RICERCHE DI AUTOMATICA
Vol. 10, No. 2, December 1979
Printed in Italy

ON THE ANALOGY BETWEEN MATHEMATICAL PROBLEMS OF NON-LINEAR FILTERING AND QUANTUM PHYSICS

S. K. Mitter

Abstract. In this paper we investigate the relationship between the mathematical structures underlying quantum physics and non-linear filtering theory.

1. Introduction.

The main thesis of this paper is that there are striking similarities between the mathematical problems of stochastic system theory, notably linear and nonlinear filtering theory, and mathematical developments underlying quantum mechanics and quantum field theory. Thus the mathematical developments of the past thirty years in functional analysis, Lie groups and Lie algebras, group representations and probabilistic methods of quantum theory can serve as a guide and indicator to search for an appropriate theory of stochastic systems. In the current state of development of linear and non-linear filtering theory, it is best to proceed by «analogy» and with care, since «unitarity» which plays such an important part in quantum mechanics and quantum field theory is not necessarily relevant to linear and non-linear filtering theory. The partial differential equations that arise in quantum

---

Manuscript received February 28, 1980.
The Author is with the Department of Electrical Engineering and Computer Science, and the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 U. S. A.
This research has been supported by the Air Force Office of Scientific Research under Grants AFOSR 77-3321, AFOSR 77-3281, and AFOSR 77-3281 B.

theory are generally wave equations, whereas the partial differential equations arising in filtering theory are stochastic parabolic equations. Nevertheless the possibility of passing to a wave equation by appropriate analytic continuation from the parabolic equation, reminiscent of the current program in euclidean field theory, should not be overlooked.

To develop these ideas, it is best to begin with a reasonably general non-linear filtering problem:

Let $(\Omega, \mathscr{A}, P)$ be the underlying probability space and let x_t denote a scalar-valued diffusion process which is the solution of the following Itô stochastic differential equation:

$$dx_t = f(x_t)\,dt + g(x_t)\,dw_t, \tag{1.1}$$

where w_t is standard Brownian motion and f satisfies appropriate assumptions so that (1.1) has a unique solution in the sense of Itô. We shall refer to (1.1) as the physical process. Let

$$z_t = h(x_t) \tag{1.2}$$

with $z_t \in L^2(\Omega, \mathscr{A}, P)$, denote the signal process, and we observe the signal in the presence of another Wiener process

$$y_t = \int_0^t z_s\,ds + \eta_t. \tag{1.3}$$

We shall refer to (1.3) as the *observation* equation. We make the assumption that (x_t, η_t) are independent. Let $\mathscr{F}_t^y$ denote the σ-field generated by $\{y_s \mid 0 \le s \le t\}$. The problem of non-linear filtering is to recursively compute

$$E[\phi(x_t) \mid \mathscr{F}_t^y], \tag{1.4}$$

where ϕ say is a bounded, continuous function. $E[\phi(\cdot) \mid \mathscr{F}_t^y]$ denotes conditional expectation with respect to $\mathscr{F}_t^y$. The solution to this problem can be obtained by Functional Integration and the Cameron-

Martin-Girsanov formula. Define a new measure $\tilde{P}$ on $(\Omega, \mathscr{A})$ by the Cameron-Martin-Girsanov transformation:

$$\frac{dP}{d\tilde{P}} = \exp\left(\int_0^t z_s\,dy_s - \frac{1}{2}\int_0^t z_s^2\,ds\right). \tag{1.5}$$

Under this new measure, the probability distribution of x_t remains invariant, but y_t and x_t are independent and y_t is standard Brownian motion.

Let π_t denote the conditional expectation operator. Then a standard application of the theory of conditional expectations gives us:

$$\pi_t(\phi) = E[\phi(x_t) \mid \mathscr{F}_t^y] = \frac{\tilde{E}\left[\phi(x_t)\frac{dP}{d\tilde{P}} \middle| \mathscr{F}_t^y\right]}{\tilde{E}\left[\frac{dP}{d\tilde{P}} \middle| \mathscr{F}_t^y\right]} \tag{1.6}$$

where $\tilde{E}$ denotes expectation with respect to the $\tilde{P}$-measure. The mapping $\phi \mapsto \pi_t(\phi)$ is defined to be the *filter* for the stochastic system (1.1)-(1.3). π_t itself can be thought of as a measure-valued stochastic process.

For what follows it is convenient to rewrite (1.6) in the form of an *input-state-output relation*. For this purpose define

$$\rho_t: \phi \mapsto \tilde{E}\left[\phi(x_t)\frac{dP}{d\tilde{P}} \mid \mathscr{F}_t^y\right]. \tag{1.7}$$

Then (1.6) may be rewritten as

$$\pi_t(\phi) = \frac{\rho_t(\phi)}{\rho_t(1)}, \tag{1.8}$$

where 1 denotes the constant function 1 for all x.

In the above ρ_t is to be thought of as the *state* of the filter and equation (1.8) as the *state-output* relation of the filter.

It is instructive to view (1.7) and (1.8) in the light of the Gelfand-Naimark-Segal construction of states and representations. The functions $x_t \mapsto \phi(x_t)$ are the *observables* of the stochastic system (1.1). The formula (1.8) computes the conditional statistics of the observables ϕ given the observation program $\{y_s \mid 0 \le s \le t\}$. The analogy with the algebraic theory of quantum mechanics is striking, the notable difference being that the idea of computing conditional statistics based on an observation program seems to be absent in physics. This viewpoint turns out to be important in the definition of generalised observables for quantum systems as probability operator-valued measures.

It can be shown that ρ_t has density q_t which satisfies a stochastic partial differential equation

$$dq_t = L_0^* q_t\, dt + L_1 \phi_t\, dy_t, \tag{1.9}$$

where L_0^* is the formal adjoint of the diffusion process generator of (1.1) and L_1 is the operator: multiplication by $h(x)$. This is the Duncan-Mortensen-Zakai equation and is the fundamental equation of non-linear filtering. The density $q_t(z, y_0^t)$ has a representation as a function space integral

$$q_t(z, y_0^t) = \int_{\mathfrak{X}} \exp\left(\int_0^t h(x_s)\, dy_s - \frac{1}{2}\int_0^t h^2(x_s)\, ds\right) d\mu^z(x) \tag{1.10}$$

where the integration is over the path space of x with $x_0 = 0$ and $x_t = z$.

Now in equation (1.9) the observation $\{y_s \mid 0 \le s \le t\}$ is given to us, and if we rewrite equation (1.9) in the following suggestive form using Stratonovich calculus:

$$\frac{dq_t}{dt} = \left(L_0^* - \frac{1}{2} L_1^2\right) q_t + \dot{y}_t L_1 q_t, \tag{1.11}$$

then equation (1.11) is the *analogue of a euclidean* (imaginary time) *quantum field* with an external force defined by the observations. We say that the analogy is to euclidean quantum field (as opposed to euclidean quantum mechanics) since q_t is a measure-valued stochastic process. This paper is concerned with a systematic investigation of this point of view. In particular, in this picture the Kalman filter occupies the role of the free quantum field.

This paper is divided into six sections. In section 2 we discuss the relationship between Dirichlet and Schrödinger operators and show that it is possible to associate a stochastic process with the ground state measure of Schrödinger operators. Section 3 is devoted to the Bayes formula on non-linear filtering and describes the Feynman point of view for non-linear filtering. Section 4 is concerned with the construction of Fock space and discussing its role in non-linear filtering.

In secion 5 we argue that the Lie algebra of operators $\mathcal{L} = LA\left\{L_0^* - \frac{1}{2}L_1^2, L_1\right\}$ has an important role to play in non-linear filtering. In section 6 we discuss the question of representation of the filter.

This is a semi-expository paper and we have tried to concentrate on the ideas involved and emphasize a certain point of view. The ideas come from constructive quantum field theory as emphasized by Nelson and Segal, recent developments in system theory and the theory of group representations. It is our hope that this paper will go towards pointing in a small way the conceptual depth of stochastic system theory which is still in its infancy.

2. The Feynman-Kac Formula, Dirichlet and Schrödinger Operators.

2.1. Introduction.

Recent work on non-linear filtering theory, euclidean quantum field theory and the stochastic mechanics of Nelson make extensive use of the Feynman-Kac formula and the interplay between Schrödinger and Dirichlet operators. In particular, a stochastic process associated with the ground-state measure of Schrödinger operators turns out to be important. In this section we give an account of these ideas and the related theory of hypercontractive semigroups.

2.2. Preliminaries.

Throughout this paper integration with respect to Wiener measure, the Brownian bridge measure and the oscillator measure will be important. The relation of Wiener measure to the Laplacian and the semigroup generated by the Laplacian will also be important. For simplicity we shall be concerned with scalar-valued stochastic processes. There is no difficulty in generalising these ideas to vector-valued stochastic processes.

Let $\Omega = C(\mathbf{R}_+; R)$ and let $W_t: \omega \to W_t(\omega) = \omega_t$(*): $\Omega \to \mathbf{R}$ be the t^{th} coordinates function. We denote by τ_t the right shift on Ω. If we denote by $\mathcal{F}_t$ the smallest σ-field with respect to which $\{W_s \mid 0 \le s \le t\}$

(*) We shall also use the notation $x(t)$ for x_t when dealing with processes.

are measurable and by $\mathcal{F}=\sigma\{W_t \mid t\geq 0\}$, then $\mathcal{F}$ is the Burel σ-field of Ω and on $(\Omega, \mathcal{F})$ there is a unique probability measure W such that $W\{W_0=0\}=0$, the random variables $W_{t_1}-W_{t_0}, \dots, W_{t_n}-W_{t_{n-1}}$ are independent, Gaussian, zero mean, with variance t_j-t_{j-1} for the j^{th} increment. This measure is *Wiener-measure*. For each $x\in\mathbf{R}$, the probability measure W_x is defined by path translation:

$$W_x(B)=W(\theta_x(B)), \quad B\in\mathcal{F},$$

where

$$\theta_x(\omega)_t=x+\omega_t, \quad x\in\mathbf{R}, \quad \omega\in\Omega, \quad t\geq 0.$$

Alternatively one could have started with a Gaussian process $\{W_t \mid t\geq 0,\ W_0=0 \text{ a. s.}\}$ which is zero-mean and has variance $|t-s|$ and stationary independent increments and constructed Wiener measure as a path space measure on $C(\mathbf{R}_+; \mathbf{R})$. For a measurable real-valued function f, define

$$(P_t f)(x)=E_{W_x}(f(W_t)) \tag{2.1}$$

where E_{W_x} denotes expectation with respect to W_x. If the right hand side makes sense,

$$(P_t f)(x)=\int_{\mathbf{R}} f(y)\, p(t,y;\ 0,x)\, dy, \tag{2.2}$$

where $p(t,y;\ 0,x)$ is the transition density of Brownian motion

$$p(t,y;\ 0,x)=\frac{1}{(2\pi t)^{\frac{-1}{2}}}\exp\left(-\frac{(x-y)^2}{2t}\right), \quad t>0, \quad x,y\in\mathbf{R}. \tag{2.3}$$

It is known that for $p\in[1,\infty]$, $(P_t \mid t\geq 0)$ is a strongly continuous contraction semigroup on $L^p(\mathbf{R})$. Its infinitesimal generator is $-\frac{1}{2}\Delta$, where Δ is the Laplacian and $\mathcal{D}(\Delta)=H^p(\mathbf{R})=\{f\in L^p \mid Df\in L^p$ in the sense of distributions$\}$.

For certain applications we shall need to do integration with respect to conditional Wiener measure. From the properties of Wiener measure and the corresponding transition density of Brownian motion, for $0<t_1<\dots<t_m$

$$W_x\{\omega \mid \omega(t_1)\in A_1, \dots, \omega(t_m)\in A_m\}=$$

$$=\int_{\mathbf{R}} \prod_{i=1}^{m} p(t_i, x_i;\ t_{i-1}, x_{i-1})\, \chi_{A_i}(x_i)\, dx_i$$

where χ_A is the characteristic function of A, $x_0=x$, $t_0=0$. We can then construct a measure $W_{x,y;t}$ on the continuous paths on $[0,t]$ with $W_0=x$, $W_t=y$ with probability 1, $0<t_1<\dots<t_m<t$

$$W_{x,y;t}\{\omega \mid \omega(t_1)\in A_1, \dots, \omega(t_m)\in A_m\}=$$

$$=\int_{\mathbf{R}} \prod_{i=1}^{m} p(t_i, x_i;\ t_{i-1}, x_{i-1})\, \chi_{A_i}(x_i)\, p(t,y;\ t_m, x_m)\, dx_i$$

$W_{x,y;t}$ is called *conditional Wiener measure*. We have

$$\int f(\omega)\, dW_x=\int dy\left(\int f(\omega)\, dW_{x,y;t}\right) \tag{2.4}$$

for f which are functions of the values the path takes on $[0,t]$.

This measure corresponds to the *Brownian bridge process*, which is the Gaussian process β_s, say on $[0,1]$ with covariance

$$E(\beta_s\beta_t)=s(1-t), \quad 0\leq s\leq t\leq 1.$$

In terms of this process we can write (2.4) as

$$\int f(\omega_s)\, dW=\int f\left(\left(1-\frac{s}{t}\right)x+\frac{s}{t}y+\sqrt{t}\,\beta_{\frac{s}{t}}\right)p(t,y;\ 0,x)\, dx\, dy\, d\mu(\beta) \tag{2.5}$$

where $d\mu(\beta)$ represents the measure on the path space of β.

The two processes which we have dealt with are not stationary Gauss-Markov processes. For this reason it is often important to deal with the *oscillator process* which is the family of Gaussian random variables $\{q_t, -\infty<t<\infty\}$ with covariance $\frac{1}{2}\exp(-|t-s|)$. It is related to the operator $L_0=-\frac{1}{2}\frac{d^2}{dx^2}+\frac{1}{2}x^2-\frac{1}{2}$, the Harmonic Oscillator Hamiltonian.

2.3. The Feynman-Kac Formula.

Both in filtering theory and in quantum physics we are required to deal with an operator on $L^2(\mathbf{R})$, say

$$H = -\frac{1}{2}\Delta + V(x)$$

where Δ is the Laplacian and we need to compute $\exp(-tH)$, the semi-group generated by $-H$. The Feynman-Kac formula provides us with a representation of the operator $T_t = \exp(-tH)$ as a Wiener Integral. The first approach exemplified in the work of Nelson uses the Trotter Product formula to make sense of the formula but requires information on the self-adjointness of the operator H. The second approach proceeds by writing the formula first, proving that it represents a strongly continuous one-parameter semigroup and then its unique infinitesimal generator is computed. The second approach uses probabilistic techniques and turns out to be the more general one.

If the potential is bounded above then a general form of the Feynman-Kac formula can be obtained using martingale methods. For this purpose and for later use we introduce certain probabilistic machinery.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t, t \geq 0)$ a non-decreasing family of sub σ-fields of $\mathcal{F}$. Let $s \geq 0$ be arbitrary and a: $[s, \infty[\times \Omega \to \mathbf{R}$ and b: $[s, \infty[\times \Omega \to \mathbf{R}$ be bounded progressively measurable functions. For any $f \in C^2(\mathbf{R})$, define $A_t f$ for $t \geq s$ by

$$(A_t(q) f)(x) = \frac{1}{2} a(t, q) \frac{\partial^2 f}{\partial x^2}(x) + b(t, q) \frac{\partial f}{\partial x} \tag{2.6}$$

If $\xi(\cdot, \cdot)$ is any progressively measurable function from $[s, \infty[\times \Omega \to \mathbf{R}$, then $(A_t(q) f)(\xi(t, q))$ defines another progressively measurable function of t and q. Then it can be shown that for any ξ, a and b satisfying the above,

$$f(t, \xi(t)) - \int_s^t \left(\frac{\partial}{\partial u} + A_u\right) f(u, \xi(u))\, du \tag{2.7}$$

is a martingale relative to $(\Omega, \mathcal{F}_t, P)$ for all $t \geq s$, $\forall f \in C^{1,2}([0, \infty[\times, \mathbf{R})$. A process ξ satisfying (2.7) will be referred to as an *Itô process* with drift b and covariance a. The Feynman-Kac formula depends on

the following observation (1).

Let ξ be an Itô process relative to $(\Omega, \mathcal{F}_t, P)$ with drift b and covariance a. Then for any progressively measurable function V: $[s, \alpha[\times \Omega \to \mathbf{R}$ which is bounded below and $f \in C^{1,2}([0, \infty[\times, \mathbf{R}])$

$$e^{\int_s^t -V(u)\,du} f(t, \xi(t)) - \int_s^t \left(\frac{\partial f}{\partial u} + A_u f - V(u) f\right)(u, \xi(u))\, e^{\int_s^t -V(\sigma)\,d\sigma}\, du \tag{2.8}$$

is a martingale after time s.

Now define the operator

$$T_t^V f(s, x) = \int e^{\int_s^t -V(u)\,du} f(t, \xi(t))\, d\mu_{sx}(\xi) \tag{2.9}$$

where $d\mu_{s,x}(\xi)$ denotes the measure on the path-space of ξ given $\xi(s) = x$. Then we can show that T_t^V is a one-parameter strongly continuous semi-group and form (2.8) we conclude that its differential generator is $A_u + V(u)$. (2.9) is the Feynman-Kac formula. When we specialize to the case when $\xi(t)$ is Brownian motion, then $A_u = \frac{1}{2}\frac{d^2}{dx^2}$ and the formula reads setting $s = 0$.

$$(T_t^V f)(x) = E_{W_x}\left[\exp\left(\int_0^t -V(W(s))\,ds\right) f(W(s))\right]. \tag{2.9'}$$

Now this formula turns out to be valid for a much larger-class of unbounded potentials and the semigroup T_t^V has strong regularity properties.

The more general class of potentials we consider are: $(H\,1)$ V is a measurable function on $\mathbf{R}$ which can be written as $V = V_1 - V_2$ with $V_2 \geq 0$, $V_2 \in L^p(\mathbf{R})$ for $p > 1$ and V_1 measurable, bounded below, such that for each compact set K in $\mathbf{R}$, there exists $q(K) \geq 2$ s. t.

$$\int_K |V_2(x)|^q\, dx < \infty.$$

For this class of potentials, using the recent estimate

(1) This argument is an exercise in Stroock-Varadhan: *Multi-dimensional Diffusion Processes*, Springer-Verlag, New York, 1979.

$$\forall t>0,\ \forall r>0,\ K(r,t)\triangleq \sup_{x\in R} E_{W_x}\left(\exp\left(-r\int_0^t V(x_s)\,ds\right)\right)<\infty$$

and their refinements, the Feynman-Kac formula holds for the class of potentials governed by $(H\,1)$. If we define the operator

$$T_t(f)(x)=E_{W_x}\left(\exp\left(-\int_0^t V(W_s)\,ds\right)f(W_t)\right),$$

we have

THEOREM 2.1. (*Carmona*). *For any* $q\in[1,\infty[$ *and* $t>0$,

(i) T_t *is a bounded operator on* L^q *and* $\|T_t\|_q\leq K(1,t)$

(ii) *If* q' *denotes the conjugate exponent of* q, *then for* $f\in L^q$, $g\in L^{q'}$

$$\int_{\mathbf{R}}(T_t f)(x)\,g(x)\,dx=\int_{\mathbf{R}} f(x)\,(T_t g)(x)\,dx$$

(iii) T_t *is a strongly continuous semi-group on* L^q

(iv) *If* $\lim_{|x|\to\infty} V_1(x)=+\infty$, *then* T_t *is a compact operator on* L^q.

2.4. SCHRÖDINGER AND DIRICHLET OPERATORS.

Let $-H$ denote the infinitesimal generator of T_t. In quantum physics we often need to show that $-H$ is the self-adjoint extension of $-\frac{1}{2}\Delta+V$, the imaginary-time Schrödinger operator. If $V\geq 0$ this is well known. For the class of potentials given by $(H\,1)$ if $V\in L^q_{loc}$ then

PROPOSITION 2.2. (*Carmona*) $C_c^\infty(\mathbf{R})\subset\mathcal{D}(H)$ *and for* $f\in C_c^\infty(\mathbf{R})$, $Hf=-\frac{1}{2}(\Delta+V)f$ *where* H *is the infinitesimal generator of the semi-group* T_t *defined on* L^q.

In mathematical physics perturbed Hamiltonians are usually defined as sums of quadratic forms. We briefly review this.

Let $\mathcal{H}$ be a Hilbert space. A *quadratic form* is a map $q\colon Q(q)\times Q(q)\to\mathbf{C}$, where $Q(q)$ is a dense linear subset of $\mathcal{H}$, called the *form domain*, such that $q(\cdot,\psi)$ is conjugate linear and $q(\phi,\cdot)$ is linear for $\phi,\psi\in Q(q)$. If $q(\phi,\psi)=\overline{q(\psi,\phi)}$ we say that q is symmetric.

If $q(\phi,\phi)\geq 0\ \forall\phi\in Q$, q is called positive and if $q(\phi,\phi)\geq -M\|\phi\|^2$, for some M we say q is semibounded.

Let A be a self-adjoint operator on $\mathcal{H}$. By passing to the spectral representation of A, A is multiplication by x on $\bigoplus_{n=1}^{\infty} L^2(\mathbf{R},\mu_n)$. Let

$$Q(q)=\{\psi_n(x)\mid \sum_{n=1}^{N}\int_{-\infty}^{\infty} x\,\|\psi_n(x)\|^2\,d\mu_n<\infty\}\ (^1)$$

for all $\psi,\phi\in Q(q)$ define

$$q(\phi,\psi)=\sum_{n=1}^{N}\int_{-\infty}^{\infty} x\,\phi_n(x)\,\psi_n(x)\,d\mu_n$$

q is called the *quadratic form* associated with A and writing $Q(A)=Q(q)$, $Q(A)$ is called the form domain of A.

Let q be a semi-bounded quadratic form, $q(\psi,\psi)\geq -M\|\psi\|^2$. q is called *closed* if $Q(q)$ is complete under the norm

$$\|\psi\|_{+1}=\sqrt{q(\psi,\psi)+(M+1)\|\psi\|^2}.$$

If q is closed and $D\subset Q(q)$ is dense in $Q(q)$ in the $\|\cdot\|_{+1}$ norm then D is called a form core for q. The following fact is important.

THEOREM: *If q is a closed semibounded quadratic form, then q is the quadratic form of a unique self- adjoint operator.*

We now define Schrödinger operators as forms sums on $L^2(\mathbf{R})$. For $f,g\in C_c^2(\mathbf{R})$ define

$$\varepsilon_0(f,g)=\frac{1}{2}\int_{\mathbf{R}}\nabla f(x)\,\overline{\nabla g(x)}\,dx.$$

Integrating by parts

$$\varepsilon_0(f,g)=\left(-\frac{1}{2}\nabla f,g\right)_{L^2}$$

and hence ε_0 is closable. The form domain $Q(\varepsilon_0)$ of ε_0 is $H^1(\mathbf{R})$. Let V be a real-valued measurable function on $\mathbf{R}$ and set

(1) The μ_n's are the corresponding spectral measures.

$$Q(V)=\{f\in L^2 \mid \int_{\mathbf{R}} |V(x)|\,|f(x)|^2\,dx<+\infty\}.$$

On $Q(\varepsilon_0)\cap Q(V)$ define

$$\varepsilon(f,g)=\varepsilon_0(f,g)+(Vf,g)$$

where

$$(Vf,g)=((\mathrm{sgn}\,V)\,|V|^{1/2}f,\,|V|^{1/2}g).$$

If we assume $V=V_1-V_2$ s. t. V_1 bounded below and $V_1\in L^1_{\mathrm{loc.}}$, $V_2\in L^1$ then ε is the form of a unique bounded below self-adjoint operator H on L^2 with form domain $Q(H)=Q(\varepsilon_0)\cap Q(V)$.

Let μ be a Borel prob. measure on $\mathbf{R}$ which satisfies

$$d\mu(x)=e^{-2h(x)}\,dx \tag{2.10}$$

where h is a real-valued, locally bounded, absolutely continuous function with first order partial derivatives in $L^2_{\mathrm{loc.}}$. For all $f,g\in C_c^\infty(\mathbf{R})$, define

$$\delta(f,g)=\frac{1}{2}\int_{\mathbf{R}} \nabla f(x)\,\overline{\nabla g(x)}\,d\mu(x)\quad (^1) \tag{2.11}$$

Integrating by parts the right hand side, we get

$$\delta(f,g)=(Df,g)_\mu, \tag{2.12}$$

where $(\cdot,\cdot)_\mu$ denotes the $L^2(\mu)$-inner product and

$$Df=-\frac{1}{2}\Delta f+\nabla h\,\nabla f. \tag{2.13}$$

From (2.12) δ is given by a symmetric operator and hence is closable. Let $\bar\delta$ denote the closure. $\bar\delta$ is referred to as the Dirichlet form of μ and D the associated Dirichlet operator.

We can prove that

Proposition 2.3. (*Carmona*). *The form domain* $Q(\bar\delta)$ *is* $H^1(\mu)$.

We show the relationship between Dirichlet forms and Quadratic forms associated with Schrödinger Operators. Let the potential V satisfy hypothesis (H 1). We assume that inf spec $(H)=E$ is an eigenvalue and let ψ be the corresponding ground state eigenfunction (1).

(1) ∇f denotes the gradient of f and Δf the Laplacian (in this case one-dimensional).

Let $h=-\operatorname{Log}\psi$ and define the Borel probability measure μ on $\mathbf{R}$ by

$$d\mu(x)=e^{-2h(x)}\,dx.$$

Define the operator $\bar D$ by

$$\bar D=C\,(H-E)\,C^{-1} \tag{2.14}$$

where C is the unitary operator from $L^2(\mathbf{R},dx)\to L^2(\mu)$ defined by

$$C\phi=\psi^{-1}\phi,\ \phi\in L^2(\mathbf{R},dx). \tag{2.15}$$

$\bar D$ is a positive self-adjoint operator in $L^2(\mu)$, 0 is a simple eigenvalue and the constant function 1 is the corresponding eigenfunction.

In fact $\bar D$ is the unique positive self-adjoint operator associated to the closed positive bilinear form $\bar\delta$ corresponding to $\varepsilon-E$ (in the unitary equivalence C) and $C_c^\infty(\mathbf{R})$ is a core for $\bar\delta$.

Since ψ is bounded and locally bounded away from 0, h is bounded and locally bounded above. Since $\psi\in Q(\varepsilon_0)$, the first order partial derivatives of h are in $L^2_{\mathrm{loc.}}$, we can associate with μ a Dirichlet form δ and a Dirichlet Operator D. Now since δ is defined as the closure of a form whose domain is $C_c^\infty(\mathbf{R})$ and $C_c^\infty(\mathbf{R})$ is a form core for $\bar\delta$, one can show that $\delta=\bar\delta$ and $D=\bar D$ and hence the Dirichlet form and Dirichlet operator are unitarily equivalent to the Schrödinger form and Schrödinger operator.

2.5. Stochastic Processes Associated with Dirichlet Forms and Dirichlet Operators.

The unitary equivalence between Schrödinger and Dirichlet operators exhibited in the previous section has an important role in Nelson's stochastic mechanics and also in non-linear filtering theory.

We consider the Schrödinger operator $H=-\frac{1}{2}\Delta+V$ defined as a sum of quadratic forms on $L^2(\mathbf{R},dx)$. We assume that inf spectrum $(H)=E$ (assumed to be 0) is an eigenvalue and the corresponding

(1) Assume $\psi>0$.

eigenfunction $\psi \geq 0$ and normalized $\int\limits_{R} \psi(x)^2 dx = 1$. Define the Borel probability measure μ by $d\mu(x) = (\psi(x))^2 dx$. Let δ be the corresponding Dirichlet form and D the corresponding Dirichlet operator. Let $h = -\log\psi$.

We want to construct a Markov diffusion process which corresponds to a stochastic differential equation with drift $-\nabla h$. There is an obvious difficulty in interpreting the stochastic differential equation in a strong sense. But we can construct a weak solution using the measure transformation technique of Girsanov and the Feynman-Kac formula.

To do this we assume the ground state $\psi(x) > 0$, $\forall x$ (by choosing a representative from an equivalence class) and we also assume that $\psi(x) \to 0$ as $|x| \to \infty$.

Now we use the unitary equivalence between the Dirichlet operator D and the operator $H = -\frac{1}{2}\frac{d^2}{dx^2} + V(x)$ to conclude that

$$(e^{-tD} f)(x) = \psi(x)^{-1} E_{W_x}\left[\exp\left(-\int_0^t V(W_s)\right)\psi(W_s) f(W_s)\, ds\right] \tag{2.16}$$

where e^{-tD} is the semi-group generated by $-D$.

For each $t > 0$, consider

$$L_t = \psi(x)^{-1}\psi(W_t)\exp\left(-\int_0^t V(W_s)\, ds\right), \tag{2.17}$$

which is a random variable, positive W_x-a. s. and $E_{W_x}(L_t) = 1$. Hence, $(\Omega, \mathcal{F}_t, L_t)$ is a martingale, where $\mathcal{F}_t$ denotes the smallest σ-algebra for which the coordinate functions W_t-are measurable. Hence for each x, $P_x = L_t \cdot W_x|_{\mathcal{F}_t}$ is a probability measure on $(\Omega, \mathcal{F})$.

Now by explicit calculation, $V(x) = \frac{1}{2}(-\Delta h(x) + |\nabla h(x)|^2)$ and since $h = -\log\psi$, we get from (2.17)

$$L_t = \exp\left[-h(W_t) + h(x) - \frac{1}{2}\left(\int_0^t \Delta h(W_s)\, ds - \int_0^t |\nabla h(W_s)|^2 ds\right)\right] \tag{2.18}$$

The idea now is to apply the Itô-differential to the function h. There is a difficulty here because the function h does not have continuous partial derivatives upto order two. But h is a convex function and for these functions the Itô-differential rule can be extended to continuous functions with first partial derivatives in $L^2_{loc}(\mathbf{R}, dx)$ and second partial derivatives in $L^1_{loc}(\mathbf{R}, dx)$. Applying the generalized Itô-differential rule to (2.18), we get

$$L_t = \exp\left[-\int_0^t \nabla h(W_s)\, dW_s - \frac{1}{2}\int_0^t |\nabla h(W_s)|^2 ds\right].$$

Therefore by the Girsanov theorem, the process $B_t = W_t - W_0 + \int_0^t \nabla h(W_s)\, ds$ is a $(\mathcal{F}_t, P_x)$ standard Brownian and calling $W_t = X_t$, the stochastic process $(X_t \mid t \geq 0)$ considered on the probability space $(\Omega, \mathcal{F}_t, P_x)$ is a unique weak solution of

$$\begin{cases} dX_t = -\nabla h(X_t)\, dt + dB_t \\ X_0 = x, \text{ a. s.} \end{cases} \tag{2.19}$$

Furthermore the measure μ defined by $d\mu(x) = \psi(x)^2 dx$ is the unique finite invariant measure of (2.19).

From the construction of the probability measure P_x, we see that it has a transition density

$$\rho(t, y; 0, x) = \psi(y)\,\psi^{-1}(x)\, E_{W_x}\left[\exp - \int_0^t V(W_s)\, ds \mid W_t = y\right] p(t, y; 0, x) \tag{2.20}$$

where p is the transition density of Brownian motion.

We remark that it is this transition density we wish to compute since this corresponds to the fundamental solution of a parabolic partial differential equation. We shall see later the importance of (2.20) for non-linear filtering problems. For non-linear filtering problems the decay properties of this transition density are also of importance.

We investigate these matters now.

Firstly, we can check that for $\forall t > 0$, $\rho(t, y; 0, x)$ is a continuous function of the pair (x, y). In fact the following estimate holds:

$$
\begin{cases}
\exists \text{ constants } c_1, c_2 \text{ such that} \\
\forall t>0, \quad \forall (x, y) \in \mathbf{R} \times \mathbf{R}, \\
\qquad \rho(t, y; 0, x) \leq c_1 e^{c_2 t} \psi(y) \psi^{-1}(x) \quad p(t, y; 0, x).
\end{cases} \tag{2.21}
$$

If the semigroup

$$
(T_t f)(x) = E_{W_x}\left[f(W_t) \exp\left(-\int_0^t V(W_s)\, ds \right)\right]
$$

is compact (see Theorem 2.1) and its spectrum is strictly bounded away from zero, then the density $\rho(t, y; 0, x)$ satisfies:

$$
\begin{cases}
\exists \text{ positive constants } c_1, c_2 \text{ such that} \\
\sup_{y \in \mathbf{R}} |\rho(t, y; 0, x) - \psi(y)^2| \leq c_1 \psi^{-1}(x) e^{-c_2 t}, \quad \forall t>0, \quad \forall x \in \mathbf{R}.
\end{cases} \tag{2.22}
$$

2.6. Hypercontractive Semigroups.

In the previous sections we have seen that for a large class of potentials, the semigroup T_t^V has a negative infinitesimal generator which coincides with a self-adjoint extension of $-\frac{1}{2}\frac{d^2}{dx^2} + V(x)$. In this section we point out that the semigroups defined by the Feynman-Kac formula are often Hypercontractive semigroups. These semigroups have played an important role in constructive quantum field theory and are likely to play an equally important role in the theory of non-linear filtering.

We follow the notation, hypotheses (specially on the potential V) and the terminology of the previous sections.

We consider the Dirichlet semi-group $(e^{-tD}, t \geq 0)$ on $L^2(\mu)$ where $d\mu(x) = \psi(x)^2\, dx$.

Definition: The semigroup $(e^{-tD}, t \geq 0)$ is said to be *hypercontractive* if for some $t>0$ and some $r>2$, e^{-tD} is a bounded operator from $L^2(\mu)$ into $L^r(\mu)$.

From our point of view the best approach towards the question of hypercontractivity is via the approach of Gross using Logarithmic Sobolev inequalities.

Definition: The operator D is called a *Sobolev generator* if for some real constants $c>0$ and γ we have

$$
\int_{\mathbf{R}} |f|^2 \operatorname{Log}|f|\, d\mu \leq c(Df, f)_{L^2(\mu)} + \gamma \|f\|^2_{L^2(\mu)} + \|f\|^2_{L^2(\mu)} \operatorname{Log} \|f\|_{L^2(\mu)} \tag{2.23}
$$

for all $f \in \mathcal{D}(D)$. The constants c and γ are called the Sobolev coefficient and the local norm of D.

Logarithmic Sobolev Inequality (Gross).

If μ is a probability measure on $\mathbf{R}$ which satisfies:

$$
\int_{\mathbf{R}} |f|^2 \operatorname{Log}|f|\, d\mu \leq c \int_{\mathbf{R}} |\nabla f|^2\, d\mu + \gamma \|f\|^2_{L^2(\mu)} + \|f\|^2_{L^2(\mu)} \operatorname{Log} \|f\|_{L^2(\mu)} \tag{2.24}
$$

for some constants $c>0$, $\gamma \geq 0$ and $\forall f$ which are bounded functions with distributional first order derivative in $L^2(\mu)$, then $\forall r \in [2, +\infty[$

$$
\int_{\mathbf{R}} |f|^2 \operatorname{Log}|f|\, d\mu \leq c(r) \int_{\mathbf{R}} \nabla f \cdot \nabla f_r\, d\mu + \gamma \|f\|^r_{L^2(\mu)} +
$$
$$
+ \|f\|^r_{L^2(\mu)} \operatorname{Log} \|f\|_{L^2(\mu)} \tag{2.25}
$$

$\forall f$ as above, where $f_r = (\operatorname{sgn} f)|f|^{r-1}$ and $c(r) = \frac{cr}{2(r-1)}$.

The salient facts about hypercontractive semigroups and Sobolev generators are the following propositions due to Carmona:

Proposition 2.4. *$(e^{-tD} \mid t \geq 0)$ is a hypercontractive semigroup if and only if D is a Sobolev generator.*

Proposition 2.5. *D is a Sobolev generator with Sobolev coefficient c if and only if $-\operatorname{Log}\psi \leq cD + b$, the above interpreted as quadratic forms on $L^2(\mu)$, for some constant b.*

Consider potentials satisfying $(H\,1)$ which have the further property:

$$
\forall x \in \mathbf{R}, \quad a_1 P(x) + b_1 \leq V_1(x) \leq a_2 P(x) + b_2
$$

where $a_1 a_2>0, b_1, b_2 \in \mathbf{R}$ and P an even polynomial, and $\liminf_{|x|\to\infty} \cdot |x|^{-2} P(x) > 0$.

PROPOSITION 2.6. *Schrödinger operators with potentials of the above class generate Dirichlet semigroups which are hypercontractive.*

EXAMPLE:

(i) The operator $D = -\frac{1}{2}\frac{d^2}{dx^2} + x\frac{d}{dx}$ generates a hypercontractive semi-group e^{-tD} on $L^2\left(\mathbf{R}, \frac{1}{\sqrt{\pi}} e^{-x^2}\,dx\right)$.

(ii) Consider a stochastic differential equation:

$$dx_t = f(x_t)\,dt + dw_t.$$

Suppose that the Riccati equation

$$\frac{df}{dx} + f^2 = V(x),$$

where $V(x)$ is an even polynomial satisfying the hypotheses of this section such that a global solution exists for the Riccati equation. Then the generator of the associated diffusion process is in fact a Dirichlet operator and generates a hypercontractive semigroup on an appropriate $L^2(\mu)$-space.

NOTES AND REFERENCES FOR SECTION 2.

(i) For general references for this section, consult:

B. SIMON: *Functional Integration and Quantum Physics*, Academic Press, New York, 1979.

M. REED, B. SIMON: *Methods of Modern Mathematical Physics*, Vols. I and II, Academic Press, New York, 1972, 1975.

D. W. STROOCK, S. R. S. VARADHAN: *Multi-dimensional Diffusion Processes*, Springer-Verlag, Berlin, New York, 1979.

(ii) In this section we followed very closely:

R. CARMONA: *Regularity Properties of Schrödinger and Dirichlet Semigroups*, J. of Functional Analysis 33, 259-296 (1979).

(iii) The material in Section 2.5 is apparently partially new and uses the Girsanov transformation to obtain a weak solution of a stochastic differential equation. For a slightly different approach see:

H. EZAWA, J. R. KLAUDER, L. SHEPP: *A Path Space Picture for Feynman-Kac Averages*, Annals of Physics 88, 588-620 (1974).

(iv) For the generalized Itô Differential rule, see:

A. BENSOUSSAN, J. L. LIONS: *Applications des Inéquations Variationnelles en Côntrole Stochastique*, Dunod, Paris 1978.

(v) The reference to stochastic mechanics is:

E. NELSON: *Dynamical Theories of Brownian Motion*, Princeton University Press, Princeton, N. J., 1976.

3. The Bayes Formula and the Unnormalized Conditional Density Equation.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, $\mathcal{F} = (\mathcal{F}_t)$, $0 \le t \le 1$ a nondecreasing family of sub σ-algebras of $\mathcal{F}$ and $(\eta_t, \mathcal{F}_t)$, $0 \le t \le 1$ a Wiener Process. Consider the signal process $z = (z_t, \mathcal{F}_t)$ and the observation equation

$$y_t = \int_0^t z_s\,ds + \eta_t. \tag{3.1}$$

Let

$$z_t = h(x_t), \tag{3.2}$$

and x_t be the solution of the Itô stochastic differential equation

$$dx_t = f(x_t)\,dt + g(x_t)\,dw_t. \tag{3.3}$$

We make the following assumptions:

A1. The functions f and g are continuous and bounded and equation (3.3) has a unique solution in the weak sense for each initial condition x_0 (later we shall have to make further smoothness assumptions).

A2. x_t and η_t are independent and $E\left(\int_0^1 h(x_s)^2\,ds\right) < \infty$.

Under the above hypotheses we can write down a Bayes formula for computing $E(\phi(x_t|\mathcal{F}_t^y)$ where $\mathcal{F}_t^y$ is the σ-field generated by $\{y_s \mid 0 \le s \le t\}$ and $E\left(\int_0^1 \phi(x_s)^2 ds\right) < \infty$. This relies on the Girsanov transformation. It is known that there is a new equivalent measure $\widetilde{P}$ under which y_t and x_t are independent, the distribution of x_t remains invariant and under this new measure y_t is Brownian motion. The Radon-Nikodym Derivative of P with respect to $\widetilde{P}$ is given by

$$\left.\frac{dP}{d\widetilde{P}}\right|_{\mathcal{F}_t^*} = \exp\left(\int_0^t h(x_s)\, dy_s - \frac{1}{2}\int_0^t h^2(x_s)\, ds\right) \triangleq L_t.$$

Hence

$$E(\phi(x_t) \mid \mathcal{F}_t^y) = \frac{\widetilde{E}(\phi(x_t) L_t \mid \mathcal{F}_t^y)}{\widetilde{E}(L_t \mid \mathcal{F}_t^y)}$$

Using the independence, it is convenient to view this conditional expectation as expectation on a product space $(\Omega \times \hat{\Omega}, \mathcal{A} \times \hat{\mathcal{A}}, P \otimes P)$ and write it explicitly as

$$\hat{\phi}(x_t) =$$

$$= \frac{\int_{\hat{\Omega}} \phi(x_t(\hat{\omega})) \exp\left(\int_0^t h(x_s(\hat{\omega}))\, dy_s(\omega) - \frac{1}{2}\int_0^t h(x_s(\hat{\omega}))^2\, ds\right) dP(\hat{\omega})}{\int_{\hat{\Omega}} \exp\left(\int_0^t h(x_s(\hat{\omega}))\, dy_s(\omega) - \frac{1}{2}\int_0^t h(x_s(\hat{\omega}))^2\, ds\right) dP(\hat{\omega})} \tag{3.4}$$

This is the Bayes Formula.

If we replace y_s by x_s for $x \in C(0, 1)$ and define

$$\rho(t, x, \hat{\omega}) = \exp\left(\int_0^t h(x_s(\hat{\omega}))\, dx_s(\omega) - \frac{1}{2}\int_0^t h(x_s(\hat{\omega}))^2\, ds\right) \tag{3.5}$$

$$g(t, x) = \int_{\hat{\Omega}} \rho(t, x, \hat{\omega})\, dP(\hat{\omega}), \tag{3.6}$$

then $\check{g}(t, x)$ is the Radon-Nikodym derivative $\frac{d\mu_y}{d\mu_\eta}(t, x)$, where μ_y is the measure induced on $C(0, 1)$ by the observation and μ_η is the Wiener measure on $C(0, 1)$ induced by η.

Let $\sigma_t(\phi)$ denote the numerator of (3.4). It will be also convenient to write the numerator of (3.4) as a Feynman-Kac formula

$$\sigma_t(\phi) = \int_{\mathcal{X}} \phi(\xi_t) \exp\left(\int_0^t h(\xi_s)\, dy_s - \frac{1}{2}\int_0^t h^2(\xi_s)\, ds\right) d\mu_x(\xi), \tag{3.7}$$

and $\mathcal{X} = C(0, 1)$.

Let P_t be the semigroup of the diffusion process corresponding to (3.3) and $L_0 = \frac{1}{2} g^2(x) \frac{\partial^2}{\partial x^2} + f \frac{\partial}{\partial x}$ be its infinitesimal generator. Then σ_t satisfies the stochastic partial differential equation

$$d\sigma_t(\phi) = \sigma_t(L_0\phi)\, dt + \sigma_t(L_1\phi)\, dy_t, \quad \forall \phi \in \mathcal{D}(L_0) \cap \mathcal{D}(L_1), \tag{3.8}$$

where L_1 is the unbounded multiplication operator h.

This is one of the fundamental results of non-linear filtering. Now using the fact that $h(x_s)$ is a semi-martingale and using the relationship between Itô and Fisk-Stratonovich integrals we get

$$\int_0^t \sigma_s(h\phi)\, dy_s = \int_0^t \sigma_s(h\phi) \cdot dy_s - \frac{1}{2}\langle \sigma(h\phi), y \rangle_t,$$

where $\langle\ ,\ \rangle_t$ denotes the quadratic variation and $\cdot$ denotes the Fisk-Stratonovich integral and hence we obtain

$$d\sigma_t(\phi) = \sigma_t(L_0\phi)\, dt - \frac{1}{2}\sigma_t(h^2\phi) + \sigma_t(h\phi) \cdot dy_t, \tag{3.9}$$

which we write symbolically as

$$\frac{d\sigma_t(\phi)}{dt} = \sigma_t\left[\left(L_0 + \dot{y}_t h - \frac{1}{2} h^2\right)(\phi)\right], \tag{3.10}$$

which we can integrate by the Feynman-Kac formula

$$\sigma_t(\phi) = E_{\sigma_0}\left[\exp\left(\int_0^t \dot{y}_t h\, ds - \frac{1}{2}\int_0^t h^2\, ds\right)\phi(x_t)\right] \tag{3.11}$$

where σ_0 is a measure-valued random variable independent of y such that $E[\sigma_0(\phi)] = \mu(\phi)$ and μ is the initial distribution.

Define

$$\Sigma(t, x_t, y_0^t) = \tilde{E}[L_t \mid \sigma\{x_t\} \times \mathcal{F}_t^y]$$

and define

$$q(t, z, y_0^t) = \Sigma(t, z, y_0^t)\, p(t, z)$$

where $p(t, z)$ is the density of the x-process. Then

$$\sigma_t(\phi) = \int_{\mathbf{R}} \phi(z)\, q(t, z, y_0^t)\, dz.$$

Now $\sigma_t(\phi) = \int \phi\, d\sigma_t$, and hence we see that σ_t has a density $q(t, z, y_0^t)$, and hence from (3.7)

$$q(t, z, y_0^t) = \int_{\mathcal{X}} \exp\left(\int_0^t h(x_s)\, dy_s - \frac{1}{2}\int_0^t h^2(x_s)\, ds\right) d\mu^z(x) \qquad (3.12)$$

where the integration is over the path space of x with $x_0 = 0$ and $x_t = z$ [1]. It can be shown that $q_t \triangleq q(t, z, y_0^t)$ satisfies the stochastic partial differential equation

$$dq_t = \left(L_0^* - \frac{1}{2} L_1^2\right) q_t\, dt + L_1\, q_t \cdot dy_t, \qquad (3.13)$$

where L_0^* is the formal adjoint of L_0.

Equation (3.13) is the Duncan-Mortensen-Zakai equation for the unnormalized conditional density.

3.1. Duncan-Mortensen-Zakai Equation and Gauge Transformations.

The study of non-linear filtering is the study of the Duncan-Mortensen-Zakai equation - its explicit solution and its group invariance properties. Equivalently it is the study of the function space representation of $\sigma_t(\phi)$ given in (3.7) and (3.12). In (3.7) we have the stochastic

(1) This function space integration is to be interpreted as integration against conditional Wiener measure or equivalently the Brownian bridge.

integral $\int_0^t h(\xi_s)\, dy_s$. Under our hypotheses we can write

$$\int_0^t h(\xi_s)\, dy_s = y_t\, h(\xi_t) - y_0(\xi_0) - \int_0^t y_s\, dh(\xi_s).$$

Using the above it is clear that σ_t can be evaluated for all $y \in C(0, 1)$ (not just a subset of full Wiener measure). It has been shown by Clark that $\frac{\sigma_t(\phi, y)}{\sigma_t(1, y)}$ is a version of the conditional expectation $\hat{\phi}$.

An equivalent way of considering this problem is to eliminate the stochastic integral from the Duncan-Mortensen-Zakai equation (3.13).

This can be done by making the observation that the operator $L_1 =$ multiplication by h is a diagonal operator, and its effect can be removed via a time-dependent Gauge transformation. To see this write $q_t = \exp(h(x)\, y_t)\, \tilde{q}_t$. Then a direct calculation shows that $\tilde{q}_t$ satisfies an ordinary partial differential equation (parametrized by y):

$$\frac{d\tilde{q}_t}{dt} = \exp(-h(x)\, y_t)\left(L_0^* - \frac{1}{2} L_1^2\right) \exp(h(x)\, y_t \cdot \tilde{q}_t). \qquad (3.14)$$

It is interesting to rewrite equation (3.14) in a form which brings out the commutation properties of $L_0^* - \frac{1}{2} L_1^2$ and L_1. If we denote by $L_2 = \left[L_0^* - \frac{1}{2} L_1^2, L_1\right]$ and $L_3 = [L_1, L_2]$ considered as formal differential operators and computed on some common invariant domain, (3.13) can be rewritten as:

$$\frac{d\tilde{q}_t}{dt} = \left(L_0^* - \frac{1}{2} L_1^2\right)\tilde{q}_t + y_t\, L_2\, \tilde{q}_t - y_t^2\, L_3\, \tilde{q}_t. \qquad (3.15)$$

Explicitly $L_2 = h_x \frac{d}{dx} + \left(\frac{1}{2} h_{xx} - f h_x\right)$, $L_3 = -h_x^2$ (assuming $g \equiv 1$ for simplicity).

When we are dealing with unbounded observation operators h it is this equation which is the easiest to deal with. The above also shows that the commutators L_2 and L_3 have an important role to play in the understanding of equation (3.14).

In (3.12) we see that evaluating q_t involves an integration over the path space of the x-process. The integration would be simplified if this could be done with respect to Wiener measure. This corresponds to removing the drift term in the operator L_0^*. For simplicity assume $g\equiv 1$. Define the operator « multiplication by $\psi(x)$ », where ψ is invertible and defined from the intertwining relationship.

$\left(\frac{d}{dx}-f\right)\psi=\psi\,\frac{d}{dx}$, which defines $\psi(x)=e^{\int_0^x f(z)\,dz}$. Then the operator L_0^* transforms to $\psi^{-1}L_0^*\,\psi=\frac{1}{2}\frac{d^2}{dx^2}-\frac{1}{2}V(x)$, where $V(x)=(f_x+f^2)$. Then $\hat{q}_t=\psi^{-1}q_t$ is the solution of

$$d\hat{q}_t=\hat{L}_0\,\hat{q}_t\,dt-\frac{1}{2}V(x)\,\hat{q}_t\,dt+h\hat{q}_t\cdot dy_t,\quad \text{where}\quad \hat{L}_0=\frac{1}{2}\frac{d^2}{dx^2}-\frac{1}{2}h^2. \quad (3.16)$$

This is also an example of a Gauge transformation. Equation (3.16) involves the (imaginary time) Schrödinger operator $-\frac{1}{2}\frac{d^2}{dx^2}+\frac{1}{2}h^2+\frac{1}{2}V(x)$. It should be noted that the Lie Algebra of operators $LA\left\{L_0^*-\frac{1}{2}L_1^2, L_1\right\}$ and $LA\{\hat{L}_0-V(x), L_1\}$ are isomorphic. Recall in section 2 we have seen how certain Schrödinger operators are unitarily equivalent to Dirichlet operators.

3.2. Integration of the Mortensen-Zakai Equation and Calculus of Variations.

In section 3 we have seen that solving the Mortensen-Zakai equation is equivalent to evaluating the function-space integral

$$q(t,z,y_0^t)=\int_{\mathcal{X}}\exp\left(\int_0^t h(x_s)\,dy_s-\frac{1}{2}\int_0^t h^2(x_s)\,ds\right)d\mu^z(x).$$

We now show how this can be done by adapting certain ideas of Feynman. Feynman's idea was to separate the classical and quantum parts of the total quantum motion and also to understand the relationship between classical and quantum mechanics. We shall attempt to separate the contribution of the observation in q_t from that of the signal process. For the class of problems considered by Benes this can be done explicitly.

For this purpose, we consider that the x_t-process is governed by $dx_t=f(x_t)\,dt+dw_t$, and the observation equation is $dy_t=x_t\,dt+d\eta_t$. Hence $L_0^*=\frac{1}{2}\frac{d^2}{dx^2}-\frac{d}{dx}f$, and $L_1=$ mult. by x.

We do a gauge transformation and attempt to solve

$$d\hat{q}_t=\hat{L}_0\,\hat{q}_t\,dt-V(x)\,\hat{q}_t\,dt+x\hat{q}_t\,dy_t,\quad \text{where}\quad V(x)=\frac{1}{2}(f_x+f^2)$$

Hence we have to compute:

$$\hat{q}(t,z,y_0^t)=E_{W^z}\left[\exp\left(\int_0^t x_s\,dy_s-\frac{1}{2}\int_0^t x_s^2\,ds-\int_0^t V(x_s)\,ds\right)\right] \quad (3.17)$$

where the integration is over Wiener-measure with $x_0=0$ and $x_t=z$ (1).

We show that if $V(x)=ax^2+bx+c$, $a\geq 0$ and $(a,b,c)\neq 0$ then a sufficient statistic for computing $\hat{q}$ is the set of equations:

$$\begin{cases} d\xi_s=-\eta_s\,ds;\ \xi_0=0,\ \xi_t=z \\ d\eta_s=-\xi_s\,ds-V_\xi(\xi_s)\,ds+dy_s. \end{cases} \ (2) \quad (3.18)$$

We call the above set of equations the deterministic *bi-characteristics* of the Mortensen-Zakai equation. To see this, we introduce the transformation $dx_s=d\xi_s+dz_s$, where ξ_s is given by the above and we do not impose the boundary conditions.

By essentially the Girsanov Theorem,

$$\frac{d\mu_z}{d\mu_x}=\exp\left(\int_0^t\eta_s\,dz_s-\frac{1}{2}\int_0^t\eta_s^2\,ds\right).$$

Now impose the boundary conditions and hence in terms of the z_s-variable

$$\hat{q}(t,z,y_0^t)=\int\exp\left(\int_0^t(z_s+\xi_s)\,dy_s-\frac{1}{2}\int_0^t(z_s+\xi_s)^2\,ds-\int_0^t V(z_s+\xi_s)\,ds\right)$$

$$\times\exp\left(\int_0^t\eta_s\,dz_s-\frac{1}{2}\int_0^t\eta_s^2\,ds\right)d\mu_z^0,$$

(1) See remark on p. 184.

(2) The interpretation is by conditioning.

where the integration is over the path space of z with $z_0=0$ and $z_t=0$.

By applying Itô's rule several times, and using the fact that $V(x)=ax^2+2bx+c$, we obtain

$$\hat{q}(t, z, y_0^t)=\exp\left(\frac{1}{2}\int_0^t \xi_s\, dy_s-\int_0^t b\xi_s\, ds+\frac{1}{2}(z\eta_t)\right)$$

$$\times\int \exp\left(-\int_0^t V(z_s)+\frac{1}{2}z_s^2\right)d\mu_z^0. \tag{3.19}$$

The terms involving the z-process can be explicitly evaluated.

It is worth remarking that (3.18) is the solution of a variational problem. To see this let us first interpret equations (3.18) in the Stratonovich sense. Then equations (3.18) are the necessary conditions of the following optimal control problem:

$$\operatorname*{Min}_u \frac{1}{2}\int_0^t u_s^2\, ds+\frac{1}{2}\int_0^t [(\xi_s-\dot{y}_s)^2+2V(\xi_s)]\, ds$$

$$\dot{\xi}_s=u_s;\ \xi_0=0,\ \xi_t=z.$$

If $y_t\equiv 0$ and $V\equiv 0$, then the equations (3.18) correspond to the imaginary time harmonic oscillator and our methods show the relation between a euclidean quantum harmonic oscillator and an imaginary time classical harmonic oscillator.

These methods could also be applied to study the limiting behavior of the Cauchy problem

$$\frac{\partial\psi}{\partial t}(t, x)=\lambda\, \nabla_x^2\psi+\frac{V(x)}{\lambda}\psi \text{ as } \lambda\to 0$$

with an initial condition $\psi_\lambda(x, 0)=\exp[-S_0(x)/\lambda]$ and show that the limiting solution satisfies the corresponding equations of classical mechanics. This would provide an alternative derivation to the results of Maslov who treated the Schrödinger equation version of this problem.

Equations (3.18) are related to the smoothing problem. It turns out, that the unnormalized conditional density could also be evaluated by the following sufficient statistic.

$$\begin{cases} d\alpha_s=(f(\alpha_s)-\beta_s)\, ds;\ \alpha_0=0\ \ \alpha_t=z \\ d\eta_s=-(f_\alpha(\alpha_s)\,\eta_s+\alpha_s)\, ds+dy_s. \end{cases} \tag{3.20}$$

These are the Euler-Lagrange equations for the following optimal control problem:

$$\operatorname*{Min}_u \frac{1}{2}\int_0^t u_s^2\, ds+\frac{1}{2}\int_0^t \{(\alpha_s-\dot{y}_s)^2+2V(\alpha_s)\}\, ds$$

$$\dot{\alpha}_s=f(\alpha_s)+u_s;\ \alpha_0=0,\ \alpha_t=z.$$

For the general Kalman filtering problem

$$\begin{cases} dx_t=Fx_t\, dt+G\,dw_t \\ dy_t=Hx_t\, dt+d\eta_t, \end{cases} \tag{3.21}$$

the equations corresponding to (3.20) would be

$$\begin{cases} d\alpha_t=F\alpha_t\, dt-GG'\beta_t\, dt;\ \alpha_0=0\ \ \alpha_t=z \\ d\beta_t=-H'H\alpha_t\, dt-F'\beta_t\, dt+H'\, dy_t. \end{cases} \tag{3.22}$$

These can be recognized to be the smoothing equations given in Hamiltonian form. Our methods show that these equations are intrinsically attached to the Mortensen-Zakai equation and play the role of bi-characteristics corresponding to Hamiltonian-Jacobi equations.

Although we do not do it in this paper it seems reasonable to believe that a perturbation theory analogous to Maslov's work could be carried out for non-linear filtering using the framework used in this section.

Notes and References for Section 3.

(i) For the derivation of the Bayes formula and the Mortensen-Zakai equations see:

E. Wong: *Stochastic Processes in Information and Dynamical Systems*, McGraw Hill, New York, 1971, and the references cited there.

(ii) The seminal paper on non-linear filtering is:

M. FUJISAKI, G. KALLIANPUR, H. KUNITA: *Stochastic Differential equations for the Non-linear Filtering Problem*, Osaka J. of Math., Vol. 9, 1972, pp. 19-40.

(iii) Equation (3.14) was first derived by:

J. M. C. CLARK: *The Design of Robust Approximations to the Stochastic Differential Equations of Non-linear Filtering*, in Communication Systems and Random Process Theory: ed. J. K. Skwirzynski, Sithoff and Noordhoff, 1978.

For more recent work see for example:

M. H. A. DAVIS: *On a Multiplicative Functional Transformation Arising in Non-linear Filtering Theory*, Z. Wahrscheinlichkeitstheorie ver. Gebiete, to appear.

Writing it in a form involving the commutators makes it clear that for certain problems these equations can be integrated using group invariance methods.

(iv) The idea of using gauge transformations in the context of non-linear filtering theory is new, although it is implicit in the work of Beneš. See also:

R. W. BROCKETT: *Classification and Equivalence in Estimation Theory*, Proceedings of the IEEE Decision and Control Conference, 1979, Ft. Lauderdale, Florida.

The gauge transformation can be introduced in a much more general setting.

(v) The variational interpretation of certain non-linear filtering problems in the form presented in Section 3.2 is new and uses certain ideas of Feynman. See for example:

R. P. FEYNMAN, A. R. HIBBS: *Quantum Mechanics and Path Integrals*, McGraw Hill, New York, 1965.

For the Girsanov transformation and absolute continuity of measures, see:

R. S. LIPSTER, A. N. SHIRYAYEV: *Statistics of Random Processes I*, Springer-Verlag, New York, 1977.

(vi) The reference to Maslov's work is:

V. P. MASLOV: *Théorie des Perturbations et Méthodes Asymptotiques*, Dunod-Gauthier-Villars, Paris, 1972.

4. Multiple Itô Integrals and Fock Space.

There is a close relationship between the theory of multiple Itô integrals and Wick polynomials. These objects also have an important role to play in Wiener's theory of homogeneous chaos and representations of the Weyl Commutation relations on Fock space. These ideas and constructions also are of importance in non-linear filtering theory.

4.1. MULTIPLE ITÔ INTEGRALS.

The multiple Itô integral of order K is a map $f \mapsto I_K(f)$: $L^2(\mathbf{R}^K) \to$ $\to L^2(\Omega, \mathcal{F}, P)$ having the following properties:

(i) $I_K(h) = \prod_{i=1}^{K} W(A_i)$ if $h = 1_{A_1} \times \dots \times 1_{A_k}$

for disjoint rectangles $A_1, \dots, A_K$.

(ii) $I_K(f+g) = I_K(f) + I_K(g)$.

(iii) If $f_i \to f$ in $L^2(\mathbf{R}^K)$ then $I_K(f_i) \to I_K(f)$ in $L^2(\Omega, \mathcal{F}, P)$.

If $k=0$ and $h \in \mathbf{R}$, define $I_0(h) = h$.

Let $\tilde{f}(t_1, \dots, t_K) = \frac{1}{K!} \sum_{\pi} f(t_{\pi(1)}, \dots, t_{\pi(k)})$ denote the symmetrization of f. The mapping $f \mapsto \tilde{f}$ is the projection of $L^2(\mathbf{R}^K)$ onto the subspace $L_s^2(\mathbf{R}^K)$ of $L^2(\mathbf{R}^K)$ spanned by the symmetric functions.

The multiple Itô integral has the following properties:

(*a*) For $f \in L^2(\mathbf{R}^K)$, $g \in L^2(\mathbf{R}^{K'})$, $I_K(f) = I_K(\tilde{f})$ and

$$E(I_K(f)\, I_{K'}(g)) = 1_{\{K=K'\}}\, K!\, (\tilde{f}, \tilde{g})_{L^2}$$

(*b*) For $\phi \in L^2(\mathbf{R})$ and $\lambda \in \mathbf{C}$,

$$\exp\left(\lambda \int_{\mathbf{R}} \phi_s\, dw_s - \frac{1}{2}\lambda^2 \int_{\mathbf{R}} \phi_s^2\, ds\right) = \sum_{k=0}^{\infty} \frac{\lambda^K}{K!} \int_{\mathbf{R}^K} \phi_{s_1} \dots \phi_{s_K}\, dw_{s_1} \dots dw_{s_K}$$

(*c*) $L^2(\Omega, \mathcal{F}, P) = \bigoplus_{K=0}^{\infty} \{I_K(f) \mid f \in L^2(\mathbf{R}^K)\} \cong \bigoplus_{K=0}^{\infty} L_s^2(\mathbf{R}^K)$.

We now consider two applications of multiple Itô integrals.

4.2. A REPRESENTATION THEOREM FOR THE BEST ESTIMATE OF A SIGNAL.

Consider the non-linear filtering problem

$$y_s = \int_0^t z_s\, ds + \eta_s,$$

ınd assume that the hypotheses of Section 3 (assumptions A 1 and A 2) ıold. Suppose that we want a representation for $\hat{z}_s = E(z_s \mid \mathcal{F}_s^y)$. It s a standard result that the innovations process

$$\nu_t = y_t - \int_0^t \hat{z}_s \, ds.$$

s standard Brownian motion. Recently we have shown $\mathcal{F}_t^y = \mathcal{F}_t^\nu \bmod P$. Hence $\hat{z}_t \in L^2(\Omega, \mathcal{F}_t^\nu, P)$ and hence from property (c)

$$\hat{z}_t = z_0 + \int_0^t k_1(t,s)\, d\nu_s + \int_0^t \int_0^{s_1} k_2(t, s_1, s_2)\, d\nu_{s_1} d\nu_{s_2} + \ldots$$

ınd a standard application of the definition of conditional expectation ıs a projection shows

$$k_1(t,s) = \frac{\partial}{\partial s} E(x_t \nu_s), \quad k_2(t, s_1, s_2) = \frac{\partial^2}{\partial s_1 \partial s_2} E(x_t \nu_{s_1} \nu_{s_2}) \text{ etc.}$$

This representation is not too useful since the computation of he innovations requires computing $\hat{z}_t$. However for the following problem it immediately leads to a finite dimensional filter. Suppose that $_s = Hx_s$ and $dx_s = Fx_s\, ds + G dw_s$. We are required to estimate x_s and $_s = P(x_s, 0 \le s \le t)$ where P is a polynomial functional of x_s with eparable kernels. Then it is not too difficult to show that the estimate s also a polynomial functional with separable kernels of the innovaions $d\nu_s = dy_s - H\hat{x}_s\, ds$ where $\hat{x}_s$ is the Kalman filter estimate.

The ideas of multiple Itô integrals also have applications to reresenting the density $q(t, z, y_0^t)$ as a multiple integral expansion in he semi-martingale y_s.

.2. Random Fields.

In quantum field theory random fields (weak distributions) and heir polynomial functions (suitably defined) have played an important ole. A Wiener integral or more generally an Itô integral is an xample of a random field. Multiple Itô integrals can be considered

to be polynomial functions of random fields, provided these polynomial functions are suitably renormalized. These are the Wick polynomials. Orthogonal polynomials and in particular the Hermite polynomials have a special role to play in this theory.

4.2.1. Basic Definitions and Properties.

Let H be a real Hilbert space, with the scalar product $\langle \cdot, \cdot \rangle$. We shall identify H with its topological dual. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. We denote by $L^0(\Omega, \mathcal{A}, \mu)$ the space of real (complex)-valued random variables on $(\Omega, \mathcal{A}, \mu)$.

Definition 4.1. A continuous (1) linear function

$$F: H \to L^0(\Omega, \mathcal{A}, \mu) \tag{4.1}$$

is termed a *random field* or *weak distribution*. Two random fields F_1 and F_2 are *equivalent* if for any $\{f_1, \ldots, f_n\} \subseteq H$, the joint distribution of $F_1(f_1), \ldots, F_1(f_n)$ and $F_2(f_1), \ldots, F_2(f_n)$ are the same.

An example of a random field which is of importance to us is one which is « generated » by the Wiener process. Let $W(t, \omega)$ denote the Wiener process for $t \ge 0$. Define $W(-t, \omega) = W(t, \omega), t \ge 0$ and thus extend the definition to all of $\mathcal{R}$. Then if we define

$$F: L^2(\mathcal{R}) \to L^0(\Omega, \mathcal{A}, \mu) \tag{4.2}$$

by $F(f) = \int f \, dW(t, \omega)$, where the right hand side is a Wiener integral, then F is a random field.

Definition 4.2. Given a random field $F: H \to L^0(\Omega, \mathcal{A}, \mu)$, the mean functional is the map

$$M_F(f) = E(F(f)) = \int F(f)\, d\mu. \tag{4.3}$$

The *mean functional* need not exist for all $f \in H$. If it exists, then M_F is a linear (not necessarily continuous) functional on H.

(1) Continuous means if $f_i \to f$ in H then $F(f_i) \to F(f)$ in probability.

The *covariance functional* C_F: $H \times H \to \mathbf{C}$ is defined by

$$C_F(f, g) = E\{[F(f) - M_F(f)]\, \overline{[F(g) - M_F(g)]}\} \tag{4.4}$$

where — denotes complex conjugation. If C_F exists for all $f, g \in H$ then C_F is a positive Hermitian bilinear form on H. The *variance functional* V_F is defined by $V_F(f) = C_F(f, f)$ and the *correlation functional* by

$$B_F(f, g) = E[F(f)\, \overline{F(g)}].$$

The random field is said to be *bounded* if there exists a $k > 0$, such that $\sqrt{B_F(f, f)} \leq k \|f\| \quad \forall f \in H$.

The *characteristic functional* ϕ_F: $H \to \mathbf{C}$ is defined by

$$\phi_F(f) = E[e^{iF(f)}] = \int e^{iF(f)}\, d\mu. \tag{4.5}$$

It is a continuous, positive definite functional and $\phi_F(0) = 1$.

REMARK: Bochner's theorem extends to this infinite dimensional situation.

The following proposition is a consequence of the Riesz representation theorem.

PROPOSITION 4.1. *Suppose F is a bounded random field. Then there exists a unique vector $f_F \in H$ (the mean) and unique bounded self-adjoint non-negative operators R_F (covariance) and S_F (correlation) such that $M_F(f) = \langle f, f_F \rangle$, $C_F(f, g) = \langle R_F f, g \rangle$ and $B_F(f, g) = \langle S_F f, g \rangle$.*

DEFINITION 4.3. A random field N such that $B_N(f, f) < \infty$ with characteristic function

$$\phi_F(f) = \exp\left[iM_N(f) - \frac{1}{2} V_N(f) \right] \tag{4.6}$$

is termed a Gaussian random field.

We shall be interested in isonormal random fields N_c with mean zero, variance parameter $c > 0$. Its covariance function is $C_N(f, g) = c\langle f, g \rangle$. If $c = 1$ then we get the unit isonormal random field. Its characteristic function $\phi_N(f) = e^{-\frac{\|f\|^2}{2}}$.

Now let us consider a bounded Gaussian random field N. Its characteristic function can be written as:

$$\phi_N(f) = \exp\left| i\langle f, f_N \rangle - \frac{1}{2} \langle R_N f, f \rangle. \right|.$$

In particular, if the bounded Gaussian random field is the isonormal field with parameter $\sigma > 0$ then

$$R_N = \sigma I.$$

Let $\mathcal{B}(H)$ denote the Borel σ-algebra of H and let μ be a probability measure on $\mathcal{B}(H)$. Define the random field $\mathcal{F}: H \to L^0(H, \mathcal{B}(H), \mu)$ by $[\mathcal{F}(f)](g) = \langle f, g \rangle$. In this way every Borel Probability measure μ on $\mathcal{B}(H)$ produces a random field. Suppose F: $H \to L^0(\Omega, \mathcal{A}, \nu)$ is a random field. Then we say F is generated by μ on $\mathcal{B}(H)$ if F is equivalent to $\mathcal{F}$. We then have

PROPOSITION 4.2. *A Gaussian random field N is generated by a probability measure on $(H, \mathcal{B}(H))$ if and only if the covariance operator R_N is trace class.*

Thus if $\dim H = \infty$, the isonormal random field with covariance operator $\sigma I, \sigma > 0$ could never be produced by a countably-additive probability measure on $(H, \mathcal{B}(H))$. By the Gross construction of « lifting », the finitely additive measure corresponding to the isonormal field can be « extended » to a countably additive measure on a separable Banach space.

$S(\mathcal{R}^n)$ will denote the Schwartz space of rapidly decreasing functions, and $S'(\mathcal{R}^n)$ its dual, the space of tempered distributions. In the sequel it will be often convenient to define a random field as a mapping F: $V \to L^0(\Omega, \mathcal{A}, \mu)$ where V is a topological space. In particular we shall have occasion to take $V = S(\mathcal{R})$. We shall often make the assumption that the random field is *determined* by $f \in \mathcal{V}$ in the sense that the smallest σ-algebra w. r. to which the random variables $\{F(f) \mid f \in V\}$ is measurable is $\mathcal{A}$. We call such a random field *full*.

Construction of the Unit Gaussian Random Field.

Let H be a separable Hilbert Space and let (e_n) be an orthonormal basis for H. Let $\dot{\mathcal{R}}$ be the one point compactification of $\mathcal{R}$ and let $\Omega = \Pi \dot{\mathcal{R}}$ be the Cartesian countable product of copies of $\mathcal{R}$. Ω is a Compact Hausdorff space in the Tychonov topology. Let $C(\Omega)$ be the set of continuous functions on Ω and let $P(\Omega)$ be the set of functions $\phi(x_1, \dots, x_n)$ in $C(\Omega)$ which depend only on a finite number of copies. With the supremum norm topology $C(\Omega)$ is a Banach space and (by the Stone-Weierstrass theorem) $P(\Omega)$ is dense in $C(\Omega)$. For $\phi \in P(\Omega)$ define

$$k(\phi) = (2\pi)^{-\frac{n}{2}} \int \phi(x_1, \dots, x_n)\, e^{-\frac{\|x\|^2}{2}\mathcal{R}^n}\, dx_1, \dots, dx_n.$$

Then $k: P(\Omega) \to \mathcal{R}$ is linear and $|k(\phi)| \le \|\phi\|$ and hence k can be extended to a continuous linear functional on $C(\Omega)$. By the Riesz Representation theorem there exists a Borel probability measure μ on $\mathcal{B}(C)$ such that $k(\phi) = \int \phi\, d\mu$, $\forall \phi \in C(\Omega)$.

Let $F(e_n)$: $\Omega \to \bar{\mathcal{R}}$ be multiplication by x_n, $n = 1, 2, \dots$ Then $F(e_n)$ is a. e. in $\mathcal{R}$ and hence $F(e_n)$: $\Omega \to \mathcal{R}$, measurable. It can be shown that $F(e_n) \in L^2(\Omega, \mathcal{B}(C), \mu)$ and $\{F(e_n) \mid n = 1, 2, \dots\}$ generate $\mathcal{B}(C)$. If $f \in H$ and $f = \Sigma\, a_n e_n$ then $\Sigma\, a_n F(e_n)$ converges in $L^2(\Omega, \mathcal{B}(C), \mu)$ to an element $F(f)$. This is the *unit full Gaussian* Random field on H.

4.2.2. Random Fields on $S(\mathcal{R}^n)$ and White Noise.

Let F: $S(\mathcal{R}^n) \to L^0(\Omega, \mathcal{A}, P)$ be a random field. The derivative F' of the random field is defined by

$$F'(f) = -F(f')$$

F' is a random field and if F is Gaussian, F' is also Gaussian. Now a random field F: $S(\mathcal{R}^n) \to L^0(\Omega, \mathcal{A}, \mu)$ is generated by a stochastic process (assumed to have square integrable sample paths) $X: \mathcal{R}^n \times \Omega \to \mathcal{R}$ if

$$[F(f)](\omega) = \int f(x)\, X(x, \omega)\, dx, \quad \forall f \in S(\mathcal{R}^n).$$

Let $W(t, \omega)$, $t \in [0, \infty]$ be the standard Wiener process and let F be the Gaussian random field generated by $W(t, \omega)$. The covariance functional of F can be computed as

$$C_F(f, g) = \int_0^\infty \int_0^\infty \min(s, t)\, f(s)\, g(t)\, ds\, dt =$$

$$= \int_0^\infty [\hat{f}(s) - \hat{f}(\infty)]\, [\hat{g}(t) - \hat{g}(\infty)]\, ds\, dt$$

where

$$\hat{f}(t) = \int_0^t f(s)\, ds \text{ and } \hat{g}(t) = \int_0^t g(s)\, ds.$$

F' the derivative of F exists (in the sense of distributions) and its covariance functional is given by

$$C_{F'}(f, g) = C_F(f', g') = \int_0^\infty f(t)\, g(t)\, dt = \langle f, g \rangle$$

where $\langle\ ,\ \rangle$ represents the natural inner product on $S(\mathcal{R})$ embedded in $L^2(\mathcal{R})$. F' is *white noise*.

4.3. Second Quantization (After Segal and Nelson).

Let H be a real Hilbert space and let F: $H \to L^0(\Omega, \mathcal{A}, \mu)$ be the *unit Gaussian full* random field. If $f_1, \dots, f_n$ are orthonormal in H and ϕ is a Bounded Baire function on $\mathcal{R}^n$, then

$$\int_\Omega \phi(F(f_1), \dots, F(f_n)) = \frac{1}{(2\pi)^{n/2}} \int_{\mathcal{R}^n} \phi(x)\, e^{-\frac{\|x\|^2}{2}}\, dx$$

For concreteness $(\Omega, \mathcal{A}, \mu)$ may be chosen to be countably infinite copies of $(\mathcal{R}, \mathcal{B}(\mathcal{R}), (2\pi)^{-1/2} e^{-x^2/2} dx)$.

If E denotes expectation on $(\Omega, \mathcal{A}, \mu)$ then

$$E(F(f_1) \dots F(f_{2n+1})) = 0 \tag{4.3.1}$$

$$E(F(f_1) \dots F(f_{2n})) = \Sigma \langle f_{i_1}, f_{j_1} \rangle \dots \langle f_{i_n}, f_{j_n} \rangle \tag{4.3.2}$$

where the sum of over all pairings of $1, \dots, 2n$ i. e. $i_1 < \dots < i_n$; $i_1 < j_1, \dots, i_n < j_n$, and $(i_1, j_1, \dots, i_n, j_n)$ is a permutation $1, \dots, 2n$.

$L^p(\Omega, \mathcal{A}, \mu)$ is denoted by $L^p(H)$ and $\Gamma(H)$ denotes $L^2(H)$. Let $\Gamma(H)_{<n}$ be the closed linear span in $\Gamma(H)$ of all elements of the form $F(f_1) \dots F(f_m)$ with $m \le n$ and let $\Gamma(H)_n$ denote the orthogonal complement of $\Gamma(H)_{\le n-1}$ in $\Gamma(H)_{\le n}$. For $f_1, \dots, f_n$ in H we define the Wick

polynomial

$$: F(f_1) \dots F(f_n):$$

to be the orthogonal projection of $F(f_1) \dots F(f_n)$ into $\Gamma(H)_n$. In the special case, where H is one dimensional and hence $\Gamma(H) = L^2(\mathcal{R}, \mathcal{B}(\mathcal{R}), (2\pi)^{-1/2} e^{-x^2/2} dx)$, $\Gamma(H)_n$ is the one dimensional sub-space spanned by the nth Hermite polynomial and $:x^n:$ is the nth Hermite polynomial normalized so that the leading coefficient is 1.

We have the formula

$$\langle :F(f_1) \dots F(f_n):, :F(g_1) \dots F(g_n): \rangle = \\ = \sum_{\pi} \langle f_{\pi(1)}, g_1 \rangle \dots \langle f_{\pi(n)}, g_n \rangle. \tag{4.3.3}$$

where the sum is over all permutations π of $1, \dots, n$. If all the f's and g's are equal, we get

$$\langle :F(f)^n:, F(f)^n \rangle = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} (:x^n:)^2 e^{-x^2/2} dx = n! \tag{4.3.4}$$

Let H_1 be the complexification of H and let H_n denote the n-fold Hilbert space symmetric tensor product of H_1 with itself. On H_n we define the inner product such that

$$\langle \mathrm{Sym}(f_1 \otimes \dots \otimes f_n), \mathrm{Sym}(g_1 \otimes \dots \otimes g_n) \rangle = \\ \sum_{\pi} \langle f_{\pi(1)}, g_1 \rangle \dots \langle f_{\pi(n)}, g_n \rangle \tag{4.3.5}$$

where

$$\mathrm{Sym}(f_1 \otimes \dots \otimes f_n) = \frac{1}{n!} \sum_{\pi} f_{\pi(1)} \otimes \dots \otimes f_{\pi(n)}. \tag{4.3.6}$$

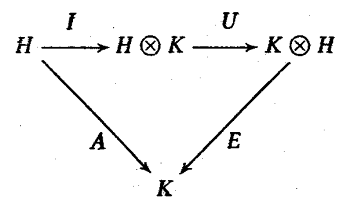
From (4.3.3) and (4.3.5), we see that the mapping $:F(f_1) \dots F(f_n): \mapsto \mathrm{Sym}(f_1 \otimes \dots \otimes f_n)$ extends uniquely to a unitary operator from $\Gamma(H)_n$ onto H_n. We use this mapping to identify $\Gamma(H_n)$ and H_n. Analogous to the fact that the one-dimensional Hermite polynomials span $L^2(\mathcal{R}, \mathcal{B}(\mathcal{R}), (2\pi)^{-1/2} e^{-x^2/2} dx)$, Segal proved

$$\Gamma(H) = \sum_{n=0}^{\infty} H_n, \text{ for arbitrary real Hilbert space } H. \tag{4.3.7}$$

$\Gamma(H)$ *is Fock Space.*

If the random field $F(f) = \int f\, dW$, where $f \in L^2(\mathcal{R}) = H$ and W is the standard Wiener process, the elements of $\Gamma(H)_n$ are multiple Wiener integrals (in the sense of Itô).

The space $\Gamma(H)$ is intrinsically attached to the structure of H as a real Hilbert space. Thus if $U: H \to K$ is an orthogonal mapping of one real Hilbert space into another, it induces a unitary mapping $\Gamma(U): \Gamma(H) \to \Gamma(K)$, where on H_n, $\Gamma(U) = \underbrace{U \otimes \dots \otimes U}_{n\text{-fold}}$. Similarly if $I: H \to K$ is an isometric embedding then it induces an isometric embedding $\Gamma(I): \Gamma(H) \to \Gamma(K)$ and similarly for an orthogonal projection $E: H \to K$. If $A: H \to K$ is a contraction then $\Gamma(A): \Gamma(H) \to \Gamma(K)$ is defined to be the direct sum of $\Gamma(A)_n: H_n \to K_n$, where $\Gamma(A)_n = \underbrace{A \otimes \dots \otimes A}_{n\text{-fold}}$. Now any contraction $A: H \to K$ can be decomposed as



where I, U and E are as above.

Hence $\Gamma(A) = \Gamma(E)\Gamma(U)\Gamma(I)$. Now $\Gamma(A)$ is *doubly Markovian* in the sense that

$$\begin{cases} \alpha \geq 0 \Rightarrow \Gamma(A)\alpha \geq 0 \\ \Gamma(A) 1 = 1 \\ E\Gamma(A)\alpha = E\alpha, \end{cases} \tag{4.3.8}$$

Any doubly Markovian operator is a contraction from L^p to L^p.

It turns out that $\Gamma(A)$ has stronger contractive properties and the precise statement of this is an important theorem of Nelson.

THEOREM 4.3.1. (*Nelson Hypercontractivity Theorem*). *Let* $A: H \to K$ *be a contraction. Then* $\Gamma(A)$ *is a contraction from* $L^q(H) \to L^p(K)$ *for* $1 \leq q \leq p \leq \infty$ *provided that*

$$\|A\| \leq \sqrt{\frac{q-1}{p-1}} \tag{4.3.9}$$

If (4.3.9) *does not hold then* $\Gamma(A)$ *is not a bounded operator from* $L^q(H) \to L^p(K)$.

4.4. ABSOLUTE CONTINUITY OF GAUSSIAN RANDOM FIELDS.

Let ϕ be a Gaussian unit random field $H \to L^2(\Omega, \mathcal{A}, \mu) = \Gamma(H)$. We consider the following related questions:

(1) Given $A: H \to H$ linear, continuous when is there a unitary map $U: \Gamma(H) \to \Gamma(H)$ so that

$$U\phi(f)U^{-1} = \phi(Af) \quad \forall f \in H?$$

(2) Given A as in (1) when is there an F in $\Gamma(H)$ so that each $\phi(f)$ is a Gaussian random variable with respect to $|F|^2\, d\mu$ but with variance $\frac{1}{2}\|Af\|^2$ instead of $\frac{1}{2}\|f\|^2$?

(3) Given two Gaussian random fields with general covariance on the same Hilbert space H when can they be realized on a single measure space but with two mutually absolutely continuous measures?

The above problems are essentially equivalent and we present an answer to (3) under the assumption that A is positive and has a bounded inverse.

THEOREM 4.4.1. (*Feldman, Segal, Shale*). *Let* $\phi: H \to L^2(\Omega, \mathcal{A}, \mu)$ *be a Gaussian unit random field and let* $A: H \to H$ *be linear, bounded, positive with bounded inverse. Then a necessary and sufficient condition for the Gaussian random field on* H *with variance* $\frac{1}{2}\|Af\|^2$ *to be realizable on* Ω *with measure* ν *equivalent to* μ *is that* $A-I$ *be a Hilbert Schmidt Operator.*

4.5. DISCUSSION.

Fock space has an important role to play in the study of the free Quantum field which can be considered to be an infinite assembly of non-interacting harmonic oscillators.

Consider the harmonic oscillator hamiltonian $H = -\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2$ as a self-adjoint operator on $L^2(\mathbf{R})$. If instead we work with $L^2(\mathbf{R}, g)$, where g is Gaussian measure then the action of H on the Hermite polynomials h_n is given by $h_n \mapsto n\, h_n$. From this it is clear that the one-parameter group e^{itH} generated by H acts as $h_n \mapsto e^{int} h_n$. If we denote by W the unitary map $f(x) \mapsto f(x)\left[\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right]^{-1/2}: L^2(\mathbf{R}) \to L^2(\mathbf{R})$ then the operators $p = i\frac{d}{dx}$ and $q = x$ on $L^2(\mathbf{R})$ transform to $p' = WpW^{-1}$ and $q' = WqW^{-1}$ and $\frac{1}{2}(p'^2 + q'^2)$ can be identified as the Harmonic Oscillator hamiltonian. By the ideas of second quantization described in section 4.3, we can extend the one-dimensional case to an infinite dimensional setting. When this is done this gives rise to the particle representation of the free field and the concept of a number operator.

Shale's theorem on the equivalence of Gaussian Random fields essentially shows that the Kalman filter is analogous to the free quantum field. This can be done by noting that the observations and innovations processes are related by $\nu = (I-K)y$ where K is a Hilbert-Schmidt operator.

NOTES AND REFERENCES FOR SECTION 4.

(i) The basic reference for multiple Itô integrals is:

K. ITÔ: *Multiple Wiener Integrals*, J. Math. Soc. Japan, 13, 1951, pp. 157-169.

(ii) References for Section 4.2 are:

D. ALLINGER, S. K. MITTER: *New Results on the Innovations Problem*, Technical Report R-964, Laboratory for Information and Decision Systems, M. I. T., January 1980. To appear in Stochastics.

S. K. MITTER, D. OCONE: *Multiple Integral Expansions for Non-linear Filtering*, presented at the 18th IEEE CDC, Ft. Lauderdale, Florida, December 1979.

D. OCONE, forthcoming Ph. D. Thesis, M. I. T. June 1980.

S. I. MARCUS, S. K. MITTER, D. OCONE: *Finite Dimensional Non-linear Estimation for a class of Continuous and Discrete Time Problems*, Proceedings of the International Conference on Analysis and Optimization of Stochastic Systems, Academic Press, 1980.

For recent results, see:

E. WONG: *Explicit Solutions to a Class of Nonlinear Filtering Problems*, to appear.

(iii) The exposition on Random fields closely follows:

B. SIMON: *The $P(\Phi_2)$-Euclidean (Quantum) Field Theory*, Princeton University Press, Princeton, N. J., 1975.

(iv) Second Quantization as presented here is due to:

E. NELSON: *The Free Markov Field*, J. Functional Analysis, 12, 1973, pp. 211-227 where the basic Hypercontractivity Theorem is proved.

(v) The idea of studying the infinite dimensional harmonic oscillator on Fock space is due to:

I. E. SEGAL: *Tensor Algebras over Hilbert Spaces*, I, Trans. Am. Math. Soc. 81, 1956, pp. 106-134.

(vi) Viewing the Kalman Filter as the Free Quantum Field is due to the author:

S. K. MITTER: *Filtering Theory and Quantum Fields*, to appear in Asterisque, 1980.

5. Some Lie Algebras of Interest in Filtering Theory.

Let us recall that in section 3 we showed that the unnormalized conditional density of x_t given the observations $\{y_s \mid 0 \le s \le t\}$ satisfies a stochastic partial differential equation

$$dq(t, x, \omega) = L_0^* q(t, x, \omega) + L_1 q(t, x, \omega)\, dy_t \tag{5.1}$$

where L_0^* is the formal adjoint of the generator of the diffusion process corresponding to the stochastic differential

$$dx_t = f(x_t)\, dt + g(x_t)\, dw_t, \tag{5.2}$$

and hence L_0^* is given by

$$L_0^* = \frac{1}{2}\frac{d^2}{dx^2}g^2 - \frac{d}{dx}f, \tag{5.3}$$

and

$$L_1 = \text{Multiplication by } h(x). \tag{5.4}$$

The Lie Algebra of operators $LA\left\{L_0^* - \frac{1}{2}L_1^2, L_1\right\}$ generated by $L_0^* - \frac{1}{2}L_1^2$ and L_1 has an important role to play in filtering theory. This comes about by writing equation (5.2) in Stratonovich form (a fact which can be justified rigorously under the hypotheses we have adopted)

$$\frac{\partial q}{\partial t}(t, x) = \left(L_0^* - \frac{1}{2}L_1^2 + \dot{y}_t L_1\right) q(t, x). \tag{5.1$'$}$$

This suggestive form of writing is meant to show the anology of the filtering problem with the interaction picture of quantum physics.

It turns out that a number of filtering problems give rise to Lie Algebras of operators which are representations of known algebras which arise in mathematical physics. We first summarise these known facts about some Lie Algebras arising in mathematical physics.

5.1. THE HEISENBERG ALGEBRA AND THE WEYL ALGEBRA.

Let $n \ge 0$, and let $(p_1, \dots, p_n, q_1, \dots, q_n, z)$ be a basis for a real vector space V. On this space we can define the structure of a Lie Algebra by defining $[p_i, q_i] = -[q_i, p_i] = z$, the other brackets between elements of the basis being zero. Let us denote this Lie Algebra by $\mathcal{H}$. The centre of $\mathcal{H}$ is $\mathbf{R}z$ and we have $[\mathcal{H}, \mathcal{H}] = \mathbf{R}z$ and hence this Lie algebra is nilpotent.

If $\mathcal{H}$ is a Lie algebra with centre c and if $c = [\mathcal{H}, \mathcal{H}]$ and $\dim c = 1$, then $\mathcal{H}$ is a Heisenberg algebra.

PROPOSITION 5.1. *Let $\mathcal{H}$ be a nilpotent algebra. Suppose that all characteristic ideals* (1) *of $\mathcal{H}$ are of dimension ≤ 1. Then either $\mathcal{H}$ is zero or a Heisenberg algebra.*

We denote the algebra defined by $2n$-generators $p_1, q_1, \dots, p_n, q_n$

(1) Let g be a Lie algebra. A derivation of g is a linear mapping. $D: g \to g$ such that $D([x, y]) = [Dx, y] + [x, Dy]$ $\forall x, y \in g$. A characteristic ideal of g is a subspace which is stable for all derivations of g.

and the bracket relations

$$[p_i, q_i] = 1.$$

$$[p_i, q_j] = [p_i, p_j] = [q_i, q_j] = 0 \text{ for } i \neq j$$

by $A_n(\mathbf{R})$ or simply A_n. These are the so-called Weyl algebras.

Elements of the form $p_1^{i_1}, q_1^{j_1}, \dots, p_n^{i_n}, q_n^{j_n} (i_1, j_1, \dots, i_n, j_n \in N)$ generate the vector space A_n.

In the vector space $E = \mathbf{R}[X_1, \dots, X_n]$ let P_i be the endomorphism $\frac{\partial}{\partial X_i}$ and Q_i the endomorphism of multiplication by X_i. We have

$$[P_i, Q_i] = 1.$$

$$[P_i, Q_j] = [P_i, P_j] = [Q_i, Q_j] = 0 \text{ for } i \neq j$$

and hence there exists a homomorphism $\rho: A_n \to \mathrm{End}(E)$ such that $\rho(p_i) = P_i$ and $\rho(q_i) = Q_i$ for all i. The elements $P_1^{i_1} Q_1^{j_1}, \dots, P_n^{i_n} Q_n^{j_n}$ are linearly independent and therefore the elements $p_1^{i_1} q_1^{j_1}, \dots, p_n^{i_n} q_n^{j_n}$ form a basis for the vector space A_n and ρ is injective. It follows that $A_n \cong A_1 \otimes, \dots, \otimes A_1$ (n-copies). The representation ρ of A_n in E is termed the standard representation. Finally E is a simple A_n-module and the set of A_n-endomorphisms of E is $\mathbf{R}$.

Let B_m be the set of linear combinations of

$$p_1^{i_1} q_1^{j_1}, \dots, p_n^{i_n} q_n^{j_n} \in A_n \text{ such that } i_1 + j_1 + \dots + i_n + j_n \leq m.$$

Then $B_m B_{m'} \subset B_{m+m'}$.

Consequently the graded algebra associated with A_n equipped with the filtration $(B_0, B_1, \dots)$ is the polynomial algebra in $2n$-variables.

PROPOSITION 5.2.

(i) *A_n is integral and noetherian.*

(ii) *The centre of A_n is $\mathbf{R}$.*

(iii) *The algebra A_n is simple.*

PROPOSITION 5.3. *Let $p_1, q_1, \dots, p_n, q_n$ be the canonical generators of A_n. Let us define vector subspaces S, T of A_n as follows:*

(i) $S = \sum_{i=1}^{n} \{(rp_i + rq_i) \mid r \in \mathbf{R}\}.$

(ii) $T = \sum_{1 \leq i, j \leq n} \left\{ \frac{r}{2}(p_i q_j + q_j p_i) + r p_i p_j + r q_i q_j \mid r \in \mathbf{R} \right\}.$

Then:

(i) *$\mathbf{R} \oplus S$ and T are Lie sub-algebras of A_n*

(ii) *$\mathbf{R} \oplus S$ is a Heisenberg algebra.*

The relationship between Heisenberg and Weyl algebras is given by

PROPOSITION 5.4.

$$A_n(\mathbf{R}) \simeq U(\mathcal{H})/J$$

where $U(\mathcal{H})$ is the enveloping algebra of the $(2n+1)$-dimensional Heisenberg Algebra $\mathcal{H}$ and J is a two-sided ideal in $U(\mathcal{H})$.

5.2. THE OSCILLATOR ALGEBRA.

The real Lie algebra with $2n+2$ generators $(h, p_1, \dots, p_n, q_1, \dots, q_n, z)$ satisfying the bracket relations

$$[h, p_i] = q_i, \; [h, q_i] = p_i, \; [p_i, q_i] = z,$$

and the other brackets being zero is defined to be the oscillator algebra. This is a solvable Lie algebra. This algebra is the semi-direct sum of the one-dimensional algebra spanned by h and the Heisenberg algebra $\mathcal{H}$.

5.3. THE POISSON BRACKET ALGEBRA.

Consider $\mathbf{R}^{2n} = \{(q, p) \mid q = (q_1, \dots, q_n), p = (p_1, \dots, p_n)\}$ with its standard symplectic structure (see later section). For $f, g \in C^\infty(\mathbf{R}^{2n})$, the Poisson bracket is given by

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i} \cdot \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \cdot \frac{\partial g}{\partial q_i} \right).$$

The real vector space of polynomials in (q, p) denoted by $\mathcal{P}_n$ is a Lie algebra under the Poisson bracket operation.

PROPOSITION 5.5. *$\mathcal{P}_n$ is generated by the two polynomials q_1 and $h = p^2 + q_1^4$ if $n = 1$ and q_1 and $h = \sum_{i=1}^{n} (p_i^2 + q_i^4) + \sum_{i=1}^{n-1} q_i q_{i+1}$ if $n > 1$.*

That is $\mathcal{P}_n$ coincides with the smallest Lie subalgebra Q_n containing q_1 and h.

PROPOSITION 5.6. *$\mathcal{P}_n/\mathbf{R}$ is simple.*

5.4. EXAMPLES OF NON-LINEAR FILTERING PROBLEMS AND THEIR LIE ALGEBRAS.

EXAMPLE 1. Consider the non-linear filtering problem

$$\begin{cases} dx_t = f(x_t)\,dt + dw_t \\ dy_t = h(x_t)\,dt + d\eta_t \end{cases} \tag{5.5}$$

and let us assume that $f, h \in C^\infty(\mathbf{R})$. In the light of section 3, the Mortensen-Zakai equation is

$$dq_t = \left(L_0^* - \frac{1}{2} L_1^2\right) q_t\,dt + L_1\,q_t \cdot dy_t$$

where

$$L_0^* = \frac{1}{2}\frac{d^2}{dx^2} - \frac{d}{dx} f$$

$$L_1 = \text{multiplication by } h(x).$$

We then have

PROPOSITION 5.7. $LA\left\{L_0^* - \frac{1}{2} L_1^2, L_1\right\}$ *is finite dimensional only in the case*

(i) $h = \alpha x + \beta$.

(ii) $f_x + f^2 = ax^2 + bx + c$,

where the Lie algebra of operators is computed on the common domain $C^\infty(\mathbf{R})$ or $\mathcal{S}(\mathbf{R})$.

In the above if we want the diffusion process to be defined globally on $\mathbf{R}$, then for the Riccati equation $f_x + f^2 = ax^2 + bx + c$ to have a global solution we need $a \geq 0$, and $(a, b, c) \neq 0$.

The example covered by Proposition 5.7 has recently been considered by Benes. From Section 3.1, we can see that we can remove the effect of the drift by a Gauge transformation. From (3.15) we see that it is enough to solve

$$d\tilde{q}_t = \hat{L}_0\,\tilde{q}_t\,dt - V(x)\,\tilde{q}_t\,dt + h\tilde{q}_t \cdot dy_t.$$

But $LA\left\{L_0^* - \frac{1}{2} L_1^2, L_1\right\} \simeq LA\{\hat{L}_0 - V(x), L_1\}$, and under the hypotheses of Proposition 5.7 the second Lie algebra has a basis consisting of the elements

$$\hat{L}_0 - \frac{1}{2} L_1^2, L_1, \left[\hat{L}_0 - \frac{1}{2} L_1^2, L_1\right] = L_2 = \alpha \frac{d}{dx}, \quad [L_1, L_2] = \alpha I.$$

This however is essentially the Harmonic Oscillator algebra and corresponds (essentially) to the filtering problem

$$\begin{cases} x_t = w_t \\ dy_t = x_t\,dt + d\eta_t \end{cases}$$

which is a Kalman filtering problem.

It should also be remarked that there is no difficulty in extending this example to the multi-dimensional situation

$$dx_t = f(x_t)\,dt + dw_t$$

$$dy_t = Hx_t\,dt + d\eta_t,$$

provided $f: R^n \to R^n$ is $\nabla F = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n}\right)'$ and $H: R^n \to R^p$. This fact is also clear from our considerations relating Schrödinger and Dirichlet operators. Finally, this problem is the analogue of the imaginary time Harmonic Oscillator problem with an external force.

EXAMPLE 2. Consider the non-linear filtering problem

$$\begin{aligned} x_t &= w_t \\ dy_t &= x_t^3\,dt + d\eta_t. \end{aligned} \tag{5.6}$$

This is the so-called cubic sensor problem. From Proposition 5.7, $LA\left\{\frac{1}{2}\frac{d^2}{dx^2} - \frac{1}{2}x^6, x^3\right\}$ is infinite dimensional. From calculations similar to that required to establish Proposition 5.7 it can be shown that the Lie algebra is isomorphic to the Weyl algebra A_1 which is simple.

EXAMPLE 3. Consider the model given by (5.5) and suppose that $f_x+f^2=V(x)$ has a global solution where $V(x)$ say is an even positive polynomial (other than the quadratic). Then from Proposition 5.7, $LA\left\{L_0^*-\frac{1}{2}L_1^2, L_1\right\}$ is infinite dimensional. Even more, these Lie algebras (modulo their centre which is R) are simple. This fact follows from calculations similar to that in Proposition 5.5.

EXAMPLE 4. Let us consider the model of (5.5) and let $f, h\in C^\infty(U)$ where U is some open set in R. Then by restricting $L_0^*-\frac{1}{2}L_1^2$ and L_1 to $C^\infty(U)$ the Lie algebra $LA\left\{L_0^*-\frac{1}{2}L_1^2, L_1\right\}$ is finite dimensional only in the (prototype) case

(i) $h=x^2$

(ii) $f_x+f^2=-\frac{1}{4}x^4$.

Note that the last equation (Riccati) has finite escape time.

Suppose we consider the slightly more general model

$$dx_t=f(x_t)\,dt+g(x_t)\,dw_t, \text{ where } f, g\in C^\infty(U)$$

U open in R and $g(x)>0$, $x\in U$. In this case we can remove the diffusion term g by a non-linear smooth change of coordinates.

Let $x_t=\alpha(z_t)$ where α is smooth and invertible. Then by the Itô differential rule

$$dx_t=\alpha_z\,dz_t+\frac{1}{2}\alpha_{zz}\,dt=f(\alpha(z_t))\,dt+g(\alpha(z_t))\,dw_t.$$

Hence

$$dz_t=(\alpha_z)^{-1}(f(\alpha(z_t))\,dt-\frac{1}{2}\alpha_{zz}(z_t))\,dt+(\alpha_z)^{-1}g(\alpha(z_t))\,dw_t.$$

Let α satisfy the differential equation

$$\frac{d\alpha}{dz}=g(\alpha(z)),$$

and let I be the maximal interval in R on which the solution is defined. Then

$$dz_t=(g(\alpha(z_t))^{-1}f(\alpha(z_t)-\frac{1}{2}g_\alpha(z_t))\,dt+dw_t=\tilde{f}(z_t)\,dt+dw_t.$$

The observation equation is

$$dy_t=\tilde{h}(z_t)\,dt+d\eta_t \text{ where } \tilde{h}=h(\alpha(z_t)).$$

Now it is easy to see that

$$LA\left\{\frac{1}{2}\frac{d^2}{dx^2}g^2-\frac{1}{2}\frac{d}{dx}f-\frac{1}{2}h^2, h\right\}\simeq LA\left\{\frac{1}{2}\frac{d^2}{dx^2}-\frac{d}{dx}\tilde{f}-\frac{1}{2}\tilde{h}^2, \tilde{h}\right\}$$

where the first Lie algebra is computed on the common invariant domain $C^\infty(U)$ and the second on the common invariant domain $C^\infty(I)$.

5.5. SIGNIFICANCE OF THE RESULTS.

If the Lie Algebra $LA\left\{L_0^*-\frac{1}{2}L_1^2, L_1\right\}$ is finite dimensional, then it should be possible to construct the filter by integrating the Lie algebra. We consider this aspect of the problem in the next section. If the Lie algebra is infinite dimensional and simple then we conjecture that it cannot be represented by a Lie algebra of vector fields with analytic coefficients on a finite-dimensional manifold. That is, we cannot represent the solution $q(t, z, y_0^t)$ of the Mortensen-Zakai equation by means of a finite dimensional sufficient statistic described by a vector differential equation

$$d\alpha_t=a(\alpha_t)\,dt+b(\alpha_t)\,dy_t. \tag{5.7}$$

Suppose we are interested only in computing unnormalized conditional statistics $\hat{\phi}(x_t)$. If we denote this by $\hat{\phi}_t$, then

$$\hat{\phi}_t=\int_R \phi(z)\,q(t, z, y_0^t)\,dz.$$

It may be possible to represent $\hat{\phi}_t$ as the output of a vector differential system:

$$\begin{cases} d\alpha_t = a(\alpha_t)\,dt + b(\alpha_t)\,dy_t \\ \phi_t = c(\alpha_t). \end{cases} \tag{5.8}$$

A theorem of Brockett says that for this to be possible there must exist a homomorphism between the Lie algebra of operators $LA\left\{L_0^* - \frac{1}{2}L_1^2, L_1\right\}$ and the Lie algebra of vector fields $a - \frac{1}{2}b_x b$ and b such that under the homomorphism $L_0^* - \frac{1}{2}L_1^2 \mapsto a - \frac{1}{2}b_x b$ and $L_1 \mapsto b$. This would suggest that the ideal structure of the Lie Algebra $\mathcal{L} = LA\left\{L_0^* - \frac{1}{2}L_1^2, L_1\right\}$ should be important. If there exist non-trivial ideals I then a candidate for such a homomorphism would be $\psi: \mathcal{L} \to \mathcal{L}/I$. The results of this section show that for a large class of problems this line of attack will not be possible, since the Lie algebra $\mathcal{L}$ is infinite dimensional and simple.

A possible approach might be to represent $\hat{\phi}_t$ as the output of a delay system, that is look for a representation of the form

$$\begin{cases} d\alpha_t = a(\alpha_t, \alpha_{t-\theta})\,dt + b(\alpha_t, \alpha_{t-\theta})\,dy_t \\ \hat{\phi}_t = \int_{-\theta}^{0} c(\alpha_t, \alpha_{t-\theta})\,d\theta. \end{cases}$$

There is some reason to believe that it might be possible to do this for the cubic sensor problem considered in Example 2.

Notes and References for Section 5.

(i) The idea of studying the Lie Algebra of operators involved in the Mortensen-Zakai equation is independently due to R. W. Brockett and myself. I was motivated by the analogy between the Kalman filter and the Free Quantum field and I discussed these ideas in a series of lectures in the University of Maryland in December 1977.

In this connection see:

R. W. Brockett: *Remarks on Finite Dimensional Non-linear Estimation*, presented in the Conference on Algebraic and Geometric Methods in System Theory, Bordeaux, France, September 1978, to appear in Asterisque, 1980.

S. K. Mitter: *Filtering Theory and Quantum Fields*, presented in the Conference on Algebraic and Geometric Methods in System Theory, Bordeaux, France, September 1978, to appear in Asterisque, 1980

and the references cited in the above two papers.

(ii) For the material on Heisenberg and Weyl Algebras see:

J. Dixmier: *Algèbres Enveloppantes*, Gauthier-Villars, Paris, 1974.

Propositions 5.5 and 5.6 on the Poisson Bracket Algebra are due to:

A. Avez, A. Heslot: *L'algebre de Lie des Polynômes en les Coordonnées Canoniques munie de Crochet de Poisson*, C. R. Acad. Sc. Paris t. 288 Serie *A*, Mai 1979, pp. 831-833.

(iii) Proposition 5.7 is due to Daniel Ocone and is related to non-linear filtering problems first introduced by V. Beneš. See D. Ocone: Doctoral Dissertation, Mathematics Department, M. I. T., June 1980 and V. Beneš: Exact finite dimensional filters for certain diffusions with non-linear drift: to appear in Stochastics.

(iv) The content of Example 2 and Example 3 are new and due to Daniel Ocone and the author. The content of Example 4 is due to Daniel Ocone (forthcoming Ph. D. dissertation) and R. W. Brockett: loc. cit, paper in Decision and Control Conference 1979.

(v) The réference to the theorem of Brockett cited in section 5.5 is contained in R. W. Brockett: loc. cit.

(vi) Recent work (unpublished) by John Baras as well as M. Hazewinkel and S. Marcus shows that the bilinear filtering has a nice ideal structure and many filterations.

6. Representation of the Filter.

The complete solution of the non-linear filtering problem requires computing the propagator of the Mortensen-Zakai equation or equivalently the unnormalized conditional transition density $q(t, z, y_0^t \mid s, x)$. We have discussed several approaches to computing the density $q(t, z, y_0^t)$. For the class of problems considered by Benes the propagator could be computed by solving the two-point boundary value problem

$$d\xi_\tau = -\eta_\tau\,d\tau;\ \xi_s = x,\ \xi_t = z$$

$$d\eta_\tau = -\xi_\tau\,d\tau - V_\xi(\xi_\tau)\,d\tau + dy_\tau$$

according to the development of section 3.2. This is the approach of functional integration.

The Lie-algebra viewpoint to computing the propagator would require that the Lie algebra be integrable. This is a difficult question as we can see by discussing the commutation relationship of quantum mechanics.

6.1. Canonical Commutation Relations and their Unitary Representations.

Consider a massive spinless, nonrelativistic particle. Its configuration space is typically $L^2(\mathbf{R})$. The position operator $q: L^2(\mathbf{R}) \to L^2(\mathbf{R})$ is defined as the operator with domain

$$\begin{cases} \mathcal{D}(q) = \{f \in L^2(\mathbf{R}) \mid xf(x) \in L^2(\mathbf{R})\} \text{ and} \\ (qf)(x) = x f(x). \end{cases} \quad (6.1)$$

The momentum operator $p: L^2(\mathbf{R}) \to L^2(\mathbf{R})$ has domain

$$\begin{cases} \mathcal{D}(p) = \{f \in L^2(\mathbf{R}) \mid f' \in L^2(\mathbf{R})\} \text{ and} \\ pf(x) = i\dfrac{df}{dx}. \end{cases} \quad (6.2)$$

The Schwartz space $\mathcal{S}(\mathbf{R}) \subset \mathcal{D}(q) \cap \mathcal{D}(p)$, $\mathcal{S}(\mathbf{R})$ is dense in $L^2(\mathbf{R})$ and is left invariant under q and p. Moreover on $\mathcal{S}(\mathbf{R})$ they satisfy the *Heisenberg Commutation Relations*

$$qp - pq = iI. \quad (6.3)$$

Since p is self-adjoint, the operator $U(a) = e^{-iap}$, $a \in \mathbf{R}$ is unitary and the operators $\{U(a) \mid a \in \mathbf{R}\}$ form a one-parameter strongly continuous unitary group. Now $\mathcal{D}(q)$ is invariant under $U(a)$, $a \in \mathbf{R}$ and it can be shown that

$$U(a)\, qU(-a) = q - aI \text{ on } \mathcal{D}(q). \quad (6.4)$$

This is the *Schrödinger form* of the Canonical Commutation Relations.

Since q is self-adjoint it generates a strongly continuous one-parameter unitary group $V(b)$ given by

$$V(b)\, f(x) = e^{-ibx} f(x). \quad (6.5)$$

It can be shown that

$$U(a)\, V(b) = e^{iab}\, V(b)\, U(a). \quad (6.6)$$

This is done by first checking it on $\mathcal{S}(\mathbf{R})$ and then extending by continuity to all of $L^2(\mathbf{R})$.

(6.6) represents the *Weyl Form* of the Commutation Relations.

It is a well known fact that the Heisenberg form of the canonical commutation relations do not give rise to a unitary group representation, while the Schrödinger and Weyl forms do. This gives rise to the following general question: Suppose we have a representation of a Lie algebra g by skew-symmetric operators defined on a common invariant domain $\mathcal{D}$ in a Hilbert space $\mathcal{H}$ and let $\mathcal{G}$ be the simply connected Lie group with Lie algebra g. When does the representation of g come from a unitary representation $\mathcal{G}$?

The answer to this question is connected with questions of essential self-adjointness and the existence of a common dense set of analytic vectors. We discuss some of these questions now.

Let $\mathcal{G}$ be a Lie group and $\mathcal{X}$ a Banach space. A *representation* T of $\mathcal{G}$ is a mapping $\mathcal{G} \to \mathcal{L}(\mathcal{X})$: $\alpha \mapsto T(\alpha)$, where $\mathcal{L}(\mathcal{X})$ is the set of bounded operators such that $T(e) = I$, e being the identity element and $T(\alpha_1 \cdot \alpha_2) = T(\alpha_1)\, T(\alpha_2)$, for all $\alpha_1, \alpha_2 \in \mathcal{G}$ and $\forall x \in \mathcal{X}$, $\alpha \mapsto T(\alpha)\, x$ is continuous (with $\mathcal{X}$ the norm topology). The representation is called *unitary* if $\mathcal{X}$ is a Hilbert space and each $T(\alpha)$ is a unitary operator.

A vector $x \in \mathcal{X}$ is an *analytic vector* for T in case the mapping $\alpha \mapsto T(\alpha)\, x$: $\mathcal{G} \to \mathcal{X}$ is analytic.

The salient facts connecting representations and analytic vectors are:

Theorem 6.1: *Let T be a representation of a Lie group $\mathcal{G}$ on a Banach space $\mathcal{X}$. Then T has a dense set of analytic vectors in $\mathcal{X}$.*

The answer to the question raised earlier in the section is contained in the following theorem and corollary due to Nelson:

Theorem 6.2: *Let g be a Lie algebra of skew symmetric operators on a Hilbert space H having a common invariant domain $\mathcal{D}$. Let $X_1, \ldots, X_d$ be a basis for g, $\Delta = X_1^2 + \ldots + X_d^2$. If Δ is essentially self-adjoint, then there is on H a unique unitary representation U of the simply connected Lie Group $\mathcal{G}$ having g as its Lie algebra such that for all X in g, $\overline{U(X)} = \bar{X}$ (bar denotes closure of an operator).*

Corollary 6.3: *Let g be a real Lie Algebra, H a Hilbert space. For each X in g let $\rho(X)$ be a skew-symmetric operator on H. Let $\mathcal{D}$ be a dense linear subspace of H such that for all X, Y in g, $\mathcal{D}$ is contained in the domain of $\rho(X)\,\rho(Y)$. Suppose that for all X, Y in g, x in $\mathcal{D}$,* *and real numbers a and b,*

$$\rho(aX+bY)\,x=a\rho(X)\,x+b\rho(Y)\,x$$

$$\rho([X,Y])\,x=(\rho(X)\,\rho(Y)-\rho(Y)\,\rho(X))\,x.$$

Let $X_1, \dots, X_d$ be a basis for g. If the restriction A of $\rho(X_1)^2+\dots+\rho(X_d)^2$ to $\mathcal{D}$ is essentially self-adjoint, then there is on H a unique unitary representation U of the simply connected Lie group $\mathcal{G}$ having g as its Lie algebra such that for all X in g, $\overline{U(X)}=\overline{\rho(X)}$.

Corollary 6.4: *Let g be a real Lie algebra with a basis $X_1, \dots, X_d$, $\mathcal{G}$ the simply connected Lie group with Lie algebra g, H a Hilbert space, $\mathcal{C}$ a dense linear subspace of H. Let ρ be a representation of g by skew-symmetric operators with domain $\mathcal{C}$. Then there is a unitary representation U of $\mathcal{G}$ such that $\mathcal{C}$ is the space of infinitely differentiable vectors for U and $U(X)=\rho(X)$ for all X in g if and only if*

$$A=\rho(X_1)^2+\dots+\rho(X_d)^2$$

is essentially self-adjoint and $\mathcal{C}=\bigcap_{n=1}^{\infty}\mathcal{D}(\bar{A}^n)$.

How do these ideas relate to the representation of the filter? Firstly, the equation we are dealing with is a stochastic parabolic equation valid for $t\geq 0$. Hence the operators L_0^* and L_1 will in general only generate semi-groups. Consider the Kalman filtering problem (or any problem Gauge equivalent to it). Then what is necessary is to give a precise meaning to the time-ordered operator product

$$e^{t_1 L_0^*}\, e^{t_2 L_1}\, e^{t_3 L_2}\, e^{t_4},\quad t_i\geq 0$$

as an evolution operator (in the sense of Kato). For the Kalman filtering this can be done using special methods. There are examples, like the estimation of a Bessel process in additive white noise where this appears not possible to do. To see the connection to unitary representations of Lie groups, it might be best to complexify and try to check the conditions of Nelson's theorem. We conjecture that if the Lie algebra representation does not extend to a unitary group representation then we shall not be able to give meaning to the time-ordered operator product considered above.

Finally, the most direct way appears to be to try to integrate equation (3.15) of Section 3

$$\frac{d\tilde{q}_t}{dt}=\left(L_0-\frac{1}{2}L_1^2\right)\tilde{q}_t+y_t\,L_2\,\tilde{q}_t-y_t^2\,L_3\,\tilde{q}_t$$

but the question of integrating the Lie algebra also appears here.

Notes and References for Section 6.

(i) For a detailed account of group representation theory relevant to this section see:

A. O. Barut, R. Raczka: *Theory of Group Representations,* PWN-Polish Scientific Publishers, Warsaw 1977.

The exposition given here closely follows:

E. Nelson: *Analytic Vectors,* Annals of Mathematics, **70**, 1959, pp. 572-615.

(ii) The details of treating the Benes and similar problems using Group Invariance and Lie Algebraic ideas will appear in a joint paper by J. Baras, S. K. Mitter, D. Ocone. See also D. Ocone, forthcoming Doctoral Dissertation, M. I. T., June 1980, for estimation problems for diffusions with boundary.

Acknowledgments.

Much of the work reported in this paper is joint work with Daniel Ocone of the Mathematics Department, M. I. T. and John Baras of the University of Maryland. The germinal ideas were discussed in a set of lectures given at the University of Maryland in December 1977. The material of the paper was discussed in detail in a one-semester seminar at M. I. T. in Fall 1979. John Baras and Eugene Wong of the University of California, Berkeley, were important participants of this seminar. I want to thank them for their cooperation. I have been profoundly influenced by the seminars of Irving Segal especially his course: Physics from a Fundamental Mathematical Standpoint offered in 1976-1977. The idea of using Gauge Transformations (and their significance) originated in conversations with M. S. Narasimham of the Tata Institute of Fundamental Research. Conversations with V. Benes of Bell Telephone Laboratories, R. W. Brockett, Harvard University, M. H. A. Davis, Imperial College (while visiting M. I. T. in Fall 1978), R. Hermann, M. I. T., and S. I. Marcus of the University of Texas, Austin, have also influenced my thinking.

NOTE ADDED IN PROOF.

(i) For a more precise development of Section 3.2 see the author's lectures given at the NATO Advanced Study Institute on Stochastic Systems, Les Arcs, France, July 1980, where the notion of stochastic bi-characteristics of the Zakai equation is developed.

(ii) The conjecture in Section 5.5 has been proved for the cubic sensor problem by M. Hazewinkel and S. Marcus.

Gubbio - Tipografia Oderisi Editrice - 1981