i}(t) \int_S^t \left[\Phi_{i1}^{-1}(\alpha)\, b_5(\alpha) \int_S^\alpha z(\beta)\, d\beta\right] d\alpha \tag{89}$$

Since $B_{21}(t) = 1$, the left-hand side of (89) is $\sum_i \Phi_{1i}(t) \int_S^t \Phi_{i2}^{-1}(\alpha)\, z(\alpha)\, d\alpha$. Integrating this by parts again, we get for the left-hand side of (89)

$$\sum_i \Phi_{1i}(t)\, \Phi_{i2}^{-1}(t)\, z(t) - \sum_i \Phi_{1i}(t) \int_S^t \frac{d}{dt}\, [\Phi_{i2}(\alpha)]^{-1} \left[\int_S^\alpha z(\beta)\, d\beta\right] d\alpha$$

The first term of the above is zero, by the definition of $\Phi$, and the second becomes, using (86) again

$$\sum_i \sum_h \Phi_{1i}(t) \int_S^t \left[\Phi_{ih}^{-1}(\alpha)\, B_{h2}(\alpha) \int_S^\alpha z(\beta)\, d\beta\right] d\alpha \tag{90}$$

But, since $B_{12}(\alpha) = b_5(\alpha)$ and $B_{22}(\alpha) = 0$, by definition, it follows that (90) is exactly the right-hand side of (89), which was to be shown.

**6.2.** It will now be shown that Wiener integrals of a general form can be expanded in an asymptotic series in powers of $\lambda$, and that the terms of this series can be evaluated term by term by methods that require nothing more than the solution of ordinary differential equations and quadratures. This asymptotic series, of course, becomes more accurate as $\lambda \to 0$. If $\lambda = 0$, then it follows that (1) no longer defines an Ito distribution, but becomes an ordinary differential equation; it follows that (27), for example, becomes $F(x(\cdot), \dot{x}(\cdot))$, where $x$ solves (1) with $\lambda = 0$. More precisely, we have the following theorem:

**Theorem 6.1.** Suppose: (1) that the function

$$D(z) = \int_S^T L\left(t, \int_S^t z(\alpha)\, d\alpha + X, z(t)\right) dt + \int_S^T f\left(t, \int_S^t z(\alpha)\, d\alpha + X, z(t)\right) dz(t)$$
$$+ \phi\left(\int_S^T z(\alpha)\, d\alpha + X, z(T)\right) - \frac{1}{2}\int_S^T [\dot{z}(t)]^2\, dt \tag{91}$$

has a unique maximizing function, $\bar{x}(t)$, over the set of all $z$'s in $C[S, Y, T]$ which are absolutely continuous and whose derivatives have the property that $\int_S^T [\dot{z}(t)]^2\, dt < \infty$.

(2) That $f(t, X, Y)$, $L(t, X, Y)$, and $\phi(X, Y)$ have $n + 2 (n \geqslant 0)$ continuous $X$ and $Y$ derivatives for all $t \in [S, T]$, $X$ and $Y$ real.

(3) That the function $F(z)$ has a Volterra series expansion about $\bar{x}(\cdot)$ out to $n + 1$ terms. (See Volterra[42] for an exposition of Volterra series.)

(4) That the matrix riccati equations

$$\dot{C}_4(t) = \left[\frac{\partial}{\partial Y} f\left(t \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right) - C_4(t)\right]^2 + \frac{\partial^2}{\partial Y^2} L\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right)$$
$$+ \frac{\partial}{\partial Y^2} f\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right) d\bar{x}(t) - 2C_5(t)$$

$$\dot{C}_5(t) = \left[\frac{\partial}{\partial Y} f\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right) - C_4(t)\right]\left[\frac{\partial}{\partial X} f\left(t, \int_S^t \bar{x}(t) + X, x(t)\right) - C_5(t)\right]$$
$$+ \frac{\partial}{\partial X\, \partial Y} L\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right)$$
$$+ \frac{\partial}{\partial X\, \partial Y} f\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right) d\bar{x}(t) - C_6(t)$$

$$\dot{C}_6(t) = \left[\frac{\partial}{\partial X} f\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right) - C_5(t)\right]^2 + \frac{\partial^2}{\partial X^2} L\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right)$$
$$+ \frac{\partial^2}{\partial X^2} f\left(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\right) d\bar{x}(t)$$

with terminal conditions

$$C_4(T) = -\frac{\partial}{\partial Y^2}\phi\left(\int_S^T \bar{x}(\alpha)\, d\alpha + X, \bar{x}(T)\right)$$
$$C_5(T) = -\frac{\partial}{\partial X\, \partial Y}\phi\left(\int_S^T \bar{x}(\alpha)\, d\alpha + X, \bar{x}(T)\right)$$
$$C_6(T) = -\frac{\partial}{\partial X^2}\phi\left(\int_S^T \bar{x}(\alpha)\, d\alpha + X, \bar{x}(T)\right)$$

have a solution.[10] Let

$$M(r) = \frac{1}{2}\sum_{j=0}^{2}\binom{2}{j}\int_S^T \frac{\partial^2}{\partial X^{2-j}\,\partial Y^j} L\left(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\right)$$
$$\times \left[\int_S^t r(\alpha)\,d\alpha\right]^{2-j} [r(t)]^j\,dt$$
$$+\frac{1}{2}\sum_{j=0}^{2}\binom{2}{j}\int_S^T \frac{\partial^2}{\partial X^{2-j}\,\partial Y^j} f\left(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\right)$$
$$\times \left[\int_S^t r(\alpha)\,d\alpha\right]^{2-j} [r(t)]^j\,d\bar{x}(t)$$
$$+\sum_{j=0}^{1}\binom{1}{j}\int_S^T \frac{\partial}{\partial X^{1-j}\,\partial Y^j} f\left(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, x(t)\right)$$
$$\times \left[\int_S^t r(\alpha)\,d\alpha\right]^{1-j} [r(t)]^j\,dr(t)$$
$$+\frac{1}{2}\sum_{j=0}^{2}\binom{2}{j}\frac{\partial^2}{\partial X^{2-j}\,\partial Y^j}\,\phi\left(\int_S^T \bar{x}(\alpha)\,d\alpha + X, \bar{x}(T)\right)$$
$$\left[\int_S^T r(\alpha)\,d\alpha\right]^{2-j} [r(T)]^j$$

for $r \in C[S, 0, T]$. Then it follows that[11]

$$E_z^w\Big\{F(\lambda z)\exp\Big[\lambda^{-2}\int_S^T L\left(t, \lambda\int_S^t z(\alpha)\,d\alpha + X, \lambda z(t)\right)dt$$
$$+\lambda^{-1}\int_S^T f\left(t, \lambda\int_S^t z(\alpha)\,d\alpha + X, \lambda z(t)\right)dz(t)$$
$$+\lambda^{-2}\,\phi\left(\lambda\int_S^T z(\alpha)\,d\alpha + X, \lambda z(T)\right)\Big]\Big|\, z(S) = Y\lambda^{-1}\Big\}$$
$$= \exp[b\lambda^{-2}](\Gamma_0 + \lambda^2\Gamma_2 + \lambda^4\Gamma_4 + \cdots + \lambda^n\Gamma_n + O(\lambda^{n+1})) \quad \text{as } \lambda \to 0 \quad (92)$$

$b$ is the maximum value of the function defined in (1) above, $\Gamma_0$ is $F(\bar{x})\exp[-\int_S^T C_4(t)\,dt]$, and the $\Gamma_i$ are Wiener integrals which can be evaluated numerically.

*Proof.*[12] We write the Wiener integral (92) in the flat-integral form and expand out all its terms in Volterra series about $\bar{x}$, the maximizing function.

---

[10] All integrals involving $d\bar{x}(t)$ or $\bar{x}(t)$ are interpreted as ordinary Stieltjes integrals. This is possible since, by hypothesis (1), $\dot{x}(t)$ exists in the ordinary sense.

[11] See also Pincus[(34)] and Varadhan.[(40)]

[12] The flat integral is used in this proof because of its intuitive appeal. The proof is, however, entirely rigorous.

Thus,

$$
\begin{aligned}
E_z^w \Big\{ & F(\lambda z) \exp \Big[ \frac{1}{\lambda^2} \int_S^T L\Big(t, \lambda \int_S^t z(\alpha)\, d\alpha + X, \lambda z(t)\Big)\, dt \\
& + \frac{1}{\lambda} \int_S^T f\Big(t, \lambda \int_S^t z(\alpha)\, d\alpha + X, \lambda z(t)\Big)\, dz(t) \\
& + \frac{1}{\lambda^2} \phi\Big(\lambda \int_S^T z(\alpha)\, d\alpha + X, \lambda z(T)\Big)\Big] \Big| z(S) = \frac{1}{\lambda} Y \Big\} \\
= \int & \Big[ \sum_{i=0}^{n} F^{(i)}(\bar{x})(\lambda z - \bar{x})^i + O(\lambda z - \bar{x})^{n+1} \Big] \exp \Big\{ \frac{1}{\lambda^2} \int_S^T \sum_{i=0}^{n+2} \sum_{j=0}^{i} \frac{1}{i!} \binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\, \partial Y^j} \\
& \times L\Big(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\Big) \Big[\lambda \int_S^t z(\alpha)\, d\alpha - \int_S^t \bar{x}(\alpha)\, d\alpha\Big]^{i-j} [\lambda z(t) - \bar{x}(t)]^j\, dt \\
& + \frac{1}{\lambda^2} O(\lambda z - \bar{x})^{n+3} + \frac{1}{\lambda} \int_S^T \sum_{i=0}^{n+2} \sum_{j=0}^{i} \frac{1}{i!} \binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\, \partial Y^j} f\Big(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\Big) \\
& \times \Big[\lambda \int_S^t z(\alpha)\, d\alpha - \int_S^t \bar{x}(\alpha)\, d\alpha\Big]^{i-j} [\lambda z(t) - \bar{x}(t)]^j\, dz(t) + \frac{1}{\lambda} O(\lambda z - \bar{x})^{n+3} \\
& + \frac{1}{\lambda^2} \sum_{i=0}^{n+2} \sum_{j=0}^{i} \frac{1}{i!} \binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\, \partial Y^j} \phi\Big(\int_S^T \bar{x}(\alpha)\, d\alpha + X, \bar{x}(T)\Big) \\
& \times \Big[\lambda \int_S^T z(\alpha)\, d\alpha - \int_S^T \bar{x}(\alpha)\, d\alpha\Big]^{i-j} \\
& \times [\lambda z(T) - \bar{x}(T)]^j + \frac{1}{\lambda^2} O(\lambda z - \bar{x})^{n+3} - \frac{1}{2} \int_S^T [\dot{z}(t)]^2\, dt \Big\}\, \delta z
\end{aligned}
$$

Change variables now by the transformation theorem 2.1, letting $\lambda z - \bar{x} = \lambda r$; then $z = (\bar{x} + \lambda r)/\lambda$ and $r(S) = 0$. Thus, substituting in the above, we find that (92) is equal to

$$
\begin{aligned}
\int & \Big[\Big( \sum_{i=0}^{n} \lambda^i F^{(i)}(\bar{x})\, r^i + O(\lambda r)^{n+1} \Big] \\
& \times \exp \Big\{ \frac{1}{\lambda^2} \int_S^T \sum_{i=0}^{n+2} \lambda^i \sum_{j=0}^{i} \frac{1}{i!} \binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\, \partial Y^j} L\Big(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\Big) \\
& \times \Big[\int_S^t r(\alpha)\, d\alpha\Big]^{i-j} [r(t)]^j\, dt + \frac{1}{\lambda^2} O(\lambda r)^{n+3} \\
& + \frac{1}{\lambda^2} \int_S^T \sum_{i=0}^{n+2} \lambda^i \sum_{j=0}^{i} \frac{1}{i!} \binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\, \partial Y^j} \\
& \times f\Big(t, \int_S^t \bar{x}(\alpha)\, d\alpha + X, \bar{x}(t)\Big) \Big[\int_S^t r(\alpha)\, d\alpha\Big]^{i-j} [r(t)]^j\, d\bar{x}(t)
\end{aligned}
$$

$$+\frac{1}{\lambda}\int_S^T \sum_{i=0}^{n+2} \lambda^i \sum_{j=0}^{i} \frac{1}{i!}\binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j}$$

$$\times f\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\Big[\int_S^t r(\alpha)\,d\alpha\Big]^{i-j}[r(t)]^j\,dr(t)$$

$$+\frac{1}{\lambda^2}\,O(\lambda r)^{n+3} + \frac{1}{\lambda}\,O(\lambda r)^{n+3}$$

$$+\frac{1}{\lambda^2}\sum_{i=0}^{n+2} \lambda^i \sum_{j=0}^{i} \frac{1}{i!}\binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j}\,\phi\Big(\int_S^T \bar{x}(\alpha)\,d\alpha + X, \bar{x}(T)\Big)\Big[\int_S^T r(\alpha)\,d\alpha\Big]^{i-j}[r(T)]^j$$

$$+\frac{1}{\lambda^2}\,O(\lambda^{n+3}) - \frac{1}{2\lambda^2}\int_S^T [\dot{\bar{x}}(t)]^2 - \frac{1}{\lambda}\int_S^T \dot{\bar{x}}(t)\,dr(t) - \frac{1}{2}\int_S^T [\dot{r}(t)]^2\,dt\Big\}\,\delta r$$

We now collect all terms in the exponent whose coefficients are $1/\lambda^2$ or $1/\lambda$. The term whose coefficient is $1/\lambda^2$ is

$$\int_S^T L\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\,dt + \int_S^T f\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\,d\bar{x}(t)$$
$$+\phi\Big(\int_S^T \bar{x}(\alpha)\,d\alpha + X, \bar{x}(T)\Big) - \frac{1}{2}\int_S^T [\dot{\bar{x}}(t)]^2\,dt. \tag{93}$$

But (93) is exactly (91) with $\bar{x}$ substituted for $z$. Since, by definition, $\bar{x}$ maximizes (91), it follows that (93) is the maximum value of (91) which was defined in the hypothesis of this theorem to be $b$.

The term whose coefficient is $1/\lambda$ is

$$\int_S^T L_X\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X\ \bar{x}(t)\Big)\int_S^t r(\alpha)\,dt + \int_S^t L_Y\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\,r(t)\,dt$$
$$+\int_S^T f_X\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\Big[\int_S^t r(\alpha)\,d\alpha\Big]\,d\bar{x}(t)$$
$$+\int_S^t f_Y\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\,r(t)\,dx(t)$$
$$+\int_S^t f\Big(t, \int_S^t \bar{x}(\alpha)\,d\alpha + X, \bar{x}(t)\Big)\,dr(t)$$
$$+\phi_X\Big(\int_S^T \bar{x}(\alpha)\,d\alpha + X, x(T)\Big)\int_S^T r(\alpha)\,d\alpha + \phi_Y\Big(\int_S^T \bar{x}(\alpha)\,d\alpha + X, \bar{x}(T)\Big)\,r(T)$$
$$-\int_S^T \dot{\bar{x}}(t)\,dr(t) \tag{94}$$

Since $\bar{x}$ maximizes (91), it follows that, if $\bar{x} + \lambda r$ [where $r$ is in the class of functions defined by assumption (1) of this theorem] is substituted for $z$ in (91) and if (91) is then differentiated with respect to $\lambda$, the result is zero. This is the standard procedure in the calculus of variations and is the way usually used to derive the Euler equation of the calculus of variations.

If this substitution and differentiation is actually carried out, then one finds that the derivative is exactly (94); it therefore follows that (94) is zero for any $r$ in the class of functions defined by assumption (1) of this theorem. It can be shown that one can go from this result to the result that (94) vanishes for almost all $r \in C[S, 0, T]$.

It therefore follows, making the substitutions for the $1/\lambda^2$ and $1/\lambda$ terms, that (91) is equal to

$$\exp\left[\frac{b}{\lambda^2}\right] \int \left[\sum_{i=0}^{h} \lambda^i F^{(i)}(\bar{x})\, r^i + O(r\lambda)^{n+1}\right]$$
$$\times \exp\left\{\sum_{i=2}^{n+2} \frac{1}{i!} \lambda^{i-2} \sum_{j=0}^{i} \binom{i}{j} \int_S^T \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j} L(t)\left[\int_S^t r(\alpha)\, d\alpha\right]^{i-j} [r(t)]^j\, dt + O(r\lambda)^{n+1}\right.$$
$$+ \sum_{i=2}^{n+2} \frac{1}{i!} \lambda^{i-2} \sum_{j=0}^{i} \binom{i}{j} \int_S^T \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j} f(t)\left[\int_S^t r(\alpha)\, d\alpha\right]^{i-j} [r(t)]^j\, d\bar{x}(t) + O(r\lambda)^{n+1}$$
$$+ \sum_{i=1}^{n+2} \frac{1}{i!} \lambda^{i-1} \sum_{j=0}^{i} \binom{i}{j} \int_S^T \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j} f(t)\left[\int_S^t r(\alpha)\, d\alpha\right]^{i-j} [r(t)]^j\, dr(t) + O(r\lambda)^{n+2}$$
$$+ \sum_{i=2}^{n+2} \frac{1}{i!} \lambda^{i-2} \sum_{j=0}^{i} \binom{i}{j} \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j} \phi(T)\left[\int_S^T r(\alpha)\, d\alpha\right]^{i-j} [r(T)]^j + O(r\lambda)^{n+1}$$
$$\left. - \frac{1}{2} \int_S^T [\dot{r}(t)]^2\, dt\right\} \Bigg|\, \delta r \tag{95}$$

In (95), the arguments of $L$, $f$, and $\lambda$ depend on $t$, $T$, and $\bar{x}$ only. They do not depend on the variable of integration $r$. Thus, we are justified in writing $L$, for example, as a function of $t$ only.

The partial derivative operators in front of these functions mean: take the partial derivatives first, and then substitute in $\bar{x}$.

It can be seen that there are no longer any negative powers of $\lambda$ in (95). We now separate the exponent in (95) into two parts. The first contains the terms which do not depend on $\lambda$, the second contains all the other terms.

It can be seen, using the Taylor expansion of $\exp[X]$ around zero, that

$$\exp[X] = \sum_{k=0}^{n-1} (X^k/k!) + R_n(X) \tag{96}$$

where

$$|\, R_n(x)| \leqslant (X^n/n!) \exp[X].$$

The exponential term containing terms which do not depend on $\lambda$ is left alone. This, it can be seen, is exactly $\exp\{M(r) - \frac{1}{2}\int_S^T [\dot{r}(t)]^2\, dt\}$, where $M(r)$ was defined in hypotheses (4) of this theorem. The second exponential term is expanded out to $n + 2$ terms, using (96). Since every polynomial in this exponent contains $\lambda$ at least to the first power, it follows that the $R_n(x)$ of (96) will have a factor of $\lambda^{n+1}$ in front of it. It can be shown that the remainder of this term multiplied by $\exp[M(r)]$ and integrated with respect to Wiener measure is finite, using assumption (4) of this theorem and Lemma 6.1. It can also be seen that (95) can be written as a Wiener integral, due to the presence of the last term, $-\frac{1}{2}\int_S^T [\dot{r}(t)]^2\,dt$, in the exponent.

Putting all of the above together, we finally arrive at the fact that (92) is equal to

$$\exp\left[\frac{b}{\lambda^2}\right] E_r^w \Bigg\{\left[\sum_{i=0}^{n} \lambda^i F^{(i)}(\bar{x})\, r^i\right]$$

$$\times\left[\sum_k^n \frac{1}{k!}\left(\sum_{i=3}^{n+2}\frac{1}{i!}\lambda^{i-2}\sum_{j=0}^{i}\binom{i}{j}\int_S^T \frac{\partial^i}{\partial X^{i-j}\,\partial Y^j} L(t)\left[\int_S^t r(\alpha)\,d\alpha\right]^{i-j}[r(t)]^j\,dt\right.\right.$$

$$+\sum_{i=3}^{n+2}\frac{1}{i!}\lambda^{i-2}\sum_{j=0}^{i}\binom{i}{j}\int_S^T\frac{\partial^i}{\partial X^{i-j}\,\partial Y^j}f(t)\left[\int_S^t r(\alpha)\,d\alpha\right]^{i-j}[r(t)]^j\,dt$$

$$+\sum_{i=2}^{n+3}\frac{1}{i!}\lambda^{i-1}\sum_{j}^{i}\binom{i}{j}\int_S^T\frac{\partial^i}{\partial X^{i-j}\,\partial Y^j}f(t)\left[\int_S^t r(t)\,d\alpha\right]^{i-j}[r(t)]^j\,dr(t)$$

$$\left.\left.+\sum_{i=3}^{n+2}\frac{1}{i!}\lambda^{i-2}\sum_{j=0}^{i}\binom{i}{j}\frac{\partial^i}{\partial X^{i-j}\,\partial Y^j}\phi(T)\left[\left(\int_S^T r(\alpha)\,d\alpha\right]^{i-j}[r(T)]^j + O(r\lambda^{n+1})\right)^k\right]$$

$$\times \exp[M(r)]\Big|\, r(S)=0\Bigg\} + O(\lambda^{n+1})$$

This expression is now expanded out over each of the sums $i, j$, and $k$ in powers of $\lambda$. The first term, corresponding to the zeroth power of $\lambda$, is

$$\exp[b/\lambda^2]\, E_r^w\{F(\bar{x})\exp[M(r)] \,|\, r(S)=0\} = \exp[b/\lambda^2]\, F(\bar{x})\, E_r^w\{\exp[M(r)] \,|\, r(S)=0\}$$

which, by Lemma 6.1, is $\exp[(b/\lambda^2) - \frac{1}{2}\int_S^T C_4(t)\,dt]\,F(\bar{x})$, as was to be shown.

The other terms are all Wiener integrals involving $\exp[M(r)]$ multiplying integrals involving powers of $\int_S^t r(\alpha)\,d\alpha$, $r(t)$, and $dr(t)$ and terms of the form $r(T)$ and $\int_S^T r(\alpha)\,d\alpha$. The first step in evaluating Wiener integrals of this type is to apply Lemma 6.1 to eliminate the $\exp[M(r)]$ term. All the coefficients are now Wiener integrals simply involving powers of Ito integrals, where the integrands of the Ito integrals are functions only of time. Lemma 2.2 shows how to evaluate this type of Wiener integral. It can be seen that, if $\lambda$ appears to an odd power in front of the $\Gamma$'s, then the Wiener integral involves $r$ to an odd power. Since, if, in lemma 2.2, $z$ appears to an odd power, the integral vanishes, it follows that all $\Gamma$'s which have $\lambda$ to an odd power multiplying them vanish; thus, the series given by this theorem involves only even powers of $\lambda$.

The reader has undoubtedly gotten the impression that the procedure advocated above is complicated—computationwise, it is, of course. However, it must be remembered that it brings within computational grasp the solutions of a large number of problems in engineering and physics. The complexity arises mostly in keeping track of a large number of loose indices which are floating around, which is not insurmountable. The differential equations to be solved are the Euler equation for the functions (91), the riccati equation of Lemma 6.1 and the linear ordinary differential equations to obtain the $\Phi$ of Lemma 6.1. Once these equations have been solved, by numerical means, if necessary, then all the $\Gamma$'s can be computed by quadratures.

The approximation here is also a physically very relevant one, because, very often, one is interested in how a small amount of noise will perturb an otherwise stable deterministic system.

This expansion is also very relevant in the quantum-mechanical case. It can even be made rigorous if one uses Cameron and Storvick's result[5] instead of the transformation theorem 2.1. While Lemma 2.2 does not, as yet, have an analog for Feynman integrals, Lemma 6.1 can be rigorously proved using present techniques.

Theorem 6.1, in this case, shows that, as $h \to 0$ (Planck's constant), the quantum theory converges to Newtonian mechanics. The expansion is also much better than all others that the author has seen, due to the fact that it is in terms of powers of $h$ and not powers of $1/h$. The convergence is therefore extremely rapid, since $h$ is very small.

Erdelyi[10] shows that asymptotic expansions are unique. *Thus, there are no expansions for the various functions discussed in this paper in terms of a small variance parameter other than the one given.*

Theorem 6.1 can be modified to handle the case of boundary conditions or of conditioning of the type given by Theorem 5.1.

For other ways of approximating Wiener integrals see Feynman and Hibbs,[11] Finlayson,[12] or Cameron and Martin.[3]

## REFERENCES


1. A. Bryson, Jr., and M. Frazier, Smoothing for linear and nonlinear systems, *Proc. of the Optimum-System Synthesis Conference*, Tech. Doc. Rep. No. ASD-TRD-63-119, Flight Control Lab., ASD AFSC Wright Patterson AFB, Ohio, February 1963.
2. R. H. Cameron and W. T. Martin, Evaluation of various Wiener integrals by the use of Sturm–Liouville differential equations, *Bull. Am. Math. Soc.* **51**, 73–89 (1945).
3. R. H. Cameron and W. T. Martin, The orthogonal development of nonlinear functions in series of Fourier–Hermite functionals, *Ann. Math.* **48**, 385–389 (1947).
4. R. H. Cameron, Nonlinear Volterra functional equations, *J. d'Analyse Math.* **5**, 470–510 (1957).
5. R. H. Cameron and D. A. Storvick, A translation theorem for analytic Feynman integrals, *Trans. Am. Math. Soc.* **125** (1), 1–6 (Oct. 1966).
6. S. Chandrasekhar, Stochastic problems in physics and astronomy, *Rev. Mod. Phys.* **15** (1) (1943); reprinted in N. Wax, ed., *Selected Papers on Noise and Stochastic Processes* (Dover Publications, New York, 1954).
7. M. Donsker, On function space integrals, in *Analysis in Function Space*, edited by W. T. Martin and A. Segal (M.I.T. Press, Cambridge, Mass., 1964).
8. J. L. Doob, *Stochastic Processes* (John Wiley and Sons, New York, 1953).
9. E. B. Dynkin, *Markov Processes*, Vols. I and II (Academic Press, New York, 1965).
10. A. Erdelyi, *Asymptotic Expansions*, (Dover Publications, New York, 1956).
11. R. T. Feynman and R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill Book Co., New York, 1965).
12. H. C. Finlayson, An approximation method for Wiener integrals of certain unbounded functionals, *Proc. Am. Math. Soc.* **18** (5), 798–803 (Oct. 1967).
13. B. Friedland and I. Bernstein, Estimation of the state of a nonlinear process in the presence of nongaussian noise and disturbances, *J. Franklin Inst.* **281** (6), 455–480 (June 1966).
14. B. Friedland *et al.*, Stability of attitude control systems acted upon by random perturbations, NASA Contractor Report CR-731.
15. I. M. Gelfand and A. M. Yaglom, Function space integration, *J. Math. Phys.* **1**, 48–69 (1960).

16. J. Gimibre, Reduced density matrices of quantum gases I, II, and III, *J. Math. Phys.* **June 1965,** 238–251, 252–262, 1432–1446.
17. I. V. Girsanov, On the transformation of a class of stochastic processes by means of an absolutely continuous change of measure (in Russian), *Teok. Beroyatnost, i Primen.* **1960,** 314–330. (There is an English translation in the Journal *Theory of Probability and its Applications.*)
18. L. Gross, Classical analysis on a Hilbert space, in *Analysis in Function Space*, edited by W. T. Martin and A. Segal (M.I.T. Press, Cambridge, Mass., 1964).
19. P. R. Halmos, *Measure Theory* (D. Van Nostrand Co., Princeton, N.J. 1950).
20. E. Hopf, On the partial differential equation $\mu_T + \mu\mu_x = \eta\mu_{xx}$, *Commun. Pure Appl. Math.* **3,** 201–230 (1950).
21. K. Ito, *On Stochastic Differential Equations, Mem. Am. Math. Soc.* **4** (1951).
22. J. H. Jeans, *The Dynamical Theory of Gases* (Dover Publications, New York, 1954).
23. M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, London, 1959).
24. M. Kac, Wiener and integration in function spaces, *Bull. Am. Math. Soc.* **75** (1/II) (Jan. 1966).
25. R. E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng. (ASME Trans.)* **82D**, 35–45 (1960).
26. R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory, *J. Basic Eng. (ASME Trans.)* **83D**, 95–108 (1961).
27. J. G. Kirkwood, The statistical mechanical theory of transport process I, *J. Chem. Phys.* **14,** 180 (1946).
28. J. G. Kirkwood, *et al.*, The statistical mechanical theory of transport processes III, *J. Chem. Phys.* **17** (10), 988–994 (1949).
29. H. J. Kushner, Dynamical equations for nonlinear filtering," *J. Diff. Equations,* 3(2), 179–190 (1967).
30. D. K. C. MacDonald, *Noise and Fluctuations, An Introduction* (John Wiley, and Sons) New York, 1962).
31. W. T. Martin, and A. Segal, *Analysis in Function Space* (M.I.T. Press, Cambridge, Mass., 1964).
32. E. Nelson, *Dynamical Theories of Brownian Motion* (Princeton Univ. Press, Princeton, N.J. 1967).
33. L. Onsager and S. Machlup, Fluctuations and irreversible processes, *Phys. Rev.* **91** (6), 1505–1510 (1953).
34. M. Pincus, Gaussian processes and Hammerstein integral equations, *Trans. Am. Math. Soc.* **134** (2), 193–214 (1968).
35. M. Schilder, Some asymptotic formulas for Wiener integrals, *Trans. Am. Math. Soc.* **25** (1), 63–85 (Oct. 1966).
36. M. Schilder, A Monte Carlo method for analyzing non linear circuits, *Intern. J. Control* **5** (2), 131–134 (1967).
37. M. Schilder, Expected values of functionals with respect to the Ito distribution, *Pacific J. Math.* **25** (2), 371–380 (1968).
38. M. Schilder, The Wiener integral calculus and some of its applications to engineering and physics, Research Report RC 68-1, General Precision Systems, Inc., Little Falls, N.J., February 2, 1968.
39. M. Schilder, Applications of Wiener integrals to fluid dynamics, A report submitted to the Foundation for the Advancement of Graduate Study in Engineering (unpublished).
40. S. R. S. Varadhan, Asymptotic probabilities and differential equations, *Commun. Pure Appl. Math.* **XIX** (3), 261–281 (Aug. 1966).
41. A. V. Skorokhod, Stochastic equations for diffusion processes in a bounded region, *Theory of Probability and Its Applications* **VI** (3), 1961.
42. V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations* (Dover Publications, New York, 1959).
43. N. Wiener, *Non Linear Problems in Random Theory* (M.I.T. Press, Cambridge, Mass., 1958).
44. N. Wiener, A. Siegel, B. Rankin, and W. T. Martin, *Differential Space, Quantum Systems and Prediction* (M.I.T. Press, Cambridge, Mass., 1966).
45. M. Wonham, Stochastic problems in optimal control, IEEE Convention Record, 1963, Part 4.
46. M. Wonham, Lecture notes on stochastic control, Lecture Notes 67-2, Division of Applied Mathematics, Brown University, February 1967.